Eye-Movements and Reasoning: Evidence for Relevance Effects and Rationalisation Processes in Deontic Selection Tasks

Linden J. Ball (l.ball@lancaster.ac.uk)
Department of Psychology, Lancaster University,
Lancaster, LA1 4YF, UK

Erica J. Lucas (erica.lucas@ntu.ac.uk)
Faculty of Economics and Social Sciences, Nottingham Trent University,
Nottingham, NG1 4BU, UK

Peter Phillips (phillips@comp.lancs.ac.uk)
Department of Computing, Lancaster University,
Lancaster, LA1 4YF, UK

Abstract

Evans' (e.g., 1996) Heuristic-Analytic theory of reasoning in Wason's selection task proposes the existence of implicit processes that direct attention to 'relevant' aspects of the problem and thereby determine card selections. This account also proposes that people pursue explicit rationalisations of relevance-determined choices. Our recent studies (e.g., Ball et al., 2003) have measured aspects of on-line attentional processing using eye-movement tracking and have supported the idea that card selections are driven by relevance and then subjected to rationalisation processes. For example, eye-movement data have revealed a reliable inspection-time imbalance between selected and non-selected cards. Our results, to date, however, have related to selection tasks with indicative contents. Here we report an eye-tracking study that involved deontic selection tasks. Various eye-movement measures revealed predicted differences in inspection times between selected and non-selected cards, with the magnitude of the effect being similar to that observed in studies of the indicative task. We discuss our results in relation to current theories of processing in the selection task.

Keywords: Wason Selection Task; eye movements; inspection times; deontic rules; relevance; rationalisation.

Introduction

Reasoning has long been of interest to cognitive scientists, and one task continues to attract research attention: Wason's four-card selection task (Wason, 1966). It has become increasingly apparent that there are, in fact, two distinct forms of selection task - 'indicative' and 'deontic' - depending on what kind of conditional rule and instructions are used in studies. Indicative conditionals express factual, scientific or common-sense knowledge, whereas deontic conditionals specify social regulations, laws and moral rules. Our interest in the present paper relates primarily to people's reasoning with deontic variants of the selection task. However, the fact that theories strive for a generic account of responding that cuts across both indicative and deontic versions means that it is valuable to begin by surveying key phenomena associated with both task variants. To this end we begin with an overview of findings relating to indicative and deontic selection tasks as well as a discussion of theories that have attempted to generalise their core assumptions across both task formats.

Indicative Selection Tasks and Matching Bias

In its typical indicative form participants are presented with an array of four cards, and are told that each card has a letter on one side and a number on the other side (only the facing sides are visible). Participants are then given a conditional rule that they are told applies to the cards, and their task is to decide which cards need to be turned over in order to determine whether the rule is true or false. For example, the rule might be *If there is an A on one side of the card then there is a 3 on the other side of the card*, and the cards might be 'A' 'J', '3' and '7'. These are referred to as the *p*, *not-p*, *q* and *not-q* cards, respectively. The logically correct response for a conditional reading of the rule is to turn the A (*p*) and the 7 (*not-q*) cards, as these could provide a letter-number combination that would show the rule to be false. Most people select either A (*p*) alone, or A (*p*) and 3 (*q*).

Evans and Lynch (1973) proposed that people's selections indicate a systematic *matching bias*, that is, they simply choose cards that are named in the given rule. Matching bias is a robust phenomenon in selection tasks employing abstract rules and generalised across various connectives, such as: *if p then q*, *q if p*, *p only if q*, and *there is not both p and q* (e.g., Evans, Clibbens, & Rood, 1996; Evans, Legrenzi, & Girotto, 1999). Despite the generality of the matching phenomenon there remains some contention as to how best to account for it. We will examine theories of the matching effects on indicative selection tasks below, but first we describe some key phenomena associated with deontic versions of the selection task that form the focus of our current experiment.

Deontic Selection Tasks and Pragmatic Effects

Johnson-Laird, Legrenzi, and Legrenzi (1972) were the first to study a deontic selection task using the rule *If a letter is sealed then it has a 50 lire stamp on it.* Participants played the role of postal workers, sorting mail and checking for rule violations. Presented 'cards' were real envelopes: a sealed letter - *p*; an unsealed letter - *not-p*, a letter with a L50 stamp - *q*, and one with a L40 stamp - *not-q*. The violating *p*, *not-q* instance was a sealed letter with a lower-value stamp on it, and 21 of 24 participants selected cards that could reveal this combination compared with 2 of 24 in the indicative letter-number version.

Although this effect was dubbed 'thematic facilitation' (as thematic content appeared to enhance logical performance) it is now known that the effect of thematic material is more complex. There are two reasons for this complexity (Evans, 1996). First, it is not the case that thematic content always alters selection frequencies. When materials are 'arbitrarily realistic' participants seem to succumb to the same matching bias as seen with indicative rules (see Manktelow & Evans', 1979, study using rule such as *Every tine I eat haddock then I drink gin*). Second, even when thematic content does change card-selection patterns it does not always produce more logical choices. Studies using 'social contract rules' with shifted perspectives have shown that a *not-p*, *q* pattern can be induced instead (e.g., Cosmides & Tooby, 1989; Gigerenzer & Hug, 1992). Manktelow and Over (1991), for example, used a rule given by a mother to a son: *If you tidy your room then you may go out to play.* The four cards represented different days - showing on one side whether the room was tidied and on the other whether the son played out. When participants adopted the role of the son checking whether the rule had been followed they selected logically correct 'Room tidied' (*p*) and 'Did not go out to play' (*not-q*) cards. However, when asked to check the rule from the perspective of the mother, the typical response was to select 'Room not tidied' (*not-p*) and 'Went out to play' (*q*). Manktelow and Over (1991) propose that people may be selecting cards because of pragmatic influences rather than because they are reasoning logically (e.g., striving to satisfy personal goals such as detecting whether cheating is arising in a social-contact situation).

Theories of the Selection Task

Three key reasoning theories have been applied to selection tasks involving both indicative and deontic conditionals.

Heuristic-Analytic Theory Evans' (e.g., 1996) *heuristic-analytic* (H-A) theory is a general theory of reasoning, It proposes that reasoning typically involves two processing stages. First, implicit, pre-conscious, heuristics determine which aspects of a task are of psychological *relevance*, thereby enabling attention to be selectively focused on these task features. Second, explicit, conscious, analytic processes are applied to these relevant task features to enable an inference or judgment to be made.

Evans' (1998) specific account of matching bias on the indicative task is that it arises from the operation of a linguistically-based 'matching heuristic' that reflects the way that negative terms are used in natural language to deny suppositions rather than to assert new information. Furthermore, any analytic processing that is applied to cards serves merely to *rationalise* decisions that have already been made on the basis of relevance (Evans, 1995). This H-A account of the indicative task is clearly *attentional* in emphasis. Linguistic cues draw attention toward certain cards and away from others; the former get selected and the latter get rejected. In this way, the H-A account can readily make sense of findings (Evans & Wason, 1976) that the retrospective verbal reports people provide when asked to explain card selections reflect attempts to justify choices in terms of either verification or falsification (depending on the nature of the rule), with no apparent insight into the logical basis of selections (see also Lucas & Ball, 2005).

What, however, about a H-A account of selection patterns on deontic variants of the selection task? Evans and Over (1996) argue that in explaining deontic task performance it is important to appreciate that the notion of heuristically-determined relevance is related to the current *goal* that the participant is pursuing. As they say, "Information will be more relevant, the greater extent to which it bears upon the chances of achieving the current goal" (Evans & Over, p. 81). Thus, in deontic tasks, any card that could reveal an outcome costing the participant something (e.g., extra money for a sealed letter or prevention from playing out) will have high 'goal relevance'. Evans and Over propose that heuristic processes mean that these costly outcomes become the focus of attention, such that individuals choose cards that could reveal whether such an outcome holds. Instructional and scenario changes can, of course, alter what the participant will see as costly outcomes for a task. Thus, perspective effects can be accommodated within this extended notion of heuristically-cued relevance. Under a H-A account of deontic task performance, rationalisation processes would be employed to justify choices in an equivalent manner to that which arises in the indicative paradigm (Evans, 1996)

Mental Models Theory Johnson-Laird and Byrne's (e.g., 1991) *mental models* theory (MMT) proposes that reasoning is based around the construction of models in which the premises of a rule are represented as being true. To explain selection-task performance with indicative rules, MMT assumes that people: (1) only think about cards that are *explicitly* represented in their models of the rule, and (2) only select those cards for which the hidden value could bear on the truth or falsity of the rule. So, for example, the failure to select the *not-q* card on *If p then q* reflects the fact that this term is not explicitly represented in the reasoner's models of the conditional, thereby resulting in the common selection of just the *p* card. Those people who represent this rule as a biconditional would select *p* and *q* (the other dominant selection combination) as both of these cards are explicitly represented in models and could bear on the rule's truth or falsity. To account for matching bias on rules that contain negations, Johnson-Laird and Byrne (1991) argue that negated components promote the expansion of models to include the affirmative counterparts of negated terms.

The MMT also accounts for why deontic selection tasks facilitate *not-q* selections (Johnson-Laird, 1995). According to MMT, selection of this card will only occur if the models of the conditional are fleshed out to represent it explicitly. It is proposed that real-world knowledge can promote such explicit modelling of the *not-q* case. This 'pragmatic modulation' effect (Johnson-Laird & Byrne, 2002) can arise from rule contents triggering memories of violations or analogous events, or through provision of a familiar deontic framework for interpreting the rule. Perspective effects with deontic rules can likewise be explained by the idea that such rules tend to get represented as bi-conditionals, and that contextual cues will determine which counterexample model the reasoner will focus their attention on.

The MMT of the selection task has some similarities to the H-A account. For example, the concept of explicit representation in models overlaps with the H-A notion of relevance (i.e., what is explicitly represented in a model is what is perceived to be relevant to the task at hand). However, MMT differs critically from the H-A account in its assumption that a degree of analytic processing *does* determine card selections (i.e., the only cards that end up being chosen are those explicitly represented ones that are deemed to bear on the rule's truth or falsity).

Optimal Data Selection Theory Oaksford and Chater's (e.g., 1994, 1996) theory of indicative selection-task performance is framed within a rational-analysis approach to human reasoning, and is referred to as the *optimal data selection* (ODS) account. Oaksford and Chater propose that selections are based on the information value of cards in relation to their potential support for the rule, estimated in the form of *expected information gain*. Their mathematical analysis of the information value of cards shows, for example, that the selection of the matching *q* card for the indicative conditional can be more useful than the selection of the non-matching (but logically appropriate) *not-q* card. In this way, the ODS model proposes that supposedly illogical matching choices on indicative rules may, in fact, be deemed to be rational in terms of a probabilistic standard.

The ODS theory presents a persuasive account of the matching effects observed on affirmative, indicative rules within the selection task. Moreover, because the ODS theory capitalises on Oaksford and Stenning's (1992) arguments that negations typically define high-probability contrast sets, it is also able to explain antecedent and consequent matching effects observed for conditional rules containing *negated* constituents (e.g., Oaksford, 2002). A version of ODS theory has also been formulated to explain effects seen with deontic selection tasks (e.g., Oaksford & Chater, 1994, although in the case the theory adopts a 'decision-theoretic' framework (cf. Manktelow & Over, 1991; and see also Evans & Over's H-A account of deontic tasks, described above). This framework refers to selections being determined by judgements of *expected utility* rather than of expected information gain. The concept of utility in decision theory reflects the value that people place on outcomes. In deontic selection tasks the instructions cue people to place a high value on instances of unfairness (e.g., taking a benefit without paying a required cost, such as posting a sealed letter without having placed a higher-value stamp on the envelope). It turns out that cards with the greatest expected utilities in deontic selection tasks are also those that tend to get picked.

Overall, the ODS account of the indicative selection task can accommodate evidence for matching bias, as well as the growing body of evidence for probabilistic influences on card selections (cf. Oaksford & Wakefield, 2003). Oaksford and Chater's decision theoretic account of deontic tasks is closely related to Evans' relevance account (although it does not embody an analytic rationalisation process), and also makes equivalent predictions to the MMT.

Tests of Selection-Task Theories

Roberts (1998) noted that only limited attempts have been made to test theories of selection-task performance using converging evidence beyond card-selection patterns. What effort has been made in this respect has primarily been by those working within the H-A tradition. For example, Evans, Ball and Brooks (1987) used computer-presented indicative tasks and recorded the order in which 'select-don't select' decisions were made about each card. As predicted by the H-A account, people made decisions about matching cards before mismatching ones, and a correlation was found between card selection frequency and card decision order (i.e., selected cards were decided about earlier than rejected ones). However, it is possible that these results merely reflect a preference for people to register 'select' before 'don't select' choices, and that what is being shown is a response bias as opposed to an attentional bias.

Evans (1996) provided stronger evidence for the H-A account in studies using a mouse-pointing methodology to record *card-inspection times*. Participants tackled computer-generated selection tasks (both indicative and deontic) and indicated a card they were 'thinking about' by holding the mouse pointer over it. Cards were selected via a mouse-click; no action was required for non-selected cards. The computer logged cumulative inspection times for each card. Evans argued that if heuristic processes were cueing card selections, then only heuristically-cued cards would be subjected to analytic rationalisation aimed at justifying their selection. Inspection times would, therefore, be higher for selected cards than rejected cards. Evans (1996) found good support for this novel prediction.

Evans' (1996) inspection-time prediction is both risky (i.e., it could have readily been falsified) and derives solely from the H-A framework (cf. Ball et al., 2003). The MMT, for example, argues that people should consider (i.e., inspect) cards that they end up rejecting because they have no bearing on the rule's truth or falsity - even though they were represented in models. Rejecting considered cards should break the link between card selection and increased inspection time, undermining any possible emergence of an inspection-time effect seen by Evans (1996). Likewise, Oaksford and Chater's probabilistic approach does not, a priori, predict an inspection-time effect, since, presumably, the computation of expected information gain values (in indicative tasks) or expected utilities (in deontic tasks) should take a similar amount of time for selected or non-selected cards. Without the addition of an analytic rationalisation component, Oaksford and Chater's probabilistic framework, although extremely compelling, seems unable to accommodate Evans' (1996) evidence for inspection-time effects in indicative and deontic tasks.

Despite this converging evidence for the H-A account, Roberts (1998) noted a need for caution in interpreting findings from the inspection-time paradigm. He argued that there are potential sources of bias inherent in this paradigm that could have led to artefactual support for the claim that people are spending time rationalising choices that have been cued by relevance-determining heuristics. For example, participants might pause the mouse pointer briefly over a card before making an *active* 'select' decision (by clicking on it). This would lead to inflated inspection times for selected cards relative to non-selected ones. Roberts (1998) manipulated the presence of such task-format biases across a series of experiments, and demonstrated that the magnitude of the inspection-time effect was closely related to the number of sources of bias present. With all sources removed the inspection-time effect was also eradicated.

Recently, Ball, Lucas, Miles and Gale (2003) critiqued inspection-time studies using mouse-pointing because of their use of an inherently insensitive technique for monitoring the second-by-second transitions in attentional processing that arise during selection-task performance. Of particular concern is the fact that mouse pointing is an indirect measure of attentional processing since participants have actively and effortfully to move the mouse pointer to cards that they are thinking about. Ball et al. advocate the use of eye-movement tracking as a more precise approach for measuring moment-by-moment attentional shifts that underlie cognitive performance with highly display-based problems like the selection task. Ball et al. report three experiments that systematically eradicated the sources of artefact discussed by Roberts (1998) by combining careful task constructions with eye-movement tracking to measure directly on-line processing. All three experiments produced good evidence for the robustness of the inspection-time effect, so supporting the predictions of the H-A account.

Despite the general support for key tenets of the H-A account of selection task that derive from Ball et al.'s (2003) eye-tracking research, it is noteworthy that their experiments focused solely on indicative selection tasks. It remains therefore, to clarify whether similar inspection-time effects can be observed in deontic tasks.

Method

Participants were 43 undergraduates at Lancaster University who received payment for taking part in the study. None had received teaching on reasoning or logic. The study involved computer-based presentation of problems. Participants were presented with general instructions as to the requirements of the study and an example of the task format based on Manktelow and Evans' (1979) *If I eat haddock then I drink gin* rule. They then tackled four target selection tasks. Each involved presentation of the task scenario and associated rule in the top third of the screen, with the cards depicted in a two-by-two arrangement in the lower two-thirds of the screen. Under each card were 'yes' and 'no' radio buttons as participants had to make decisions for all cards using the mouse pointer. The location of cards was randomised for each presentation, as was the order of tasks for participants. To avoid confounding our results we standardised the amount of information appearing on all cards across all rules so that it was exactly two items. For each task an on-screen reminder was also given of the task requirements.

The tasks used in the study were derived from Evans (1996) as follows: (1) A 'weak facilitator' task based on a 'town and transport' problem - *If I go to Burton, then I go by train* – and showing the cards 'To Burton', 'To Ashby', 'By Train' and 'By Car'; (2) A 'strong facilitator' task based on a 'drinking age' rule - *If a person is drinking alcohol, then that person must be over 18 years old* - with cards showing the words 'Drinking Beer', 'Drinking Coke', '22 Years', and '16 Years'; (3) A 'social contract' rule - *If you tidy your room, then you may go out to play* – with the request to reason from the child's perspective and the cards showing the words 'Room Tidied', 'Room Left', 'Played Out', and 'Kept In'); and (4) a 'switched social-contact' rule - *If a customer spends over £100, then they may take a free gift* – with the request to reason from a store-detective's perspective and with the cards showing the words 'Took Gift', 'Left Gift', 'Spent £120', and 'Spent £90'.

Whilst participants tackled the computer-presented selection tasks their eye movements were recorded using an LC Technologies EyeGaze system that determines gaze direction using the pupil-centre/corneal-reflection method. The tracker consists of a standard desktop computer running Windows NT/2000, an infrared camera mounted beneath the monitor, and software to process the resulting data. The tracker is accurate to within 0.45 degrees of visual angle, which, at 50 cm from the screen, covers approximately 0.38cm (12.8 pixels). Eye movements are sampled at 60 Hz and participants used a chin-rest to minimise data loss. Fixations were detected at a threshold of 100 ms or above.

Results and Discussion

Card Selection Frequencies

Our first concern was whether card selection frequencies (Table 1) conformed to established selection patterns seen in the literature. For the weak facilitator we would expect relatively high *p* and *q* selections, and this is exactly what was observed. For the strong facilitator and the social contract rules a *p*, *not-q* selection pattern should dominate; this indeed seemed to be the case for the strong facilitator. However, for the social contract rule, although selection of *not-q* was higher than that seen for the weak facilitator, it was the *q* card that dominated consequent choices. This is not entirely consistent with the literature, and suggests a remaining susceptibility to matching bias on this rule. Finally, on the switched social contract task we would expect an *not-p, q* choice pattern; this is what was observed. Overall, these tasks produced broadly similar response patterns to those found by others (e.g., Evans, 1996).

Table 1. Percentage frequencies of selections across rules.

Rule	Logical case			
	p	not-p	q	not-q
Weak facilitator	100	9	61	37
Strong facilitator	91	12	21	77
Social contact	81	33	64	49
Switched contract	35	70	93	16

Card Inspection Times

Cumulative inspection-time scores were positively skewed such that inferential analysis was conducted on log-transformed data. The critical prediction underpinning our analysis was that there should be a difference in the mean inspection times for participants' selected versus non-selected cards, with the former showing longer inspections times than the latter. This prediction derives from the H-A account of relevance effects and rationalisation processes in the selection task (Evans, 1996). We assessed this prediction at the participant level (i.e., by comparing participants' mean inspection times for cards that they selected versus those that they rejected). This participant-level analysis affords the advantage of increased test power (Roberts, 1998; see also Ball et al., 2003; Lucas & Ball, 2005) relative to an equivalent item-based analysis that Evans (1996) pursued in his original mouse-based inspection-time study.

Table 2: Mean inspection times (in seconds) by participants for selected and non-selected cards.

	Selected		Non-selected	
	Mean	SD	Mean	SD
ND	4.14	1.64	3.62	1.55
TD(log)	0.57	0.15	0.51	0.17
TD	3.52		3.04	

Note. ND = natural data; TD(log) = transformed data (in $\log_{10}$ units); TD = transformed data (in original units).

Mean inspection times (before and after transformation) are presented in Table 2. ANOVA revealed a significant inspection-time difference for selected versus non-selected cards, $F(1, 42) = 13.33$, $MSE = 0.08$, $p < .001$, lending good support to H-A predictions in the context of the present thematic and deontic selection tasks. It is noteworthy, however, that the magnitude of the inspection time imbalance is small, at around 0.5 s. This inspection time effect is, in fact, very similar to that seen in studies of the indicative selection task (Ball et al., 2003), and Lucas and Ball (2005) have argued that rationalisation may actually take place rapidly as the explicit consideration of what may be on the reverse sides of to-be-selected cards seems also to be cued by relevance processes (e.g., secondary matching bias can guide rationalisation in indicative selection tasks).

Other Measures of On-Line Processing

In addition to logging cumulative card inspection times we were also able to use the eye-tracker to monitor aspects of on-line processing relating to the number of fixations on cards. Of particular interest in the context of the H-A theory of the selection tasks were measures of: (1) the frequency of fixations on each card; and (2) the frequency of 'revisitations' to a previously fixated card. Such measures have not previously been analysed in the study of the selection task, although one would predict that they should reveal a similar imbalance between selected versus non-selected cards to that observed in the case of the standard inspection-time measure. This is because it should only be those cards (i.e. selected ones) that are subjected to rationalisation that should receive more fixations (and, potentially, more revisitations) during the process of justifying heuristically-determined choices.

Table 3: Mean fixation frequency and revisitation frequency by participants for selected and non-selected cards.

	Selected		Non-selected	
	Mean	SD	Mean	SD
Fixation frequency				
ND	13.50	4.47	12.05	4.50
TD(log)	1.10	0.13	1.05	0.14
TD	11.59		10.22	
Revisitation frequency				
ND	5.75	1.88	5.37	1.89
TD(log)	0.78	0.12	0.76	0.12
TD	5.03		4.75	

Note. ND = natural data; TD(log) = transformed data (in $\log_{10}$ units); TD = transformed data (in original units).

Mean data for these fixation-based measures are shown in Table 3. Again, since scores were positively skewed for all cards we pursued ANOVAs on log-transformed data. These analyses revealed a reliable difference for selected versus non-selected cards on the fixation frequency measure, $F(1, 42) = 9.01$, $MSE = 0.0063$, $p = .005$, and a near significant difference on the revisitation frequency measure, $F(1, 42) = 3.37$, $MSE = 0.004$, $p = .074$. These data seem interesting in that they indicate that processing differences between selected and rejected cads impact not only on the duration of card processing but also on the tendency for people to re-fixate and revisit cards that end up being selected. This observation, again, can perhaps best be interpreted as arising from the application of rationalisation processes that serve to justify heuristically-cued selection decisions.

Conclusion

Our eye-movement data have demonstrated that the inspection-time imbalance between selected and rejected cards that is observed with indicative selection tasks (Ball et al., 2003) generalises to deontic versions of the task. This inspection-time effect is directly as predicted under Evans' (e.g., 1986) heuristic-analytic account of the selection task, where implicit heuristic processes direct attention to relevant aspects of the problem and determine card selections, whilst analytic processes only serve to rationalise choices that have been cues by relevance. Although we believe that mental-models theory can explain such an effect in relation to deontic selection tasks we remain concerned about its capacity to explain the effect in indicative tasks (cf. Lucas & Ball, 2005). The probabilistic approach advanced by Oaksford and Chater (1994, 2003), whilst very impressive in its capacity to predict selection patterns across a range of indicative and deontic manipulations, seems curiously silent as to how to account for inspection-time effects. We would welcome the development of this powerful theory to accommodate such observations (see Lucas & Ball, 2005, for some initial steps in this direction).

References

Ball, L. J., Lucas, E. J., Miles, J. N. V., & Gale, A. G. (2003). Inspection times and the selection task: What do eye-movements reveal about relevance effects? *Quarterly Journal of Experimental Psychology, 56A,* 1053-1077.

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

Evans, J. St. B. T (1995). Relevance and reasoning. In S.E. Newstead and J. St. B. T. Evans (Eds.), *Perspectives on thinking and reasoning: Essays in honour of Peter Wason* (pp. 147-172). Hove: Lawrence Erlbaum Associates.

Evans, J. St. B. T. (1996). Deciding before you think: Relevance and reasoning in the selection task. *British Journal of Psychology, 87,* 223-240.

Evans, J. St. B. T. (1998). Matching bias in conditional reasoning: Do we understand it after 25 years? *Thinking and Reasoning, 4,* 45-82.

Evans, J. St. B. T., Ball, L. J., & Brooks, P. G. (1987). Attentional bias and decision order in a reasoning task. *British Journal of Psychology, 78,* 385-394.

Evans, J. St. B. T., Clibbens, J., & Rood, B. (1996). The role of implicit and explicit negation in conditional reasoning bias. *Journal of Memory and Language, 35,* 392-409.

Evans, J. St. B. T., Legrenzi, P., & Girotto, V. (1999). The influence of linguistic form on reasoning: The case of matching bias. *Quarterly Journal of Experimental Psychology, 52A*, 185-216.

Evans, J. St. B. T., & Lynch, J. S. (1973). Matching bias in the selection task. *British Journal of Psychology, 64,* 391-397.

Evans, J. St. B. T., & Over, D. E. (1996). *Rationality and reasoning.* Hove: Psychology Press.

Evans, J. St. B. T., & Wason, P. W. (1976). Rationalization in a reasoning task. *British Journal of Psychology, 67,* 479-486.

Gigerenzer, G., & Hug, K. (1992). Domain-specific reasoning: Social contracts, cheating and perspective change. *Cognition, 31,* 127-171.

Johnson-Laird, P. N. (1995). Inference and mental models. In S. E. Newstead and J. St. B. T. Evans (Eds.), *Perspectives on thinking and reasoning: Essays in honour of Peter Wason* (pp. 115-146). Hove: Lawrence Erlbaum Associates.

Johnson-Laird, P. N., & Byrne, R. M .J. (1991). *Deduction.* Hove: Lawrence Erlbaum Associates.

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

Johnson-Laird, P. N., Legrenzi, P., & Legrenzi, M. M. (1972). Reasoning and a sense of reality. *British Journal of Psychology, 63,* 395-400.

Lucas, E. J., & Ball, L. J. (2005). Think-aloud protocols and the selection task: Evidence for relevance effects and rationalisation processes. *Thinking and Reasoning, 11,* 35-66.

Manktelow, K. I., & Evans, J. St. B. T. (1979). Facilitation of reasoning by realism: Effect or non-effect? *British Journal of Psychology, 70,* 477-488.

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

Oaksford, M. (2002). Contrast classes and matching bias as explanations of the effects of negation on conditional reasoning. *Thinking and Reasoning, 8,* 13-151.

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. (1995). Information gain explains relevance which explains the selection task. *Cognition, 57,* 97-108.

Oaksford, M. & Chater, N. (1996). Rational explanation of the selection task. *Psychological Review, 103,* 381-391.

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

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

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

Roberts, M. J. (1998). Inspection times and the selection task: Are they relevant? *Quarterly Journal of Experimental Psychology, 51A,* 781-810.

Roberts, M. J., & Newton, E. J. (2001). Inspection times, the change task, and the rapid response selection task. *Quarterly Journal of Experimental Psychology, 54A,* 1031-1048.

Wason, P. C. (1966). Reasoning. In B. M. Foss (Ed.), *New horizons in psychology, Vol. I.* Harmondsworth: Penguin.