Does the selection task detect cheater detection?

Dan Sperber, Vittorio Girotto

DOES THE SELECTION TASK DETECT CHEATER-DETECTION?

Dan Sperber[a], Vittorio Girotto[b]

[a]*Institut Jean Nicod (CNRS, EHESS and ENS), 1bis avenue de Lowendal, 75007 Paris*

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

Evolutionary psychology—in its ambitious version well formulated by Cosmides and Tooby (e.g., Cosmides & Tooby 1987, Tooby & Cosmides 1992) —will succeed to the extent that it causes cognitive psychologists to rethink central aspects of human cognition in an evolutionary perspective, to the extent, that is, that psychology in general becomes evolutionary. The human species is exceptional by its massive investment in cognition, and in forms of cognitive activity—language, meta-representation, abstract thinking—that are as unique to humans as echolocation is unique to bats. The promise of evolutionary psychology is thus to help explain not just traits of human psychology that are homologous to those of many other species, but also traits of human psychology that are genuinely exceptional and that in turn help explain the exceptional character of human culture and ecology.

However, most of the work done in evolutionary psychology so far is on aspects of human psychology that are not specifically human except in their details. Showing, for instance, how human preferences in mate choice are fine-tuned in the way the theory of evolution would predict is of great interest (see e.g., Buss 1994) but it can be done on the basis of a relatively shallow psychology.

This makes work on distinctly human adaptations involving higher cognition of particular importance for defenders of a psychologically ambitious evolutionary psychology. What is often presented (e.g., Pinker, 1997) as the signal achievement of cognitive evolutionary psychology in this respect is the experimental testing of Cosmides' (1989) hypothesis that there exists an evolved competence to deal with social contracts, and, in particular to detect cheaters. We want to argue that, because of faulty methodological choices—the quasi-exclusive reliance on the four-cards selection task—, the hypothesis has in fact not yet been tested.

The plan of this chapter is as follows: We begin, with a short presentation of Cosmides's social contract hypothesis, of Wason selection task, and of Cosmides's reasons to use the task in order to test the theory. In a second section, we present the relevance-theoretic analysis of the selection task proposed by Sperber, Cara and Girotto (1995) which cast doubts on the appropriateness of the task to study reasoning. In a third section, we present Fiddick, Cosmides and Tooby's (2000) defense of the use of the selection task as a tool to test evolutionary theories of reasoning, and argue that it is methodologically flawed. In a fourth section, we

present three experiments designed to test contrasting predictions deriving from the two approaches. In the conclusion, we come back to Cosmides's hypothesis and reflect on how it might be really tested.

1. Social contract theory and the selection task

1.1. The theory

Cosmides' (1989) argued that, for cooperation to have stabilized among humans in evolutionary time, and not to have been made impossible by excessive free-riding or "cheating", humans must have evolved domain-specific cognitive capacities that allowed them to detect cheaters. She argued that the cognitive capacities in question consisted in a "social contract algorithm" specialized in reasoning about social contracts and allowing, in particular, to detect parties that were not abiding by the terms of the contract. It might be argued that cheater detection need not to be achieved by a mechanism so calibrated. It could have been that the solution to the problem of free-riding was already provided by other, more general cognitive mechanisms, or that such more general mechanisms were the best solution available to biological evolution at the time, even if the pressure then was just to solve this particular problem. It could be that cheaters' detection is achieved by several more specialized or diversely specialized mechanisms acting conjointly, and so on. Still, the hypothesis that there might be a cognitive mechanism calibrated to deal precisely with a problem that caused a specific selection pressure is, arguably, the best point where to start, the first hypothesis to test.

In non-human cases, cheating consists in failing to reciprocate in a standard kind of exchange that is part of the behavioral repertoire of the species. The ability to detect cheaters boils down to that of detecting quite specific behavioral features. For instance, in the famous case of vampire bats, detecting cheaters seems to be done by detecting individuals who, in spite of having a distended stomach, fail to regurgitate for the benefit of others some of the blood they have swallowed (Wilkinson 1984). In the human case on the other hand, there is an endless variety of forms or reciprocal interactions based on explicit or implicit contractual agreements. Therefore the detection of cheating is a cognitive challenge that cannot be based on

simple behavioral cues and that must rely on an understanding of the contracts involved.

Cosmides (1989) analyzed contracts as involving the commitment to meet a given requirement and the entitlement to receive a given benefit, and cheating as taking the benefit without meeting the requirement. No doubt, we would be surprised if we found non-human animals capable of deploying such a complex abstract notion. Humans however use and deploy in verbal communication a great variety of complex notions, many of them without particular evolutionary significance. Are the concepts of a social contract and that of a cheater exceptionally complex, or are they used with much greater ease than equally complex concepts? Does their effective deployment require ad-hoc domain-specific capacities? Cosmides believes that they are and set out to demonstrate this experimentally. Before formulating criticisms of Cosmides experimental approach, let us point out how innovative it was. She brought together evolutionary considerations with the most intensively used experimental paradigm in research on reasoning, the Wason (1966) selection task. Whereas much work with the selection task had modest theoretical aims (if any), she used the task for a novel and ambitious theoretical purpose.

1.2. **The task**

The standard version of the Wason selection task reads as follows:

There is a deck of cards, all of which have a letter on one side and a number on the other side. Four cards are taken from the deck and placed on the table. You can see the four facing sides as follows

$$\boxed{A} \quad \boxed{B} \quad \boxed{2} \quad \boxed{5}$$

The following rule applies to these four cards and may be true or false: "If there is an A on the one side of a card, then there is a 2 on the other side".
Which of these cards would you need to turn over to decide whether the rule is true or false?

In a typical deontic version, the selection task reads as follows (see Griggs & Cox, 1982)

Imagine that you are police officer on duty. The cards in front of you have information about four people sitting in a bar. On the one side of a card is a person's age, and on the other side is what a person is drinking:

Beer	Coke	16 year	24 years
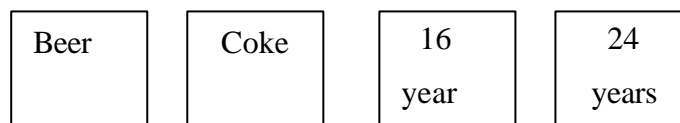
Which of these cards would you need to turn over to decide whether these people have obeyed the rule "If a person is drinking beer, then the person must be over 18 year of age"?

As these example illustrate, participants in a selection task 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 each an individual item. 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 items combining the P and the not-Q features, the logically correct selection is that of the P and of the not-Q cards, either of which could turn out to provide a counter-example to (or a violation of) the rule. However, with standard descriptive versions, more than 80% of the participants fail to solve the problem (and tend to select the P and the Q cards, or just the P card). With standard deontic versions, on the other hand, more than 50% of the participants make the correct P and not-Q selection (for a recent review, see Manktelow, 1999).

1.3. Using the task to test the theory

What does this have to do with cheater detection? Following Cosmides's (1989) analysis, what you need to know, in order to determine whether individuals who have entered into a contract are cheating is whether or not they have "taken the benefit,"

and whether or not they have "met the requirement" involved. These features can be treated as special cases of the P and Q of the selection task.

Cosmides then built deontic versions of the selection task where the cards represented individuals, and where features P/not-P were of the form "has/has not taken the benefit" and features Q/not-Q were of the form "has/has not met the requirement" (with the specific benefit and requirement mentioned on the cards determined by the narrative introduction to the task). Cosmides showed that most participants presented with such a task and asked who, among the four individuals represented, might be in violation of a rule of the form "if you take the benefit, then you must meet the requirement" correctly choose the individual known to have taken the benefit (and who might not have met the requirement) and the individual known not to have met the requirement (and who might have taken the benefit), that is, the two possible cheaters. For example, in one of her selection task scenarios, the male members of a tribe were described as seeing eating cassava root, an aphrodisiac, as a desirable benefit, but as having for this to meet a requirement: to have a tattoo on their face. When presented with the rule "If a man eats cassava root, then he must have a tattoo on his face" and four cards representing men of the tribe and indicating respectively "is eating cassava root" (P card), "is eating molo nuts" (not-P card), "has a tattoo on his face" (Q card) and "has no tattoo on his face" (not-Q card), most participants correctly selected the P and the not-Q cards. Other researchers have found similar results with a variety of deontic rules, and with both adults and children participants (e.g. Cheng & Holyoak, 1985; Girotto, Light & Colbourn, 1988; Manktelow & Over, 1991; Gigerenzer & Hug 1992).

Cosmides (1989), however, argued that not all deontic rules elicit correct selections, but only those which are processed by means of an underlying evolved algorithm such as the social contract algorithm. She concluded that whereas people do not have the general logical abilities required to solve the Wason selection task in every version, and not even in every deontic version, they have a specialized ability to reason about contracts, and in particular to identify cheaters.

To evolutionary psychologists who are otherwise unfamiliar with the psychology of reasoning, this experimental work may seem like a brilliant empirical confirmation of sound theoretical claims, but the fact is that, for good or bad reasons, specialists of reasoning in general and of the selection task in particular have not been swayed (for

a notable exception, see Gigerenzer & Hug. 1992). Among the bad reasons, may have been hostility to evolutionary approaches. Among the good reasons is the fact that these data are open to a variety of interpretations, among which Cosmides' (1989) own is not uniquely compelling (as pointed out by Atran, 2000). In general, no selection task data have ever been seen as compelling enough to convince the scholarly community of any theoretical claim about human reasoning. The task seems to be used again and again, almost compulsively, not because it ever provides conclusive evidence on anything theoretical, but because, on the contrary, it lends itself to endless controversies.

2. Is the selection task an appropriate tool for studying human reasoning?

Sperber, Cara and Girotto 1995 (henceforth SCG) have argued that the selection task is altogether an inappropriate tool for studying human reasoning (whether domain-general or domain-specific) because in all version of the task, people tend to be guided not by any form of reasoning but by context-sensitive intuitions of relevance (see also Evans, 1989). Intuitions of relevance are activated by the pragmatic mechanism involved in comprehending the task (just as they are by any comprehension process). People trust these intuitions and select cards accordingly. Different wordings and contexts yield different intuitions and therefore different selections.

SGG's approach is in no way hostile to evolutionary psychology. In fact, the relevance-guided comprehension mechanism involved in the selection task is viewed as an evolved module specialized for the comprehension of communicative intentions, and more specifically as a sub-module of a Theory-of-Mind mechanism (Sperber 2000; Sperber & Wilson 2002). This specialized comprehension module pre-empts any downstream domain-general or domain specific reasoning mechanism people might have otherwise appealed to, including a cheater detection module, if there is one.

What happens, SCG claimed, is that, in standard descriptive versions of the task, the intuitions provided by the comprehension module are misleading, whereas in most deontic versions, including those used by Cosmides (1989), they point to the correct selection. SCG 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. The recipe aims at making the P-and-not-Q cases contextually more relevant than the P-and-Q cases. For this, all you have to do is:

1. Choose P and Q so as to render the P-and-not-Q cases easier to mentally represent than the P-and-Q cases.
2. Create a context where knowing whether there are P-and-not-Q cases would have greater cognitive effects than knowing whether there are P-and-Q cases.
3. Present the rule "if P, then Q" in a pragmatically felicitous manner.

SCG argued that all previous genuine versions of the selection task where a majority of participants had given a correct response happened to conform to this recipe. They produced several descriptive versions of the task that elicited a higher percentage of correct responses than had ever been found before with such versions.

Consider as an illustration the following problem:

> Imagine that you are a journalist and that you are preparing a piece on the small principality of Bagurstan. The reigning Prince studied Economics in Oxford and has imposed a radical form of liberalism upon his country. In Bagurstan, people retire at 65, students get a salary, but they have no social security, no right to work, no unemployment benefits, no minimum salary, no maternity leave. Yet, the Prince claims that in Bagurstan there are no serious economic or social problems. Economic mechanisms alone allow everyone to find a satisfactory solution. For instance , the Prince asserts, in my country, if a person is of working age, then this person has a job
>
> Before you, on the desk of the Bagurstani civil servant that you are interviewing, are four information cards corresponding to four Bagurstani subjects. The top part of each card indicates the age. The bottom part indicates whether the person has a job. Unfortunately, some papers conceal the bottom of two of these cards, as well the top of two other cards.
>
> Which cards should be uncovered to determine whether what the Prince says (**if a person is of working age, then this person has a job)** is true, at least as far as these four Bagurstani are concerned.
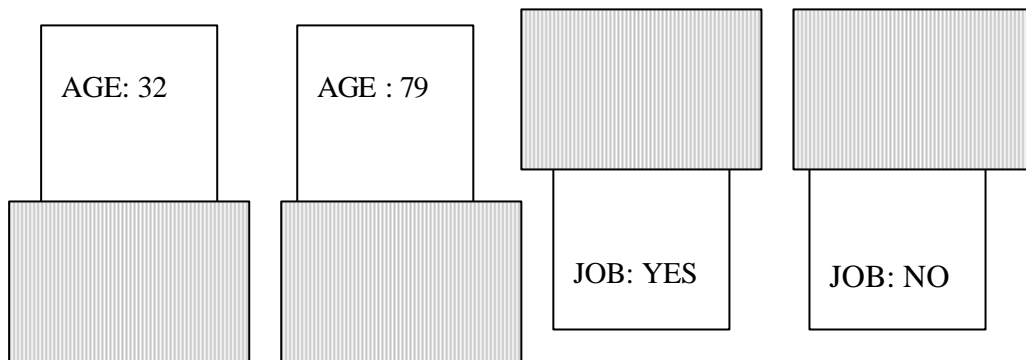
This version of the task does follow the relevance recipe: The conditional statement is uttered in a context in which it is likely to be understood as a denial of the existence of people of working age and without a job. Moreover, such a combination of features is easily representable, given that it is a highly salient and lexicalized category ("unemployed"). As predicted by SCG, most participants (70%) solved the task, by correctly selecting the individual of working age (and who might not have a job), i.e., the P card, and the individual with no job (and who might be of working age) i.e., the not-Q card.

By contrast, only 25% of the participants solved a control condition identical in every respect except that the Prince's assertion was "Of course, if a person is older than 65, then this person is without a job". With such a conditional rule, the problem does not anymore conform to the recipe: the case of a person older than 65 and with a job, a combination of features which is neither salient nor lexicalized, is not easier to represent than that of a person older than 65 and without a job, who normally falls in the well-known category of retired people. Moreover, the Prince's statement is understood as granting that there are people without a job in his country, namely people of retirement age, and not as denying that there might be a few people of that age that are still working.

Girotto, Kemmelmeir, Sperber, and van der Henst (2001) showed how deontic versions could equally well be manipulated so as to obtain at will correct selections or incorrect selections (of a type usually found with descriptive version, i.e. P and Q selections). In a within-subject design, these authors used the very same rule four times with different contexts, twice descriptively, and twice deontically. The rule was:

"If a person travels to any East African country, then that person must be immunized against cholera". In the first deontic condition, participants were informed of the existence of such a rule and asked which cards it was necessary to turn over in order to check whether clients of a travel agency represented by four card (with their country of destination on one side and their immunization on the other) had obeyed the rule. In the second deontic condition, participants were told that the very same rule was not anymore in force and were asked to check whether misinformed clients of the agency had nevertheless obeyed it. When the rule was described as being in force, 62% of the participants selected the usual P and not-Q cards. When the rule was described as not being in force anymore, only 15% of the participants selected the logically correct P and not-Q cards, whereas 71% chose the P and Q cards, representing people who might have obeyed the rule.

Findings similar to those of SCG have been reported (e.g., Green & Larking, 1995; Hardman, 1998; Liberman & Klar, 1996; Love & Kessler, 1995). While alternative or complementary interpretations of SCG's data have been proposed (see Johnson-Laird & Byrne, 1995; Oaksford & Chater, 1995) the only extensive challenge of SCG based on experimental data is the article by Fiddick, Cosmides, and Tooby 2000 (henceforth FCT). In the following sections, we discuss this challenge in detail.

3. Examining Fiddick, Cosmides & Tooby's defense of the use of the selection task as a tool to test evolutionary theories of reasoning

3.1. The principle of pre-emptive specificity

FCT properly note that SCG's proposal "is not intended to deny that evolved domain-specific inference mechanisms, such as social contract algorithms, exist, or that they play an important role in mental life. [...] The authors simply argue that, in explaining content effects on the Wason selection task, once relevance effects are subtracted, either (1) there is nothing else to explain, or (2) relevance effects inevitably confound experiments, rendering the Wason task useless as a means of detecting the operation of content-specialized inference mechanisms" (p. 4). FCT then go on to try and defend their use of the selection task against SCG deflationary

methodological claims. They do so by means of a variety of theoretical arguments and new experimental results.

FCT's main theoretical argument is based on what they call the "principle of pre-emptive specificity" according to which "human cognitive architecture should be designed so that, when two or more alternative inference systems are activated in a competing manner by the same input, the more specialized system will tend to override the more general one, dominating performance for that input" (FCT, p. 5). They agree that a relevance-based comprehension mechanism may well handle selection task problem in the general case, when no more specific mechanism is available. But in the cases where the Wason task can be interpreted as a cheater detection problem, then the more specialized cheater detection mechanism should pre-empt the more general comprehension mechanism.

The "principle of pre-emptive specificity" is an interesting and plausible hypothesis about cognitive architecture, but does it apply here? More specifically, are the comprehension module and the cheater detection module activated in a competing manner by the same input? As we pointed out, a human cheater detection mechanism must be activated not just by perception of the behavior of possible cheaters but also by a conceptual representation of the contractual relationship between the parties. In Cosmides' selection tasks, participants must comprehend the text of the problem in order to have access to information about the social contract involved. In other words, comprehension processes will anyhow take place before the task can be recognized as a cheater detection one, if it ever is. SCG's claim was that, because of the general structure of the selection task, mere comprehension of the problem automatically provides participants with intuitions of relevance about the cards. Since the question asked is about the very relevance of the cards, participants have in mind what feels like an answer to the problem before any further reasoning mechanism (whether a general reasoning ability if there is one, or a cheater detection module if there is one) gets fed the terms of the problem. Contrary to FCT's argument, the comprehension mechanism and the cheater detection mechanism cannot be "activated in a competing manner by the same input" since the input of the cheater detection mechanism has to be the output of the comprehension mechanism.

At this point we may invoke another general principle of good cognitive design: human cognitive architecture should be designed so that, when a problem has

already been given a solution by one mechanism, no other mechanism is activated to solve it a second time. Suppose however that, for some reason, this principle does not apply to the case at hand. Then, participants will have two solutions activated in their mind, one by comprehension, the other by cheater detection, which may be different or identical. SCG explained why, in general, the comprehension mechanism makes the P and not-Q cards stand out as the relevant ones in standard deontic selection problems, just as would a cheater detection mechanism. Thus, they claimed, there is indeed the risk of a confound between the comprehension and the hypothetical cheater detection mechanism (if the latter is activated at all). FCT granted that there might be such cases of confound. They pointed out that, "However, the confound works both ways: content specific cognitive adaptations may be doing the work attributed to conversational pragmatics on deontic versions of the selection task. Unconfounded experiments are needed to demonstrate which effect dominates" (FCT, p. 24) Their article presents several such experiments, that is, experiments where relevance theory on the one hand and social contract theory on the other hand should, according to them, predict different selection of cards. Such experiments might in principle help decide which of the comprehension module and the cheater detection module pre-empts the other.

3.2. Instructional blindness or pragmatic interpretation?

There are, however, major methodological problems with FCT's use of the selection task, and therefore with their data. To explain the problems—and how they might have arisen in the first place—let us underscore a peculiarity of the choice of the selection task by Cosmides and her collaborators as a means to establish the existence a cheater detection mechanism. According to FCT, "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" (FCT, p. 15). Now, in a standard deontic Wason selection task, participants are asked to select the cards that could represent violations of the conditional rule. Depending on the formulation of the rule, these

cards may or may not also represent potential cheaters. Cosmides' (1989) idea is that the evolved look-for-cheater strategy is so powerful (and attention to explicit instructions and logical properties so powerless) that participants will select the cards that represent potential cheaters, whether or not they constitute the correct solution. In other words, the selection task is being used to show that people are unable to follow its instructions. In standard descriptive tasks where no domain-specific module is activated, people fail to obey the instructions, whether they pay attention to them or not. In social contract problem, they follow the mental "instructions" of their cheater-detection module and just ignore the explicit instructions. Even when they seem to succeed at the task, participants are in fact doing something else.

It is peculiar to use, as one's experimental paradigm of choice, a task whose instructions, one believes, participants are unable to obey. However, if this allowed establishing that, presented with a Wason selection task, participants suffer from a kind of "instruction blindness" compensated by "looking-for-cheater" intuitions, it might be of genuine and indeed striking cognitive interest. This is clearly FCT's intent, who claim to show that "the inference mechanisms activated by social contracts and precaution rules short-circuit any logical SIAs [spontaneous inferential abilities] and content-general pragmatic factors that may exist, causing domain-appropriate inferences to be made even when these violate rules of logical inference" (pp. 62-63). Is this, however, what is shown by the experiments of Cosmides and her collaborators, and in particular by the experiments of FCT? There are two problems with this interpretation, one serious and one critical. Let us first evoke the merely serious problem.

It is one thing to point out the conceptual difference between cheating and logical violation; it is another to establish experimentally whether people tend spontaneously to look for cheaters rather than for violators. To do this by means of the selection task, it is necessary to find versions of the task where looking for violators and looking for cheaters results in different selection of cards. According to Cosmides (1989), this occurs in the case of "switched rules." Switched rules are obtained by switching the antecedent and the consequent of the original (unswitched) conditional rule (with possible minor linguistic adjustments, e.g. of tense). With an original rule of the form "if A, then B" , the swiched rule is of the form "if B, then A". This relationship corresponds to what in classical logic is called "conversion" . Since Aristotle, it is

known that a conditional and its converse are not logically equivalent. For instance, from a logical point of view, when B is true, "If A then B" is necessarily true, whether or not A is also true, whereas "If B then A" may be false, if A happens to be false. On the other hand, as we will soon illustrate, an original rule and its converse or "switched" counterpart may, in some contexts, be understood as expressions of the same social contract. In such cases, the same actions may constitute cheating both with switched and unswitched rules. Therefore, switching the rule may allow one to dissociate logical violation, which necessarily changes with switching, from acts of cheating, which may remain the same. Comparing participants' selection of cards with switched and unswitched rules should demonstrate whether they are looking for logical violations or for cheaters. If participants are looking for violations, then they should select different cards in the two conditions; if they are looking for cheaters, they should select the same cards in both conditions, or so the argument goes.

Here is FCT's own example of a switched rule:

> " … assuming 'you' are the potential cheater, 'If I give you $10, then you give me your watch' would be standard, and 'If you give me your watch, then I'll give you $10' would be switched. To detect cheaters, one needs to choose the benefit accepted card and the requirement not satisfied card, regardless of their logical category. Choosing these cards results in a logically correct answer, 'P & not-Q', for a standard social contract. But for a switched social contract, the same choice results in a logically incorrect (but adaptively sound) answer, 'not-P & Q'" (FCT, Fig. 1, p. 17).

But how should participants understand the instruction to look for potential violators of the rule "If you give me your watch, then I'll give you $10"? Is there any sensible logical and literal interpretation of this instruction? Interpreted literally and logically, the only behavior that could violate the rule is that of the individual who has made the contractual conditional promise and who could break it by taking the watch and not giving the ten dollars. The promisees, on the other hand, are not in a position to break the promise. However, in FCT's example, the promisor is not represented by any of the cards. The cards represent rather four promisees. So, the logically correct response to the instruction "Indicate only those card(s) you definitely need to turn over to see if any of these people are violating the rule" (FCT's own formulation in Table 1, p. 9), contrary to what FCT assert would not be the P card (has given the watch) and the not-Q card (has not received the $10), but no card at all. Again, the people represented by the cards cannot logically violate a promise they have not

made. Why then do participants choose the not-P card (has not given the watch) and the Q card (has received the $10) with such switched rules? For obvious pragmatic reasons: They interpret the rule and the instruction so as to make good sense of the problem. This is a textbook case of the very common phenomenon of "invited inference" and more specifically of "conditional perfection" (Geis & Zwicky, 1971; van de Auwera, 1977; Horn, 2000) where a conditional statement is interpreted bi-conditionally, as implying in this case, "if I give you 10$, you give me your watch". This bi-conditional interpretation is even reinforced by the very instructions of the task. It is only with this enriched interpretation that participants can make sense of the instruction to identify potential violators out of four people to whom a promise has been made. Then, the logical selection based on the pragmatically interpreted rule and the selection of potential cheaters are achieved by the selection of the same not-P card (has not given the watch) and Q card (has been given the $10).

Switched rules were meant to provide crucial evidence for the social contract theory, by prying apart the identification of violation and that of cheating. What we have just shown (see also SCG p. 86; Sperber & Girotto, in press) is, at the very least, that switched rules are pragmatically infelicitous and awkward and that results obtained with such rules are open to a quite sensible pragmatic interpretation different from FCT's. Even if FCT were to stand by their interpretation, the fact remains that the evidence obtained with switched rules data can be sensibly interpreted in a manner that takes away its evidential force in favor of social contract theory.

With social contract rules, separating cheating from formal violation of the rule is not easy and cannot be achieved by mere switching of the rule. One case where this separation has been genuinely achieved is Girotto et al. (2001). As reported above, these authors showed that participants select the usual P and not-Q cards, when the deontic rule "If a person travels to any East African country, then that person must be immunized against cholera" is presented as being in force. The cards selected represent both potential "logical" violators and potential cheaters, the two being equivalent. However, if, in an otherwise similar scenario, the same rule is presented as being not in force anymore, participants incorrectly select the P and Q cards. Note that, in this condition, the P and not-Q cards are the logically correct selection but do

not represent potential cheaters: a rule that is not in force can still, logically, be obeyed or disobeyed, but disobeying it is not a case of cheating.

These results (which were replicated in a between-participants design) are quite consistent with social contract theory, since participants show themselves to be interested in cheaters and not in violators per se. However they are also quite consistent with the relevance approach, since, when the rule is in force, what is relevant is the case of violators who may incur sanction, whereas when the rule is not in force, what is relevant is the case of people who have undergone an unnecessary immunization and may complain. Moreover, if both the social contract and the relevance approach can explain why participants select P and not-Q cards in the first and not in the second condition, only the relevance approach explains why they specifically select the P and Q card in the second condition (rather than answering at random).

3.3. Selection task vs. categorization task

Let us now turn to the other, not just serious but critical problem raised by FCT's methodology. To show that people are blind to the "logical" instructions of the selection task and obey rather the mental instructions of their social contract algorithm, it is essential that they be presented only with the standard instructions of the task and in particular that they not be explicitly asked to look for cheaters. If they are explicitly asked to look for cheaters and if that is what they do, the most parsimonious explanation of their behavior is that they have understood the instruction and have followed it. There is no obvious reason to invoke a dedicated mental mechanism. In general, in a deontic Wason selection task, participants should not be given any description of the cards to be selected, other than the fact that these cards must be relevant to establishing whether or not the rule is obeyed. Otherwise, it cannot be decided whether they are trying to follow the Wason task standard instructions, or whether they are selecting cards meeting the additional description. 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) have failed to respect this obviously sound methodological principle.

Far from heeding SCG's warning, FCT are going one dramatic step further. Rather than just adding to a standard Wason task instruction a direct question about cheating, as Cosmides (1989) had done, they altogether drop the standard instruction and just ask participants which of the four cards might represent a cheater (or a person endangering oneself in their Experiments 2 and 3 which are about precaution and not about social contracts). This is not a Wason selection task anymore. It is what we will call an "FCT selection task". In spite of superficial similarities, the two selection tasks, the Wason and the FCT one, are cognitively quite different. In the true Wason task, the 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, treating the rule as the major premise and information displayed on the visible side of the card as the minor premise in a conditional syllogism. What made the task apparently relevant to the study of reasoning is its surprising difficulty. More specifically, participants' failure to select the not-Q card could be seen as revealing their inability to perform a modus tollens syllogism, and their tendency to select the Q card could be seen as a revealing a disposition to commit the affirmation-of-the-consequent fallacy.

In FCT's task, a conditional statement may be part of the narrative, but it is not mentioned in the question. The question is of the form: which of the cards could represent an item falling under a given category (that of a cheater or that of a person taking risks, for instance)? This is a categorization task and not a conditional reasoning one. If, for instance, the category of a cheater is defined by the co-occurrence of the two features: "has taken the benefit", and "has not fulfilled the requirement", selecting the cards that might represent cheaters requires just picking those cards that display either one of these two features (and that might have the other defining feature on the other side).

More generally, one could construct FCT selection tasks with any category characterized by two features. Is there any reason to think that such tasks would be 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 investigated the matter in Sperber and Girotto (in press). We compared three FCT tasks, one where the

relevant category was that of a cheater, one where it was that of a glider which can be defined by the two features: "flying vehicle", and "with no engine", and a third arbitrary category which we could not resist calling "Wason selection" and which was characterized by the two features "food" and "non-Italian". All three categories involved a positive and a negative feature so that the correct selection was that of a P and a not-Q card.[1] Only one of the categories we used, that of a cheater, might plausibly be subserved by a evolved specialized mechanism. Whereas the defining features of a cheater and of a "Wason selection" were explicitly mentioned in the problem text, those of a glider were not, which might have made the selection of the cards potentially representing a glider more difficult. In spite of these differences, we predicted that participants would perform quite well in all three conditions.

Our prediction was confirmed: 68% of the participants made the correct selection in the cheater condition, 73% in the glider condition, and 91% in the "Wason selection" condition. We concluded that the FCT selection task is trivially easy. Participants do as they are told, whether the category is an evolutionarily significant one or not, whether it is one already known to the participants or not, and whether they are reminded of the defining features of the category or not. FCT tasks look interesting only if they are mistaken for genuine Wason tasks, which are puzzlingly difficult. When a Wason task condition is compared with an FCT task condition, as FCT do in their Experiment 2, the difference in performance is great, but this should not be more surprising than, say, the difference in performance of young children in an adding vs. multiplying the same two numbers. In both cases, the simplest algorithm for one task is objectively simpler than the simplest algorithm for the other. Only a difference in the opposite direction would be surprising and of genuine psychological interest.

---

[1] We had done another experiment with two arbitrary features both positive. The text of the problem read as follows: "There is a game consisting of a deck of cards, each of which has a letter on one side and a number on the other side. In this game the cards with a vowel on one side and an even number on the other are called "jokers". Below you find four cards. Two cards show the side with the letter. The two other cards show the side with the number." The cards displayed were: E, G, 7, and 4. The instructions were: "Indicate only the card or cards you definitely need to turn over to see whether among these cards there are any jokers." 100% of the 35 participants correctly chose the card with an E and that with a 4. The experiments reported in the text suggest that having the relevant category defined by a positive and a negative feature introduces a modicum of difficulty, but not enough to make the task really interesting.

In another experiment, we replicated FTC Experiment 1 and added two conditions to ascertain again whether or not, in FCT selection tasks, cheaters are more easily detected than items of any other category. FCT used two versions of a scenario where a South American farmer meets neighboring villagers who offer to give some of their corn for some of his potatoes. In one "conditional" condition, each villager expresses his offer conditionally: "If you give me some potatoes, then I will give you some corn." In another "want" condition, the offer is implicit: The villagers say, "I want some potatoes", and the farmer replies "I want some corn." The card represent four villagers and indicate: "You gave this person potatoes" (P card), "You gave this person nothing" (not-P card), "This person gave you corn" (Q card), and "This person gave you nothing" (not-Q card). Participants were then asked to "Indicate only the card or the cards you definitely need to turn over to see whether any of these people have cheated you."

In both conditions the majority did as they were told and selected the P and not-Q cards that represent possible cheaters. What was the point of this experiment? To show that the presence of a conditional rule is not necessary for cheater detection to operate. When the situation describes an exchange, and when moreover participants are instructed to take the perspective of a party to the exchange, then, according to FCT, the social contract algorithm is activated and participants look for cheaters, with or without a conditional expression of the contract. We believe that, since this is an FCT selection task, the presence or absence of a conditional statement is irrelevant, provided that the narrative is intelligible. We believe however that participants select the cards corresponding to potential cheaters just because they are asked to do so. There is no need to invoke a look-for-cheater strategy at work. To confirm our interpretation, we replicated FTC's experiment and added two conditions which differed from theirs only in the last words of the instructions. Our instructions were: "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." We predicted that participants, in these two conditions, would select the P and Q cards, that is, the two possible honest exchangers.

The result confirmed our prediction. In the replications of FCT two conditions, the most frequent selection was that of the possible cheaters (33% in the conditional version and 43% in the "want" version). In our versions, the most frequent selection was that of the potential honest exchangers (33% in the conditional version and 53% in the "want" version). With such an exchange scenario, participants seem to be no less disposed to look for an honest exchanger than to look for a cheater, provided that this is what they are asked to do (in all conditions, and in particular in the "conditional" ones, the rate of good performance is rather low, possibly because the scenario was not as clear to our participants as it may seem to be).

As we said, and hope now to have demonstrated, FCT's experimental evidence suffers from an major methodological problem. The task used is not a Wason but an FCT selection task. FCT selection task experiments are irrelevant to assess SCG's explanation of the Wason selection task. Moreover it is unclear what they might be relevant to.

Still, social contract theory is based not so much on these new experiments as on rich earlier experimental data. Apart from the methodological issues we have just discussed, there are also more empirical differences between FCT and us, regarding the expected outcome of various selection task scenarios. Since we dispute that selection task data is of relevance to the psychology of reasoning, we do not see these differences in prediction as having significant theoretical import. However, we believe that testing predictions may help decide between FCT's and our interpretation and evaluation of the Wason selection task. We do this with in the following experiments.

4. New experimental tests of Fiddick, Cosmides & Tooby's predictions

4.1. Experiment 1

FCT raise an interesting issue in their discussion of SCG. SCG had explained the fact that most people correctly select the P and not-Q cards in most deontic versions of the Wason task, by arguing that realistic deontic rules such as social laws achieve relevance by singling out and forbidding specific states of affairs represented by this selection of cards. For instance the drinking age rule: "If a person drinks beer, then

that person must be over 19 years old" (Griggs and Cox 1982) implicitly distinguishes four cases that there would be little reason to distinguish in the absence of the rule:

> drinking beer and being over 19,
> not drinking beer and being over 19,
> drinking beer and being under 19,
> not drinking beer and being under 19.

Out of these four cases, the rule singles out the third one—drinking beer and being under 19—by forbidding it. The three other cases are not in violation of the rule and are of no particular relevance. In general, social laws make relevant not behavior that conforms to them but behavior that violates them. Girotto et al. (2001) showed how this too can be manipulated by pragmatic means.

FCT (p. 30) point out, however, that this reasoning "does not extend easily to the class of private exchanges (trades) as distinct from social laws. In formulating their argument that the communicative intent is to forbid, SCG analyzed social laws, where their argument more plausibly fits […]. But social contracts expressing a trade lack key features upon which SCG's analysis of deontic rules is based. Their purpose is to enable mutual access to goods through a trade, not to forbid access to a good, and this has implications for what subjects should find relevant. […] a consistent application of relevance theory to private exchanges should lead subjects to answer 'P and Q' or 'P, Q, and not-Q' for trades, rather than 'P and not-Q'." We are grateful to FCT for this observation, which, we think, is essentially correct. There are interesting exceptions, of course; in some cases of trade, the risk of being cheated is more salient than the opportunity of obtaining goods, as we will illustrate. More generally, we believe that any type of rule can be manipulated to give logically correct or incorrect selections. Still, we accept that our account of the Wason selection task commits us to predict that most subjects will select the P and Q cards, or the P, Q and not-Q cards in standard cases of exchange. Now, it might seem that, if not FCT's new results, at least earlier results of Cosmides (1989) and others show that this prediction is dead wrong. Experiments have repeatedly shown that the majority of participants select P and not-Q in such cases, just as they do with social laws.

At this point, we must either concede that something is wrong with the relevance theory approach to the selection task, or else argue that something is wrong with the experimental evidence. In general, experimental evidence trumps theoretical claims; this, after all, is the reason for seeking experimental evidence in the first place! However, in this case, we have already demonstrated that many experiments performed by Cosmides, Tooby, and their collaborators either are not Wason selection tasks but FCT selection tasks, or are an unfortunate mix of the two (which we will call "FTC/Wason tasks"), with participants being simultaneously asked two questions, a Wason one and an FCT one, as if it was the same question asked twice. In order to clarify the issue, one should therefore compare genuine Wason selections tasks and mixed Wason/FCT tasks with private trade scenarios. We predict, as FCT correctly pointed out we should, a high rate of P and Q, or P, Q and not-Q selections with the genuine Wason tasks; we predict, on the other hand, a high rate of P and not-Q selections with mixed Wason/FCT tasks asking for cards representing possible cheaters, as has been found previously in the literature. We believe that when, in these Wason/FCT tasks, participants choose the cards that could represent cheaters, they are answering the easy FCT question about cheaters, and not the more difficult Wason question as to whether or not the rule is being respected. If this is correct, the responses should be different when only the Wason question is posed.

In Experiment 1, we compared two conditions:

(i) A genuine Wason selection task, where participants are asked which cards should be turned over to see whether the conditional rule expressing the conditions of a private trade was respected. Even though the trade has the canonical requirement-benefit structure, we did not expect participants to look for cheaters, given that they were not asked to do so. We expected them rather to look for potential instances of exchange.

(ii) A mixed Wason/FCT selection task with exactly the same scenario as in the preceding condition, except that participants were asked both whether some cheating had taken place and which cards should be turned over to see whether the rule was respected. This is of course methodologically flawed, but it reproduce a flaw found in previous experiments, such as Cosmides's (1989) Big Kiku experiment discussed in Sperber and Girotto (in press). We predicted that

we would find results similar to those found in these experiments, that is, a high rate of P and not-Q selection.

Method

Participants

The participants in all our experiments were humanities undergraduates at Udine University (Italy), who participated in the experiment voluntarily. In Experiment 1, 60 were tested. They were randomly assigned to one of two equal-sized groups (N= 30): "private trade (Wason version)", "private trade (Wason/FCT version")".

Material and procedure

The study was carried out in Italian with native speakers of the language. Each participant had to solve one problem. The participants were tested in groups, but they worked on the problem individually, at their own pace. The text was identical in both the Wason version and the Wason/FCT version, except for the words in brackets in the final instructions, which appeared only in the Wason/FCT version. The problem read as follows:

A group of tourists visits the island of the Boro-Boro people. Some tourists and natives agree to perform the following exchange: If a tourist gives a native a can of tomato sauce, the native gives him or her a bracelet.
Below you find four cards, each of which represents a different encounter between a given tourist and a given Boro-Boro native. On the one side, the card indicates whether the tourist has given the native a can of tomato sauce. On the other side, the card indicates whether the native has given the tourist a bracelet.

Mr. Rossi: has given Kiku a can of tomato sauce	Mr. Bianchi: has given Kapa nothing
Kanesa: has given Mr. Neri a bracelet	Kori: has given Mr. Verdi nothing

Indicate only the card or the cards you definitely need to turn over to see whether [there are tourists who have been cheated by the natives, and], as far as these four

encounters are concerned, the rule "If a tourist gives a native a tomato sauce can, the native gives him/her a bracelet" has been respected.

Results and Discussion

Table 3 presents the percentages of the main selection patterns in the two conditions.

Table 3

Percentage of the main selection patterns in the two conditions of Experiment 1 ($\underline{N}$ = 30)

Condition		
Pattern	Private trade (Wason version)	Private trade (Wason/FCT version)
P, not-Q	0	57
P, Q	63	0
P,Q,not-Q	10	0
P	13	0
Q	7	0
not-Q	0	23
Other	7	20

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

As predicted, a majority of participants selected P and Q, or P, Q and not-Q cards in the Wason version of the task, whereas no one selected these cards in the Wason/FCT version (χ^2(1, $\underline{N}$=60) = 19.9, $\underline{p}$ < .001). Similarly, a majority of participants selected the P and not-Q card in the Wason/FCT version, whereas no one selected these cards in the Wason version (χ^2(1, $\underline{N}$=60) = 23.3, $\underline{p}$ < .001).

It could be argued, however, that the direct exchange described in the scenario did not leave much room for cheating, and that this is why, in the Wason version,

participants did not select cards representing potential cheaters. Note that the same argument could be leveled against FCT's Experiment 1 (replicated by Sperber & Girotto forthcoming and discussed above), but to make sure anyhow, we performed a modified version of Experiment 1 with a scenario where the risk of one party being cheated by the other was a genuine and obvious one.

4.2 Experiment 2A

Experiment 2 came in two parts. In Experiment 2A, the method was exactly the same as in Experiment 1, with two groups of 30 participants assigned to a genuine Wason task and to a mixed Wason/FCT task. We used the same scenario as in Experiment 1except that we replaced the rule "If a tourist gives a native a can of tomato sauce, the native gives him or her a bracelet" with the rule "if a native gives a tourist a bracelet, the tourist will send him or her a picture from Italy". We modified the cards accordingly. The instructions read (with the added part of the FCT/Wason version in square brackets): "Indicate only the card or the cards you definitely need to turn over to see whether [there are natives who have been cheated by the tourists, and], as far as these four encounters are concerned, the rule " If a native gives a tourist a bracelet, the tourist will send him or her a picture from Italy" has been respected."

In this scenario, there is a genuine and obvious risk that a tourist who had accepted a bracelet might, out of negligence or dishonesty, fail to send a picture from Italy. Nevertheless, in the Wason version, we assumed that the very fact that such exchanges would have taken place between tourists and Boro-Boro natives might be more relevant to the majority of the participants than the fact that some of the tourists might not have performed their part of the deal, and that they would therefore select the P and Q card. We also expected that, when, in the Wason/FCT version, participants were asked directly about cheating, then they would select the P and not-Q cards.

The results corroborated our predictions. In the Wason version of the task, 50% selected the P and Q cards, whereas only 13% made this selection in the Wason/FCT version (χ^2(1, $\underline{N}$=60) = 7.70, $\underline{p}$ < .01). Similarly, in the latter version, 47% selected the P and not-Q cards whereas only 17% selected these cards in the former (χ^2(1, $\underline{N}$=60) = 4.93, $\underline{p}$ < .05). On the other hand, contrary to what one might

also have expected, there were no P, Q and not-Q selections; that is, participants were interested either in the exchanges or in the cheatings, but not in both. Overall, as we anticipated, the results of Experiment 2A were very similar to those of Experiment 1, in spite of the much greater plausibility of cheating in the modified scenario. This confirms the prediction, suggested to us by FCT, that private trades in true Wason tasks elicit more incorrect P and Q selections than correct P and not-Q selection. The impression from earlier data (e.g., Cosmides, 1989) that private trades, with their paradigmatic requirement-benefit structures, elicit selections of cards representing potential cheaters (P and not-Q in the regular versions, and not-P and Q in the "switched" version) was based on the methodologically unsound use of mixed Wason/FCT versions of the task.

4.3. Experiment 2B

From the pragmatic point of view we are defending, we did not doubt that it is possible to devise a genuine Wason selection task based on a private trade such that participants select the P and not Q cards, that is, the cards representing potential cheaters. All that needs to be done is create a context where the risk of cheating particularly salient and is more relevant than the possibility of a regular exchange. Consider for instance the following problem:

A new cure for insomnia is being advertised by the Morpheus drug Company. The Company's offer is: "If a person sends 40 euro to the Morpheus Company, then the company sends that person a bottle of the cure."
An survey has been done on people who buy drugs against insomnia. Below you find four cards, each of which represents one of these persons. On one side, the card indicates whether the person represented by the card has sent 40 euro to the Morpheus Company. On the other side, the card indicates whether the Morpheus Company has sent a bottle of the cure to the person represented by the card.

Mr. Rossi	**Mr. Bianchi**
Has he sent 40 euro to the Morpheus Company?	Has he sent 40 euro to the Morpheus Company?
Yes	No

Mr. Neri Has the Morpheus Company sent him the cure? Yes	Mr. Verdi Has the Morpheus Company sent him the cure? No

Indicate only the card or cards you definitely need to turn over to see whether, as far as these four people are concerned, the rule "If a person sends 40 euro to the Morpheus Company, then the company sends that person a bottle of the cure" has been followed.

Given that this describes buying of goods by customers from a commercial company, the occurrence of actual buyings is more banal and less relevant than that of one-time barters between individuals. Moreover, given the lack of confidence (at least in Europe, at present) regarding both mail trade and the selling of miracle drugs, the risk of consumers being cheated is particularly salient, and the offer of the Morpheus company is likely to be read as implying a commitment not to fail to send the cure to those who send the money. Because of these factors, such a problem would be more likely to elicit the selection of cards representing possible cheatings than most examples of private trade. We tested this prediction by asking a new sample of participants ($\underline{N}$ = 25) to perform the Morpheus Company selection task. As predicted, and as found in the Wason/FCT versions of the Boro-Boro problems, most participants (44%) correctly selected the P and not-Q cards, whereas only 16% selected the P and Q cards. This shows that, even when asking whether the rule has been followed rather than whether it has been violated, even without instructing participants to adopt the perspective of a potential victim of cheating, and of course without instructing them to look for possible cheaters, they do so in contexts in which the possibility of cheating is sufficiently salient and plausible to be more relevant than the occurrence of actual exchanges. However, it is not the case that private trades in general provide such a context.

4.4. Experiment 3

There is another difference in prediction between FCT and SCG. FCT (p. 41, n. 24; Appendix A) argue that not all social contracts, but only those where the requirement comes in compensation of a benefit should yield the selection of the P card (benefit received) and the not-Q card (requirement not fulfilled). SCG predict

that any statement of a social law that achieves relevance by forbidding violations (rather than by promoting conformity) should trigger the selection of the P and not-Q cards whether or not the requirement it expresses comes in compensation of a benefit. As shown by Girotto, Blaye and Farioli (1989), even children do solve selection problems presenting such social laws (e.g. "If a bee is sick, then it must stay out of the hive"). Repressive social laws are cases in point. For instance, we expect that a Wason selection task using the rule "If a person has been condemned by a tribunal, then that person must always carry his/her I.D. card" would yield a high rate of P and not-Q selections. Yet, the requirement of carrying one's I.D. does not come as a prize paid for the benefit of having been condemned.[2]

We tested this prediction in Experiment 3. The method was exactly the same as in Experiment 1 and 2, except that we used the following scenario:

In your country, a new law has been introduced, in order to improve public security. According to this law:
"If a person has been condemned by a tribunal, then that person must always carry his/her I.D. card."
An investigation has been done on randomly selected people. Below you find four cards, each of which represents one of these people. On the one side, the card indicates whether the person has been condemned by a tribunal. On the other side, the card indicates whether the person carries his/her I.D. card.

Mr. Rossi. Has he been condemned by a tribunal: Yes	Mr. Bianchi. Has he been condemned by a tribunal: No
Mr. Neri. Does he carry his I.D. card: Yes	Mr. Verdi. Does he carry his I.D. card: No

---

[2] FCT also suggest that there is another evolved mental mechanism dedicated to hazard management. This mechanism will cause people to select the P and not-Q cards in a selection task with a rule of the general form "if you engage in a dangerous activity, then take the appropriate precaution." However carrying one's I.D. is not a precaution taken against the "activity" of having been condemned. The only danger for the people to whom the rule applies is that of being in violation of the rule. Surely the rule can not be seen as a precaution against a danger that is created by the very existence of the rule itself.

Indicate only the card or cards you definitely need to turn over to see whether, as far as these four people are concerned, the law "If a person has been condemned by a tribunal, then that person must always carry his/her I.D. card" has been followed.

Note that this is a genuine Wason selection task. Even though the question was about the law being followed rather than being violated, even though participants were not instructed to take the perspective of a law-enforcing character, and even though, of course, the requirement expressed by the social law does not come in compensation of any benefit, 47% of the participants selected the P and not-Q cards, and none of them selected the P and Q cards.

This result confirms that, in genuine deontic Wason selection tasks, what drives selection of cards is not the requirement-benefit structure of the rule but the relative relevance of conformity vs. violation.

4.5. Discussion of the experiments

In some deontic Wason selection tasks, predictions based on relevance theory differ from FCT's predictions. The present results suggest that, in those cases, the relevance based predictions are the correct ones. Impression to the contrary are based on a faulty methodology using FCT tasks, or mixed Wason/FCT tasks.

Since we don't believe that the Wason selection task is a good tool to investigate anything other than the pragmatic interpretation of conditional statements (and it is not such a good tool for that either), we emphatically do no intend these results to serve as evidence for any claim about the mental representation of private trades or social laws. We are well aware—this has been the whole point of SCG and Girotto and al. (2001)—that any kind of rule can be contextualized so as to favor many kinds of selections. We see these results as showing just three things:

(i) It is not the case that social contracts with a requirement-benefit structure making cheating possible automatically elicit a look for cheater procedure.

(ii) It is not the case that social laws without such a requirement-benefit structure necessarily fail to elicit a search for violations.

(iii) Mixing true Wason and FCT instructions is likely to yield trivial results based just on the FCT part of the instructions.

6. Conclusion

What does all this mean regarding Cosmides' (1989) hypothesis about the existence of a social contract algorithm? It just means that this hypothesis has not been tested yet! The search for evidence went the wrong way. Of course, the fact that the search for evidence was centered on the selection task influenced the way the hypothesis itself was formulated and refined. So, time has come to rethink and retest the hypothesis. Let us just make some suggestions as to how this might be done.

Being able to conceive of, and enter in contractual agreements involves not just higher cognitive abilities but also higher communicative abilities. For all except the simpler kinds of contractual agreements, it requires a language. Although it might be an interesting line to explore, no one is suggesting that human communicative abilities and in particular human language have evolved as adaptations specifically geared to the passing of contracts. Everybody assumes on the contrary that, from the start, the contents that could be linguistically communicated were quite diverse, implying that at least a good part of the cognitive abilities needed to conceive of these contents were contemporaneously available.

A standard and plausible view is that the cognitive abilities in evidence in linguistic communication involve the possession of conceptual primitives and the capacity to combine these conceptual primitives into more complex concepts. Conceptual primitives may need cognitive inputs to be activated and fine-tuned, but they are each based on a specific evolved disposition. Complex concepts resulting from the combination of primitive ones need not have a distinct evolutionary history. New combinations can be devised all the time. This, after all, is what combinatory ability is for. Still, there may exist evolved dispositions to combine some primitive concepts in specific ways if the resulting complex concept is of particular evolutionary significance and is to be used in special ways.

Cosmides does not discuss the issue of the primitive or complex character of the concept of a contract and of the associated concept of a cheater. She just treats both concepts as complex ones, combining two simpler ones, that of a right (the benefit) and that of a duty (the requirement). A good case can be made for considering the concepts of right and duty to be primitives (see Holyoak & Cheng, 1995; Politzer &

Nguyen-Xuan, 1992), and part of the evolved conceptual repertoire of humans (and arguably only of humans).

What Cosmides's (1989) approach implies is that the complex concepts of a contract and of a cheater are not just a combination of primitives concepts of rights and duties, but that their deployment is based on a specific evolved disposition.[3] Given the evolutionary significance of these concepts, this is not implausible. But in which way is the deployment of these concepts more than the deployment and combination of their primitive parts? Cosmides hypothesis is that in the use of these concepts, humans are capable of drawing specific kinds of inferences that are not available to them in the case of other complex concepts. These inferences may mimic the "modus tollens" form in their special domain, but they are not genuine domain general modus tollens inferences. She takes her experimental data to provide crucial evidence for such a hypothesis.

At this point we doubly disagree. We believe that what makes modus tollens inferences easy or hard is not the specific subject matter (e.g., contracts versus colors) but the pragmatics of the utterances in which these notions are used. People spontaneously perform modus tollens inference in all cognitive domains when this is the way to give a relevant interpretation to the utterance they are processing (see Politzer, forthcoming). Moreover, we have argued that, when they identify potential cheaters by selecting individuals who are known either to have taken the benefit or not to have met the requirement, Cosmides' (1989) experimental participants merely perform simple Boolean combinations of concepts. The ability to use such simple combinations of features (including negative features) for categorization is found across domains: a bitch is a female dog, a Barbarian was a non-Greek human being, a stool is a seat without a back, a glider is an aircraft without an engine. We have shown (Sperber & Girotto, in press) that detecting potential instances of such categories, whether it be an evolutionary important category such as that of cheaters, or an unimportant one such as that of gliders, is trivially easy.

---

[3] In fact, it would not be inconsistent for Cosmides (1989) to argue that the concepts of a contract and that of a cheater are not acquired by combining simpler concepts but are themselves primitives and are linked to the concepts of right and duties by meaning postulates or inference rules.

This not the end of the story. Even if our criticisms are justified, it would be wrong to dismiss Cosmides' (1989) basic hypothesis that there may be cognitive adaptations dedicated to the detection of people cheating on social contract. Even if our criticisms were mistaken and Cosmides and her collaborators had actually uncovered such an adaptation in reasoning about cheaters, one should consider the possibility that the risk of being cheated put selective pressure not just on reasoning with the concept of a cheater, but also, and maybe more importantly, at other levels.

To begin with, cheaters have a vested interest in doing their best to hide the fact that they are cheating, or, at least, to hinder as much as possible access to evidence of their cheating. This can be done in a variety of ways: by making vague or ambiguous contractual agreements, by making the benefit seem less valuable, by making the requirement seem more costly, by invoking special circumstances such that failure to meet the requirement does not count as cheating, and so on. Detecting cheaters may requires forms of strategic reasoning that are much richer than what is needed in order to put together "has taken the benefit" and "has not met the requirement".

As Cosmides and Tooby (1989, 1992) themselves have noted, the risk of cheating is particularly high when there is a time gap between the taking of the benefit and the fulfillment of the requirement. The potential victim of cheating must keep track of past transactions, which is not always easy. To give but one typical example, among the Dorze of Southern Ethiopia, a society where one of us has done ethnographic fieldwork, people had to remember dozens or even hundred of "gifts" made or received over several years, and which were all meant to be reciprocated when the proper circumstances would arise. Until recently, they were doing so without the help of writing. As this case illustrates, what makes detecting cheaters difficult in many case, is the fact that clear evidence of cheating is hard to find or to keep track of [4].

---

[4] The difficulty to keep track of past events may be the source of errors in other reasoning domains. In particular, Girotto and Gonzalez (2001, 2002) pointed out that individuals may produce erroneous frequency predictions in situations in which they acquire information through a sequence of actual observations (e.g., Gluck & Bower, 1988). These results contrast with the finding that individuals tend to make correct frequency predictions in verbal problems presenting numerical symbols, rather than actual events, and with the claim that the human mind contains an evolved mechanism to make accurate frequency predictions (e.g., Cosmides & Tooby, 1996; Gigerenzer & Hoffrage, 1995). Along with the present

Given all this, there may be more than one type of cheater-detection mechanism. One such mechanism particularly likely to have evolved would result in an automatically high level of attention to, and of recall of, details of interaction that might be evidence of cheating. The prediction that evidence of cheating is more carefully attended and better remembered than others is experimentally testable (for an example, see Mealey, Daood, & Krage, 1996). Such a prediction would be worth testing across cultures. Other likely sources of evidence are to be found in experimental developmental psychology and in the study of moral development and of moral emotions in particular.

There is a variety of methods by means of which evolutionary psychology hypotheses such as Cosmides' (1989) might be seriously tested. Experimenting with the Wason selection task is not one of them. Investing massively in a single experimental paradigm is anyhow unsound (and don't count using patients with brain lesions [see Adolphs, 1999] with Wason or FCT selection tasks as a genuinely different paradigm). Cognitive psychology and neuro-psychology have developed a rich battery of sophisticated experimental tools. If evolutionary psychologists want to exert a major influence on the field, they must share in, and contribute to this sophistication, and not overuse an oldish experimental paradigm. Moreover, the use of diverse and sophisticated methods would be likely to contribute to the enrichment and refinement of evolutionary psychology hypotheses.

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(2), 187-193.

Buss, D.M. (1994). *The evolution of desire: Strategies of human mating.* New York: Basic Books.

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

remarks, these results warn against testing evolutionary hypotheses about reasoning abilities by using standard verbal problems and suggest developing alternative methods.

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. (1987). From evolution to behavior: Evolutionary psychology as the missing link. In J. Dupré (Ed.), *The latest on the best: Essays on evolution and optimality*. Cambridge MA: MIT Press, 277-306.

Cosmides, L. & Tooby, J. (1989). Evolutiona ry 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.* New York: Oxford University Press. 163-228.

Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition, 58*, 1-73

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.

Geis, M.C., & Zwicky, A.M. (1971). On invited inferences. *Linguistic Inquiry, 2*, 561-566.

Gigerenzer, G., & Hug, K. (1992). Reasoning about Social Contracts: Cheating and Perspective Change, *Cognition, 43*, 127-171

Gigerenzer, G., & Hoffrage, U. (1995). How to improve bayesian reasoning without instruction: Frequency format. *Psychological Review, 102*, 684-704.

Gigerenzer, G., & Hug, K. (1992). Domain specific reasoning: Social contracts, cheating, and perspective change. *Cognition, 43*, 127-71

Girotto, V., Blaye, A., & Farioli, F. (1989). A reason to reason: Pragmatic basis of children's search for counterexamples. *European Bulletin of Cognitive Psychology, 9*, 297-321.

Girotto, V., & Gonzalez, M. (2001). Solving probabilistic and statistical problems: A matter of question form and information structure. *Cognition*, 78, 247-276

Girotto, V. & Gonzalez, M. (2002). Chances and frequencies in probabilistic reasoning: Rejoinder to Hoffrage, Gigerenzer, Krauss and Martignon. *Cognition*, in press.

Girotto, V., Kemmelmeir, M., Sperber, D., & van der Henst, J.B. (2001). Inept reasoners or pragmatic virtuosos? Relevance and the deontic selection task. *Cognition*, *81*, 69-76.

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

Gluck, M. A., & Bower, G. H. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General, 117,* 227-247.

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.

Horn, L. (2000). From IF to IFF: Conditional Perfection as Pragmatic Strengthening, *Journal of Pragmatics 32,* 289-326.

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

Liberman, N., & Klar, Y. (1996). Hypothesis testing in Wason's selection task: Social exchange, cheating detection or task understanding. *Cognition, 58,* 127-56

Light, P.H., Girotto, V., & Legrenzi, P. (1990). Children's reasoning on conditional promises and permissions. *Cognitive Development, 5,* 369-83

Love, R., & Kessler, C. (1995). Focussing in Wason's selection task: Content and instruction effects. *Thinking and Reasoning, 1,* 153-82

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

Mealey, L., Daood, C. & Krage, M. (1996). Enhanced memory for faces of cheaters. *Ethology and Sociobiology: 17,* 119-128

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

Pinker, S. (1997). *How the mind works.* New York: Norton.

Politzer, G., (forthcoming). Reasoning, judgement and pragmatics. In I. Noveck & D. Sperber (Eds.), *Experimental pragmatics.* London: Palgrave.

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,* 401-12

Sperber, D. (2000) Metarepresentations in an evolutionary perspective. In D. Sperber (Ed.). *Metarepresentations.* New York: Oxford University Press. 117-137.

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

Sperber, D., & Girotto, V. (in press) Use or misuse of the selection task? Rejoinder to Fiddick, Cosmides, and Tooby. *Cognition*

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and Cognition. Second edition.* Oxford: Blackwell

Tooby, J. & Cosmides, L. 1992. The psychological foundations of culture. In J. Barkow, L. Cosmides, & J. Tooby, (Eds.), *The Adapted Mind: Evolutionary Psychology and the Generation of Culture.* New York: Oxford University Press, 19-136.

van der Auwera, J. (1997). Pragmatics in the last quarter century: The case of conditional perfection. *Journal of Pragmatics,* 27, 261-74

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

Wilkinson, G.S. (1984). Reciprocal food sharing in the vampire bat. *Nature. 308*, 181-184.