Cognitive Science 29 (2005) 681–718

Deontic Reasoning With Emotional Content: Evolutionary Psychology or Decision Theory?

Nick Perham, Mike Oaksford

School of Psychology, Cardiff University, Wales

Received 9 January 2004; received in revised form 22 April 2005; accepted 2 May 2005

Abstract

Three experiments investigated the contrasting predictions of the evolutionary and decision-theoretic approaches to deontic reasoning. Two experiments embedded a hazard management (HM) rule in a social contract scenario that should lead to competition between innate modules. A 3rd experiment used a pure HM task. Threatening material was also introduced into the antecedent, *p,* of a deontic rule, if *p* then must *q.* According to the evolutionary approach, more HM responses (Cosmides & Tooby, 2000) are predicted when *p* is threatening, whereas decision theory predicts fewer. All 3 experiments were consistent with decision theory. Other theories are discussed, and it is concluded that they cannot account for the behavior observed in these experiments.

Keywords: Deontic reasoning; Evolution; Central cognitive processes; Decision theory

1. Introduction

Deontic reasoning, that is, reasoning about what one should and should not do, is central to human societies in regulating people's behavior to achieve mutual goals. It permeates most areas of human conduct. For example, if avoiding infectious diseases is the goal, then it is a good idea to enforce rules such as, *if you want to enter the country, then you must have an inoculation against cholera.* Moreover, if the goal is to avoid unnecessary work injuries, then it is a good idea to enforce rules such as, *if you feel pain, then you must stop working for a break.* Since the pioneering work on deontic reasoning (Cheng & Holyoak, 1985, 1989; Cosmides, 1989; Manktelow & Over, 1987, 1990, 1991), this area of reasoning research has become increasingly important. This is due, in no small measure, to the strong claims made by Cosmides and her colleagues that deontic reasoning may be under the control of various innate, domain-specific, reasoning modules (Cosmides, 1989; Cosmides & Tooby, 1989, 1992, 2000; Fiddick, Cosmides, & Tooby, 2000). This research has therefore come to provide some of the

Requests for reprints should be sent to Mike Oaksford, now at School of Psychology, Birkbeck College London, Malet Street, London, WC1E 7HX, UK. E-mail: mike.oaksford@bbk.ac.uk

principal evidence for an evolutionary psychology approach to high-level cognitive function. Many of these functions, or "central processes" (Fodor, 1983), it is argued are under the control of such evolved cognitive modules.

However, there are other approaches to deontic reasoning. In pragmatic reasoning schema theory, people have various domain-specific learned schemata for reasoning in deontic and in other domains (Cheng & Holyoak, 1985, 1989; Cheng, Holyoak, Nisbett, & Oliver, 1986; Holyoak & Cheng, 1995a, 1995b, 1995c). Decision-theoretic approaches argue that deontic reasoning involves a subjective expected utility calculation, where the utilities may be provided by learning or by innate preferences (Chater & Oaksford, 1996; Kirby, 1994a, 1994b; Manktelow & Over, 1987, 1990, 1991; Manktelow, Sutherland, & Over, 1995; Oaksford & Chater, 1994, 1995a, 1995b). Mental models theory argues that deontic reasoning involves the identification of the relevant deontic possibilities in just the same way as standard conditional reasoning may involve identifying the relevant logical possibilities (Johnson-Laird & Byrne, 1992, 1995, 2002).

Moreover, some researchers have argued that the whole evidential basis for specific (or general) deontic reasoning mechanisms is illusory (Fodor, 2000, 2002; Girotto, Kemmelmeier, Sperber, & van der Henst, 2001; Sperber, Cara, & Girotto, 1995; Sperber & Girotto, 2002). Despite recent exceptions (Fiddick, 2004), the evidence for specific deontic reasoning mechanisms derives almost completely from research using the Wason selection task (Wason, 1966, 1968). However, Sperber and Girotto argued that this task does not require participants to reason, that is, to draw conclusions from premises. All it requires is for participants to form a linguistic interpretation of the conditional rule. Consequently, the selection task cannot inform researchers about any subsequent reasoning processes that can only be invoked once the rules have been interpreted. Sperber and Girotto (2002) provided interesting evidence for this position that we look at later.

The purpose of the research we report here was to set up a critical test of the evolutionary and the decision-theoretic approaches. The way we achieved this was to investigate what happens when putative domain-specific reasoning modules must compete because the stimulus materials provide information that could potentially activate more than one module. Moreover, recently, Cosmides and Tooby (2000) argued that the emotions play a crucial role in activating appropriate innate modules. In particular, they suggested that fear-provoking or threatening stimuli may selectively activate an innate reasoning module responsible for managing hazardous situations. We therefore also varied the threat values associated with the antecedent (*p*) of a conditional rule, *if* p *then* q. In the work injury rule we mentioned previously, "pain" is threatening. Consequently, one might expect it to evoke a fear reaction and so lead to activation of a module responsible for reasoning in this domain. Certainly one would expect less activation of this module if the antecedent of the work injury rule read, *if you have worked for 90 minutes, you must stop working for a break.* Although we focus on the contrasting predictions of the evolutionary approach and the decision-theoretic approach, we discuss how other approaches could address our results in the General Discussion.

We first introduce the evolutionary approach where we outline the specific tasks used in these experiments. We then go on to discuss the predictions made by the decision-theoretic approach. All the experiments used the deontic version of Wason's (1966, 1968) selection task. For example, participants might be told that their job is to enforce the regulation that *if some-*

one is allowed to enter the country (p) then he or she must be inoculated against cholera (q) (Cheng & Holyoak, 1985). They are presented with four cards and told that each card represents a person, with information about whether he or she was allowed to enter on one side, and their inoculations on the other. Four cards would be presented showing, for example, enter (*p* card), denied (*not-p* card), inoculated against cholera (*q* card), and not inoculated against cholera (*not-q* card); (in the actual experiment, for the *q* cards participants saw a list of inoculations including cholera [*q* card] or not including cholera [*not-q* card]).[1] Participants' task is to indicate which cards must be turned over to determine if the person is cheating or violating the rule. Participants predominantly select the *p* and the *not-q* cards, which is the correct selection to identify cheaters, but is also the logical response. This contrasts with standard abstract versions of this task where as few as 4% of participants make this response (Johnson-Laird & Wason, 1970).

2. The evolutionary approach

Two innate reasoning modules have been explicitly investigated. One concerns reasoning about social contracts (SCs), and the other concerns reasoning about hazard management (HM).

2.1. Social contracts

SC algorithms have putatively evolved to deal with situations involving social exchange. These are situations where "to be entitled to receive a benefit from another individual or group, an individual is obligated to satisfy a requirement of some kind" (Fiddick et al., 2000, p. 13). SCs embodying this relation have the form, "*If you accept the benefit, then you must satisfy the requirement,* or equivalently, *if you satisfy the requirement, then you are entitled to take the benefit*" (Fiddick et al., 2000, p. 13). It would appear that people are sensitive to cheaters, that is, people who take the benefit without paying the cost, rather than to logical violations of the conditional rules used to express SCs. As we saw for the immigration rule, a cheater, that is, a *p, not-q* instance, corresponds to the logical response. However, these SC rules can be reversed: *If you have an inoculation against cholera (p), then you may be allowed to enter the country (q)*. With respect to this rule a cheater is a *not-p, q* instance—that is, still someone who enters the country without being inoculated against cholera. Cosmides (1989) showed that people do make the *not-p, q* response for such a "reversed SC." This switch in the pattern of card selections would not appear to be predictable from a purely logical interpretation of the conditional. Moreover, a similar switch can be achieved by another manipulation.

Manktelow and Over (1991) showed that the perspective participants adopt can also affect responses. When acting as the enforcer of a regulation, participants are concerned to detect cheaters who are trying to take the benefit without meeting the requirement. However, someone trying to enter the country will have different concerns. They will feel cheated if although they have an inoculation against cholera they are still denied entry into the country. From this new perspective, with respect to the original regulation, people will pick the card showing denied (*not-p*) and the card with cholera on it (*q*). With respect to the switched SC, they will pick the card showing denied (*not-q*) and the card with cholera on it (*p*). Notice that in both perspectives people are looking for the cases where someone is being cheated. However, what counts as a cheater and who is doing the cheating, switches between perspectives.

2.2. Hazard management

Evolutionary psychologists have also proposed that people possess an evolved cognitive module for HM (Cosmides & Tooby, 1997). This module has, "its own distinct architecture, representational format, and licensed inferential procedures" (Fiddick et al., 2000, p. 15). HM algorithms include precaution rules such as, "If a valued entity is subjected to a hazard, then taking an appropriate precaution lowers the risk of harm" (Fiddick et al., 2000, p. 17). An example precaution rule was mentioned in the first paragraph: *If you feel pain, then you must stop working for a break.* Workers are valued entities who are at hazard when suffering a work injury, and taking the precaution of making them stop work if they feel pain lowers their risk of harm. Such rules reliably lead to the selection of cards corresponding to people who are at risk, that is, those who are subject to a hazard but do not take the precaution (Manktelow & Over, 1991). For example, a worker who is in pain (*p*) but who does not take a break (*not-q*) violates the previously mentioned precaution rule.

2.3. Social contracts about HM

The experiments we report investigated what happens if the input conditions for an SC algorithm and an HM algorithm are satisfied at the same time. Creating SCs about HM situations is quite straightforward because these situations are ubiquitous in society. As we argued in the section on HM, the works injury rule clearly satisfies the conditions for a precaution rule. This situation is also embedded in an SC. The workers are obliged only to take breaks when indicated by the HM rule and not to take breaks when they are not in pain. So their situation is also defined by an SC (Fiddick et al., 2000). For example, the workers are cheating with respect to this rule if they take a break although they are in not in pain.

Suppose that participants are told to adopt the perspective of someone checking to ensure that the workers are not cheating. The SC algorithm indicates that they should select people who are taking a break (*q*) although they are not in pain (*not-p*). However, if the precaution rule activates the HM algorithm, then participants should check people at risk, that is, people who are in pain (*p*) but who do not take a break (*not-q*). This means that this task version could lead to either of the two main response patterns. Fiddick et al., (2000) also created similar scenarios, but only ever observed a response consistent with one module or another but not both. As we shall see, whether this is a prediction of the evolutionary account depends on the control mechanisms that must choose between modules that are not specified in the evolutionary approach. Moreover, Fiddick et al. (2000) did not vary emotional content in the rules.

2.4. Emotions and the fear module

Cosmides and Tooby (2000) also considered the role of the emotions in reasoning, which could decide what module is activated. They argued that, "instead of emotions activating or depressing 'thinking' in general, the specific emotion program activated should selectively activate appropriate specialized inferential systems" (Cosmides & Tooby, 2000, pp. 16–17). In particular, they are "conducting research to see whether, as predicted, fear influences precautionary reasoning." It has recently been proposed that many of the most basic phobic reactions are evolutionarily preprogrammed and that threatening stimuli automatically activate an evolved fear module (Mineka & Oehman, 2002; Oehman & Mineka, 2001). Some of the most basic phobic reactions are to animals, such as spiders and snakes, and to blood (Sarlo, Palomba, Angrilli, & Stegagno, 2002).

Although the fear reaction to these stimuli may be part of everyone's genetic inheritance, it does not mean that everyone goes on to develop debilitating phobias. Nonetheless, according to Oehman and Mineka (2001), these stimuli will activate the fear module even in "normal" people. Moreover, there is good reason to believe that this reaction should generalize to threatening words. For example, recent research shows that threatening words produce attention disengagement deficits even in nonanxious normal participants (Fox, Russo, Bowles, & Dutton, 2001). Moreover, a similar attentional bias to threat words is associated with lack of positive affect, again in normal participants (Lundh, Wikstroem, & Westerlund, 2001). More relevantly, in human reasoning, the inclusion of emotional words in standard conditional reasoning tasks significantly affects performance (Blanchette & Richards, 2004) in a similar way to being in the same emotional state (Oaksford, Morris, Grainger, & Williams, 1996). Consequently, precaution rules such as our work injury rule, which contains the threatening word *pain,* should activate the fear module. This should in turn selectively activate the appropriate specialized inferential system, that is, the HM algorithm. Activating this algorithm tells people what precautions they or others should take to avoid further harm.

One might argue that these considerations are only relevant if someone is personally at risk. The supervisor who is checking the cards and whose perspective participants are asked to adopt, is not personally at risk. So, even if a fear reaction is evoked by the threatening word, why should his (and by assumption the participants') HM module be activated? Peoples' reactions to horror films seem to show that activating appropriate avoidance responses, even if one is not personally at risk, may be close to being obligatory, at least when young. The urge to run away, even though not actually at risk, is very strong, and it may take many years to learn to overcome this reaction. Moreover, in the context of performing this task, the only HM rule in operation is the work injury rule. Thus, it is contextually appropriate for this rule to be activated. Finally, although the supervisor's primary task is to detect cheating, he should still be concerned that his workforce does not sustain injuries. Thus there is no reason to suppose that responses according to this secondary goal will be completely inhibited.

In the experiments we report in this article, we also included rules that were not as threatening; for example, *if you work for 90 minutes, then you must stop working for a break*. It is implausible to suggest that people have a phobic reaction to working for 90 minutes. One might initially be skeptical that such a small manipulation—the only difference between conditions is that these phrases ("are in pain" and "work for 90 minutes") were switched—could be predicted to produce observable effects. However, familiarity with the cognition and affect literature soon dispels any such skepticism. For example, the simple positive mood manipulation of giving participants a small gift (a voucher for a free burger) can lead to highly significant effects on decision making (e.g., Isen & Patrick, 1987).

2.5. Predictions

The evolutionary approach is computationally underspecified. There is no specification of the control mechanisms that must choose between evolved modules when, as in these experiments, more than one module could be active. To make clear predictions, we therefore consider three possible modes of controlling these cognitive processes. These are shown in Fig. 1. There are a few preliminary points to make before considering each mode of control. First, according to the evolutionary approach, general reasoning (GR) mechanisms should be preempted (Fiddick et al., 2000). This is shown in the diagram by the links connecting to the SC and HM modules before the GR mechanism. Second, according to most accounts of the control of cognition (e.g., Anderson, 1983; Minsky, 1977), the closer the match to the input conditions of a production or innate module, the more activated it should be. In our Experiments 1 and 2, the precaution rule is embedded in a quite complex scenario indicating how participants are to interpret the task situation as an SC. Consequently, the context should provide a more complete match to the SC algorithm, which will therefore be more strongly activated. This is shown in Fig. 1 by the thicker arrow going into the SC module. Third, according to the evolutionary approach, the more threatening rule should lead to greater activation of the fear module. This should in turn lead to greater activation of the HM algorithm. This is shown in Fig. 1 by the link from the input to the fear module (F) that then feeds in to the HM module.

The three possible control mechanisms we considered were winner takes all (WTA), processing tree model (PTM; Batchelder & Riefer, 1999), and an activation-level model (ACT). We consider each in turn. In a WTA mode of operation, if a module is active it attempts to switch off other modules that are competing to respond. This is shown in Fig. 1 by the mutually

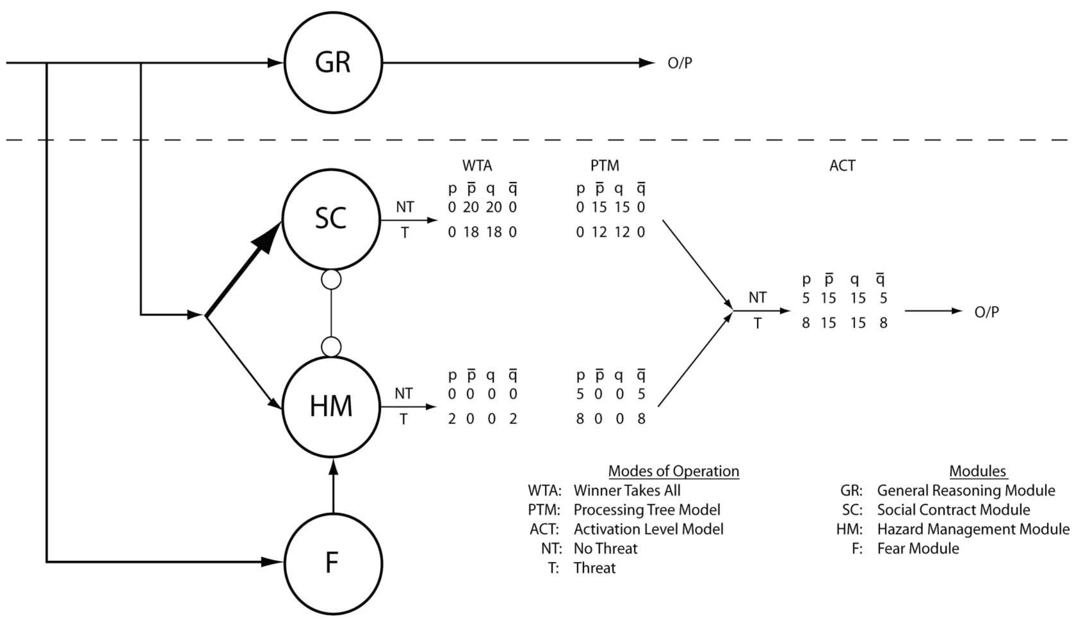

Fig. 1. Modes of operation for the evolutionary approach.

inhibitory link between the SC and HM modules. In this mode of operation, because the SC module has the best match, it is being activated more strongly and so should win the competition and hence switch off the HM module. Consequently, participants should make only SC, *not-p, q* responses. However, if the input from the fear module were sufficiently strong, then in the threat condition some participants may switch to the HM, *p, not-q,* response. For the WTA and PTM modes, the numbers following each card corresponds to the number of participants out of 20 that would be predicted to turn each card for the threat and no-threat conditions.

In a PTM the links to each module specify the probability that a certain module will be activated. This mode of operation involves ignoring the inhibitory link between the HM and SC modules. Ignoring the fear module for the moment, this option makes predictions at the group and individual level. Taken at the group level, a certain proportion of people should make HM responses and a certain proportion should make SC responses. But there will be more making the SC response. Taken at the individual level, that is, one person having several attempts at the task or related tasks, someone would be expected to make HM and SC selections in the same proportions as the group. The input from the fear module should increase the probability that the HM module will be activated and so decrease the probability that the SC module will be activated. This would predict more HM responses and fewer SC responses for threatening material.

In the ACT mode, each module activates representations of the corresponding cards dependent on its level of activation. Consequently, *not-p* and *q* will be activated quite strongly by the SC module, whereas *p* and *not-q* will be activated by the HM module. On this mode of operation (ACT), the numbers in Fig. 1 must be interpreted as activation levels not as the frequency of participants selecting a card. Threat will increase the activation of the HM module, thus increasing the activation of *p* and *not-q* without causing a reduction in the activation of *not-p* and *q*. On this view, there should be more HM selections for threatening words but no changes in SC selections. Moreover, card selection will depend on where individuals place their criterion activation level. If someone's criterion is quite low then they may select all the cards in both the threat and no-threat conditions. However, because of the higher activations in the threat condition, more selections of all the cards should occur in this condition than in a no-threat condition.

There may be other possible modes of operation, but those we have looked at seem to exhaust the most obvious. Over these three, some consistent predictions emerge. First, the SC response, *not-p* and *q,* will predominate according to all modes of operation. Second, according to the WTA and the PTMs, there should be fewer SC responses and more HM responses, *p* and *not-q,* in the threat condition. Third, in the ACT mode, however, there should be fewer SC selections in the threat condition, whereas HM selections remain constant.

This account of the control processes involved in deploying innate modules implies that only three responses are possible, the SC (*not-p* and *q*), HM (*p* and *not-q*) or the all-cards response. However, it is reasonable to suppose that other possible control mechanisms could introduce noise or error into the decision to select cards. In this task people may make errors of commission or errors of omission. So for a task where people should select *not-p* and *q* (SC), it is an error of commission to select the inappropriate *p* or *not-q* cards and an error of omission not to select *not-p* or *q*. To capture these considerations, we used a version of the Pollard index (Pollard & Evans, 1987) as the dependent variable in these experiments (see also, Evans, 1983;

Oaksford et al., 1996; Oaksford & Stenning, 1992).[2] These indexes are computed by adding 1 for selecting the two cards predicted to be selected and by deducting 1 for selecting the two cards predicted not be selected. So in these experiments, we added 1 for selecting the *not-p* and *q* cards and deducted 1 for selecting the *p* or *not-q* cards, that is, (*not-p* + *q*) – (*p* + *not-q*). For example, someone selecting the *p, not-p,* and *q* cards would score 1 [(1 + 1) – (1 + 0)]. This index ranges between –2 and +2, where –2 indicates a pure HM response and +2 indicates a pure SC response. The main prediction of the evolutionary account is that this index should be lower in the threat condition than in the no-threat condition because more HM responses should lead to lower scores.[3]

3. Decision-theoretic approaches

According to the decision-theoretic approach, people's judgments in the deontic selection task are based on calculating the expected utility associated with each card (Kirby, 1994a, 1994b; Manktelow & Over, 1987, 1990, 1991, 1995; Oaksford & Chater, 1994, 1995a, 1995b). Instructions indicating that people are to detect cheaters do not activate domain-specific, evolved modules but tell people to attach a high utility to instances of people who cheat.

3.1. The decision-theoretic model

Most of the research into decision-theoretic accounts of the deontic selection task has employed the concepts of decision theory in a qualitative way (see Manktelow & Over, 1991; Manktelow et al., 1995; for an exception, see Kirby, 1994b). However, formal accounts have been proposed by Oaksford and Chater (1994) and more recently by Klauer (1999). The core of these accounts is the contingency table in (A) and the payoff matrix in (B):

$$\text{(A)}\quad \begin{array}{ccc} & q & \neg q \\ p & a(1-e) & ae \\ \neg p & b-a(1-e) & 1-b-ae \end{array} \qquad \text{(B)}\quad \begin{array}{ccc} & q & \neg q \\ p & C_A & B_H \\ \neg p & B_S & C_D \end{array}$$

The contingency table (A) shows the joint probability distribution for the condition (*p:* pain/working for 90 min) and the action (*q:* taking a break); *a* is the probability of pain/working for 90 min, $P(p)$; *b* is the probability of taking a break, $P(q)$; and *e* is the conditional probability of not taking a break given the worker has a pain/has worked for 90 min, *P*(*not-q*|*p*) (Hattori, 2002; Oaksford, 2002; Oaksford & Chater, 2003a, 2003b, 2003c, 2003d, 2003e; Oaksford, Chater, & Larkin, 2000; Oaksford & Wakefield, 2003).

In the payoff matrix (B), B_H is the benefit associated with detecting someone at risk, that is, someone in pain who has not taken a break (*p, not-q*). B_S is the benefit associated with detecting cheating, that is, someone who is not in pain who takes a break (*not-p, q*). Costs can arise for two reasons. First, the task instructions make it clear that there is a cost associated with turning any card and Klauer (1999) suggested manipulating this cost in a standard selection task. Second, costs can arise for failing to find a cheat or someone at risk. Kirby (1994b) manipulated these costs by, for example, telling participants, who were asked to act as if they were enforcers of a regulation, that they should be careful not to miss anyone who might be violating the regulation. In the payoff matrix B, these costs are represented by C_A in the *p, q* cell and by C_D in the *not-p, not-q* cell. The subscripts derive from the normal cell labels used in the analysis of 2 × 2 contingency tables (B and C correspond to the *p, not-q* and the *not-p, q* cells, respectively).

The utilities in the payoff matrix (B) relate to acts of detection. In Experiments 1 and 2, people were asked to adopt the perspective of someone checking that the rule was being applied appropriately. That is, they should be concerned that no one was cheating by taking the benefit of a break (*q*) without feeling in pain (*not-p*). This means they should attach a high utility to acts of detecting these instances, that is, B_S should be high. But as the rules relate to HM, B_H may also take on a positive but low value.

People's acts of detection also bring about a state of the world that people may find threatening. People have an innate fear reaction to threatening stimuli such as spiders, snakes, or blood (Davidson, 1993; Mineka & Oehman, 2002; Oehman & Mineka, 2001). Such stimuli invoke an avoidance response (Davidson, 1993). Such a fear-and-flight response may also be innate as it made adaptive sense in our ancestral environment (Mineka & Oehman, 2002; Oehman & Mineka, 2001). As we observed in the section on the evolutionary approach, these reactions generalize to the words used to describe these stimuli. Given the work injury rule, the threatening stimulus "pain" (*p*) may be found for either of the states of the world described in the A cell (*p, q*) or the B cell (*p, not-q*). In calculating generalized expected utility (Zeelenberg, van Dijk, Manstead, & van der Pligt, 2000), a regret term is subtracted from the expected utility of the act of detection if the state of the world brought about leads to a negative emotion such as fear. Consequently, the cost associated with the A cell, C_A, should increase and/or the benefit associated with the B cell, B_H, should decrease. As in prospect theory (Kahneman & Tversky, 1979), costs and benefits may be differentially affected. However, either way we would expect the ratio B_H/C_A to decrease for threatening material.

It is important to note that we expect these effects only when the primary goal is to detect cheaters, that is, the perspective participants were asked to adopt in Experiments 1 and 2. From this perspective, it is not participants' goal to detect people at most risk. So, as in Zeelenberg et al.'s (2000) account of generalized expected utility, we expect the fear-provoking stimulus, for example, the words *pain* or *blood* in these experiments, to modify the utilities of acts of detection in the way we have described.

However, in Experiment 3 we investigated the situation where participants adopt an HM perspective. From this perspective, they should be concerned to detect people who are at risk, that is, who are not taking a break (*not-q*), although they feel in pain (*p*). This means they should attach a high utility to acts of detecting these instances, that is, B_H should be high. Moreover, in calculating generalized expected utility (Zeelenberg et al., 2000), the regret term now needs to be added to the expected utility of the act of detection because it is important to detect those most at risk. Negative regret in regret theory is "rejoicing" (Bell, 1982, 1985; Loomes & Sugden, 1982, 1986, 1987), in this case at detecting someone at greater risk. Consequently, in the threat condition, the cost associated with the A cell, C_A, should decrease and/or the benefit associated with the B cell, B_H, should increase.

In sum, there are two factors that may affect the level of B_H. First, when the goal is to detect cheaters, threatening material should lead to fear avoidance, which should make B_H fall. This is the situation we investigated in Experiments 1 and 2. Second, when the goal is to detect people at risk, then this goal in itself should lead participants to value detecting people in the most threatening situation, which should make B_H rise. We investigated this situation in Experiment 3. We now examine how these factors should affect behavior in the deontic selection task.

3.2. Predictions

In deriving predictions for the deontic selection task, we make no assumptions about the probability distribution in A. When modeling the standard selection task, Oaksford and Chater (1994) introduced a rarity assumption, that is, the probabilities of the antecedent, $P(p)$, and of the consequent, $P(q)$, were both low. However, they also observed that, “for deontic rules it is not reasonable to prejudge rarity” (Oaksford & Chater, 1994, p. 623). Here we assume that participants are maximally uncertain, and so the cell values in the probability matrix A are equiprobable, that is, $P(p) = P(q) = P(not\text{-}q|p) = .5$. This is the maximum entropy prior which makes the fewest assumptions about people’s state of knowledge (Jaynes, 1978). This simplifies the model because we can effectively ignore the probabilities.

As in Klauer (1999), we can derive expected utilities for each of the 16 possible card combinations. For each card, one possible outcome leads to a cost, C_A or C_D, and the other to a benefit, B_H or B_S. For example, if the card showing that the worker felt fine (*not-p*) is selected, then finding that the worker had taken a break (*q*) leads to B_S, whereas finding that they continued work (*not-q*) leads to C_D. For any combination of *n* selected cards, *c,* the expected utility, $EU(c)$ is

$$EU(c) = \sum_{i=1}^{n} B_i - C_i \qquad (1)$$

where B_i is the benefit, B_H or B_S, and C_i is the cost, C_A *or* C_D, associated with each card, *i.* (We do not scale these utilities by the probabilities of the different outcomes because they are all equal at .5.)

From the supervisor’s perspective, concern centers on finding cases of cheating, that is, B_S is large relative to B_H and the costs. In the pure SC version of the deontic selection task $B_H = C_A$. This is because the rule no longer relates to an HM situation and so finding *p, not-q* cases counts as a cost. In this case, if there were no costs associated with card choice, that is, $C_A = C_D = 0$, then picking any card combination that included feeling “fine” (*not-p*) and “break” (*q*) would maximize expected utility. Consequently, one may as well select all the cards. However, when costs are introduced, only selecting these two cards maximizes expected utility. This remains true for this case of an SC about an HM rule as long as the benefit of detecting someone at risk is less than the lowest cost associated with turning a card (i.e., failing to detect a cheater or someone at risk), that is, $B_H < min(C_A, C_D)$. As B_H tends toward C_A, the expected utilities of the other card combinations that include *not-p* and *q* will increase. When the benefit of detecting someone at risk is higher than the greatest cost associated with turning a card (i.e., failing to detect a cheater or someone at risk), that is, $B_H > max(C_A, C_D)$, maximizing expected utility will be associated with turning all the cards.

In the literature on rational choice it is rare to observe maximizing behavior; rather people tend to show some form of probability matching. That is, choice probabilities tend to match the relative probabilities of reward. However, it is recently becoming apparent that under some conditions some participants do optimize (e.g., Shanks, Tunney, & McCarthy, 2002). Capturing this variability in choice behavior can be achieved by treating response probabilities as proportional to expected utilities, using a response function such as

$$P(R_i) \propto e^{\beta . EU(R_i)} \quad (2)$$

where $P(R_i)$ is the probability of making response i and $EU(R_i)$ is the expected utility associated with that response. Whether utility matching or optimizing behavior is predicted depends on the value of the β parameter. If β is zero, then responding is random and independent of expected utility. If β is very small but still greater than zero, then the response probabilities will be approximately proportional to the expected utilities. If β is sufficiently large, then responding is predicted to be quite close to optimizing, that is, a step function will result with only the responses with the highest expected utility on the respond side.

The factors that would be expected to determine whether matching or optimizing behavior is observed are related to the spread of expected utilities across the response options. One would expect to observe behavior that was closer to optimization—the greater the spread and the more distinct the response with the highest expected utility. To illustrate how different predictions may arise, in Table 1 we have substituted in numerical values for B_S, B_H, C_A, and C_D. The first column indicates the relevant card combination. The second and third columns show the expected utilities associated with each card combination for a pure SC task ($B_H = C_A$) and for a pure HM task ($B_S = C_D$). The fourth, fifth, and sixth columns show the expected utilities for an SC about an HM rule when B_H is high ($B_H = 3$) and when it is low ($B_H = 1$) but *min*(C_A,

Table 1
The expected utilities for each card combination for different values of B_S and B_H

Card Combination	Benefits $B_H = C_A$	$B_S = C_D$	$B_H = 3$	$B_H = 1$	$B_H = 6$
¬*p*,*q*	12.5	–15.5	12.5	12.5	12.5
¬*p*,*q*,¬*q*	5	–9.5	11.5	9.5	14.5
p,¬*p*,*q*	5.5	–9	12	10	15
p,¬*p*,*q*,¬*q*	–2	–3	11	7	17
q	6.5	–7.5	6.5	6.5	6.5
¬*p*	6	–8	6	6	6
q,¬*q*	–1	–1.5	5.5	3.5	8.5
¬*p*,¬*q*	–1.5	–2	5	3	8
p,*q*	–0.5	–1	6	4	9
p,¬*p*	–1	–1.5	5.5	3.5	8.5
p,*q*,¬*q*	–8	5	5	1	11
p,¬*p*,¬*q*	–8.5	4.5	4.5	0.5	10.5
None	0	0	0	0	0
¬*q*	–7.5	6	–1	–3	2
p	–7	6.5	–0.5	–2.5	2.5
p,¬*q*	–14.5	12.5	–1.5	–5.5	4.5

Note. B_S was set to 10, C_D to 4, and C_A to 3.5, except when $B_S = C_D$, when B_H was set to 10.

$C_D) > B_H$, and when it is high ($B_H = 6$) and $B_H > max(C_A, C_D)$. In all cases $C_A = 3.5$ and $C_D = 4$, and either B_S (Columns 2, 4, and 5) or B_H (Column 3) was set to 10.

As we observed previously, threat material is expected to reduce the ratio B_H/C_A when $B_S > B_H$. Consequently, the $B_H = 1$ column indicates the distribution of expected utilities for the threat rule (pain) and the $B_H = 3$ or $B_H = 6$ columns indicate the distribution of expected utilities for the no-threat rule (works for 90 min). As long as $min(C_A, C_D) > B_H$, then the overall spread and the distinctness of the highest expected utility response increase as B_H decreases, that is, as B_H/C_A decreases or the rule is more threatening. For example, when $B_H = 3$, the overall spread from minimum to maximum expected utility is 14, and the response with the highest expected utility is .5 units higher than the next highest response. However, when $B_H = 1$, the overall spread from minimum to maximum expected utility is 18, and the response with the highest expected utility is 2.5 units higher than the next highest response. In sum, when the threat value is higher the decision-theoretic account predicts that participants are likely to show less matching and more maximizing behavior. Consequently, we would predict that when compared to the other response options, the probability of selecting the maximum utility choice should be higher in the threat conditions.

In comparing the decision-theoretic model to the evolutionary account in terms of the Pollard indexes we introduced previously, we considered the associated expected utilities. Selecting the *not-p* and *q* cards contribute additively to the index, whereas the *p* and *not-q* cards contribute subtractively. In the threat condition (e.g., $B_H = 1$) the average expected utility associated with selecting the *not-p* or *q* cards (where "or" is inclusive) is 5.58. The corresponding figure for selecting the *p* or *not-q* cards is 2.58. That is, the expected utility for selecting the *p* or *not-q* cards is less than half that (.46) for selecting the *not-p* or *q* cards. However, in the no-threat condition (e.g., $B_H = 3$) the average expected utility associated with selecting the *not-p* or *q* cards (where "or" is inclusive) is 7.58. The corresponding figure for selecting the *p* or *not-q* cards is 5.25. That is, the expected utility for selecting the *p* or the *not-q* card is now considerably higher at .69 times that for selecting the *not-p* or the *q* card. If participants approximately probability match, in the threat condition the Pollard index ([*not-p* + *q*] – [*p* + *not-q*]) is predicted to increase. This is because the expected utility associated with selecting the *p* or *not-q* cards decreases relative to the expected utility associated with selecting the *not-p* or *q* cards. This prediction is the exact opposite to the prediction of the evolutionary approach that this index should decrease in the threat condition.

To illustrate these effects, Fig. 2 shows the relation between the expected value of the Pollard index (PI) and B_H and C_A. In calculating this graph B_S was set to 10, C_D to 4, and β to .5. The probability of each of the 16 responses was calculated by dividing the transformed expected utilities (as in Equation 2) by the sum of these values over the 16 responses. The expected PI was then calculated by multiplying these probabilities by the PI for that response and summing. The graph shows that reducing B_H or increasing C_A, that is, decreasing the ratio B_H/C_A, leads to increases in PI. This is the central prediction tested in Experiments 1 and 2.

A possible objection to the decision-theoretic model is that it is overparameterized. For example, if we were predicting simply the selection frequencies over the four cards, then the model would have as many parameters as there are data points. However, following Klauer's (1999) optimal Bayesian model of the standard selection task, which included five free param-

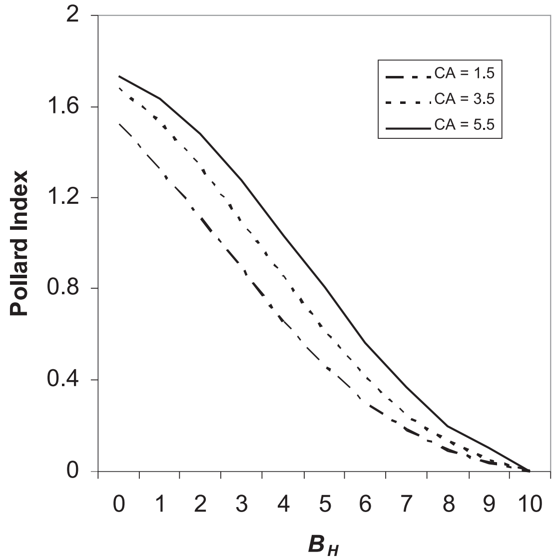

Fig. 2. The relation between B_H, C_A, and the expected value of PI. In computing values for the graph B_S was set to 10, C_D to 4, and β to .5.

eters, we make predictions for all 16 possible four-card combinations. That is, with respect to this data the decision-theoretic account is not overparameterized. In analyzing our results we fitted the model to the data in a similar way to Klauer. The only differences were that he used rank order correlations, whereas we used the product moment correlation and the response function in Equation 2.

4. Experiment 1

In this experiment, we used scenarios such as the work injury task we introduced previously to investigate the situation where the benefits associated with detecting cheating should be greater than the benefits of detecting someone at risk, that is, $B_S > B_H$. The presence of a threat word was a between subject manipulation. In each of the threat and no-threat groups, participants performed three deontic selection tasks, one of which used the work injury scenario we have used as an example in the Introduction. The other two tasks concerned rules governing blood injuries in sport (the sport scenario) and health and safety regulations in a butcher's shop (the butcher's scenario). The scenarios and cards for the latter two tasks are described in Appendix 1. We used multiple scenarios because, as we have suggested, different people might view the utilities differently in different situations—they are subjective utilities, after all. We hoped to maximize the chances of seeing effects by providing multiple scenarios.

According to Sperber and Girotto's (2002) relevance approach, the "look for cheaters" instructions—used in the task by most researchers, in particular evolutionary psychologists—mean that no reasoning is involved in this task. Participants simply look for instances of the category "cheater" defined by, for example, the combination no pain, take break. Thus Sperber and Girotto argued that all people need to do is interpret the materials; they engage in no reasoning. However, according to our interpretation of the decision-theoretic model (Oaksford & Chater, 1995a), it provides an alternative and richer interpretative framework for determining relevance. If participants interpret the materials in terms of the decision-theoretic model, they choose the card combinations with the highest expected utilities. In both cases, all people need to do is interpret the task to determine the most relevant selections. In our Experiments 1 and 2, we did not use what Sperber and Girotto refer to as "proper selection tasks," where people are asked whether the materials violate the rule. This is because our aim was to test the mutually inconsistent predictions of the evolutionary and decision-theoretic accounts. If we had used different instructions to those used by researchers in the evolutionary approach, they could rightly object that our data were due to our nonstandard instructions. Moreover, these instructions were constant across the threat and no-threat conditions. Sperber and Girotto (2002) therefore provided no reason to expect that the threat manipulation should make the identification of the appropriate category easier (in line with the predictions of the decision-theoretic approach) or harder (in line with the predictions of the evolutionary approach). Consequently, if threat effects are observed, then it is unlikely that people are simply identifying the category "cheater" when interpreting these materials.

4.1. Method

4.1.1. Participants

Sixty-four undergraduate students from Cardiff University participated in return for course credit. They were randomly assigned to the threat and no-threat conditions so that there were 32 participants in each group.

4.1.2. Materials

The materials consisted of three newly created deontic selection tasks (see Appendix 1). Each task had two versions, threat and no threat. Both versions used the same neutral description. The only difference concerned the *p* card. In the threat condition, in all three tasks, this card showed a threat word. All the other cards were exactly the same in each condition (there was one exception in the work injury scenario, see Procedure).

4.1.3. Design

The experiment was a 4 (cards: *p, not-p, q, not-q*) × 2 (threat condition: threat, no threat) mixed design.

4.1.4. Procedure

Participants were tested individually or in pairs. In the laboratory there were computers in separate cubicles for each participant. The selection tasks were presented on Power Macintosh 4/440 computers using the PsyScope software (Cohen, MacWhinney, Flatt, & Provost, 1993) to control the presentation of stimuli and to record responses. The selection tasks were presented in random orders. For each task, participants first saw a screen showing the following information:

> Thank you for agreeing to take part in this experiment. The experiment is looking at the decisions people make. You will be presented with some text passages and asked to make decisions about them.
>
> If you have any questions, please ask the experimenter. Please press the space bar when you are ready to continue.

On pressing the space bar the scenario for each selection task was presented. We use the factory scenario in the threat condition as an illustration. In the no-threat scenario the words in italics were replaced by the material that follows them in brackets.

> You are the director of a factory. Most of the factory involves manual labor, and workers cannot be allowed to get injured. Because of this you have the following rule:
>
> IF YOU FEEL PAIN (WORK FOR 90 MIN), THEN YOU MUST STOP WORKING FOR A BREAK.
>
> Although you want to increase profits you cannot allow your workers to feel pain (become tired). However, you suspect that four particular workers may be taking breaks when they have not felt pain (worked for 90 min). One day you ask a supervisor to record if these workers were feeling pain (had worked for 90 min) and whether they were working or not. At the end of the day the supervisor leaves you a note with some information about each worker on it. However, it has had coffee spilled on it, and you can only see one piece of information about each worker: whether the worker felt pain (worked for 90 min) or whether he or she had stopped working.
>
> On the next page you will be briefly presented with the piece of information you can see from the four forms.

After viewing the scenario participants were shown all four cards for a period of 6 sec in the following order: *p, q, not-p,* and *not-q*. Participants were then shown each of the four cards one at a time in random order and were asked to make a decision whether to turn a card over:

> After viewing all four pieces together, you will see each piece individually. During this time you will be asked whether you think you must ask for the other piece of information to determine whether the workers are taking more breaks than they are entitled to.
>
> If you think you must ask for the other information then press “Y.” If you do not think you must ask for the other information then press “N.”
>
> If you have any questions please ask the experimenter. Press the space bar to continue.
>
> PAIN FINE STOP CONTINUE
>
> PAIN Do you think you must ask whether the worker stopped work?
>
> Yes or no?
>
> FINE Do you think you must ask whether the worker stopped work?
>
> Yes or no?

STOP Do you think you must ask how the worker felt?

Yes or no?

CONTINUE Do you think you must ask how the worker felt?

Yes or no?

For the no-threat condition, the word *pain* was replaced by *90* and the word *fine* was replaced by *70.* This was the one scenario in which a card, other than the *p* card, changed between threat conditions. Participants pressed the *Y* button for "Yes" and the *N* button for "No." On completion participants were debriefed and thanked for participating.

It could be argued that being presented with the cards one at a time to evaluate could lead to nonstandard responses. However, this procedure was constant across threat and no-threat conditions. Moreover, it was only introduced to facilitate the collection of reaction times that proved not to be interesting. In a further unpublished experiment (Perham, 2003), identical results were found for these materials using the standard four-card presentation.

4.2. Results and discussion

Table 2 shows the percentage of each card combination selected in Experiment 1. Our first analysis directly tested the main predictions of the evolutionary approach and the deci-

Table 2
The percentage of selections for each card combination for each scenario and condition (Threat vs. No-Threat) in Experiment 1 ($B_S > B_H$)

Card Combination	Task: Sports T	Sports NT	Factory T	Factory NT	Butcher T	Butcher NT	Overall T	Overall NT
¬p,q	46.9	46.9	56.3	9.4	37.5	21.9	46.9	26.0
¬p,q,¬q	3.1	9.4	0	18.8	9.4	3.1	4.2	10.4
p,¬p,q	15.6	15.6	15.6	9.38	0	9.4	10.4	11.5
p,¬p,q,¬q	9.4	6.3	9.4	46.9	15.6	12.5	11.5	21.9
q	0	0	0	0	6.3	0	2.1	0.0
¬p	6.3	0	0	3.1	0	0	2.1	1.0
q,¬q	3.1	3.1	3.1	6.3	0	6.3	2.1	5.2
¬p,¬q	0	0	0	0	0	0	0.0	0.0
p,q	6.3	6.3	6.3	3.13	6.3	12.5	6.3	7.3
p,¬p	0	0	0	0	0	0	0.0	0.0
p,q,¬q	9.4	6.3	6.3	3.13	15.6	21.9	10.4	10.4
p,¬p,¬q	0	0	0	0	0	0	0.0	0.0
None	0	0	0	0	0	0	0.0	0.0
¬q	0	0	3.1	0	0	0	1.0	0.0
p	0	3.1	0	0	6.3	3.1	2.1	2.1
p,¬q	0	3.1	0	0	3.1	9.4	1.0	4.2

Note. $N = 64$. T = threat; NT = no threat.

sion-theoretic approach. On the one hand, according to evolutionary psychology, the threat condition should lead to a decrease in the Pollard index ([*not-p* + *q*] – [*p* + *not-q*]) as threat material should lead to more HM (*p, not-q*) responses. On the other hand, according to the decision-theoretic approach, the threat condition should lead to an increase in the Pollard index as threat material should reduce the utility associated with the *p, q* case and increase the cost associated with the *p, not-q* case. For each participant, we therefore computed the Pollard index for each scenario and then averaged these across the three scenarios. This index was significantly positive in one-sample *t* tests in both the threat, $t(31) = 5.69$, $p < .0001$, and the no-threat conditions, $t(31) = 4.00$, $p < .0005$ (see following for mean and standard errors). Thus the SC interpretation dominated in both conditions. However, in general these indexes do not meet the requirements for parametric tests, so the comparison between conditions was made using a one-tailed Mann–Whitney test (Siegel & Castellan, 1988, pp. 128–137). The mean Pollard index was significantly higher in the threat condition (mean = .97, standard error [*SE*] = .14) than in the no-threat condition (mean = .54, *SE* = .17), $z = 2.05$, $p < .025$. Consequently the results using the Pollard indexes were consistent with the decision-theoretic model but not with the predictions of evolutionary psychology. In sum, even allowing for possible errors of omission and commission in the evolutionary approach, the results of Experiment 1 were only consistent with the decision-theoretic approach.

However, it is also clear from Table 2 that there were quite large differences between scenarios. We therefore performed similar analyses on the Pollard indexes by scenario. For the factory scenario, this index was significantly higher in the threat condition (mean = 1.19, *SE* = .19) than in the no-threat condition (mean = .47, *SE* = .13), $z = 3.09$, $p < .001$. This was replicated, albeit not significantly, for the butcher's scenario, $z = 1.49$, $p = .068$ (threat condition: mean = .63, *SE* = .23; no-threat condition: mean = .13, *SE* = .23). However, for the sport scenario, there was no effect, $z = .09$, *ns* (threat condition: mean = 1.09, *SE* = .18; no-threat condition: mean = 1.03, *SE* = .20). It was clear from these analyses that the threat effects observed in Experiment 1 were largely attributable to a single scenario, the factory scenario. Although, if the scenarios were regarded as independent tests such that the results could be combined meta-analytically (Wolf, 1986), then the hypothesis that PI should increase in the threat condition could not be rejected using Stouffer's combined test (sum of the *z* scores divided by the square root of the number of tests; see Wolf, 1986, p. 20), $z = 2.70$, $p < .005$.

Nonetheless, we would have more confidence in the reliability of these results if they could be demonstrated across scenarios. We therefore checked the between-scenario reliability of our results. We computed Spearman's rank order correlations between each scenario within conditions, using the 16 possible responses as the unit of analysis (six comparisons in all). We used rank order correlations to remove the effects of influential data points. The average correlation is the index of reliability. In Experiment 1, all six rank order correlations were significant at at least the .05 level, and the average value was .73 (range: .50–.90). It is difficult to say whether this is good or bad, but it is satisfactory by the usual standards of reliability. Nonetheless, it would be desirable for the between-scenario reliability to be close to the test–retest reliability for these scenarios and conditions. We therefore repeated Experiment 1 in Experiment 2.

However, before turning to our second experiment, we report the results of fitting the decision-theoretic model to the results in Table 2. The expected utilities should be able to predict the frequency with which the different card combinations are selected. The expected utilities for each card combination depend on the costs and benefits, B_H, B_S, C_A, and C_D. We used a similar procedure to Klauer (1999) who fitted a nonsequential Bayesian model to the data from the standard selection task. He estimated best fitting parameter values by maximizing the rank order correlation between the expected losses predicted by the model and the overall selection frequencies for each card combination. We used a similar procedure but sought parameter values that maximized the product moment correlation between the exponential transformation of the expected utilities in Equation 2 and the overall selection frequencies. β was a free parameter that was constrained to take the same value across the threat and no-threat conditions.

We maximized the correlation coefficient using a steepest descent search implemented in Mathematica's (Wolfram, 1991) MultiStartMin function (Loehle, 2000). This function supplements the Newton–Raphson method with a grid search procedure to ensure a global maximum. The absolute values of the utilities are of course arbitrary; what matters are their relative values. We therefore set C_D to 4 and B_S to 10 and constrained the values of B_H and C_A to vary between 0 and 10, as in the examples shown in Table 1. We predicted that the ratio B_H/C_A should be lower for the threat condition than for the no-threat condition. We fitted the model to the overall selection frequencies shown in Table 2.

For the threat condition, the fit to the data was good, $r(14) = .95$. The parameter values were $B_H = .004$, $C_A = 2.61$, and $\beta = .63$. Therefore in the threat condition, $B_H/C_A = .002$. In the no-threat condition, the fit was also good, $r(14) = .84$; $B_H = 3.59$, $C_A = 3.89$, and $B_H/C_A = .92$. As predicted, the B_H/C_A ratio was much lower in the threat condition. Thus the threat manipulation has affected the parameters of the model in the way predicted. Threat has reduced the benefit associated with detecting people at risk who do not take precautions, that is, *p, not-q* cell, although the cost associated with detecting people at risk and who do take the precaution did not rise. As predicted, the probability of choosing the maximum utility choice (*not-p, q*) in the threat condition (.47) compared to the other choices was higher than making this choice in the no-threat condition (.26), albeit not quite significantly, $\chi^2(1, N = 64) = 3.33$, $p = .07$.

By way of comparison, we developed a simple processing tree implementation of the evolutionary approach assuming reasonable errors of commission and omission. Only three parameters were needed, p_s: the probability of an SC interpretation (so $1 - p_s$ is the probability of the HM interpretation); p_o: the probability of making an error of omission; and p_c: the probability of making an error of commission. It is straightforward to derive expressions for the probability of making each of the 16 possible responses. For example, the probability of selecting just the *not-p* card is as follows:

$$P(not - p) = p_s(1 - p_c)^2 p_o(1 - p_o) + (1 - p_s)p_c(1 - p_c)p_o{}^2$$

If participants adopt the SC interpretation (first expression in the sum), then to make this selection, they have to avoid making two errors of commission $[(1 - p_c)^2]$ and one error of omission $(1 - p_o)$, and make one error of omission (p_o). The same reasoning applies if they adopt the HM interpretation (second expression in the sum), but because they are predicted to select different cards what counts as an error of omission or commission changes.

Again seeking parameter values that maximize the product moment correlation, this model provided very good fits to the data: threat condition, $r(14) = .95$; no-threat condition, $r(14) = .87$. However, as could be predicted from the experimental results, the effect of the threat manipulation on the parameter values ran counter to theoretical prediction. In each condition there were always two possible solutions where the parameters were mirror images of each other. In the threat condition, one solution was $p_s = .9998$, $p_o = .0005$, $p_c = .1121$, and the other was $p_s = .0002$, $p_o = .8879$, $p_c = .9995$. In the no-threat condition, one solution was $p_s = .7641$, $p_o = .0959$, $p_c = .4353$, and the other was $p_s = .2359$, $p_o = .5647$, $p_c = .9041$. That is, either interpretation—SC or HM—can provide a good fit to the data as long as a high level of error is tolerated. But the levels of error cannot be so high that they are more likely than the responses predicted by theory. Consequently the "high error" second solutions for each condition were not serious contenders as they revealed the inability of the responses predicted by evolutionary psychology to explain the data. The first "low error" solutions show that the probability of the HM interpretation was less ($1 - p_s = .0002$) in the threat condition than in the no-threat condition ($1 - p_s = .2359$), not more as the evolutionary account predicts. Moreover, intuitively, errors of commission are worse than errors of omission. Consequently, they would be expected to be less likely than errors of omission, not more likely as in these model fits. In sum, an evolutionary psychology account plus a reasonable errors model fails to account for these data.

This experiment confirmed many of the predictions of the decision-theoretic model. However, there were some problems with reliability that we addressed in Experiment 2. Moreover, we assumed that *pain* and the other threatening terms used in these scenarios are more threatening than "working for 90 (or 70) min" or the other nonthreatening words. However, we did not test whether this is the case. In our next experiment, we therefore asked participants to rate these words or expressions for threat value.

5. Experiment 2

This experiment repeated Experiment 1 but also included a rating task to obtain ratings of the threat value of the words used in the task rules. We also increased the sample size.

5.1. Method

5.1.1. Participants

Eighty undergraduate students from Cardiff University participated in return for course credit. They were randomly assigned to the threat and no-threat conditions so that there were 40 participants in each group.

5.1.2. Materials

The materials were the same as in Experiment 1. Participants also completed a 30-item rating task (the task instructions are shown in Appendix 2). They had to rate 30 words as being either threatening or nonthreatening on a 5-point Likert-type scale. All 15 words used on the cards in the selection tasks were included and 15 other words acted as filler items.

5.1.3. Design

The design was the same as Experiment 1.

5.1.4. Procedure

The selection task phase of the experiment was the same as Experiment 1. After completing the selection tasks participants performed the ratings task. On completing the ratings task participants were debriefed and thanked for participating.

5.2. Results and discussion

We first report the results of the ratings task, which was used to demonstrate that the threat words were more threatening than the no-threat words. The mean threat ratings for the five cards used in each scenario are shown in Table 3. A series of one-way within-subjects analyses of variance (ANOVAs) were carried out for each scenario and the pairwise differences assessed using Scheffé multirange tests. The threat word was always more threatening than all the other words, at least the .0001 level. However, there were also significant differences in threat value between the nonthreatening words. We therefore looked at the average effect sizes for comparisons between threat and no-threat words and between the no-threat words. On average the threat words led to a 1.30 *SD* unit increase in threat value over the nonthreat words. However, the average increase between any two no-threat words was only .29 *SD* units. We can therefore conclude that the threat words were relatively more threatening than the no-threat words. Note that differences between the no-threat words could not affect our predictions because the threat word on the *p* cards was always significantly more threatening than the no-threat word. Moreover, for the *not-q* cards, exactly the same no-threat words were used in both conditions.

Table 4 shows the percentage of each card combination selected in Experiment 2. The first thing to observe is that the overall results of Experiments 1 and 2 are very similar. We looked at the overall test–retest reliability for all six conditions between Experiments 1 and 2. All six rank order correlations were significant at at least the .025 level, and the average value was .71 (range: .65–.81). Consequently, the reliability between scenarios in Experiment 1, at .73, com-

Table 3
Mean threat ratings and standard deviations for all cards used in these experiments by task

	Card									
	p(T)		*p*(NT)		*not-p*		*q*		*not-q*	
Task	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Sports	68[a]	19	33[cd]	22	25[d]	19	51[b]	21	36[c]	21
Factory	78[a]	17	37[c]	26	30[c]	20	56[b]	20	53[b]	22
Butcher	68[a]	19	29[c]	21	28[c]	22	31[c]	22	41[b]	22

Note. The specific words used in each task appear in the Methods sections and in Appendix 1. For each task, where superscripts differ, the means differ at least at the .0001 level in post hoc Scheffé tests. Range = 0–100 for all cards used in these experiments by task.

p(T) = the *p* card in the threat condition; *p*(NT) = the *p* card in the no-threat condition.

Table 4
The percentage of selections of each card combination for each scenario and condition (Threat vs. No-Threat) in Experiment 2 ($B_S > B_H$)

Card Combination	Task: Sports T	Sports NT	Factory T	Factory NT	Butcher T	Butcher NT	Overall T	Overall NT
¬*p*,*q*	50.0	22.5	40.0	22.5	30.0	15.0	40.0	20.0
¬*p*,*q*,¬*q*	10.0	5.0	2.5	17.5	17.5	2.5	10.0	8.3
p,¬*p*,*q*	5.0	22.5	12.5	2.5	5.0	2.5	7.5	9.2
p,¬*p*,*q*,¬*q*	7.5	20.0	15.0	17.5	12.5	25.0	11.7	20.8
q	5.0	2.5	7.5	0	2.5	0	5.0	0.8
¬*p*	0	2.5	0	0	0	2.5	0.0	1.7
q,¬*q*	2.5	5.0	0	10.0	0	10.0	0.8	8.3
¬*p*,¬*q*	2.5	2.5	2.5	10.0	5	2.5	3.3	5.0
p,*q*	2.5	5.0	5.0	0	2.5	7.5	3.3	4.2
p,¬*p*	0	2.5	2.5	5.0	0	0	0.8	2.5
p,*q*,¬*q*	7.5	5.0	5.0	10.0	12.5	22.5	8.3	12.5
p,¬*p*,¬*q*	2.5	0	0	2.5	2.5	2.5	1.7	1.7
None	0	0	0	0	2.5	0	0.8	0.0
¬*q*	2.5	5.0	2.5	2.5	0	2.5	1.7	3.3
p	0	0	2.5	0	0	0	0.8	0.0
p,¬*q*	2.5	0	2.5	0	7.5	5.0	4.2	1.7

Note. T = threat; NT = no threat.

pared favorably with the test–retest reliability between Experiments 1 and 2. We also tested the reliability between scenarios in Experiment 2. All six rank order correlations were significant at at least the .05 level, and the average value was .67 (range: .57–.84). The reliabilities between scenarios within experiments were very close to the test–retest reliability between experiments. It therefore seems that we can have some confidence that the three scenarios within each experiment are producing largely the same behavior, and hence it makes sense to collapse scenarios in analyzing these results.

Experiment 2 replicated the results of Experiment 1. The mean Pollard index was significantly positive in one-sample t tests in both the threat, $t(39) = 6.47$, $p < .0001$, and the no-threat conditions, $t(39) = 3.74$, $p < .001$ (see the following for means and standard errors). Thus the SC interpretation again dominated in both conditions. The mean Pollard index was significantly higher in the threat condition (mean = .82, SE = .13) than in the no-threat condition (mean = .39, SE = .11), $z = 2.88$, $p < .0025$ (all tests were one-tailed). In Experiment 2, the results for the Pollard indexes also replicated at the level of the individual scenarios: factory scenario, $z = 1.69$, $p < .05$ (threat: mean = .88, SE = .16; no-threat: mean = .50, SE = .18); sports scenario, $z = 1.83$, $p < .05$ (threat: mean = 1.03, SE = .19; no-threat: mean = .68, SE = .15); butcher's scenario, $z = 2.14$, $p < .025$ (threat: mean = .55, SE = .20; no-threat: mean = .00, SE = .17). In sum, these results show that the threat effect is robust and reliable.

The decision-theoretic model again provided good fits to the data: threat condition, $r(14)$ = .95; no-threat condition, $r(14)$ = .78. The best fit parameter values for the threat condition were

$B_H = .003$, $C_A = 4.75$, $\beta = .31$, and $B_H/C_A = .001$. In the no-threat condition, $B_H = 4.42$, $C_A = 3.82$, and $B_H/C_A = 1.16$. Replicating Experiment 1, B_H/C_A was much lower in the threat condition. In this experiment, threat has reduced the benefit associated with detecting people at risk who do not take precautions, that is, the *p, not-q* cell, and increased the cost associated with detecting people at risk who do take the precaution. In this experiment, B_H was greater than C_A or C_D in the no-threat condition. This was because in this condition selecting all the cards dominated, and for this to happen it must be the case that $B_H > max(C_A, C_D)$.

Again, the probability of choosing the maximum utility choice (threat condition: *not-p, q*) in the threat condition (.40) compared to the other choices was higher than making this choice (no-threat condition: all cards) in the no-threat condition (.21), albeit not quite significantly, $\chi^2(1, N = 80) = 3.81$, $p = .051$. However, collapsing across Experiments 1 and 2 revealed a highly significant effect, $\chi^2(1, N = 144) = 7.11$, $p < .01$.

The results of fitting the evolutionary psychology plus reasonable errors model to these data produced similar results to Experiment 1. The fits were again good: threat condition, $r(14) = .96$; no-threat condition, $r(14) = .82$. For the threat condition, $p_s = .9270$, $p_o = .0002$, $p_c = .1801$, and for the no-threat condition, $p_s = .7370$, $p_o = .2320$, $p_c = .4820$ (we do not report the high error mirror-image solutions). However, again, and contrary to prediction, the probability of the HM interpretation was lower in the threat condition (.073) than in the no-threat condition (.263).

There was one minor problem with the factory scenario. The phrase, “You suspect that four particular workers may have a break when they have not felt pain (worked for 90 min)” was included in the instructions for this scenario in both Experiments 1 and 2. It could be argued that this phrase cued participants to the *not-p, q* response. However, this phrase was present in both the threat and no-threat scenarios and so was a constant factor. Consequently, this potential confound could not explain the significant effect of the threat manipulation, which was also significant for the other scenarios where this potential confound was absent.

Experiment 2 replicated the results of Experiment 1. Moreover, in Experiment 2 all the scenarios revealed a similar pattern of between-condition differences (threat vs. no-threat). Our final experiment tested a manipulation that would be expected to remove or reverse the effect of threatening material.

6. Experiment 3

These experiments have introduced novel materials that appear to reveal all the predicted decision-theoretic effects. However, the novelty of the materials may suggest that the effects we have observed are not likely to generalize. One way to allay these concerns is to demonstrate that these materials reveal similar effects to other materials used in the deontic selection task. In the Introduction we mentioned that in the deontic selection task, changes in perspective are predicted to alter the pattern of responding. In Experiments 1 and 2 participants were asked to adopt an SC perspective where the supervisor was concerned to enforce the regulation that workers only took a break if they were in pain. However, it is just as easy to ask participants to adopt an HM perspective where the supervisor is concerned that workers take a break if they are in pain to prevent injury. Participants should now be interested in the case where a worker is in pain but does not take a break, that is, the *p, not-q* case. This creates a pure HM task (see the $B_S = C_D$ column in Table 1) where avoiding injury is the focus of attention.

From this perspective it seems unlikely that increasing the level of threat would reduce the tendency to select items indicating that people are at risk. As before, according to the evolutionary approach, the fear module should be more activated by a threatening stimulus, and so there should be more HM responses the greater the threat. Moreover, if one's explicit goal is to manage injuries, then one might imagine that more attention should be paid to those in the most threatening situation, that is, a higher utility should be assigned to the more threatening case. Of course, this explicit goal has to compete with a possibly innate fear response, which could serve to cancel any effects on people's subjective utilities. Thus the decision-theoretic approach either makes the same prediction as the evolutionary approach or predicts no effects of threat material. So, according to both approaches there should be more *p, not-q* response in the threat condition, although no such differences would also be consistent with the decision-theoretic approach. Consequently, the Pollard index should be lower in the threat condition according to the evolutionary approach. But no effect would also be consistent with decision theory.

In this experiment we also removed the cheater detection instruction, that is, the sentence mentioning defining dimensions of the cheater category was not present. Instead violation instructions were used. In the butcher's scenario participants were asked to select cards "to determine whether the employee was using the correct hand-wear." In the sports scenario participants were told to check the cards, "to determine whether players receive the correct treatment." In the factory scenario they were told to check cards, "to determine whether the workers are potentially damaging their health."

6.1. Method

6.1.1. Participants

Sixty-four undergraduate students from Cardiff University participated in return for course credit. They were randomly assigned to the threat and no-threat conditions so that there were 32 participants in each group.

6.1.2. Materials

The materials were the same as in Experiments 1 and 2, apart from the shift in perspective, which is exampled for the factory scenario in the Procedure section.

6.1.3. Design

The experiment was a 4 (cards: *p, not-p, q, not-q*) × 2 (threat condition: threat, no-threat) mixed design.

6.1.4. Procedure

The procedure was the same as in Experiments 1 and 2. However, after reading the initial instructions, the scenarios with the HM perspective were presented. Again, the factory scenario is used as an illustration. The relevant sections of the instructions were replaced with the following (i.e., up to sentence beginning, "One day you ask a supervisor … "):

> You are the director of a factory. Most of the factory involves manual labor and workers cannot be allowed to get injured (work for long periods of time). Because of this you have the following rule:
>
> > IF YOU FEEL PAIN (WORK FOR 90 MIN), THEN YOU MUST STOP WORKING FOR A BREAK.
>
> Employees have to reach a certain quota every day. If they exceed this then they get paid more. However, some workers do not stop for breaks when they should. You realize that this is potentially bad for their health, and thus you could have a reduced workforce.

Participants were also asked to select cards "to determine whether the workers are potentially damaging their health." Similar changes were made to the other two scenarios, and these are shown in Appendix 1.

6.2. Results and discussion

The results of Experiment 3 are shown in Table 5. The reliability between scenarios in Experiment 3, at .63 (range: .55–.84, all significant at least .05), was comparable to that found in Experiments 1 and 2, and the test–retest reliabilities between Experiments 1 and 2.

The mean Pollard index was significantly negative in one-sample t tests in both the threat, $t(31) = 13.55$, $p < .0001$, and in the no-threat conditions, $t(31) = 8.84$, $p < .0001$ (see following for means and standard errors). Thus the HM interpretation now dominated in both conditions. These results strongly confirm the success of the perspective manipulation in this experiment and shows that these novel scenarios are susceptible to the same manipulations as other deontic scenarios used in the literature.

The mean Pollard index showed no significant differences between the threat (mean = –1.25, SE = .09) and the no-threat conditions (mean = –1.06, SE = .12), $z = 1.00$, $p = .16$ (all tests were one-tailed). This lack of an effect was repeated for the factory scenario, $z = 1.28$, $p = .10$ (threat: mean = –1.06, SE = .16; no-threat: mean = –.75, SE = .17) and for the sports scenario, $z = .59$, $p = .28$ (threat: mean = –.84, SE = .16; no-threat: mean = –.97, SE = .16). However, for the butcher's scenario there was a significant effect, z = 1.93, $p < .05$, such that the index was lower in the threat condition (mean = –1.84, SE = .08) than in the no-threat condition (mean = –1.47, SE = .16). This pattern of effects is consistent with both the evolutionary psychology and the decision-theoretic approaches. Now that the participants' goal, to prevent injury, is consistent with the threat manipulation, participants either show no effect of threat, or it increases their tendency to make the HM interpretation.

Fitting the decision-theoretic model to the data involved setting $B_S = C_D = -4$, to reflect the fact that in a pure HM task selecting the case where the person is not at risk counts as a cost. The decision-theoretic model again provided good fits to the data: threat condition, $r(14) = .97$; no-threat condition, $r(14) = .96$. Even with B_H free to vary, the best fit parameter values for the threat condition were, $B_H = 10$, $C_A = .0003$, and $\beta = .43$. In the no-threat condition, $B_H = 10$ and

Table 5
The percentage of selections of each card combination for each scenario and condition (Threat vs. No-Threat) in Experiment 3 ($B_S = C_D$)

Card Combination	Scenario: Sports T	Sports NT	Factory T	Factory NT	Butcher T	Butcher NT	Overall T	Overall NT
¬*p*,*q*	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
¬*p*,*q*,¬*q*	3.1	3.1	0.0	3.1	0.0	0.0	1.0	2.1
p,¬*p*,*q*	3.1	3.1	3.1	3.1	0.0	12.5	2.1	6.3
p,¬*p*,*q*,¬*q*	25.0	0.0	6.3	9.4	0.0	0.0	10.4	3.1
q	0.0	0.0	0.0	3.1	0.0	0.0	0.0	1.0
¬*p*	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
q,¬*q*	3.1	9.4	6.3	6.3	0.0	0.0	3.1	5.2
¬*p*,¬*q*	0.0	0.0	3.1	12.5	0.0	6.3	1.0	6.3
p,*q*	3.1	12.5	9.4	6.3	0.0	0.0	4.2	6.3
p,¬*p*	0.0	0.0	0.0	0.0	3.1	0.0	1.0	0.0
p,*q*,¬*q*	25.0	18.8	15.6	15.6	9.4	3.1	16.7	12.5
p,¬*p*,¬*q*	6.3	6.3	9.4	3.1	0.0	6.3	5.2	5.2
None	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
¬*q*	0.0	12.5	3.1	6.3	0.0	3.1	1.0	7.3
p	3.1	3.1	6.3	3.1	0.0	0.0	3.1	2.1
p,¬*q*	28.1	31.3	37.5	28.1	87.5	68.8	51.0	42.7

Note. $N = 64$. T = threat; NT = no threat.

C_A = .0027. As is clear from these parameter values, there were virtually no differences between conditions. Relative to the cost of looking at people not at risk, that is, C_D, there was a very low cost associated with detecting someone at risk who is taking the precaution, that is, C_A. Moreover, as makes intuitive sense, this cost was lower when the risk was greater, that is, in the threat condition. This effect had the consequence that the ratio B_H/C_A was 10 times greater in the threat condition (33,333) than in the no-threat condition (3,704).

We also fitted the evolutionary psychology plus reasonable errors model to these data. The fits were again good: threat condition, $r(14) = .97$; no-threat condition, $r(14) = .96$. For the threat condition, $p_s = .0006$, $p_o = .0166$, $p_c = .1813$, and for the no-threat condition, $p_s = .0006$, $p_o = .0630$, $p_c = .1424$ (we do not report the high-error mirror-image solutions). Consequently, the probability of the HM interpretation was high and the same in both conditions (.9994).

Only the butcher's scenario led to a lower Pollard index in the threat condition, indicating higher levels of *p, not-q* responses. Consistent with Sperber and Girotto (2002), it could be argued that this is due to the violation instruction making this case particularly relevant in this scenario. However, in general it would seem more consistent with Sperber and Girotto to expect lower levels of *p, not-q* responses in this experiment than *not-p, q* responses in Experiments 1 and 2. They argued that violation instructions provide a genuine selection task and that the high *p, not-q* and *not-p, q* selection rates observed in deontic selection tasks only occur due to the cheater detection instructions. This argument suggests that there should be fewer of the deontic selection patterns in Experiment 3 than in Experiments 1 and 2. We therefore compared the absolute values of the PI scores in Experiment 2 and Experiment 3. The absolute values of the PI scores were higher in Experiment 3 than in Experiment 2 for the threat, $t(70) = 2.65, p < .01$, and the no-threat conditions, $t(70) = 4.22, p < .0001$. These results shows that the deontic selection patterns were more in evidence, given the violation instruction than the cheater detection instruction, not less as Sperber and Girotto would seem to have to predict.

However, perhaps according to relevance theory, good performance on the butcher's task could be expected if the violating case was particularly relevant. Even if there were such an argument for the butcher's scenario, this hypothesis still could not explain the significant differences between conditions we observed for this scenario. Given that the overall reliability between scenarios was comparable to that found in Experiments 1 and 2, we prefer not to speculate further on between-scenario differences in Experiment 3. Moreover, it is worth noting that similar unexplained differences were also observed by Sperber and Girotto (2002). In their Experiment 1, "look for cheater" condition, 68% of participants made the *p, not-q* selection. In contrast, in their Experiment 2, "look for cheater (conditional)" condition, only 33% of participants made this selection. This difference is directly comparable to the difference between the butcher's scenario, "no-threat" condition (69%), and the sports scenario, "no-threat" condition (31%), in Experiment 3. Clearly, there are further material effects that remain to be explained in these tasks.

It might further be argued that the factory scenario introduces other violations. For example, participants were told as part of the background that the workers are required to fulfill a quota. Of course getting paid without filling the quota has nothing to do with whether people take more breaks than they are entitled to, which was the explicit violation participants were asked to look for and which was the only possible violation that could appear on the cards. So it is highly unlikely that this background information could have introduced a confound.

This experiment has shown that these materials readily produce perspective effects similar to those previously observed using different materials. Moreover, consistent with the decision-theoretic approach, changing perspective so that a participant's goal is to detect people at risk, removed or reversed the effect of threatening material. The effect was not as strong as in Experiments 1 and 2 because the culturally specified task of detecting people at greater risk has to compete with the possibly innate tendency to avoid more threatening situations.

7. General discussion

The experiments reported in this article investigated how people respond when they are confronted by an HM rule embedded in an SC situation when the severity of the hazard, or threat value, varies. The evolutionary approach and the decision-theoretic approach make divergent predictions for how people should respond in this situation. The most striking finding in Experiments 1 and 2 was that when the threat value of the antecedent increased in an SC situation there were fewer HM responses, even when errors were taken into account. This finding was consistent with the decision-theoretic approach but not with the evolutionary psychology approach. Moreover, when the threat manipulation was aligned with participants' primary goal (to detect people at risk), as in Experiment 3, there was only a small effect of threat in the opposite direction to that observed in Experiments 1 and 2. In general, these experiments confirmed the predictions of the decision-theoretic model. In this section, we discuss how our findings relate to other theories of deontic reasoning.

7.1. Pragmatic reasoning schemas

Cheng and Holyoak (1985) were the first to fully articulate the deontic nature of facilitation in the selection task. The theory of pragmatic reasoning schemas that they put forward in that article was part of a larger psychological theory of inductive reasoning (Holland, Holyoak, Nisbett, & Thagard, 1986). Pragmatic reasoning schemas contained learned rules relating the contents of deontic regulations. In later developments (Holyoak & Cheng, 1995b) it was argued that these knowledge structures can be motivated by appeal to ideas from jurisprudence and that they were compatible with a decision-theoretic approach. Pragmatic reasoning schemas were the knowledge structures that provided information about what constituted a violation of someone's rights in a particular domain. Once identified, people's actual decision to turn or not turn cards is determined by the utilities people associate with the various outcomes.

This interpretation was questioned by Oaksford and Chater (1995b), who argued that the pragmatic reasoning schemas represented an unnecessary processing stage in deontic reasoning. However, this argument was misguided for at least two reasons. First, processing issues are strictly beyond the scope of the decision-theoretic analysis because it is only defined at the computational level. Second, in their response, Holyoak and Cheng (1995b) rightly observed that the cognitive system requires a way of determining which outcomes violate someone's rights. In the decision-theoretic model this is just assumed as a given. More generally, at the processing or algorithmic level, the decision-theoretic account requires a specification of where the probabilities and utilities that figure in deontic reasoning come from and how they are computed (probably only to an approximation). Pragmatic reasoning schemas may provide at least one of the knowledge structures required to implement a decision-theoretic model at the algorithmic level. Consequently, like Holyoak and Cheng (1995c), we see no necessary incompatibility between these approaches.

7.2. Text processing and relevance

Recently it has been argued that perspective effects in the selection task are a result of normal text processing and are compatible with standard logical interpretations (Almor & Sloman, 1996, 2000). So for example, in our Experiments 1 and 2, it would be argued that the dominant *not-p, q* response arises as a result of participants interpreting the task rule not as stated, that is, if you feel pain, then you must stop working for a break, but as, if you stop work for a break, then you may feel in pain. Of course the *not-p, q* response to the original rule is the logical *p, not-q* response with respect to this reinterpreted rule. This account of perspective effects is the same as that proposed by relevance theory (Girotto et al., 2001; Sperber et al., 1995; Sperber & Girotto, 2002) and so is subject to the same criticisms. We therefore address both theories by offering a critique of the relevance approach. In particular, we concentrate on the grounds Sperber and Girotto advanced for discontinuing research using the Wason selection task.

Sperber and Girotto (2002; see also Girotto et al., 2001; Sperber et al., 1995) argued that responses on the selection task are driven by a comprehension mechanism that includes spontaneous inferential abilities for computing logical equivalence. Because comprehension must precede reasoning, they argue that any more specific (or indeed more general) reasoning mechanisms are preempted. Consequently, the only information that can be gleaned from the selection task is how to push the comprehension mechanism toward particular interpretations by making them more relevant. Sperber et al. showed that manipulations of relevance can indeed bring about interpretations that lead to correct responding on the abstract version of the selection task. They moreover argued that similar principles apply to the deontic task. Girotto et al. and Sperber and Girotto (2002) showed how this could be achieved. They also pointed out that the tasks used by Fiddick et al. (2000) were not really selection tasks. Rather, participants look for cheaters because the task instructions explicitly tell them to.

Although we find many of Sperber and Girotto's (2002) arguments compelling, should they also compel us to abandon the selection task? We agree that it is an overworked paradigm. However, we cannot agree with the grounds that Sperber and Girotto and their colleagues provide for abandoning it. These grounds are really no more than the conditional claim that if their theory of the selection task is correct, then further research using this task can tell us nothing about specific or GR mechanisms. However, we suggest that the selection task may yet yield useful information about human reasoning. Oaksford and Chater (1995a) argued that all the experimental evidence in Sperber et al. (1995) could be reinterpreted in terms of manipulations of probabilities. Moreover, they argued that for the standard selection task, their information gain measure (Oaksford & Chater, 1994) was a measure of relevance. When Sperber et al's experimental manipulations were reinterpreted as manipulating probabilities, information gain provided as good an account of their results as relevance theory. Oaksford and Chater (1995a, p. 105) also observed that, "a further quantitative measure of relevance based on expected utilities … provides excellent fits to the data on the deontic selection task." That is, calculations of expected utility were considered part of the automatic interpretative process, which in turn guides selection task performance. In sum, we view the decision-theoretic approach as an alternative and richer theory of the interpretative component. Consequently, if the selection task can provide evidence for Sperber et al.'s theory of interpretation, it can provide evidence for the decision-theoretic account. Moreover, a richer theory of the interpretative component is clearly needed as neither text processing theory nor Sperber et al.'s account provide any mechanisms that can deal with emotional effects such as those we have described here.

7.3. Mental models theory

Mental models theory has long been applied to the deontic selection task (Johnson-Laird & Byrne, 1992, 1995, 2002). The idea behind the theory is that people make a distinction between what is factually possible or impossible, given a conditional is true, and what is deontically permissible given a deontic rule is in force. As in all mental models of the logical connectives, this involves representing the logical possibilities for a conditional, if *p* then *q*. Note that from the perspective participants were asked to adopt in Experiments 1 and 2, the *not-p, q* case is deontically impermissible according to an SC interpretation. Another way of thinking about this in the mental models illustrated in the following paragraphs is that we are considering the rule if *q* then *p.*

Logical			*Deontic*
	p	q	
Impossible (BC)	p	$\neg q$	Impermissible (HM)
Impossible (C&BC)	$\neg p$	q	Impermissible (SC)
	$\neg p$	$\neg q$	

From the perspective participants adopted in Experiments 1 and 2, a deontic SC rule suggests that the *not-p* and *q* case, for example, no pain but takes a break, is deontically impermissible. With respect to the rule if *q* then *p,* this is also the case that should be logically impossible on a conditional interpretation (C). On a biconditional interpretation (BC) the *p* and *not-q* is also logically impossible. According to the relevance and text processing theories people reinterpret the rule in this way and so people's responses are based on a logical interpretation. However, for mental models theory people are not simply reinterpreting the rules, they are engaging in a reasoning process specific to deontic rules. That is, they seek the *not-p* and *q* cases because they view it as deontically impermissible (Johnson-Laird & Byrne, 2002), rather than logically impossible.[4]

The first and most obvious criticism of this approach is that it would appear to be circular. As a definition of the meaning of deontic conditionals, deontic concepts appear on both sides of the definition. That is, "if *p* then (deontic) must *q* " means that the *p, not-q* case is labeled deontically impermissible (Manktelow & Over, 1995). Manktelow and Over (1995) provided a noncircular definition in terms of utilities such that "if *p* then (deontic) must *q* " holds if the expected utility of *q* given *p* is strictly higher than the expected utility of *not-q* given *p.* Such an account is defined with respect to the goals of society, rather than the goals of the enforcers or actors whose roles participants adopt in deontic reasoning tasks. So if the conditions Manktelow and Over (1995) outlined are met, it makes sense to enforce particular deontic rules. However, as they also point out, the utilities that are considered when checking for cheating are the same as those we have outlined.

Johnson-Laird and Byrne (1992) also used this mode of explanation to account for the occurrence of "all" responses in Manktelow and Over's (1991) experiments. They argued that both deontically impermissible cases, *p, not-q* (HM) and *not-p, q* (SC), may be active. Moreover, they argued that if one of these cases was deemed more serious, then just that one case would be active. If they were both considered of moderate seriousness, then card combinations compatible with looking for both impermissible cases may be chosen, that is, all cards. From the perspective participants are asked to adopt in Experiments 1 and 2, the SC case is more serious, and this is true in both the threat and the no-threat conditions. Nonetheless, presumably as the HM case becomes more serious, that is, more threatening, the closer in seriousness the two cases become. The closer they are in seriousness, the more likely it is that people will select all the cards. Moreover, given that the range of responses this theory can predict are binary and mutually exclusive (*not-p, q* or "all"), this should be accompanied by a reduction in *not-p, q* responses. Therefore, the mental models theory seems to predict the opposite effect to that observed in Experiments 1 and 2. That is, this pattern of responding must predict that the Pollard index should be lower in threat condition, not higher. Moreover, talk of the "seriousness" of the various impermissible outcomes allows values to be attached to possibilities just like utilities. However, seriousness does not capture the fact that threat invokes an avoidance response, and hence negative utility, thus mental models get this prediction the wrong way round.

Recent extensions of mental models theory have also suggested that people may annotate their mental models with probabilities over which they may perform some mental calculations (Johnson-Laird, Legrenzi, Girotto, Legrenzi, & Caverni, 1999). So to account for our results, perhaps mental models could also be annotated with utilities. This would also serve to make the utilities of these different possibilities explicit.

		Probabilities	Utilities
p	q	.25	C_A
p	$\neg q$	.25	B_H
$\neg p$	q	.25	B_S
$\neg p$	$\neg q$	.25	C_D

However, each row in this mental model corresponds directly to the cell values of the contingency table (A) and the payoff matrix (B) (assuming the maximum entropy distribution for the probabilities). Consequently, this extension would amount to simply implementing the decision-theoretic model directly in the mind and proposing that people perform the calculations in Equation 1 over this mental representation. We suspect that the relation between our simple decision-theoretic computational-level model and the mental representations and processes that implement it in the human mind are likely to be more complex than this.

In sum, just like relevance theory and text processing theory, mental models theory seems to make the wrong predictions concerning the effects of threat material. Moreover, an obvious extension to accommodate decision-theoretic effects seems to amount to an implausibly direct implementation of the decision-theoretic computational-level model.

8. Conclusion

The results of these experiments seem to show that people do attend to the emotional or threat values associated with possible outcomes in a way that is predictable from a decision-theoretic standpoint but not from the standpoint of current evolutionary psychology. Moreover, other theories of deontic reasoning also seem incapable of explaining these results. This is not to say, however, that evolutionary considerations are irrelevant to explaining the patterns of performance in tasks like this. The factors that affect the calculation of relevant utilities may well include innate as well as learned preferences along with socially and culturally related expectations, probably to do with rights and duties (Holyoak & Cheng, 1995c). However, it does seem that some form of general purpose mechanism may be needed to coordinate these different sources of constraint on deontic thought. Only further research will show.

Notes

1. The precise wording we have used for this rule is slightly different from that which appears in the literature. This was so that we could use it to illustrate all the variants of the deontic task. In the actual task as, for example, investigated by Cheng and Holyoak (1985), people were described as attempting to enter the country or as "in transit." Here we describe a situation where they have been permitted to enter or been denied entry. This was important to illustrate perspective shifts.
2. We thank Laurence Fiddick for suggesting the use of this index as the dependent variable.
3. Another possibility concerning the effect of threat arises from Cosmides and Tooby (1992). They argued that as rationed benefits are removed from an SC situation, SC responses should decline. So as the perceived benefit of the action in the antecedent decreases, the fewer SC responses should occur. Such an effect would predict fewer SC responses in the threat condition and so a lower Pollard index. That is, such an account, where the threat affects the input conditions to the SC module, makes exactly the same predictions as the control mechanisms we have considered in the text.
4. Johnson-Laird and Byrne (2002) did not distinguish between SCs and HM situations, so our inclusion of the HM case represents a minor extension of their account.

Acknowledgments

Nick Perham, School of Psychology, Cardiff University, Park Place, Cardiff, CF10 3AT, Wales, UK. Mike Oaksford, School of Psychology, Birbeck College London, Malet Street, London, WC1E 7HX, UK.

We thank Larry Fiddick, Vittorio Girotto, Keith Holyoak, David Over, Josh Tenenbaum, and an anonymous reviewer for their very useful and constructive comments on this article.

References

Almor, A., & Sloman, S. A. (1996). Is deontic reasoning special? *Psychological Review, 103,* 374–380.

Almor, A., & Sloman, S. A. (2000). Reasoning versus text processing in the Wason selection task: A nondeontic perspective on perspective effects. *Memory and Cognition, 28,* 1060–1070.

Anderson, J. R. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial processing tree modeling. *Psychonomic Bulletin & Review, 6,* 57–86.

Bell, D. E. (1982). Regret in decision making under uncertainty. *Operations Research, 30,* 961–981.

Bell, D. E. (1985). Disappointment in decision making under uncertainty. *Operations Research, 33,* 1–27.

Blanchette, I., & Richards, A. (2004). Reasoning about emotional and neutral materials: Is logic affected by emotion? *Psychological Science, 15,* 745–752.

Chater, N., & Oaksford, M. (1996). Deontic reasoning, modules and innateness: A second look. *Mind and Language, 11,* 191–202.

Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology, 17,* 391–416.

Cheng, P. W., & Holyoak, K. J. (1989). On the natural selection of reasoning theories. *Cognition, 33,* 285–313.

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

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments and Computers, 25,* 257–271.

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition, 31,* 187–276.

Cosmides, L., & Tooby, J. (1989). Evolutionary psychology and the generation of culture: II. Case study: A computational theory of social exchange. *Ethology and Sociobiology, 10,* 51–97.

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. H. Barkow & L. Cosmides (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 163–228). Oxford, England: Oxford University Press.

Cosmides, L., & Tooby, J. (1997). Dissecting the computational architecture of social inference mechanisms. In F. Ciba (Ed.), *Characterizing human psychological adaptations* (pp. 132–161). New York: Wiley.

Cosmides, L., & Tooby, J. (2000). Evolutionary psychology and the emotions. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of the emotions* (2nd ed., pp. 91–115). New York: Guilford.

Davidson, R. J. (1993). Parsing affective space: Perspectives from neuropsychology and psychophysiology. *Neuropsychology, 7,* 464–475.

Evans, J. S. B. T. (1983). Linguistic determinants of bias in conditional reasoning. *Quarterly Journal of Experimental Psychology, 35,* 635–644.

Fiddick, L. (2004). Domains of deontic reasoning: Resolving the discrepancy between the cognitive and moral reasoning literatures. Quarterly *Journal of Experimental Psychology, 57(A),* 447–474.

Fiddick, L., Cosmides, L., & Tooby, J. (2000). No interpretation without representation: The role of domain-specific representations and inferences in the Wason selection task. *Cognition, 77,* 1–79.

Fodor, J. (1983). *Modularity of mind.* Cambridge, MA: MIT Press.

Fodor, J. (2000). Why we are so good at catching cheaters. *Cognition, 75,* 29–32.

Fodor, J. (2002). Reply to Beaman. *Cognition, 83,* 221.

Fox, E., Russo, R., Bowles, R., & Dutton, K. (2001). Do threatening stimuli draw or hold visual attention in subclinical anxiety? *Journal of Experimental Psychology: General, 130,* 681–700.

Girotto, V., Kemmelmeier, M., Sperber, D., & van der Henst, J. B. (2001). Inept reasoners or pragmatic virtuosos? Relevance and the deontic selection task. *Cognition, 81,* B69–B76.

Hattori, M. (2002). A quantitative model of optimal data selection in Wason's selection task. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 55A,* 1241–1272.

Holland, J. H., Holyoak, K. J., Nisbett, R. E., & Thagard, P. R. (1986). *Induction: Processes of inference, learning and discovery.* Cambridge, MA: MIT Press.

Holyoak, K. J., & Cheng, P. W. (1995a). Pragmatic reasoning about human voluntary action: Evidence from Wason's selection task. In S. E. Newstead & J. S. B. T. Evans (Eds.), *Perspectives on thinking and reasoning: Essays in honour of Peter Wason* (pp. 67–89). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Holyoak, K. J., & Cheng, P. W. (1995b). Pragmatic reasoning from multiple points of view: A response. *Thinking and Reasoning, 1,* 373–389.

Holyoak, K. J., & Cheng, P. W. (1995c). Pragmatic reasoning with a point of view. *Thinking and Reasoning, 1,* 289–314.

Isen, A. M., & Patrick, R. (1987). The effect of positive feelings on risk taking: When the chips are down. *Organisational Behavior and Human Decision Processes, 31,* 194–202.

Jaynes, E. T. (1978). Where do we stand on maximum entropy. In R. D. Levine & M. Tribus (Eds.), *The maximum entropy formalism* (pp. 15–118). Cambridge, England: Cambridge University Press.

Johnson-Laird, P. N., & Byrne, R. M. (1992). Modal reasoning, models, and Manktelow and Over. *Cognition, 43,* 173–182.

Johnson-Laird, P. N., & Byrne, R. M. J. (1995). A model point of view. *Thinking and Reasoning, 1,* 339–350.

Johnson-Laird, P. N., & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review, 109,* 646–678.

Johnson-Laird, P. N., Legrenzi, P., Girotto, V., Legrenzi, M. S., & Caverni, J. P. (1999). Naive probability: A mental model theory of extensional reasoning. *Psychological Review, 106,* 62–88.

Johnson-Laird, P. N., & Wason, P. C. (1970). Insight into a logical relation. *Quarterly Journal of Experimental Psychology, 22,* 49–61.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decisions under risk. *Econometrica. 47,* 263–291.

Kirby, K. N. (1994a). False alarm: A reply to Over and Evans. *Cognition, 52,* 245–250.

Kirby, K. N. (1994b). Probabilities and utilities of fictional outcomes in Wason's four card selection task. *Cognition, 51,* 1–28.

Klauer, K. C. (1999). On the normative justification for information gain in Wason's selection task. *Psychological Review, 106,* 215–222.

Loehle, C. (2000). Global optimization 4.0 [Computer software]. Naperville, IL: Loehle Enterprises.

Loomes, G., & Sugden, R. (1982). Regret theory: An alternative of rational choice under uncertainty. *Economic Journal, 92,* 805–824.

Loomes, G., & Sugden, R. (1986). Disappointment and dynamic consistency in choice under uncertainty. *Review of Economic Studies, 53,* 271–282.

Loomes, G., & Sugden, R. (1987). Testing for regret and in choice under uncertainty. *Economic Journal, 97,* 118–129.

Lundh, L. G., Wikstroem, J., & Westerlund, J. (2001). Cognitive bias, emotion, and somatic complaints in a normal sample. *Cognition and Emotion, 15,* 249–277.

Manktelow, K. I., & Over, D. E. (1987). Reasoning and rationality. *Mind and Language, 2,* 199–219.

Manktelow, K. I., & Over, D. E. (1990). Deontic thought and the selection task. In K. Gilhooly, M. Keane, R. Logie, & G. Erdos (Eds.), *Lines of thought: Reflections on the psychology of thinking* (pp. 153–164). London: Wiley.

Manktelow, K. I., & Over, D. E. (1991). Social roles and utilities in reasoning with deontic conditionals. *Cognition, 39,* 85–105.

Manktelow, K. I., & Over, D. E. (1995). Deontic reasoning. In S. E. Newstead & J. S .B. T. Evans (Eds.), *Perspectives on thinking and reasoning: Essays in honour of Peter Wason* (pp. 91–114). Hove, Sussex, England: Lawrence Erlbaum Associates, Inc.

Manktelow, K. I., Sutherland, E. J., & Over, D. E. (1995). Probabilistic factors in deontic reasoning. *Thinking and Reasoning, 1,* 201–220.

Mineka, S., & Oehman, A. (2002). Phobias and preparedness: The selective, automatic, and encapsulated nature of fear. *Biological Psychiatry, 51,* 927–937.

Minsky, M. (1977). Frame system theory. In P. N. Johnson-Laird & P. C. Wason (Eds.), *Thinking: Readings in cognitive science* (pp. 355–376). Cambridge, England: Cambridge University Press.

Oaksford, M. (2002). Contrast classes and matching bias as explanations of the effects of negation on conditional reasoning. *Thinking and Reasoning, 8,* 135–151.

Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review, 101,* 608–631.

Oaksford, M., & Chater, N. (1995a). Information gain explains relevance which explains the selection task. *Cognition, 57,* 97–108.

Oaksford, M., & Chater, N. (1995b). Two and three stage models of deontic reasoning. *Thinking and Reasoning, 1,* 350–357.

Oaksford, M., & Chater, N. (2003a). Computational levels and conditional reasoning: Reply to Schroyens and Schaeken (2003). *Journal of Experimental Psychology: Learning, Memory & Cognition, 29,* 150–156.

Oaksford, M., & Chater, N. (2003b). Conditional probability and the cognitive science of conditional reasoning. *Mind and Language, 18,* 359–379.

Oaksford, M., & Chater, N. (2003c). Modelling probabilistic effects in conditional inference: Validating search or conditional probability? *Revista Psychologica, 32,* 217–242.

Oaksford, M., & Chater, N. (2003d). Optimal data selection: Revision, review and re-evaluation. *Psychonomic Bulletin and Review, 10,* 289–318.

Oaksford, M., & Chater, N. (2003e). Probabilities and pragmatics in conditional inference: Suppression and order effects. In D. Hardman & L. Macchi (Eds.), *Thinking: Psychological perspectives on reasoning, judgment and decision making* (pp. 95–122). London: Wiley.

Oaksford, M., Chater, N., & Larkin, J. (2000). Probabilities and polarity biases in conditional inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 883–899.

Oaksford, M., Morris, F., Grainger, B., & Williams, J. M. G. (1996). Mood, reasoning, and central executive processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 476–492.

Oaksford, M., & Stenning, K. (1992). Reasoning with conditionals containing negated constituents. *Journal of Experimental Psychology: Learning, Memory & Cognition,* 18, 835–854.

Oaksford, M., & Wakefield, M. (2003). Data selection and natural sampling: Probabilities do matter. *Memory and Cognition, 31,* 143–154.

Oehman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review, 108,* 483–522.

Perham, N. R. (2003). *Anticipated and experienced emotion in deontic reasoning.* Unpublished doctoral dissertation, Cardiff University, Cardiff, Wales.

Pollard, P., & Evans, J. St.B. T. (1987). Content and context effects in reasoning. *American Journal of Psychology, 100,* 41–60.

Sarlo, M., Palomba, D., Angrilli, A., & Stegagno, L. (2002). Blood phobia and spider phobia: Two specific phobias with different autonomic cardiac modulations. *Biological Psychology, 60,* 91–108.

Shanks, D. R., Tunney, R. J., & McCarthy, J. D. (2002). A re-examination of probability matching and rational choice. *Journal of Behavioral Decision Making, 15,* 233–250.

Siegel, S., & Castellan, N. J. (1988). *Non-parametric statistics for the behavioural sciences.* New York: McGraw-Hill.

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

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

Wason, P. C. (1966). Reasoning. In B. Foss (Ed.), *New horizons in psychology* (pp. 135–151). Harmondsworth, Middlesex, England: Penguin.

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

Wolf, F. M. (1986). *Meta-analysis: Quantative methods for research synthesis.* Newbury Park, CA: Sage.

Wolfram, S. (1991). Mathematica [Computer software]. Reading, MA: Addison-Wesley.

Zeelenberg, M., van Dijk, W. W., Manstead, A. S. R., & van der Pligt, J. (2000). On bad decisions and disconfirmed expectancies: The psychology of regret and disappointment. *Cognition and Emotion, 14,* 521–541.

Appendix 1

This appendix contains the descriptions of the sports and butcher's scenarios and the instructions used in these experiments. The scenarios were identical between the two experiments except where indicated. Comments on these instructions are in italics.

Sports Scenario (Actor's Perspective: Experiments 1 and 2)

> You are the manager of your local team and are in the final of the cup. You are confident about your team's commitment and ability but have reservations about the referee. In the past this particular referee has made some incorrect decisions concerning injuries to players: Players have been made to leave the pitch when they did not need to. Obviously if your players are made to leave the pitch when they should not, it weakens your team. As a result you ask a member of your staff to record the referee's decisions. If the referee has not been fair you can have a word with the match officials. A rule the referee is supposed to follow is:

IF THERE IS A BLOOD INJURY THEN THE PLAYER MUST LEAVE THE PITCH.

You are just about to go out for the second half when the member of staff says that they only noted four decisions. However, as you are rushing out you only hear one piece of information about each decision—the type of injury or whether the player left the pitch.

On the next page you will be briefly presented with the piece of information you can see from the four forms.

After viewing all four pieces together, you will see each piece individually. During this time you will be asked whether you think you must ask for the other piece of information to determine whether the referee has been weakening your team unnecessarily.

If you think you must ask for the other information then press "Y". If you do not think you must ask for the other information then press "N".

If you have any questions please ask the experimenter. Press the space bar to continue.

In the no-threat scenario the word "blood" in the task rule was replaced with "head," otherwise the tasks were identical. After viewing the scenario participants were shown all four cards for a period of 6 sec in the following order: p, q, not-p, and not-q:

BLOOD ANKLE LEAVE STAY

One might argue that envisioning the cheating case, "leaving the pitch without a blood injury," is made difficult by the use of "ankle injury" as the exemplar of not having a blood injury. How could someone with an ankle injury stay on the pitch? In football and rugby, injuries, other than blood injuries, can be treated on the pitch. When this happens, the match is stopped, and most often the player continues to play after treatment. This means that the best player stays on the pitch for the duration of the match. However, when there is a blood injury, the player is sent off and the game continues without that player, or a second-string player is bought on as a "blood" substitute. That is, there is a clear and obvious cost involved if players are sent off without a blood injury. This issue does not arise for the remaining scenarios.

Participants are then shown each of the four cards one at a time in random order and are asked to make a decision whether to turn a card over:

BLOOD	Do you think you must ask whether the player left the pitch?
	Yes or no?
ANKLE	Do you think you must ask whether the player left the pitch?
	Yes or no?
LEAVE	Do you think you must ask what injury the player had?
	Yes or no?
STAY	Do you think you must ask what injury the player had?
	Yes or no?

They pressed the Y button for "Yes" and the N button for "No." On completion participants were debriefed and thanked for participating.

A further issue with these materials is that it could be argued that the no-threat condition, where "blood" in the antecedent is replaced with "head," may be more threatening than the threat condition. This is because, intuitively, head injuries may be viewed as more threatening than blood injuries. However, as we have already indicated, from the SC perspective, people's goal is not to detect people at risk of injury but to detect cheaters. Thus considering the relative risks involved with blood and head injuries is irrelevant to their task. In contrast, as we have argued, the automatic fear response to the word "blood," would be expected to modulate B_H in a way that "head," which does not evoke a fear response, could not.

Sports Scenario (Enforcer's Perspective: Experiment 3)

You are the manager of your local team and are in the final of the cup. You are confident about your team's commitment and ability but have reservations about the referee. In the past this particular referee has made some incorrect decisions concerning injuries to players. Obviously, if your players are injured they should receive the appropriate treatment. As a result you ask a member of your staff to record the referee's decisions. If the referee has not been fair, you can have a word with the match officials. A rule the referee is supposed to follow is

Participants were also asked to select cards "to determine whether players receive the correct treatment."

The rest of this scenario was the same as in Experiments 1 and 2.

Butcher's Scenario (Actor's Perspective: Experiment 1 and 2)

(This scenario used the same rule as Manktelow & Over, 1991.)

You are in charge of a butcher's department in a superstore, and recently you have been under pressure to reduce your expenditure. One way you have found to reduce this is by limiting the amount of gloves used for cleaning. However, because of hygiene regulations your employees have to follow the rule:

IF YOU CLEAN UP BLOOD THEN YOU MUST WEAR GLOVES.

One day you ask your supervisor to record employees cleaning as you are concerned that some workers may be using the gloves unnecessarily. Just as you are about to see the general store manager about the effectiveness of your reduction, you are told the information about four employees. Unfortunately you only hear one piece of information about each person: what was being cleaned up or what was being worn on their hands.

On the next page you will be briefly presented with the piece of information you heard.

After viewing all four pieces together you will see each piece individually. During this time you will be asked whether you think you must ask for the other piece of information to determine whether the employee was unnecessarily using the gloves.

If you think you must ask for the other information then press "Y". If you do not think you must ask for the other information then press "N".

If you have any questions please ask the experimenter. Press the space bar to continue.

BLOOD WATER GLOVES BARE

BLOOD Do you think you must ask what the employee was wearing on his or her hands?

Yes or no?

WATER Do you think you must ask what the employee was wearing on his or her hands?

Yes or no?

GLOVES Do you think you must ask what the employee was cleaning?

Yes or no?

BARE Do you think you must ask what the employee was cleaning?

Yes or no?

For the no-threat condition, the word "blood" was replaced by the word "packaging" in both the rule and the response array.

Butcher's Scenario (Enforcer's Perspective: Experiment 3)

You are in charge of a butcher's department in a superstore and are soon to have a visit from Health and Safety. Because of this you implement some new hygiene policies. One rule your employees have to follow is

IF YOU CLEAN UP BLOOD THEN YOU MUST WEAR GLOVES.

One day you ask your supervisor to record employees cleaning as you are concerned that some workers may not be following the rule. Just as you are about to see the general store manager about the effectiveness of your implementations, you are told the information about four employees. Unfortunately you only hear one piece of information about each person: what was being cleaned up or what was being worn on their hands.

On the next page you will be briefly presented with the piece of information you heard.

After viewing all four pieces together you will see each piece individually. During this time you will be asked whether you think you must ask for the other piece of information to determine whether the employee was using the correct hand-wear.

The rest of this scenario was the same as in Experiments 1 and 2.

Appendix 2

The following instructions were given before the threat ratings task. Participants indicated responses on a 5-point Likert-type scale; these responses were then rescaled as a percentage.

Please, could you fill in the following pieces of information:

AGE____GENDER____

Have you ever been taught logic or reasoning to such an extent that it could have influenced your responses on the previous tasks? YES NO

Please, could you rate the following words on the basis of how threatening you think they are.

For example, when presented with the word *table,* if you think that this word is quite threatening to you, then you indicate that on the following scale (as indicated by the letter *x*):

I——————I——————I——————I-x——————I

Definitely not threatening	Not very threatening	Neutral	Quite threatening	Very threatening

If, however, you think that this word is not very threatening to you then you indicate that on the following scale (as indicated by the letter *x*):

I——————I-x——————I——————I——————I

Definitely not threatening	Not very threatening	Neutral	Quite threatening	Very threatening

If you have any questions about this procedure please ask the experimenter.