Brief article

Inept reasoners or pragmatic virtuosos? Relevance and the deontic selection task

Vittorio Girotto[a,b,*], Markus Kemmelmeier[c], Dan Sperber[d],
Jean-Baptiste van der Henst[d]

[a]*LPC, CNRS-University of Provence, Aix-en-Provence, France*
[b]*University of Trieste, Trieste, Italy*
[c]*University of Michigan, Ann Arbor, Michigan, USA*
[d]*Institut Jean Nicod, EHESS, Paris, France*

Abstract

Most individuals fail the selection task, selecting P and Q cases, when they have to test descriptive rules of the form "If P, then Q". But they solve it, selecting P and not-Q cases, when they have to test deontic rules of the form "If P, then must Q". According to relevance theory, linguistic comprehension processes determine intuitions of relevance that, in turn, determine case selections in *both* descriptive and deontic problems. We tested the relevance theory predictions in a within-participants experiment. The results showed that the same rule, regardless of whether it is tested descriptively or deontically, can be made to yield more P and Q selections or more P and not-Q selections. We conclude that the selection task does not provide a tool to test general claims about human reasoning. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Reasoning; Relevance theory; Selection task; Deontic inferences; Information representation

1. Introduction

The Wason (1966) Selection Task has been the most commonly used tool in the psychology of reasoning (see Manktelow, 1999). In the task, people are presented

* Corresponding author. Dipartimento di Psicologia, Via S. Anastasio 12, 34127 Trieste, Italy. Fax: +39-040-312272.
   E-mail address: girotto@univ.trieste.it (V. Girotto).

with a conditional rule of the form *If an item has the property P, then it has the property Q* (descriptive versions) or *If an item has the property P, then it should have the property Q* (deontic versions), and with four cards representing individual items. Only half of the information these four cards contain is visible, showing that the four items represented have respectively the P, not-P, Q and not-Q property (the cards are accordingly called the P, not-P, Q, and not-Q cards). The full information can be made visible by turning over the card in order to find out whether or not the P and not-P cards also have the Q property, and whether or not the Q and not-Q cards also have the P property. Participants are asked which cards it is necessary to turn over to determine whether the rule is true or false (descriptive version) or obeyed or disobeyed (deontic versions). Since the rule is true (or obeyed) unless there are items combining the P and the not-Q properties, the logically correct selection is that of the P and the not-Q cards, each of which could turn out to provide a counter-example to (or a violation of) the rule.

Work on the selection task has been the basis of several general claims about human reasoning and rationality. In particular, two such claims have received much attention. The first claim is that most individuals do not reason in accordance with the rules of logic, not even the elementary rules of propositional calculus, as shown by their failure to select the P and the not-Q cards in descriptive versions of the task (e.g. Cheng & Holyoak, 1985; Griggs & Cox, 1982). The second claim is that most individuals are better at reasoning on deontic problems (or some subclass of them) than on descriptive ones, as shown by their selection of the P and the not-Q cards in deontic versions of the task (e.g. Cheng & Holyoak, 1985; Cosmides, 1989; Girotto, Light, & Colbourn, 1988).[1] Indirect evidence for these two claims is provided by the fact that people asked to solve a series of selection tasks, some deontic, some descriptive, show no transfer from one task to the next (e.g. Johnson-Laird, Legrenzi, & Legrenzi, 1972).

Does the selection task really provide a tool to test general claims about human reasoning? Evans (1989) maintained that participants understand the task as one of identifying the *relevant* cards, and use for this heuristic cues of relevance rather than deductive reasoning. Extending this insight, Sperber, Cara, and Girotto (1995) argued that participants' poor performance on the selection task is best explained by considering that (1) the very process of linguistic comprehension provides participants with intuitions of relevance (see Sperber & Wilson 1995), (2) these intuitions, just as comprehension generally, are highly content- and context-dependent, and (3) participants trust their intuitions of relevance and select cards accordingly. In standard descriptive versions of the task, these intuitions are misleading, whereas in standard deontic versions, they point to the correct selection. Sperber et al. (1995) discussed how to manipulate content and context and thereby affect comprehension, intuitions of relevance, and selection of cards. They provided and tested a recipe to construct versions of the task where a majority of participants would give the logically correct response. They argued that all previous genuine versions of the

---

[1] Sperber and Girotto (forthcoming) discuss more specifically the Cosmides (1989) version of this second claim and the response by Fiddick, Cosmides, and Tooby (2000) to Sperber et al. (1995).

selection task where a majority of participants had given a correct response happened to conform to this recipe. One highly deflationary methodological implication of this work was that these previous versions did not provide a proper test for whatever theoretical claim about reasoning had been guiding the researchers, since the relevance-based pragmatic approach always provided a more general explanation of the experimental results. If, in the selection task, pragmatic comprehension mechanisms indeed pre-empt the use of whatever domain-general or domain-specific reasoning mechanisms people are endowed with, the task cannot be a good tool for the study of these reasoning mechanisms.

Sperber et al. (1995) produced several *descriptive* versions of the task that elicited a higher percentage of correct responses than had ever been found before with such versions. They showed that – contrary to what was generally believed at the time – good performance is not restricted to deontic versions.[2] In this article, we provide further evidence for the relevance approach by demonstrating how it can be used to manipulate *deontic* versions of the task and obtain at will either the common correct P and not-Q selections or incorrect P and Q selections (more commonly found in descriptive versions).[3] Our participants are asked to solve four versions of the selection task, two deontic and two descriptive ones, linked in a single narrative. Although all four versions involve *the same rule and the same pattern of cards, thus requiring the same logical solution*, participants change their answer from one problem to the next.

Every deontic rule has a descriptive reading. Consider the following rule (adapted from Cheng & Holyoak, 1985): "If a person travels to any East African country, then that person must be immunized against cholera." This rule expresses both a deontic and a descriptive claim. The rule applies as an obligation to people travelling to East Africa. These travelers are in a position to obey or disobey the rule, but not to make it true or false. The rule also makes a descriptive claim about East African countries. These countries are in a position to make the rule true (by all enforcing it) or false, but not to obey or disobey it. To use the rule in a deontic selection task, participants must be presented with cards representing a traveler to an East African country (P), a traveler to a country from another region (not-P), a traveler immunized against cholera (Q), and a traveler not so immunized (not-Q). Participants must be asked which cards must be selected to see whether these four travelers have obeyed the rule. To use the rule in a descriptive selection task, participants must be presented with cards representing an East African country (P), a country from another region (not-P), a country requiring immunization against cholera (Q), and one not requiring such an immunization (not-Q). They must be asked then which cards must be selected to see whether the rule is true, at least as far as the four countries repre-

---

[2] Other studies have confirmed this (e.g. Green & Larking, 1995; Hardman, 1998; Johnson-Laird & Byrne, 1995; Liberman & Klar, 1996; Love & Kessler, 1995).

[3] Patterns other than P and Q selection can be obtained in deontic problems by manipulating participants' perspectives (e.g. Gigerenzer & Hug, 1992; Holyoak & Cheng, 1995; Light, Girotto, & Legrenzi, 1990; Manktelow & Over, 1991; Politzer & Nguyen-Xuan, 1992) or the plausibility of the violations (e.g. Girotto, Mazzocco, & Cherubini, 1992; Liberman & Klar, 1996; Light et al., 1990; Mosconi & D'Urso, 1974). See Sperber et al. (1995) for a pragmatic analysis of these results.

sented are concerned.[4] The standard expectation, based on previous experiments (e.g. Cheng & Holyoak, 1985; Gigerenzer & Hug, 1992; Manktelow & Over, 1991), would be that most participants make the correct P and not-Q selection in the deontic case and incorrect selections (in particular P and Q) in the descriptive case. However, if the relevance approach is correct, it should be possible to obtain opposite results by means of appropriate alterations of context.

2. Experiment

2.1. Method

Thirty-four undergraduates at Trieste University volunteered to participate in the experiment. They received a four-page booklet containing four versions of the cholera problem, presented as in the narrative below.

2.1.1. Condition 1: "True descriptive"

The list of immunizations for international travels has changed over the years. Imagine that you are in 1979, and that you work in a travel agency. A client comes and tells you, 'I would like to visit East Africa, but I am allergic to the cholera immunization'. You answer, 'I'm sorry, but any East African country requires the cholera immunization'. In order to convince the client, you decide to show her that you are right. The agency has cards, each of which represents a country. On the upper part of each card, the name of the country is indicated. On the lower part of each card, the immunizations required to enter that country are indicated. Four of these cards, representing four countries, happen to be in front of you. However, other papers cover the lower part of two of the cards and the upper part of the two other cards. Which of these cards must have their hidden part uncovered in order to find out whether it is true that: "If a person travels to any East African country, then that person must be immunized against cholera."

Four cards indicated "Somalia", "Sweden", "Immunizations required: Cholera" and "Immunizations required: None", respectively. In such a context, the rule achieves relevance to the audience (i.e. to the client of the agency) by implying that any given East African country would require immunization, and that among

---

[4] In several studies, participants had to ascertain whether a deontic rule was in force (e.g. Cosmides, 1989; Gigerenzer & Hug, 1992; Kroeger, Cheng, & Holyoak, 1993; Noveck & O'Brien, 1996). In all these cases, however, the cards represented not possible counter-examples to the claim that the rule was in force, but possible violations of the rule. Unlike standard selection tasks, these tasks do not have a demonstratively valid solution. If some people do not act in accordance with a rule, this may provide evidence that the rule is not in force, but the inverse is also true. A priori, all four cards could be relevant to an inductive inference regarding the actuality of the rule. The experiment we present below is the first to test a deontic rule from a descriptive point of view with cards appropriate to a descriptive task (Conditions 1 and 3).

countries requiring immunization are East African countries. We predicted therefore that participants, in order to test these implications, would choose the P (an East African country) and the Q (a country requiring cholera immunization) cards.

2.1.2. Condition 2: "True deontic"

The narrative went on to say that the rule had been indeed confirmed. The agency's boss then asked the employee to check that clients of the agency had obeyed the rule "If a person travels to any East African country, then that person must be immunized against cholera" by examining cards representing these clients and their immunizations. The four cards indicated "Mr. Neri. Destination: Ethiopia", "Mr. Verdi. Destination: Canada", "Immunizations done: Cholera" and "Immunizations done: None", respectively. In this context, the relevance of the rule is to prevent people without cholera immunization from travelling to East African countries. We predicted therefore that participants would choose the P (a traveler to an East African country) and the not-Q (a person without cholera immunization) cards. These two first predictions were in keeping with previous evidence, and the only novelty so far was to have the same participants perform twice with the same rule used once descriptively and once deontically.

2.1.3. Condition 3: "False descriptive"

The second part of the narrative involved the same employee in the same agency, but 20 years later, i.e. in the present. The employee him/herself is now allergic to the cholera immunization. He or she believes, however, that this immunization is no longer required to go to an East African country, and he or she is planning to travel there him/herself. The agency's boss disagrees and asserts that "If a person travels to any East African country, then that person must be immunized against cholera." To test the boss's assertion, participants have to make a selection from four cards similar to those used in Condition 1 (with the P and not-P cards indicating "Kenya" and "Ireland", respectively). The boss's assertion achieves relevance by implying that any given East African country would require immunization and that there is no East African country that does not require it. We predicted accordingly that in order to test this implication, participants would select the P (an East African country) and the not-Q (a country not requiring cholera immunization) cards.

2.1.4. Condition 4: "False deontic"

The narrative stated that the employee was right and that the cholera immunization is not required anymore in East Africa. The boss is now worried that she may have misinformed clients and caused them to follow a rule that is no longer in force. She then asks the employee to see whether or not clients have obeyed the rule "If a person travels to any East African country, then that person must be immunized against cholera" by looking at cards similar to those used in Condition 2 (with the P and not-P cards indicating "Mr. Rossi. Destination: Eritrea" and "Mr. Bianco. Destination: France", respectively). In this context, what is relevant is that some clients may have followed the false rule and that they may have been immunized unnecessarily (and might now sue the agency). However, in this context the fact that

Table 1
Percentage of the main selection patterns in the four versions ($N = 34$)

Pattern	Version			
	True descriptive	True deontic	False descriptive	False deontic
P and Q	65	26	15	71
P and not-Q	9	62	47	15
P, Q and not-Q	9	9	24	6
Other[a]	17	3	14	8

[a] Each of the patterns indicated in the Other cells was produced by fewer than four participants.

there might be clients who ignored the rule is no longer relevant. We predicted therefore that participants would select the P (a traveler to an East African country) and the Q (a person with cholera immunization) cards. Note that this prediction is non-standard but follows from the relevance-based explanation of the selection task.

In sum, we predicted that participants would select cards P and Q in both true descriptive and false deontic conditions, and cards P and not-Q in both true deontic and false descriptive conditions.

2.2. Results

Table 1 presents the percentages of the main selection patterns in the four versions. As predicted, in the true deontic version, participants selected the P and not-Q combination more frequently, and the P and Q combination less frequently than in the true descriptive and false deontic versions (sign tests: $P < 0.002$). In the false descriptive version, participants selected the P and not-Q combination more frequently, and the P and Q combination less frequently than in the true descriptive and false deontic versions (sign tests, for the former comparisons: $P < 0.002$ and $P < 0.025$, respectively; for the latter comparisons, $P < 0.002$). Only two participants (binomial, $z = 5.32$, $P < 0.00003$) gave the same response (in both cases, P and Q) to all four problems.

It could be suggested that the within-participants design conveyed to participants that different responses were expected of them. However, this would not explain why only four (binomial, $z = 4.63$, $P < 0.00003$) gave the same response to the two descriptive problems, and only nine (binomial, $z = 2.91$, $P < 0.01$) gave the same response to the two deontic problems. It could also be suggested that our results are confounded with an order-of-test effect. However, this is ruled out by a between-participants replication that we performed with four different narratives, each corresponding to one of the four episodes of the continuous narrative used in the repeated measures study. Each participant was presented with only one narrative and therefore had to solve only one selection task. Nevertheless, we got the same pattern of responses as in the repeated measures study. In the true deontic condition, participants selected the P and not-Q combination more often (77%) and the P and Q combination less often (3%) than in the false deontic condition (respectively: 33%,

$\chi^2(1, N = 64) = 10.8$, $P < 0.01$; and 52%, $\chi^2(1, N = 64) = 16.1$, $P < 0.001$. In the false descriptive condition, participants selected the P and not-Q combination more often (40%) and the P and Q combination less often (27%) than in the true descriptive condition (respectively: 6%, $\chi^2(1, N = 62) = 8.3$, $P < 0.01$; and 63%, $\chi^2(1, N = 62) = 6.7$, $P < 0.01$).

3. Conclusions

The present results corroborate our predictions and confirm the analysis of Sperber et al. (1995). We showed that the same rule, whether it is tested descriptively or deontically, can be made to yield more P and Q selections or more P and not-Q selections by acting on intuitions of relevance. In particular, we confirmed the non-standard prediction that most individuals, when testing a false deontic rule, will select the P and Q cards, corresponding to the possibility that some people may have obeyed the false rule, which is much more relevant than the possibility that some people may have disobeyed it. These results support the relevance approach according to which the very process of linguistic comprehension of the problem yields intuitions of relevance that, in turn, determine card selection. These intuitions of relevance can be manipulated through alterations of content and context. Differences in performance observed over the years with various descriptive or deontic versions of the task reflect such alterations, whether performed wittingly or unwittingly.

Participants' good performance with standard deontic versions of the task, being explainable in terms of relevance factors, provides no evidence for or against the existence of domain-specific mechanisms to reason about deontic rules (Cheng & Holyoak, 1985) or some evolutionarily significant subset of these rules (Cosmides, 1989). Our results are of course compatible with the hypothesis that there are such mechanisms. However, given that the contents of selection tasks are verbally communicated to the participants, that therefore their comprehension mechanism is necessarily activated, and that this mechanism predicts observed performance, the effects of whatever further domain-specific mechanisms may exist are pre-empted or confounded. It is unlikely therefore that we will improve our understanding of deontic reasoning by yet again investigating versions of the selection task.

Regarding the common view that selection task results show that people do not reason in accordance with the rules of logic, the fluctuating results we obtained with a within-participant design are not so much evidence of a lack of transfer, than of a downright refusal to generalize. This suggests either that people are even worse reasoners than was claimed, or, more plausibly, that they are pragmatic virtuosos.

References

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

Evans, J. St. B. T. (1989). *Bias in human reasoning: causes and consequences*. Hillsdale, NJ: Erlbaum.

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

Gigerenzer, G., & Hug, K. (1992). Domain specific reasoning: social contracts, cheating, and perspective change. *Cognition*, *43*, 127–171.

Girotto, V., Light, P. H., & Colbourn, C. J. (1988). Pragmatic schemas and conditional reasoning in children. *Quarterly Journal of Experimental Psychology*, *40A*, 342–357.

Girotto, V., Mazzocco, A., & Cherubini, P. (1992). Pragmatic judgements of relevance in reasoning: a reply to Jackson and Griggs. *Quarterly Journal of Experimental Psychology*, *45A*, 547–574.

Green, D. W., & Larking, R. (1995). The locus of facilitation in the abstract selection task. *Thinking and Reasoning*, *1*, 183–199.

Griggs, R. A., & Cox, J. R. (1982). The elusive thematic-materials effect in Wason's selection task. *British Journal of Psychology*, *73*, 407–420.

Hardman, D. (1998). Does reasoning occur in the selection task? A comparison of relevance-based theories. *Thinking and Reasoning*, *4*, 353–376.

Holyoak, K. J., & Cheng, P. W. (1995). Pragmatic reasoning with a point of view. *Thinking and Reasoning*, *1*, 289–313.

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

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

Kroger, K., Cheng, P. W., & Holyoak, K. J. (1993). Evoking the permission schema: the impact of explicit negation and a violation-checking context. *Quarterly Journal of Experimental Psychology*, *46A*, 615–635.

Liberman, N., & Klar, Y. (1996). Hypothesis testing in Wason's selection task: social exchange, cheating detection or task understanding. *Cognition*, *58*, 127–156.

Light, P. H., Girotto, V., & Legrenzi, P. (1990). Children's reasoning on conditional promises and permissions. *Cognitive Development*, *5*, 369–383.

Love, R., & Kessler, C. (1995). Focussing in Wason's selection task: content and instruction effects. *Thinking and Reasoning*, *1*, 153–182.

Manktelow, K. I. (1999). *Reasoning and thinking*. Hove: Psychology Press.

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

Mosconi, G., & D'Urso, V. (1974). *The selection task from the standpoint of the theory of double code*. Paper presented at the Conference on The Selection Task, University of Trento, Trento.

Noveck, I. A., & O'Brien, D. P. (1996). To what extent do pragmatic reasoning schemas affect performance on Wason's selection task? *Quarterly Journal of Experimental Psychology*, *49A* (2), 463–489.

Politzer, G., & Nguyen-Xuan, A. (1992). Reasoning about conditional promises and warnings: Darwinian algorithms, mental models, relevance judgements or pragmatic schemas? *Quarterly Journal of Experimental Psychology*, *44*, 401–412.

Sperber, D., Cara, F., & Girotto, V. (1995). Relevance theory explains the selection task. *Cognition*, *52*, 3–39.

Sperber, D., & Girotto, V. (forthcoming). Response to Fiddick, Cosmides, and Tooby.

Sperber, D., & Wilson, D. (1995). *Relevance: communication and cognition* (2nd ed.). Oxford: Blackwell.

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