

Use or misuse of the selection task?
Rejoinder to Fiddick, Cosmides and Tooby

Dan Sperber^a, Vittorio Girotto^b

^a*Institut Jean Nicod (CNRS, EHESS and ENS), Ibis avenue de Lowendal, 75007 Paris*

^b*LPC, CNRS and University of Provence, Aix-en-Provence; Department of Psychology,
University of Trieste, Trieste*

Abstract : Sperber, Cara, and Girotto (1995) argued that, in Wason's selection task, relevance-guided comprehension processes tend to determine participants' performance and pre-empt the use of other inferential capacities. Because of this, the value of the selection task as a tool for studying human inference has been grossly overestimated. Fiddick, Cosmides, and Tooby (2000) argued against Sperber et al. that specialized inferential mechanisms, in particular the "social contract algorithm" hypothesized by Cosmides (1989), pre-empt more general comprehension abilities, making the selection task a useful tool after all. We rebut this argument. We argue and illustrate with two new experiments, that Fiddick et al. mix the true Wason selection task with a trivially simple categorization task superficially similar to the Wason task, yielding methodologically flawed evidence. We conclude that the extensive use of various kinds of selection tasks in the psychology of reasoning has been quite counter-productive and should be discontinued.

Keywords : Cheater detection, reasoning, evolutionary psychology, relevance theory, Wason selection task.

1. INTRODUCTION

Why has Wason's selection task (Wason 1966)¹ been, for almost forty years, so extensively used in psychology of reasoning? Because, it has a simple, logically compelling solution, and yet, in most versions, most participants fail to solve it. Philosophers have seen this as highly relevant evidence in assessing human rationality (e.g., Stein 1996).

Psychologists have found ways of improving participants' performance, in particular by changing the narrative content of the task and have offered various interpretations of these results. Selection task data has thus been garnered in support of various general claims about human reasoning.

In particular, Leda Cosmides, John Tooby and their collaborators have, over the past twenty years, performed a variety of original selection task experiments to establish the existence of evolved domain-specific reasoning mechanisms. Their most famous, best developed hypothesis concerns the existence of a "social contract algorithm" one sub-component of which is a cheater-detection device (Cosmides 1989; Cosmides & Tooby 1989, 1992, 1997, Fiddick, Cosmides & Tooby, 2000). They define a social contract as a situation in which one party is obligated to satisfy a requirement in order to be entitled to receive a benefit from another party, and they define cheating as the taking of the benefit without satisfying the requirement. A social contract situation can be depicted in a selection task by means of cards representing on one side whether or not the benefit has been taken and on the other side whether or not the requirement has been satisfied. A cheater-detection device should favor the

¹ In the Wason selection task, participants are presented with a conditional rule of the form *If an item has the feature P, then it has the feature Q* (descriptive versions) or *If an item has the feature P, then it should have the feature 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 feature (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 feature Q, and whether or not the Q and not-Q cards also have feature P. 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, as far as these four cards are concerned, the rule is true (or obeyed) unless there are cards combining the P and the not-Q features, the logically correct selection is that of the P and the not-Q cards, either of which could turn out to provide a counter-example to (or a violation of) the rule.

selection of the "benefit taken" and of the "requirement not satisfied" cards, either of which could turn out to correspond to a case of cheating. This is indeed what a majority of participants select in the social contract situations used in the experiments of Cosmides, Tooby and their collaborators. The selection task thus seem to provide crucial data in favor of the evolutionary psychology approach of which Tooby and Cosmides have been the most articulate exponents.

Although many researchers clearly believe that the selection task provides appropriate evidence for general claims about human reasoning, no such claim has ever been accepted by the scientific community on the basis of such evidence. Rather, what seems to drive the continuous production of selection task experiments is that their interpretations can be endlessly contested by means of further experiments with the task. This in itself would be reason enough to question the reasonableness of the proliferation of research based the selection task of which the work of Cosmides, Tooby and their collaborators is a striking example. Moreover, Sperber, Cara and Girotto (1995, henceforth SCG) have put forward an analysis of the task itself that casts principled doubts on its scientific utility.

SCG argued that, in the selection task, relevance-guided comprehension mechanisms tend to pre-empt the use of whatever domain-general or domain-specific reasoning mechanisms people are endowed with.² In support of this claim, they showed how to manipulate relevance factors in descriptive versions of the task so as to elicit more than 50% correct responses (a rate of success normally found with deontic rather than descriptive versions). Girotto, Kimmelmair, Sperber and van der Henst (2001) provided further evidence by manipulating relevance factors in deontic versions so as to elicit more than 80% incorrect responses (a rate of failure normally found with descriptive rather than deontic versions). The only challenge to SCG purporting to contain counterevidence has been Fiddick, et al. (2000 - henceforth FCT). What we aim to do here is rebut FCT's central theoretical argument (what they call the "principle of pre-emptive specificity") and demonstrate a methodological flaw in their experimental evidence. We hope that this will be enough to vindicate and indeed reinforce our earlier warnings against basing any important claim in the psychology of reasoning on Wason selection task data (other issues raised by FCT are discussed in Sperber & Girotto (in press); see also Atran 2000).

² See also Evans (1989), for an earlier and different relevance-based account of the task, and Sperber and Wilson (1995), for relevance theory.

1.1. Comprehension, reasoning, and the order of pre-emption

According to most of modern pragmatics, comprehension is an inferential process that takes as input an utterance and contextual information, and that produces as output an interpretation of the speaker's meaning. Comprehension is an attribution of a mental state to the speaker, a form of mindreading. According to relevance theory in particular, this comprehension process consists in developing an interpretation that satisfies the expectation of relevance raised by the utterance itself. This process is specific to comprehension. If mindreading is viewed as a module of the human mind, then comprehension can be viewed as a sub-module (Sperber, 2000, Sperber & Wilson 2002). Cheater detection too is presented by FCT as a sub-module of the social contract algorithm. Let us accept, for the sake of discussion, that the normally developed human mind is equipped with both a comprehension mechanism and a cheater detection mechanism. What happens when participants in an experiment are verbally presented with a selection task involving a social contract, that is, with an input that could, in principle, activate both the comprehension and the cheater-detection mechanisms? This is where SCG and FCT give almost diametrically opposed answers.

SCG argued that, in Wason's selection task, pragmatic comprehension mechanisms tend to pre-empt the use of whatever domain-general or domain-specific reasoning mechanisms people may be endowed with (and this is what makes the selection task an inappropriate tool to study reasoning). FCT claim on the contrary that the more specialized mechanism will pre-empt the less specialized mechanism and in particular cheater detection will pre-empt inferential comprehension.

FCT's (p. 24) main argument is this: "a well-engineered problem-solving system should deploy, to the extent possible, the most specialized problem-solving machinery that is activated by the problem at hand, because on average, it will be more knowledgeable than the alternative, more general problem-solvers that also apply [...] This principle of cognitive design – what we will call the *principle of pre-emptive specificity* – should be expressed in design features throughout the cognitive architecture. It applies to the problems herein. Social contract algorithms and hazard management algorithms are more content-specialized than relevance mechanisms, whose domain is all content that arrives via communication from an agent". While this argument has obvious merits, it ignores an essential design factor. Some mechanisms accept as input the output of other mechanisms. Whatever their relative degree of

specialization, the receiver mechanism cannot prevent the feeder mechanism from performing its operations. In fact, without the output of the feeder mechanism, the receiver mechanism would have no input to process.

In order to recognize a verbally presented reasoning problem, whatever the domain-general or domain specific reasoning mechanism they may ultimately bring to bear on its solution, participants must first comprehend the text of the problem and, for this, use their comprehension mechanism. According to relevance theory, comprehending a text involves grasping its intended relevance, and this involves a meaning construction process that often goes beyond or away from strict literal meaning. With standard reasoning problems, participants are presented with verbal premises and are explicitly asked to infer or evaluate conclusions. Their understanding of the premises may be biased by comprehension factors (for a review, see Evans, Newstead & Byrne, 1993; for illustrations, e.g. Begg & Harris, 1982; Politzer, 1990), but, at least, it is clear that they must engage in a reasoning effort, over and above comprehension proper. The selection task is not a standard reasoning problem. Participants are not asked to reason from premises to conclusions, and they are not even told that the question they are asked can best be answered by making use of deductive reasoning. Participants are just asked to evaluate the relevance of one part of the problem (the cards) to another part of the problem (the conditional rule). What SCG argued is that the intuitions of relevance arrived at in the process of comprehending a selection task problem provide what looks like an intuitive solution to the problem itself. So, it is not just that comprehension will take place anyhow. It is that, in the particular case of the selection task, there is no incentive, for most participants, to engage in further active reasoning of either a general or a domain-specific kind, in order to find the solution to the problem.

Our answer to FCT is therefore, firstly, that there is no way that a cheater detection mechanism could pre-empt a comprehension mechanism in the processing of a selection task problem, since, anyhow, the problem must first be comprehended. Secondly, the comprehension mechanism automatically provides what looks like a solution to the selection task. It may well—and does for a majority of participants—pre-empt any further reasoning processes.

1.2. Distinguishing Wason selection task from FCT selection task

As SCG (p. 42) pointed out, a number of past experiments with deontic tasks (e.g. Cosmides, 1989; Light, Girotto & Legrenzi, 1990; Manktelow & Over, 1991) suffered from a methodological defect. Participants were simultaneously asked, not one, but two questions: the standard Wason task question whether the conditional rule had been obeyed or violated and the direct question whether some cheating has taken place (or worse, they were asked just the question about cheating). Answering correctly the first question is performing successfully on Wason's selection task, but answering correctly the second question is not: it demonstrates just the ability to identify a cheater on the basis of the information provided in the problem narrative. FCT (p. 15) themselves point out the difference. They write: 'Note that the definition of cheating *does not map onto the logical definition of violation* (the latter being a true antecedent paired with a false consequent). Cheating is a content-dependent concept: there must be an illicitly taken *benefit*. This, and only this, counts as cheating. Logical categories and definitions of violation form an orthogonal representational dimension.' This should make them wary of using as evidence data obtained with the two questions, that of cheating and that of violation posed simultaneously.

For instance, in one of the problems used in Cosmides's (1989) Experiment 2 and 4, participants were asked "Did Big Kiku get away with cheating any of these four men [a question directly about cheating]? Indicate only those card(s) you definitely need to turn over to see if Big Kiku has broken his word to any of these four men [an indirect question about the respect or violation of the conditional rule uttered by Big Kiku]" (Cosmides, 1989; p. 265). This kind of formulation suggests that both questions have the same correct answer and, in fact, in Cosmides's Experiment 2, where the conditional rule was Big Kiku's statement: "if you get a tattoo on your face, then I'll give you cassava root", they do. Big Kiku got away with cheating a man just in case Big Kiku's did not do what he said he would do. This correct answer was given by 71% of the participants, but there is no way to decide whether they based their response on the first question ("Did Big Kiku get away with cheating"?), or the second (is it the case that "Big Kiku has broken his word"?), or paid attention to both questions.

In Experiment 4, however, the conditional rule was "switched" to "If I give you cassava root, then you must get a tattoo on your face" (Cosmides, 1989, p. 217). This time, to begin with, both questions become harder to interpret. Strictly understood, Big Kiku's word express not a conditional contractual promise (i.e., a commitment to do something conditional on the

other party's action), but a requirement that the other party do something (see Legrenzi, Politzer, & Girotto, 1996). Therefore, Big Kiku is not in a position to either cheat or break his word. However—as has long been known in pragmatics (Geis & Zwicky, 1971)—many conditionals statement and in particular conditional promises, are commonly understood as implying their converse (see e.g., Fillembaum, 1975; Newstead, Ellis, Evans & Dennis, 1997). Here Big Kiku's statement can be understood as implying that, if his interlocutor got a tattoo, then he would give him cassava root, making sense of the narrative and of the ideas that Big Kiku might cheat. At this point, of course, the question about cheating, and the Wason task question about whether the conditional rule was violated or not have different answers (that is, if the conditional rule is strictly understood). The correct answer to the Wason task question would be, as always, to select the P and not-Q cards, whereas the correct answer to the cheating question would be to select the not-P and Q cards, which 75% of Cosmides's (1989), participants did. She took this to show that people use a cheater detection Darwinian algorithm to solve such a Wason selection task. We suggest that this shows that participants answered the cheating question and not the Wason task question, or that they reinterpreted Big Kiku's statement as meaning its converse (in which case both questions admit of the same answer).

What is it that makes the Wason task question significantly different from the question about cheating? The Wason task question is about the truth or falsity (in the descriptive versions) or the respect or violation (in the deontic versions) of a conditional rule. The only failsafe way of answering it involves applying conditional reasoning to four hypothetical cases. The cheating question is not a conditional reasoning question but a categorization question. As explained in detail by Cosmides and her collaborators, cheating is commonly understood as the co-occurrence of the taking of a benefit and the failure to fulfill a requirement, in particular of paying a cost. It is, in other terms, characterized by the conjunction of these two features. In order to answer the cheating question, then, all that participants have to do is select the cards that exhibit one of these two features (and that might have the other characteristic feature on the other side).

More generally, one could devise a categorization task that would use material similar to that used in the Wason selection task, and in particular cards exhibiting one of two possible values for a pair of traits (e.g. *benefit taken* vs. *benefit not taken*, and *requirement fulfilled*, vs. *requirement not fulfilled*; or *flying vehicle* vs. *non-flying vehicle*, and *has an engine* vs. *does*

not have an engine). Out of the four possible combinations of such trait values, one and only one would determine a given category. For instance the combination *benefit taken + requirement not fulfilled* would determine the category of cheaters, or the combination *flying vehicle + does not have an engine* would determine the category of gliders. Participants would then be asked which of the four cards could turn out to represent an instance of the category. This is what FCT have actually done, devising a task that is superficially similar to the Wason selection task, but that tests for a different cognitive capacity. As we indicated above, in order to distinguish it from the genuine Wason task—something that FCT totally failed to do—, we will call this categorization task the "FCT selection task." When both the Wason and the FCT questions are posed in the same problem, as in the Big Kiku example, we speak of a Wason/FCT task.

Is there any reason to think that the FCT selection task is of particular psychological interest? Should we expect people to generally fail at the FCT task, as they do with the Wason task, and to succeed when the material used activates domain-specific evolved mechanisms? We believe not. The FCT task is quite trivial. As we will show, participants perform it without difficulty, even with totally artificial material. It looks interesting only when it is confused with the Wason task, and used without adequate control conditions. Just as Wason selection task, the FCT selection task is inadequate to test Cosmides's (1989) interesting hypothesis that humans are endowed with an evolved mental module to reason about social contracts, or other similar evolutionary psychology hypotheses.

2. EXPERIMENT 1

The goal of our first experiment was to check whether the FCT selection task is, as we claimed, trivially easy. Can individuals, if asked to do so, correctly select the cards that might represent an instance not just of cheating, but of any category defined by the combination of a positive and a negative trait?

We devised three conditions: a first condition where the relevant category was explicitly that of cheating defined by the combination of "cost paid" and "benefit not received"; a second condition, with material strictly parallel to the first condition but where the relevant category was an arbitrary one called "Wason's selection" defined by the combination of "food item" and "non-Italian"; and a third condition where the relevant category was that of a glider. Whereas, in the two former conditions, the relevant categories were explicitly defined as the

combination of a positive and a negative trait, participants did not receive the definition of “glider”. In this experiment, we were comparing a condition of a kind assumed by FCT to provide evidence for the existence of an evolved mental cheating detector, with two conditions where what had to be “detected” was either a made-up arbitrary category, or an ordinary language category of no evolutionary significance and the relevant properties of which were not even mentioned in the text. We predicted that participants would do well in all conditions.

2.1. Method

Participants

A total of 100 humanities undergraduates at Trieste University (Italy) participated in the experiment voluntarily. They were randomly assigned to one of three groups: “look for cheater”, “look for Wason's selection”, and “look for glider”.

Procedure and Materials

The study was carried out in Italian with native speakers of the language. The participants were tested in groups, but they worked on the problems individually, at their own pace. In the "look for cheater" condition, the problem read as follows:

Paolo often buy things through the Internet, but he is fearful of being cheated. For each order, he fills a card. On one side of the card, he indicates whether he has received the item ordered. On the other side of the card, he indicates whether he has paid for the item ordered.

Paolo puts in a box labeled “Risk of cheating” the cards indicating that he has paid for the item and has not received the item.

Below are four cards, each of which represents a different order. Two cards show the side that indicates whether Paolo has received the item ordered. The two other cards show the side that indicates whether Paolo has paid for the item.

7th December 2000:
Item paid for

14th December 2000:
Item not paid for

29th October 2000:
Item received

21st November 2000:
Item not received

Indicate only the card or cards you definitely need to turn over to see whether, among these cards, there are some cards that Paolo should put in the “Risk of cheating” box.

In the "look for Wason's selection" condition, the problem read as follows:

Paolo often buy things through the Internet. For each order, he fills a card. On one side of the card, he indicates whether the item ordered is food. On the other side of the card, he indicates whether the item ordered is Italian.

Paolo puts in a box labeled “Wason’s selection” the cards indicating a food item that it is not Italian.

Below you find four cards, each of which represents a different order. Two cards show the side that indicates whether the items ordered are food or not. The two other cards show the side that indicates whether the items are Italian or not.

7th December 2000: Food item

14th December 2000: Non food item

29th October 2000: Italian item

21st November 2000: Non Italian item

Indicate only the card or the cards you definitely need to turn over to see whether among these cards there are some cards that Paolo should put in the “Wason’s selection” box.

In the "look for gliders" condition, the problem read as follows:

Paolo collects pictures of vehicles. For each vehicle, he fills a card. On one side, he indicates where the vehicle movements take place. On the other side, he indicates whether the vehicle has an engine or not.

Paolo puts in a box labeled “Gliders” the cards representing a glider. Below are four cards, each of which represents a different vehicle. Two cards show the side that indicates where the vehicle movements take place. The two other cards show the side that indicates whether the vehicle has an engine or not.

Vehicle 15
Where does it move?
In the air

Vehicle 8
Where does it move?
On rails

Vehicle 34
Does it have an engine?
Yes

Vehicle 27
Does it have an engine?
No

Indicate only the card or cards you definitely need to turn over to see whether, among these cards, there are some cards that Paolo should put in the “Gliders” box.

2.2. Results and Discussion

Table 1 reports the percentages of the selection patterns obtained in the three conditions.

Table 1

Percentage of the selection patterns in the three conditions of Experiment 1

Pattern	Condition		
	Look for cheater (N=35)	Look for Wason (N=35)	Look for glider (N=30)
P, not-Q	68	91	73
P	6	3	-
P, Q	-	-	13
not-Q	20	-	10
not-P,Q	-	3	3
P, not-P, not-Q	6	-	-
All	-	3	-

As predicted, participants performed well in all conditions. In particular, in the condition in which they were required to look, not for cheaters, but for the arbitrary category of Wason's selection, 91% of them selected the cards corresponding to the P and not-Q features. This rate of correct performance is significantly higher than the one (68%) obtained

in the "look for cheater" condition ($\chi^2(1, N = 70) = 4.37, p < .05$), and marginally higher than in the look for glider condition ($\chi^2(1, N = 65) = 2.7, p < .10$),

The fact that participants' performance is lower in the "look for glider" than in the look for Wason's selection" condition is readily explained by the fact that, the two defining traits of a "Wason's selection" were mentioned in the text of the problem, whereas participants had to know and remember that a glider is a vehicle that moves in air without an engine. This interpretation was confirmed by a replication of the "look for glider" condition, in which we asked participants (20 humanities undergraduates at Trieste University) to justify their selections. The rate of correct performance (65%) was similar to the one obtained in the previous "look for glider" condition, and significantly lower than the one obtained in the "look for Wason" condition ($\chi^2(1, N = 55) = 4.33, p < .05$). All but one participant who failed selected only the P card ("vehicle moving on air"). All of them explained that they did not select not-Q card ("engineless vehicle") because they were not sure whether a glider has an engine or not. Presumably, then, if we had mentioned in the text of the problem the two defining traits of the relevant category, the performance in the "look for glider" condition would have been higher.

On the other hand the fact that participants' performance is lower in the "look for cheater" than in the "look for Wason's selection" condition and is similar to that in the "look for glider" condition is somewhat puzzling, since in the "cheater" condition the two defining traits *were* mentioned. Whatever the explanation of this less-than-perfect performance in the "look for cheater" condition, and whatever the remedies that could be found to improve this performance, the fact remains that, in an FCT task, cheaters are not more easily detected than arbitrary items such as Wason's selections, or items devoid of evolutionary significance such as gliders. As we predicted, participants did well in all three conditions. We would expect participants to do well with any category whatsoever, provided that they knew or were told its defining traits.

3. EXPERIMENT 2

Experiment 1 showed that the FCT selection task is trivially easy, and therefore is quite unlike the Wason selection task, in spite of being modeled on it. In their Experiment 1, FCT argue that, in the Wason selection task, people do not reason from the logical form expressed in the conditional rule, as shown by the fact that they perform as well when the logical

connective "if...then" is removed. From a pragmatic point of view, we would anyhow agree that logical connectives are not necessary for people to give a conditional interpretation to a text or dialogue. For instance in a dialogue such as:

Child: I want to go out and play!

Mother: You must put on your coat!

the mother's reply is normally understood as meaning implicitly: *If* you want to go out and play, *then* you must put on your coat. The logical form of an utterance need not be wholly explicit, and there is no reason why a proper Wason selection task could not be performed with an implicit rather than an explicit conditional rule. This, however, was not FCT's point.

FCT assumed that if people, in a task describing a social exchange, performed equally well with or without an explicit conditional, this showed that they didn't reason on the conditional logical form, but just on the "logic of social exchange." However, FCT used as evidence not a Wason selection task but an FCT selection task. We would argue that the conditional form, whether it is implicitly retrieved by the participants or not, is anyhow irrelevant to successful performance of an FCT selection task. More specifically, FCT asked participants to indicate the cards they had to turn over "to see whether any of these people [represented by the cards] have cheated you". They compared two conditions, one where the exchange was explicitly described in conditional form, the other where it was not. Since this is an FCT task, we would have predicted that participants would perform equally well in both conditions, as indeed they did. Our claim is that, when people are explicitly instructed to identify possible cheaters, they have no problem understanding the instruction and following it. FCT do not consider this simple possibility and argue that when people recognize an exchange situation, and moreover adopt the perspective of one of the two parties to the exchange, then they automatically look for cheaters. If FCT are so confident that this is so, it is hard to understand why they mar their experimental evidence by explicitly asking participants to look for cheaters.

To help decide whether people look for cheaters because they are asked to look for cheaters, or because they are asked to reason about an exchange situation, we performed an FCT experiment with two conditions replicating FCT's own and two other conditions identical in every point, except that participants were not asked to look for instances of cheating, but for instances of exchange. If it is the exchange situation that triggers the selection of the P and not-Q cards representing possible instance of cheating, then this modification should have

little or no effect. If, as we claim, participants in an FCT task trivially do just what they are asked to do, then, we should expect participants in the new conditions to select the P and Q cards, which represent possible instances of exchange, more often than the P and not-Q cards, which represent possible instance of cheating.

3.1. Method

Participants

The participants were 120 humanities undergraduates from Trieste University. They were randomly assigned to one of four equal-sized groups: "look for exchange (conditional)", "look for exchange (want)", "look for cheater (conditional)" and "look for cheater (want)" (N= 30).

Procedure and Materials

The same procedure as in Experiment 1 was used. Each participant had to solve one problem. The "look for cheater (conditional)" and "look for cheater (want)" conditions replicated, respectively, the "conditional" and "want" problems of FCT Experiment 1. They read as follows:

You are a South American farmer. At the end of the harvest you find you have more potatoes than you need so you pack up some of them and travel to the neighboring village. When you get to the village four different people approach you, and though you don't speak the same dialect, you recognize that each of them is telling you:

"If you give me some potatoes, then I will give you some corn." [conditional version]

"I want some potatoes." You, in turn, know a little bit of their dialect, and tell them "I want some corn." [want version]

The cards below represent four people who approached you. One side of the cards tells whether or not you gave the person any potatoes and the other side of the cards tells whether or not that person gave you any corn.

You gave this person
potatoes

You gave this person nothing

This person gave you corn

This person gave you nothing

Indicate only the card or the cards you definitely need to turn over to see whether any of these people have cheated you.

The "look for exchange (conditional)" and "look for exchange (want)" conditions presented exactly the same problems, excepted that the final instruction read as follows:

Indicate only the card or the cards you definitely need to turn over to see whether any of these people have made an exchange with you.

3.2. Results and Discussion

Table 2 presents the percentages of the main selection patterns in the four versions.

Table 2
Percentage of the main selection patterns in the four conditions of Experiment 2 ($N = 30$)

Pattern	Condition			
	Look for cheater (conditional)	Look for cheater (want)	Look for exchange (conditional)	Look for exchange (want)
P, not-Q	33	43	10	0
P, Q	7	13	33	53
P	30	10	30	20
not-Q	13	17	13	0
Other	17	17	13	27

Note: Each of the patterns indicated in the Other cells was produced by fewer than four participants.

In both "look for cheater" conditions, the most frequently selected pattern was the correct P and not-Q pattern. In the want version, the proportion of participants who made such a selection was similar to the one obtained in FCT's study (43% vs. 50%, respectively), and significantly higher than the one obtained in the corresponding "look for exchange" condition, in which no participant made this selection ($\chi^2(1, N = 60) = 14.14, p < .001$). In the conditional version, however, our participants made this selection less often than FCT's ones (33% vs. 66%, respectively). We have no confident explanation for this failure to replicate their results. As we predicted, in both "look for exchange" conditions the most frequently selected pattern was the correct P and Q pattern. Participants produced more such P and Q

selections than in the “look for cheater” conditions (conditional versions: $\chi^2(1, N = 60) = 5.10, p < .05$; want versions: $\chi^2(1, N = 60) = 9.07, p < .01$).

These results confirm our main claim that, in an FCT task, people do as they are told and, in particular, look for a cheater when they are told to do so, just as they look for an exchange when this is what they are told to do. They disconfirm FCT's claim that, when people recognize a situation as an exchange and moreover take the perspective of a party, they automatically look for cheaters.

5. CONCLUSION

Let us be quite clear: nothing we have said, none of the evidence we have mustered implies that Cosmides' (1989) hypothesis about the existence of a specific competence to deal with social exchanges is wrong. What we have tried to show is that this hypothesis, in spite of having been at the center of heated debates in the past twelve years, has not yet been properly experimentally tested, since almost all the evidence is based on the Wason or the FCT selection task, which are inappropriate for this purpose, or on a mixed Wason/FCT task, which is methodologically unsound. Incidentally, using patients with brain lesions (see Adolphs, 1999) with Wason or FCT selection tasks to test hypotheses about specialized mechanisms for social information processing is just further and more costly use of an inappropriate methodology. In no ways, does this constitute independent evidence for the claims based on standard uses of the tasks. There is a variety of methods by means of which Cosmides' hypothesis might be seriously tested (see Sperber & Girotto, in press). Experimenting with the Wason or the FCT selection task is not one of them.

More generally, much of the work done with the selection task should be considered a sunk cost in the history of the psychology of reasoning and further investments of research effort and journal pages in uses of the task should be discouraged.

Acknowledgements. The research reported in this study has been supported in part by 1.14-60%(2000-2001) grants. We thank Renata Piacentini for her help in data collection.

REFERENCES

- Adolphs, R. (1999). Social cognition and the human brain. *Trends in Cognitive Sciences*, 3, 469-479.
- Atran, S. (2000) A cheater-detection module? Dubious interpretations of the Wason Selection Task and Logic. *Evolution and Cognition*, 7, 187-193.
- Begg, I., & Harris, G. (1982). On the interpretation of syllogisms. *Journal of Verbal Learning and Verbal Behavior*. 21, 595-620.
- 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, Part II. Case study: A computational theory of social exchange. *Ethology & Sociobiology*, 10, 51-97
- Cosmides, L. & Tooby, J. (1992). Cosmides, L. & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 163-228). New York: Oxford University Press.
- Cosmides, L. & Tooby, J. (1997). Dissecting the computational architecture of social inference mechanisms. In: *Characterizing human psychological adaptations* (Ciba Foundation Symposium #208) (pp. 132-156). Chichester: Wiley.
- Evans, J.St.B.T. (1989). *Bias in Human Reasoning: Causes and Consequences*. Hillsdale, NJ: Erlbaum.
- Evans, J.St.B.T., Newstead, S.R, Byrne, R.M.J. (1993). *Human reasoning. The psychology of deduction*. Hove: 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.
- Fillenbaum, S. (1975). If: some uses. *Psychological Research*, 37, 245-260.
- Geis, M.C., & Zwicky, A.M. (1971). On invited inferences. *Linguistic Inquiry*, 2, 561-566.
- Giroto, V., Kimmelmeir, M., Sperber, D, & van der Henst, J.B. (2001). Inept reasoners or pragmatic virtuosos? Relevance and the deontic selection task. *Cognition*, 81, 69-76.

- Legrenzi, P., Politzer, G., & Girotto, V. (1996). Contract proposals: A sketch of a grammar. *Theory and Psychology*, 6, 247-265.
- Light, P.H., Girotto, V., & Legrenzi, P. (1990). Children's reasoning on conditional promises and permissions. *Cognitive Development*, 5, 369-83
- Manktelow, K.I., & Over, D.E. (1991). Social rules and utilities in reasoning with deontic conditionals. *Cognition*, 39, 85-105
- Newstead, S.R., Ellis, M.C., Evans, J.St.B.T., & Dennis, J. (1997). Conditional reasoning with realistic material. *Thinking and Reasoning*, 3, 49-76.
- Politzer, G. (1990). Immediate deduction between quantified sentences. In K.J. Gilhooly, M.T.G. Keane, R.H. Logie, & G. Erdos (Eds.), *Lines of thinking: Reflections on the psychology of thought*. London: Wiley.
- Sperber, D. (2000) Metarepresentations in an evolutionary perspective. In Dan Sperber (Ed.). *Metarepresentations* (pp. 117-137). New York: Oxford University Press.
- Sperber, D., Cara, F., & Girotto, V. (1995). Relevance theory explains the selection task. *Cognition*, 52, 3-39.
- Sperber & Girotto (in press). Does the selection task detect cheater detection? In Fitness, J. & Sterelny, K. (Eds.), *New directions in evolutionary psychology*, Macquarie Monographs in Cognitive Science, Psychology Press.
- Sperber, D., & Wilson, D. (1995). *Relevance: Communication and Cognition. Second edition*. Oxford: Blackwell. [1st edition 1986]
- Sperber, D., & Wilson, D. (2002). Pragmatics, modularity and mind-reading. *Mind and Language*. 17. 3-23.
- Stein, E. (1996). *Without good reason. The rationality debate in philosophy and cognitive science*. Oxford: Clarendon Press.
- Wason, P.C. (1966). Reasoning. In B. M. Foss. (Ed.) *New Horizons in Psychology*. Harmondsworth: Penguin.