Chapter 6
Social Contracts and the Wason Selection Task:
Experiments

The game-theoretic conditions governing reciprocation and non-reciprocation in social exchange are too complex and too important to leave to the vagaries of trial-and-error learning. Humans should have evolved inferential procedures that make them very good at detecting cheating on social contracts. The "look for cheaters" procedure predicted in Chapter 5 can be expected to generate a quite specific and unusual pattern of responses on the Wason selection task when its content involves social exchange.

In a social exchange situation for which a subject has incomplete information, a "look for cheaters" procedure should draw attention to any person who has NOT paid the required cost (has he illicitly absconded with the benefit?) and any person who has accepted the benefit (has he paid the required cost?).

The Wason selection task is a paper and pencil problem that invites a subject to see if a conditional rule of the form "If P then Q" has been violated by any one of four instances (represented by cards) about which the subject has incomplete information (see Chapter 2). By presenting one term of a conditional rule as a rationed benefit, and the other term as a cost/requirement, one can create a Wason selection task that instantiates a social contract. This can be used to see how people reason about social contracts in the face of incomplete information. Figure 6.1 shows the cost/benefit structure of such a Wason selection task.
It is your job to enforce the following law:

**Rule 1 — Standard Social Contract (STD-SC):** "If you take the benefit, then you pay the cost."

\[(P \rightarrow Q)\]

**Rule 2 — Switched Social Contract (SWC-SC):** "If you pay the cost, then you take the benefit."

\[(P \rightarrow Q)\]

The cards below have information about four people. Each card represents one person. One side of a card tells whether a person accepted the benefit and the other side of the card tells whether that person paid the cost.

Indicate only those card(s) you definitely need to turn over to see if any of these people are breaking this law.

<table>
<thead>
<tr>
<th>Rule 1 — STD-SC:</th>
<th>Rule 2 — SWC-SC:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benefit Accepted</td>
<td>Benefit NOT Accepted</td>
</tr>
<tr>
<td>Cost Paid</td>
<td>Cost NOT Paid</td>
</tr>
</tbody>
</table>

 Irrespective of logical category, a "look for cheaters" procedure would cause the subject to:

1. Choose the "cost NOT paid" card and the "benefit accepted" card. These cards represent potential cheaters.
2. Ignore the "cost paid" card and the "benefit NOT accepted" card. These cards represent people who could not possibly have cheated.

As Figure 6.1 shows, the logical category to which each card corresponds varies, and is determined by where the costs and benefits to the potential cheater are located in the "If-then" structure of the rule. For a "standard" social contract (STD-SC) -- one with the benefit in the "If" clause and the cost in the "then" clause -- the two chosen cards correspond to the logical categories 'not-Q' (cost NOT paid) and 'P' (benefit accepted). However, for a "switched" social contract (SWC-SC) -- one with the cost in the "If" clause and the benefit in the "then" clause -- the same two cards correspond to the logical categories 'not-P'
(cost NOT paid) and 'Q' (benefit accepted). In Figure 6.1, Rule 1 is a STD-SC and Rule 2 is a SWC-SC.

The correct formal logic response is 'P & not-Q', regardless of content. Therefore, a subject using a "look for cheaters" procedure would appear to be reasoning logically -- by choosing 'P & not-Q' -- on STD-SC rules, yet appear to be reasoning illogically -- by choosing 'not-P & Q' -- on SWC-SC rules. In the literature, a rule is said to have produced a content effect when it elicits a higher percentage of logically falsifying, 'P & not-Q' responses than an abstract problem. By this definition, a STD-SC should elicit a content effect but a SWC-SC should not.

A "look for cheaters" procedure should ignore the "cost paid" card and "benefit NOT accepted" card. These correspond to 'P & not-Q' (the correct formal logic response) for a SWC-SC rule, and to 'not-P & Q' for a STD-SC rule. Thus social contract theory predicts that the dominant response to a STD-SC problem will be very rare on a SWC-SC problem, and vice versa.

To sum up, the correct SC answers to STD-SC and SWC-SC rules differ from the logically correct answers:

<table>
<thead>
<tr>
<th>SC answers</th>
<th>Logical answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>'P &amp; not-Q'</td>
<td>'not-P &amp; Q'</td>
</tr>
<tr>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>yes</td>
<td>no</td>
</tr>
</tbody>
</table>

Therefore, by comparing performance on STD-SC and SWC-SC rules one can tell if reasoning is governed by a logical procedure or a "look for cheaters" procedure.

If people do, in fact, have Darwinian algorithms governing
how they reason about social exchange, these ought to function, in part, as frame-builders that structure new experiences. This means they should operate in unfamiliar situations. No matter how unfamiliar the relation or terms of a rule, if the subject perceives the terms as representing a rationed benefit and a cost requirement -- that is, if the subject recognizes the situation as one of social exchange -- a "look for cheaters" procedure should produce the above pattern of responses. Non-social contract rules, either descriptive or prescriptive, should not show this particular pattern of variation, regardless of their familiarity. In general, they can be expected to elicit the same low levels of 'P & not-Q' and very low levels of 'not-P & Q' typically found in the literature for non-SC problems.

Previous results on the Wason selection task are consistent with a social contract interpretation (see Chapter 2 for a detailed review). Robust and replicable content effects are found only for rules that relate terms that are recognizable as benefits and costs in the format of a standard social contract. No thematic rule that is not a social contract has ever produced a content effect that is both robust and replicable. For thematic content areas that do not express social contracts, either no content effect is found (e.g., the food problem), or there are at least as many studies that do not find content effects as there are studies that do (transportation and school problems). Moreover, most of the content effects reported for non-SC rules are either weak (Gilhooly & Falconer, 1974; Pollard, 1981), clouded by procedural difficulties (Bracewell & Hidi, 1974; Van Duyne, 1974), or have some earmarks of a social
contract problem (Van Duyne, 1974). All told, for non-SC thematic problems, three experiments have produced a substantial content effect (transportation: Wason & Shapiro, 1971; Bracewell & Hidi, 1974; school: Van Duyne, 1974), two have produced a weak content effect (transportation: Gilhooly & Falconer, 1974; Pollard, 1981), and 14 have produced no content effect at all (transportation: Bracewell & Hidi, 1974; Manktelow & Evans, 1979; Yachanin & Tweney, 1982; Griggs & Cox, 1982. food: Manktelow & Evans, 1979 (4 experiments); Brown et al., 1982; Reich & Ruth, 1982; Yachanin & Tweney, 1982; school: Yachanin & Tweney, 1982. non-SC post office: Golding, 1980; Griggs & Cox, 1982). The few effects that were found did not replicate. In contrast, sixteen out of sixteen experiments with standard social contracts elicited substantial content effects. These include the Drinking Age Problem, the Post Office Problem, and the Sears Problem. In this extensive literature, STD-SC rules are the only thematic content rules to elicit strong, replicable content effects on the Wason selection task.

However, none of these studies tested switched social contract rules -- rules for which the correct social contract answer is 'not-P & Q.' Moreover, most of them contrasted familiar STD-SC rules with unfamiliar non-SC rules (descriptive non-SC rules: Johnson-Laird, Legrenzi & Legrenzi, 1972; Cox & Griggs, 1982; Griggs & Cox, 1982; Griggs & Cox, 1983. prescriptive non-SC rules: D'Andrade, 1981; Golding, 1981; Cox & Griggs, 1982). Hence, these studies do not allow one to choose directly between a social contract explanation and the explanation most prevalent in the literature, "availability."
The alternative hypothesis: Availability

"Availability" theory comes in a variety of forms with some important theoretical differences, but common to all is the notion that the subject's actual past experiences create associational links between terms mentioned in the selection task (see Chapter 3 for detailed explanations). The more exposures a subject has had to, for example, the co-occurrence of P and Q, the stronger that association will be, and the easier it will come to mind -- become "available" as a response. A subject is more likely to have actually experienced the co-occurrence of 'P & not-Q' for a familiar rule, therefore familiar rules are more likely to elicit logically falsifying responses than unfamiliar rules. However, if all the terms in a task are unfamiliar, the only associational link available will be that created between P and Q by the conditional rule itself, because no previous link will exist among any of the terms. Thus 'P and Q' will be the most common response for unfamiliar rules. Falsifying responses will be rare for all unfamiliar rules, whether they are social contracts or not.

Testing social contract theory against availability theory

At present, it is widely believed that some variant of availability theory accounts for all content effects on the Wason selection task. In contrast, social contract theory proposes that for content involving social exchange, a social contract algorithm is the primary regulator of responses. Thus, for social contract theory, the major determinant of responses is
whether a rule is a social contract (SC) or descriptive (D).*

For availability theory, the major determinant of responses is whether a rule is familiar (F) or unfamiliar (U). Because these two variables are orthogonal, one can create an array of four problem types: Unfamiliar-Social Contract (U-SC), Familiar-Social Contract (F-SC), Unfamiliar-Descriptive (U-D), Familiar-Descriptive (F-D):

<table>
<thead>
<tr>
<th>'Availability dimension'</th>
<th>familiar</th>
<th>unfamiliar</th>
</tr>
</thead>
<tbody>
<tr>
<td>descriptive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>'SC dimension'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>social contract</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F-D</td>
<td>U-D</td>
</tr>
<tr>
<td></td>
<td>F-SC</td>
<td>U-SC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(U-STD-SC, U-SWC-SC)</td>
</tr>
</tbody>
</table>

Moreover, there are two kinds of U-SC problems: unfamiliar standard social contracts (U-STD-SC) and unfamiliar switched social contracts (U-SWC-SC). All but the F-SC, which confounds familiarity with being a social contract, can be used to construct critical tests disentangling the following two hypotheses:

**AV:** Availability is the sole determinant of performance on Wason selection tasks of varying content. This is the null hypothesis from the standpoint of the existing literature.

* Actually, the non-SC rule can be either descriptive or prescriptive. All SC rules are prescriptive, but not all prescriptive rules are SC rules. Most of the non-SC thematic problems tested in the literature were descriptive.
SC: Humans have social contact algorithms that are the major
determinant of performance on Wason selection tasks whose
content involves social exchange.

It would be difficult to believe that availability has no effect
on familiar problems. The SC hypothesis is silent on this point.
Indeed, any effect availability might have in eliciting
falsifying responses to familiar descriptive problems can be used
as a metric for judging the size of a social contract effect. SC
algorithms can be said to be a major determinant of responses for
problems involving social exchange if there are more SC responses
(STD-SC: 'P & not-Q'; SWC-SC: 'not-P & Q') to unfamiliar social
contract problems than falsifying responses to familiar
descriptive problems.

In the experiments that follow, story context was used to
transform an unfamiliar (U) rule -- like, "If a man eats cassava
root, then he must have a tattoo on his face" -- into either a
U-SC or U-D problem. By embedding the same unfamiliar rule in
two different stories, one can contextually define that rule as
either a social contract (U-SC) or a descriptive rule (U-D). A
social contract story contextually defines one term of the
unfamiliar rule as a rationed benefit that must be earned and the
other term as a cost/requirement. A descriptive story does not
define the terms as costs and benefits, but it does contextually
link them to familiar concepts and tie them together by a
familiar relation. In these experiments, the U-D problems invoke
what should be one of the most familiar relations according to
standard associationism: spatio-temporal co-occurrence.

Switching the position that a U-SC's terms occupy in the
"If-then" structure of the rule transforms its theoretical status from standard (U-STD-SC) to switched (U-SWC-SC), or vice versa. Let's say a story portrays cassava root as a rationed benefit and having a tattoo as a cost requirement. Then:

"If a man eats cassava root then he has a tattoo on his face" (If a man takes the benefit, then he pays the cost)

has a standard SC format, whereas:

"If a man has a tattoo on his face then he eats cassava root" (If a man pays the cost, then he takes the benefit)

has a switched SC format. Switching the position of a U-D's terms does not change its theoretical status.

Experiments 1 through 4 are the most important from a theoretical point of view. They compare performance on unfamiliar social contract problems to performance on unfamiliar and familiar descriptive problems. These experiments permit six critical tests -- comparisons for which hypotheses AV and SC make radically different predictions. These tests address the following questions:

1. Does an unfamiliar standard social contract elicit the predicted SC response, 'P & not-Q'?

2. Are there more SC responses to an unfamiliar standard social contract than falsifying responses to a familiar descriptive problem?

3. Does an unfamiliar switched social contract elicit the predicted SC response, 'not-P & Q'?

4. Are there more SC responses to an unfamiliar switched social contract than falsifying responses to a familiar descriptive problem?

5. Is the correct SC response to a standard social contract ('P & not-Q') very rare for a switched social contract?

6. Is the correct SC response to a switched social contract ('not-P & Q') very rare for a standard social contract?
Experiment 5 tests whether an abstract problem with a social contract story context can elicit SC responses; Experiment 6 shows that the social contract effect is replicable with familiar content.

**Experiment 1**

The purpose of Experiment 1 was to see whether an unfamiliar STD-SC problem would elicit the predicted SC response, 'P & not-Q'. A high percentage of "falsifying" responses on a U-STD-SC is predicted only by social contract theory; availability theory predicts a low percentage of falsifying responses on the U-STD-SC because it is unfamiliar. Each subject was asked to solve four Wason selection tasks, which were presented in random order.

Theoretically, the problem types can be described as follows:

- **U-STD-SC**: Unfamiliar - Standard Social Contract
- **U-D**: Unfamiliar - Descriptive
- **AP**: Abstract Problem
- **F-D**: Familiar - Descriptive

The AP was a non-SC prescriptive rule; it was included because it is commonly used as a standard for assessing availability (Wason, 1983).

Table 6.1 shows the relative percentages of 'P & not-Q' and 'not-P & Q' responses expected in Experiment 1, assuming that responses are determined by either SC algorithms or availability, but not both.

**Subjects.**

Twenty-four undergraduates from Harvard University participated in Experiment 1; they were paid volunteers recruited by advertisement (13 females, 11 males; mean age: 19.4 years). 

205
Table 6.1 Predictions, Experiment 1: Social contract theory versus availability theory.

<table>
<thead>
<tr>
<th></th>
<th>Social Contract (STD-SC) v. Descriptive (D) problems.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>'P &amp; not-Q'</td>
</tr>
<tr>
<td></td>
<td>'not-P &amp; Q'</td>
</tr>
<tr>
<td></td>
<td>Social contract</td>
</tr>
<tr>
<td>U-STD-SC:</td>
<td>high</td>
</tr>
<tr>
<td>U-D:</td>
<td>low</td>
</tr>
<tr>
<td>AP:</td>
<td>low</td>
</tr>
<tr>
<td>F-D:</td>
<td>low</td>
</tr>
</tbody>
</table>

Table 6.1 Relative percentages of 'P & not-Q' and 'not-P & Q' responses expected for Exp 1, assuming that responses are solely determined by either SC algorithms or availability.

'P & not-Q' responses: Social Contract Predictions: Rationale for U-STD-SC described in text. Because SC algorithms will not be structuring responses to non-SC problems (U-D, AP, F-D), these should not elicit high levels of SC responses. They can be expected to elicit the same low levels of 'P & not-Q' responses (and very low levels of 'not-P & Q', see below) typically found in the literature for non-SC problems. SC is silent on whether availability will exert an independent effect on F-D problems. Availability Predictions: The F-D transportation problem could elicit a somewhat higher percentage than the unfamiliar problems (U-STD-SC, U-D, and AP) because it is likely that at least some subjects living in the Boston area have a pre-existing or dominant 'P & not-Q' association (see Chapter 3).

'not-P & Q' responses: This is a very rare response on Wason selection tasks. It involves the failure to choose the P card, which is almost universally chosen, and which even availability theorists concede is guided by at least a rudimentary understanding of logic (Evans & Lynch, 1973; Pollard, 1979). In addition, when chosen with Q, the substitution of not-P for P violates ordinary notions of contingency as expressed in English (why say "If P then Q" if you mean "If not-P then Q"?). No availability theorist has ever predicted this response. Both hypotheses predict a very low percentage for all problems other than a U-SWC-SC, for which, according to social contract theory alone, it is the predicted response.
Materials and Procedures.

Each subject received a sealed booklet with instructions on the first page, followed by four Wason selection tasks, one per page. Each selection task was embedded in a brief story. Each booklet contained a U-STD-SC, a U-D, an AP, and an F-D. The order of the four problems was randomized across subjects. Experiment 1 had a within subjects design.

All stories were phrased so as to activate a "detective set" (Van Duyne, 1974), and all asked subjects to look for violations of the rule. There were two versions ('A' and 'B') of unfamiliar problems; the rules used and the two cultures described therein were fictitious. A booklet either contained the 'A' version of the U-D problem and the 'B' version of the U-STD-SC, or vice versa. Figure 6.2 shows the 'A' versions of the unfamiliar problems; Figure 6.3 shows the 'B' versions. The 'A' version of the U-D and U-STD-SC varied only in surrounding story context; the rules used were identical. The same was true of the 'B' version problems. Figure 6.4 compares the unfamiliar problems used in Experiments 1 and 2.

I wanted to use any effect availability might have in eliciting falsifying responses to the F-D problem as a metric for judging the size of a social contract effect. For this reason I used a transportation problem as the F-D problem; the transportation problem had been the most successful non-SC problem in the literature (see Chapter 2). Various versions (using different terms) of the F-D and AP problems were randomized with respect to each other and the unfamiliar problems. Figure 6.5 shows examples of the F-D and AP problems used.
You are an anthropologist studying the Kaluame people, a Polynesian culture found only on Maku Island in the Pacific. Before leaving for Maku Island you read a report that says some Kaluame men have tattoos on their faces, and that they eat either cassava root or molo nuts, but not both. The author of the report, who did not speak the language, said the following relation seemed to hold:

"If a man eats cassava root, then he must have a tattoo on his face."

You decide to investigate your colleague's peculiar claim. When you arrive on Maku Island, you learn that cassava root is a starchy staple food found on the south end of the island. Molo nuts are very high in protein, and grow on molo trees, which are primarily found on the island's north shore.

You also learn that bachelors live primarily on the north shore, but when men marry, they usually move to the south end of the island. When a Kaluame man marries, he gets a tattoo on his face; only married men have tattoos on their faces. A facial tattoo means that a man is married, an unmarked face means that a man is a bachelor.

Cassava root is a powerful aphrodisiac — it makes the man who eats it irresistible to women. Moreover it is delicious and nutritious — and very scarce.

Unlike cassava root, molo nuts are very common, but they are poor eating — molo nuts taste bad, they are not very nutritious, and they have no other interesting "medicinal" properties.

Although everyone craves cassava root, eating it is a privilege that your people closely ration. You are very sensual people, even without the aphrodisiacal properties of cassava root, but you have very strict sexual mores. The elders strongly disapprove of sexual relations between unmarried people, and particularly distrust the motives and intentions of bachelors.

Therefore, the elders have made laws governing rationing privileges. The one you have been entrusted to enforce is as follows:

"If a man eats cassava root, then he must have a tattoo on his face."

Cassava root is so powerful an aphrodisiac, that many men are tempted to cheat on this law whenever the elders aren't looking. The cards below have information about four young Kaluame men sitting in a temporary camp at the center of the island. Each man is eating either cassava root or molo nuts which he has brought with him from home. Each card represents one man. One side of a card tells which food a man is eating, and the other side of the card tells whether or not the man has a facial tattoo.

Your job is to catch men whose sexual desires might tempt them to break the law — if any get past you, you and your family will be disgraced. Indicate only those card(s) you definitely need to turn over to see if any of these Kaluame men are breaking the law.

A. no tattoo
B. tattoo
C. eats cassava root
D. eats molo nuts

Unfamiliar Standard Social Contract (U-STD-SC)

PAGE

You are a Kaluame, a member of a Polynesian culture found only on Maku Island in the Pacific. The Kaluame have many strict laws which must be enforced, and the elders have entrusted you with enforcing them. To fail would disgrace you and your family.

Among the Kaluame, when a man marries, he gets a tattoo on his face; only married men have tattoos on their faces. A facial tattoo means that a man is married, an unmarked face means that a man is a bachelor.

Cassava root is a powerful aphrodisiac — it makes the man who eats it irresistible to women. Moreover it is delicious and nutritious — and very scarce.

Unlike cassava root, molo nuts are very common, but they are poor eating — molo nuts taste bad, they are not very nutritious, and they have no other interesting "medicinal" properties.

Although everyone craves cassava root, eating it is a privilege that your people closely ration. You are very sensual people, even without the aphrodisiacal properties of cassava root, but you have very strict sexual mores. The elders strongly disapprove of sexual relations between unmarried people, and particularly distrust the motives and intentions of bachelors.

Therefore, the elders have made laws governing rationing privileges. The one you have been entrusted to enforce is as follows:

"If a man eats cassava root, then he must have a tattoo on his face."

Cassava root is so powerful an aphrodisiac, that many men are tempted to cheat on this law whenever the elders aren't looking. A tray filled with cassava root and molo nuts has just been left for them. Each card below have information about four young Kaluame men sitting in a temporary camp at the center of the island. Each man is eating either cassava root or molo nuts which he has brought with him from home. Each card represents one man. One side of a card tells which food a man is eating, and the other side of the card tells whether or not the man has a facial tattoo.

Your job is to catch men whose sexual desires might tempt them to break the law — if any get past you, you and your family will be disgraced. Indicate only those card(s) you definitely need to turn over to see if any of these Kaluame men are breaking the law.

A. eats cassava root
B. no tattoo
C. eats molo nuts
D. tattoo

Unfamiliar Descriptive (U-D)
You are an anthropologist studying the Namka, a hunter-gatherer culture living in the deserts of southwest Africa. You are particularly interested in whether Namka boys obey the laws of their people.

Every full moon there is a special feast in which a duiker — a small antelope — is slaughtered and eaten. Duiker meat is quite scarce and delicious — a real treat. Eating duiker meat is a privilege that must be earned.

For boys, this privilege is governed by the following law:

"If you eat duiker meat, then you have found an ostrich eggshell."

Finding ostrich eggshells is a sophisticated and difficult task which takes a boy years to learn. Having found an ostrich eggshell on your own is therefore a sign that you have mastered the most difficult skills of hunting. For the Namka, it represents a boy's transition into manhood.

You wonder if Namka boys cheat on this law when nobody is looking. You decide to hide behind some bushes and watch. During the course of the feast of the full moon, you see four different boys approach the roasted duiker while no one else is looking.

The cards below have information about these four boys. Each card represents one boy. One side of a card tells whether a boy has ever found an ostrich eggshell and the other side of the card tells whether that boy took any of the roasted duiker meat.

The smell of the roasting duiker is truly tempting to the boys. You want to know if any of them cheated on the law. Indicate only those card(s) you definitely need to turn over to see if any of these boys have broken the law.

: some: : an ostrich: : eat any: : has never found:
: duiker meat: : eggshell: : duiker meat: :
: ..........: : ..........: : ..........: : ..........:

You are an anthropologist studying the Namka, a hunter-gatherer culture in the deserts of southwest Africa. Over and over again, you hear various Namka repeat the following saying:

"If you eat duiker meat, then you have found an ostrich eggshell."

Duikers are small antelopes found in the eastern part of the Namka's home range. Both duiker meat and ostrich eggshells are sought by the Namka. They eat the meat and use the eggshells as canteens because they are light and hold lots of water. Furthermore, duikers frequently feed on ostrich eggs.

As an anthropologist, you don't know if this saying is metaphorical, referring, for example, to clan territories or ritual practices, or if the saying reflects a real relationship the Namka use to guide their foraging behavior. Does it mean that if you find the first you find the second? This is what you are trying to find out.

Is it fact or folklore? Do the Namka mean eggshells and duiker meat, or are these things merely symbols for something else entirely? Unfortunately, you don't know their language well enough to ask them. So you decide to investigate whether the rule stated in this saying has any factual basis.

Many species of birds populate the area, and in your wanderings you have come across several caches of eggs of various sorts. The cards below have information about four different locations with egg caches. Each card represents one location, and each location has the tracks of one mammal associated with it. One side of a card tells what kind of eggshell you found at a location, and the other side of the card tells which mammal's tracks you found there.

Perhaps the Namka's saying has no factual basis. Indicate only those card(s) you definitely need to turn over to see if your finds at any of these locations violates the rule expressed in the Namka's saying.

: ..........: : ..........: : ..........: : ..........: : ..........: : : : : 

Unfamiliar Standard Social Contract (U-STD-SC) PAGE

Unfamiliar Descriptive (U-D) PAGE
Figure 6.4 Experiments 1 and 2: Comparison of Unfamiliar Problems

'A' VERSIONS OF UNFAMILIAR PROBLEMS

Rule A:
(Exp 1): "If a man eats cassava root, then he must have a tattoo on his face."
(Exp 2): "If a man has a tattoo on his face, then he eats cassava root."

Cards: Logical Category:
(Exp 1 & 2) (Exp 1) (Exp 2)
"eats cassava root" P Q not-P
"eats molo nuts" not-P P not-P
"tattoo" Q P not-Q
"no tattoo" not-Q not-P

A:Common Story Elements: All 'A' problems involve a fictional Polynesian people called the "Kaluame". A facial tattoo means a man is married. no tattoo means he is a bachelor.

A:O-D Story Summary: You (the subject) are an anthropologist studying the Kaluame. Cassava root and molo nuts are foods found on two different parts of their island; men eat one or the other, but not both. Married men and bachelors tend to live on two different parts of the island, where cassava root and molo trees grow, respectively. You read a report asserting that Rule A seems to hold. If so, perhaps men are simply eating foods which are most available to them. But Rule A may not be true. You want to investigate for yourself by seeing if any of four men (each card represents one man) are breaking Rule A.

A:O-STD, U-SWC Story Summary: You are a Kaluame who has been entrusted to enforce your people's laws. Cassava root is a powerful but scarce aphrodisiac and a delicious food source; eating it is a rationed privilege, governed by Rule A. Molo nuts are a common, undesirable food source with no interesting "medicinal" properties. The Kaluame have very strict sexual mores, and disapprove of sexual relations between unmarried people. Your job is to catch any of four men (each card represents one man) who might have cheated on Rule A.

**must** is left out of the first clause because it violates common English usage36.

'B' VERSIONS OF UNFAMILIAR PROBLEMS

Rule B:
(Exp 1): "If you eat duiker meat, then you have found an ostrich eggshell."
(Exp 2): "If you have found an ostrich eggshell, then you eat duiker meat."

Cards: Logical Category:
(Exp 1 & 2) (Exp 1) (Exp 2)
"duiker" "eats some duiker meat" P Q not-P
"weed" "does not eat any duiker meat" not-P P P
"ostrich eggshell" "has found an ostrich eggshell" Q P Q
"quail eggshell" "has never found an ostrich eggshell" not-Q not-P

B:Common Story Elements: In all 'B' problems, you (the subject) are an anthropologist studying a (fictional) southwest African hunter-gatherer group called the "Namka". Duikers are antelopes whose meat the Namka eat.

B:O-D Story Summary: You want to know if Rule B is fact or folklore. The whereabouts of both duikers (hunted for their meat) and ostrich eggshells (used as canteens) are of interest to the Namka. Duikers frequently feed on ostrich eggs, so you guess Rule B reflects a real relationship that the Namka use to guide their foraging behavior. You don't know the Namka's language well enough to ask, so you decide to see if Rule B has any factual basis. To do this, you can investigate any of four locations (each card represents one location). Associated with each location is the egg cache of one bird species and the tracks of one mammal.

B:O-STD, U-SWC Story Summary: You are interested in whether Namka boys obey the laws of their people. Finding ostrich eggshells is a difficult, sophisticated task which represents a boy's transition into manhood. Duiker meat is a scarce and prized food; eating duiker meat at feasts is a privilege that must be earned, and is regulated by Rule B. You want to know if any of four boys at the feast cheated on this law when no one but you was looking (each card represents one boy).
Familiar Descriptive (F-D: Transportation Problem)

Part of your new job for the City of Cambridge is to study the demographics of transportation. You read a previously done report on the habits of Cambridge residents which says:

"If a person goes into Boston, then he takes the subway."

The cards below have information about four Cambridge residents. Each card represents one person. One side of a card tells where a person went and the other side of the card tells how that person got there.

Indicate only those card(s) you definitely need to turn over to see if any of these people violate this rule.

A. : : : 
   : subway : 
   : : :

B. : : : 
   : Arlington : 
   : : :

C. : : : 
   : cab : 
   : : :

D. : : : 
   : Boston : 
   : : :

Abstract Problem (AP)

Part of your new clerical job at the local high school is to make sure that student documents have been processed correctly. Your job is to make sure the documents conform to the following alphanumeric rule:

"If a person has a 'D' rating, then his documents must be marked code '3'."

You suspect the secretary you replaced did not categorize the students' documents correctly. The cards below have information about the documents of four people who are enrolled at this high school. Each card represents one person. One side of a card tells a person's letter rating and the other side of the card tells that person's number code.

Indicate only those card(s) you definitely need to turn over to see if the documents of any of these people violate this rule.

A. : F :
   : : ::

B. : D :
   : : ::

C. : 3 :
   : : ::

D. : 7 :
   : : ::
The instructions (reproduced in Figure 6.6) asked subjects to do the four tasks in order, without rereading any previous story or reviewing or changing any previous answers. The instructions were read aloud to subjects; in addition, subjects were given as much time as they wanted to read the instructions over to themselves before breaking the seal and beginning the experiment. Although most subjects completed the experiment in about 10 minutes, they were told they could take as much time as they wanted.

Results.

The percent of subjects choosing 'P & not-Q' for each problem closely matches the social contract predictions shown in Table 6.1. Hypotheses AV and SC both predict that the percent of

<table>
<thead>
<tr>
<th></th>
<th>P not-Q</th>
<th>not-P</th>
</tr>
</thead>
<tbody>
<tr>
<td>U-STD-SC:</td>
<td>75</td>
<td>0</td>
</tr>
<tr>
<td>U-D:</td>
<td>21</td>
<td>0</td>
</tr>
<tr>
<td>AP:</td>
<td>25</td>
<td>0</td>
</tr>
<tr>
<td>F-D:</td>
<td>46</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 6.2 shows the percent of subjects who chose either 'P & not-Q' or 'not-P & Q'. Residual responses: These two categories do not, of course, exhaust all possible combinations of card choices -- there are sixteen in all. Of the 6 responses to the U-STD-SC that were not full SC answers, 5 were half correct, "sins of omission": 2 'P' responses (omitted not-Q) and 3 'not-Q' responses (omitted P). The 'not-Q' response is most interesting because not-Q is the card people routinely forget to choose on Wason selection tasks, and it is rarely chosen alone. If one counts 'not-Q' as also correct for both the U-STD-SC and the F-D, the magnitude of the difference between them increases (87.5% v. 54%).
Figure 6.6 Instructions

PLEASE DON'T LOOK THROUGH THIS BOOKLET YET
(But go ahead and read these instructions)

This is an experiment, completely optional, which will take about 15 minutes.

I am interested in how people think about different social situations. You will be reading some brief stories and then evaluating sentences like:

"If a fruit is an apple, then it must be red."

You will be evaluating the sentence with respect to information on four cards -- actually, pictures of four cards, like the pictures below.

In this example, each card would represent a fruit. Each card would have the name of a fruit on one side and that fruit's color on the other side, for example:

A. .....................
   : apple :
   : .....................

B. .....................
   : blue :
   : .....................

C. .....................
   : red :
   : .....................

D. .....................
   : pear :
   : .....................

Of course, since they are only pictures, you will only be able to see one side of each card. The cards need not correspond to the way the world really is; for example, the "pear" card could say "yellow" or "red" or "purple" on the back, the "red" card could say "apple" or "banana" or "blueberry" on the back, and so on. For each story, you will be asked to indicate only those card(s) you definitely need to turn over to see if any of them violate the relation stated in the sentence. Circle the letter(s) (A,B,C, or D) which is next to the card(s) you want to turn over.

There are four different stories on four different pages. Don't look ahead at any stories: Read the first story and answer the question, then read the second story and answer its question, and so on, in the order the stories appear in this booklet. Please read a story in its entirety before you answer the question. Once you have finished answering the question associated with a story and gone on to the next story, do not go back and reread any previous stories or review or change any previous answers.

Never try to answer a question without first having read the entire story carefully. There are no "trick" questions. If a sentence or story seems ambiguous, use your common sense -- I am interested in what you would really do if faced with these situations in real life.

In sum: Pretend you really do have to investigate the situation described in a story. Then, for each card, ask yourself: "Would I need to see the information on the other side of this card in order to make a judgment?" If the answer is "yes", then circle the letter, A,B,C, or D, corresponding to that card.

Don't start until I have finished reading the instructions aloud. Take your time and have fun!

What year are you? Fresh __, Soph __, Jr __, Sr __, Other ___ (please specify)

Are you Female ___ or Male ___ ? Age ____ ?
subjects choosing 'not-P & Q' will be very low for these four problems; indeed, no one made this response in Experiment 1 (see Table 6.2).

Critical Tests

Predictions for critical tests 1 and 2 are taken from Table 6.1; for critical tests 3 and 4, they are taken from Table 6.3. Predictions for critical tests 5 and 6 are derived from both tables.

Critical Test 1: Does an unfamiliar standard social contract elicit the predicted SC response, 'P & not-Q'?

To answer this question, responses to the two unfamiliar problems must be compared; these problems use the same rule, but the story surrounding one rule makes it a social contract whereas the story surrounding the other makes it a descriptive rule.

Percentage 'P & not-Q' responses:

<table>
<thead>
<tr>
<th>Social Contract Prediction:</th>
<th>Availability Prediction:</th>
</tr>
</thead>
<tbody>
<tr>
<td>U-STD-SC v. U-D.</td>
<td>U-STD-SC &gt; U-D</td>
</tr>
<tr>
<td>high</td>
<td>low</td>
</tr>
<tr>
<td>low</td>
<td>low</td>
</tr>
</tbody>
</table>

75% > 21%

AV does not predict, and cannot account for, a wide discrepancy in falsifying ('P & not-Q') responses between these two unfamiliar problems. Yet a highly significant 54 point discrepancy occurred, just as SC predicts (75% v. 21%: \( F = 27.18, p << .001, r = .74^* \)). U-STD-SC also produced a

\* r is an effect size, which varies between zero and one (Rosenthal & Rosnow, 1984).
significant "content effect" when measured against the AP
(75% v. 25%: $F_{1,23} = 13.80, p < .005, r = .61$). The U-D and AP
both elicited the same low levels of falsifying responses (21% v.
25%: $F_{1,23} = 0.138, n.s.$).

Critical Test 2: Are there more SC responses to an unfamiliar
standard social contract than falsifying responses to a familiar descriptive problem?

All problems asked subjects to detect potential violations
of the rule. Therefore, any effect availability has in eliciting falsifying
responses to familiar non-SC problems can be used as a metric for judging the size of the social contract effect.

Percentage 'P & not-Q' responses:

<table>
<thead>
<tr>
<th>Social Contract Prediction</th>
<th>Availability Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>U-STD-SC v. F-D</td>
<td>U-STD-SC &gt; F-D</td>
</tr>
<tr>
<td>high</td>
<td>low*</td>
</tr>
<tr>
<td></td>
<td>high</td>
</tr>
<tr>
<td></td>
<td>low</td>
</tr>
<tr>
<td></td>
<td>mid-low</td>
</tr>
</tbody>
</table>

75% > 46%

As social contract theory predicts, an unfamiliar social contract (U-STD-SC) with which no subject could have had any actual experience elicited significantly more "falsifying" responses (SC responses) than a familiar relation (F-D) with which subjects were likely to have had experience (75% v. 46%: $F_{1,23} = 5.24, p < .05, r = .44$).

direction. Counting rare residuals (see legend, Table 6.2) for both problems magnifies U-STD-SC's advantage (87.5% v. 54%: $F_{1,23} = 5.41, p < .05, r = .44$).

* On Table 6.1, SC predictions regarding magnitudes for F-D problems assume no effect of availability; actually, SC is silent on whether or not availability exercises an independent effect on familiar problems.
Experiment 2

Experiment 2 was identical to Experiment 1, except the unfamiliar rules were switched rather than standard. The four problems fell into the following theoretical categories:

U-SWC-SC: Unfamiliar - Switched Social Contract
U-D: Unfamiliar - Descriptive
AP: Abstract Problem
F-D: Familiar - Descriptive

Table 6.3 shows the relative percentages of 'P & not-Q' and 'not-P & Q' responses expected in Experiment 2, assuming that responses are determined by either SC algorithms or availability, but not both.

<table>
<thead>
<tr>
<th>Social Contract</th>
<th>'P &amp; not-Q'</th>
<th></th>
<th>Social Contract</th>
<th>'not-P &amp; Q'</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Social Contract</td>
<td>Availability</td>
<td></td>
<td>Social Contract</td>
</tr>
<tr>
<td>U-SWC-SC:</td>
<td>very low</td>
<td>low</td>
<td></td>
<td>high</td>
</tr>
<tr>
<td>U-D:</td>
<td>low</td>
<td>low</td>
<td></td>
<td>very low</td>
</tr>
<tr>
<td>AP:</td>
<td>low</td>
<td>low</td>
<td></td>
<td>very low</td>
</tr>
<tr>
<td>F-D:</td>
<td>low</td>
<td>middling to low</td>
<td></td>
<td>very low</td>
</tr>
</tbody>
</table>

Table 6.3 Relative percentages of 'P & not-Q' and 'not-P & Q' responses expected for Experiment 2, assuming that responses are determined by either SC algorithms or availability, but not both.

'P & not-Q' responses: Social Contract Predictions: These are the cards SC algorithms should ignore on switched social contract problems; hence the percentage of "falsifying" responses should be very low on the U-SWC-SC. The rationale for the other predictions is the same as that presented in Table 6.1 for Experiment 1.

'not-P & Q' responses: Both hypotheses predict a very low percentage for all problems other than the U-SWC-SC, for which, according to social contract theory, it is the predicted response. See rationale in Table 6.1.
Subjects.

Twenty-four undergraduates from Harvard University participated in Experiment 2; they were paid volunteers recruited by advertisement (11 females, 13 males; mean age: 19.0 years).

Materials and Procedure.

The procedure was identical to that for Experiment 1. The materials were also identical with one exception: for unfamiliar rules, the propositions were switched. Thus, the 'A' version rule for the U-D and U-SWC-SC problems was: "If a man has a tattoo on his face then he eats cassava root,"* and the 'B' version rule was "If you have found an ostrich eggshell, then you eat duiker meat."

Results.

Table 6.4 shows the percent of subjects choosing 'P & not-Q' and 'not-P & Q' for each problem: these figures closely match the social contract predictions (see Table 6.3).

Critical Test 3: Does an unfamiliar switched social contract elicit the predicted SC response, 'not-P & Q'?

Adaptive inference diverges sharply from logical inference for SWC-SC problems. 'Not-P & Q' is completely at variance with formal logic and a very rare response on Wason selection tasks -- AV predicts it will be rare for all problems. However, 'not-P & Q' is the correct SC response to a switched social contract, no matter how unfamiliar. Therefore, Critical Test 3 requires that the U-SWC-SC be compared to all the other rules; 'not-P & Q' is the predicted response only for the U-SWC-SC.

* "must" was left out of the "If" clause because it violates common English usage.
Table 6.4 shows the percent of subjects who chose either 'P & not-Q' or 'not-P & Q' in Experiment 2. Residual responses: Of the 8 responses to the U-SWC-SC that were not full SC answers, 6 were "sins of omission": 1 'not-P' (omitted Q) and 5 'Q' (omitted not-P). Both are rare answers on Wason selection tasks, and the frequency of 'Q' responses is much higher than its expected value based on the other problems in Exps 1 and 2 (z=4.00, p<.00003). If one counts rare, half-correct residuals as also correct ('Q' for U-SWC-SC, 'not-Q' for F-D), the magnitude of the difference between U-SWC-SC and F-D increases (87.5% v. 54%), exactly matching the U-STD-SC v. F-D figures in Exp 1 for the equivalent categorization scheme.

The large and significant 63-67 point difference between the U-SWC-SC and all other problems (67% v. 4%, 0%, 0%, L = +3,-1,-1,-1: $F = 116.26$, $p << .001$, $r = .79$) is predicted only by SC.

'Not-P & Q' was chosen only once on any non-SWC-SC problem in Experiment 2, and was not chosen by anyone in Experiment 1.
Critical Test 4: Are there more SC responses to an unfamiliar switched social contract than falsifying responses to a familiar descriptive problem?

As in Critical Test 2, one can use the percentage of falsifying responses to the F-D problem as a metric for judging the size of the social contract effect. This requires that the proportion of SC responses ('not-P & Q') to the U-SWC-SC be compared to the proportion of falsifying responses to the F-D.

Percentage 'not-P & Q' responses to U-SWC-SC,
Percentage 'P & not-Q' responses to F-D:

<table>
<thead>
<tr>
<th>Social Contract Prediction:</th>
<th>Availability Prediction:</th>
</tr>
</thead>
<tbody>
<tr>
<td>U-SWC-SC v. F-D.</td>
<td></td>
</tr>
<tr>
<td>high</td>
<td>low</td>
</tr>
<tr>
<td>U-SWC-SC &lt; F-D</td>
<td>very low</td>
</tr>
<tr>
<td></td>
<td>mid-low</td>
</tr>
</tbody>
</table>

As SC predicts, SC responses to the U-SWC-SC outstripped falsifying responses to the F-D (67% v. 50%: F = 2.09, n.s.).

The difference is not significant; however AV predicts an inequality in the opposite direction. When rare residuals are counted for both problems (U-SWC-SC: 'Q'; F-D: 'not-Q'), the difference is magnified to the exact proportions found for the parallel U-STD-SC v. F-D comparison in Experiment 1, and is significant (87.5% v. 54%: F = 8.36, p < .01, r = .52). This supports the contention that SC algorithms are a major determinant of responses to problems involving social exchange, even when those problems are unfamiliar.
Experiment 1 versus Experiment 2

Critical Test 5: Is the correct SC response to a standard social contract ('P & not-Q') very rare for a switched social contract?

Because the U-STD-SC and U-SWC-SC are both unfamiliar problems, AV predicts they should both elicit low levels of 'P & not-Q' responses. The social contract prediction could not be more different. For a STD-SC, P represents the "benefit accepted" card and not-Q represents the "cost NOT paid" card, the cards that a "look for cheaters" procedure should choose.

However, for a SWC-SC, P represents the "cost paid" card and not-Q represents the "benefit NOT accepted" card, the cards a "look for cheaters" procedure should ignore because they represent people who could not possibly have cheated (see Figure 6.1). What logical category these cards fall into is simply irrelevant from a social contract perspective. Cards should be chosen on the basis of their cost/benefit category, not their logical category.

Percentage 'P & not-Q' responses:

<table>
<thead>
<tr>
<th>Social Contract Prediction</th>
<th>Availability Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>very low</td>
</tr>
<tr>
<td>low</td>
<td>low</td>
</tr>
</tbody>
</table>

75% >> 4%

The large and significant 71 point discrepancy in 'P & not-Q' responses between U-STD-SC and U-SWC-SC problems is predicted only by SC (75% v. 4%: Z = 5.02, p < .0000005, phi = .72*).

Furthermore, the SC prediction that the dominant, SC response to
the U-STD-SC will be very rare on the U-SWC-SC was borne out.
Only one subject gave the STD-SC answer, 'P & not-Q', in response
to the U-SWC-SC -- and this was one of only two subjects in Exp 2
to give falsifying answers to all of the three other problems.

Critical Test 6: Is the correct SC response to a switched social
contract ('not-P & Q') very rare for a standard
social contract?

Critical Test 6 is simply the flip side of Critical Test 5.
The predicted SC response to a SWC-SC is 'not-P & Q': not-P
represents the "cost NOT paid" card and Q represents the "benefit
accepted" card (see Figure 6.1). But for a STD-SC, not-P
represents the "benefit NOT accepted" card and Q represents the
"cost paid" card -- the cards a "look for cheaters" procedure
should ignore, regardless of their logical category. Hence, SC
predicts that the correct SC answer to a SWC-SC, 'not-P & Q',
will be very rare for a STD-SC. In contrast, AV predicts that
the percentage of subjects choosing 'not-P & Q' on the U-STD-SC
and the U-SWC-SC will be about equal, and very low (see Table 6.1).

Percentage 'not-P & Q' responses:

<table>
<thead>
<tr>
<th>Social Contract Prediction:</th>
<th>Availability Prediction:</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>very low</td>
</tr>
<tr>
<td>very low</td>
<td>very low</td>
</tr>
<tr>
<td>very low</td>
<td>very low</td>
</tr>
<tr>
<td>67% &gt;&gt; 0%</td>
<td></td>
</tr>
</tbody>
</table>

The large and significant 67 point discrepancy in 'not-P & Q'
responses between U-SWC-SC and U-STD-SC problems is predicted
only by SC (67% v. 0%: Z = 4.90, p < .0000005, phi = .71).

* phi is an effect size, which varies between zero and one
(Rosenthal & Rosnow, 1984).

216
Furthermore, the SC prediction that the dominant, SC response to the U-SWC-SC will be very rare on the U-STD-SC was borne out: no one gave the SWC-SC answer, 'not-P & Q', in response to the U-STD-SC.

Summary, Critical Tests for Experiments 1 and 2.

Because the AV and SC hypotheses make very different predictions regarding six comparisons between problems, critical tests between these two hypotheses can be constructed from the predictions of Table 6.1. The results for each of the six tests verify the SC prediction and falsify the AV prediction (see Figure 6.7 for summary of critical tests and results).

Social Contract Tests

Critical tests only address the question: Is the data better explained by social contract theory or availability theory? However, there are other questions one can ask of this data that are specific to social contract theory. Because these questions involve a comparison of results from Experiments 1 and 2, for convenience, the data from Tables 6.2 and 6.4 are combined into Table 6.5 below.

Are the logically distinct SC answers to standard and switched SC problems produced by the same algorithms?

The correct SC answers for standard and switched social contracts are very different from a logical point of view: 'P & not-Q' for U-STD-SC v. 'not-P & Q' for U-SWC-SC. However, the proportions of SC answers to the U-STD-SC (75%) and U-SWC-SC (67%) are not significantly different ($Z = 0.63$), just as one
Figure 6.7

CRITICAL TESTS: SOCIAL CONTRACT THEORY VERSUS AVAILABILITY THEORY

<table>
<thead>
<tr>
<th>CRITICAL TEST</th>
<th>SOCIAL CONTRACT PREDICTION (SC):</th>
<th>AVAILABILITY PREDICTION (AP):</th>
<th>WHICH HYPOTHESIS DO THE DATA SUPPORT?</th>
<th>CONCLUSION</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CRITICAL TEST 1:</strong> Does an unfamiliar STANDARD social contract elicit the predicted SC response, &quot;P &amp; not-Q&quot;?</td>
<td>U-STD-SC &gt; U-D high</td>
<td>U-STD-SC = U-D low</td>
<td>SC: 75% &gt; 21%, ( F_{1,27}=27.18, p&lt;.001, r=.74^* )</td>
<td>SC verified, ( AV ) falsified.</td>
</tr>
<tr>
<td><strong>CRITICAL TEST 2:</strong> Are there more SC responses to a standard UNFAMILIAR social contract problem than falsifying responses to a FAMILIAR descriptive problem?</td>
<td>U-STD-SC &gt; F-D high</td>
<td>U-STD-SC = F-D low</td>
<td>SC: 75% &gt; 46%, ( F_{1,25}=5.24, p&lt;.05, r=.43 )</td>
<td>SC verified, ( AV ) falsified.</td>
</tr>
<tr>
<td>percentage 'P &amp; not-Q' responses:</td>
<td>U-STD-SC &gt; F-D high</td>
<td>U-STD-SC = F-D low</td>
<td>SC: 87.5% &gt; 54%, ( F_{1,31}=5.41, p&lt;.05, r=.44 )</td>
<td>SC verified, ( AV ) falsified.</td>
</tr>
<tr>
<td>2a. U-STD-SC v. F-D:</td>
<td>same as 2a</td>
<td>same as 2a</td>
<td>SC: 87.5% &gt; 54%, ( F_{1,31}=5.41, p&lt;.05, r=.44 )</td>
<td>SC verified, ( AV ) falsified.</td>
</tr>
<tr>
<td>2b. U-STD-SC v. F-D:</td>
<td>same as 2a</td>
<td>same as 2a</td>
<td>SC: 87.5% &gt; 54%, ( F_{1,31}=5.41, p&lt;.05, r=.44 )</td>
<td>SC verified, ( AV ) falsified.</td>
</tr>
<tr>
<td><strong>CRITICAL TEST 3:</strong> Does an unfamiliar SWITCHED social contract elicit the predicted SC response, &quot;not-P &amp; Q&quot;?</td>
<td>U-SWC-SC &gt; U-D,AP,F-D high</td>
<td>U-SWC-SC = U-D,AP,F-D low</td>
<td>SC: 67% &gt; 4%, 0%, 0%</td>
<td>SC verified, ( AV ) falsified.</td>
</tr>
<tr>
<td>percentage 'not-P &amp; Q' responses:</td>
<td>U-SWC-SC &gt; U-D,AP,F-D high</td>
<td>U-SWC-SC = U-D,AP,F-D low</td>
<td>SC: 67% &gt; 50%, correct direction, but n.s. ( (F_{1,2}=.09) )</td>
<td>SC verified, ( AV ) falsified.</td>
</tr>
<tr>
<td>4a. U-SWC-SC v. F-D:</td>
<td>same as 4a</td>
<td>same as 4a</td>
<td>SC: 87.5% &gt; 54%, ( F_{1,31}=8.36, p&lt;.01, r=.52 )</td>
<td>SC verified, ( AV ) falsified.</td>
</tr>
<tr>
<td>plus rare residuals (U-SWC-SC: 'Q': F-D: 'not-Q'):</td>
<td>same as 4a</td>
<td>same as 4a</td>
<td>SC: 87.5% &gt; 54%, ( F_{1,31}=8.36, p&lt;.01, r=.52 )</td>
<td>SC verified, ( AV ) falsified.</td>
</tr>
<tr>
<td>4b. U-SWC-SC v. F-D:</td>
<td>same as 4a</td>
<td>same as 4a</td>
<td>SC: 87.5% &gt; 54%, ( F_{1,31}=8.36, p&lt;.01, r=.52 )</td>
<td>SC verified, ( AV ) falsified.</td>
</tr>
<tr>
<td><strong>CRITICAL TESTS 5 and 6:</strong> Is the correct SC response to a STANDARD social contract very rare for a SWITCHED social contract, and vice versa?</td>
<td>U-STD-SC &gt;&gt; U-SWC-SC high</td>
<td>U-STD-SC = U-SWC-SC low</td>
<td>SC: 75% &gt;&gt; 4%, ( Z=5.02, p&lt;.0000005, \text{phi}=.72^* )</td>
<td>SC verified, ( AV ) falsified.</td>
</tr>
<tr>
<td>6. U-SWC-SC v. U-STD-SC:</td>
<td>same as 4a</td>
<td>same as 4a</td>
<td>SC: 67% &gt;&gt; 0%, ( Z=4.90, p&lt;.0000005, \text{phi}=.71 )</td>
<td>SC verified, ( AV ) falsified.</td>
</tr>
</tbody>
</table>

AVAILABILITY ASSESSED: EXPERIMENTS 1 AND 2

| **AVAILABILITY TEST:** Does availability have any effect at all on familiar problems? (standard test in literature: F-D > AP) | | |
| percent. 'P & not-Q' responses: | F-D > AP, U-D mid-low low | Exp 1: 46% > 25%, 21% | Availability has a marginal effect on familiar problems |
| Exp 1: F-D v. AP: NO EFFECT \( (F_{1,25}=4.02, \text{n.s}) \) | | |
| Exp 2: F-D v. AP: EFFECT \( (F_{1,25}=5.31, p<.05) \) | | |

Exp 2: F-D v. AP: EFFECT \( (F_{1,25}=13.80, p<.005) \) | | |
Exp 2: F-D v. AP: EFFECT \( (F_{1,25}=13.80, p<.005) \) | | |
Table 6.5 Experiments 1 and 2: Percent of subjects choosing 'P & not-Q' or 'not-P & Q' for each problem

<table>
<thead>
<tr>
<th></th>
<th>Experiment 1 (n=24)</th>
<th>Experiment 2 (n=24)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P not-Q</td>
<td>Q not-P</td>
</tr>
<tr>
<td>U-STD-SC:</td>
<td>75*</td>
<td>0</td>
</tr>
<tr>
<td>U-D:</td>
<td>21</td>
<td>0</td>
</tr>
<tr>
<td>AP:</td>
<td>25</td>
<td>0</td>
</tr>
<tr>
<td>F-D:</td>
<td>46</td>
<td>0</td>
</tr>
</tbody>
</table>

*predicted SC response to social contract problems.

would expect if the same SC algorithms were producing these two, logically distinct, responses. When rare residuals (see legend, Tables 6.2 and 6.4) are added in, the percentage of SC answers on these two problems is identical -- 87.5%. Using percent falsifying answers to the U-D (same rule as U-SC) as a baseline for comparison, the relative advantage SC status gave in producing SC answers is almost identical for both SC problems: 54 points between U-STD-SC and its U-D, 55 points between U-SWC-SC and its U-D.

Table 6.6 shows the frequencies with which individual cards were selected in Experiments 1 and 2.

Table 6.6 Experiments 1 and 2: Selection frequencies for individual cards, sorted by logical category and social contract category

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>21</td>
<td>19</td>
<td>23</td>
<td>23</td>
<td>22</td>
<td>Benefit Accepted</td>
<td>20</td>
<td>21</td>
</tr>
<tr>
<td>not-P</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>4</td>
<td>Benefit NOT Accepted</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Q</td>
<td>9</td>
<td>12</td>
<td>8</td>
<td>14</td>
<td>1</td>
<td>Cost Paid</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>not-Q</td>
<td>6</td>
<td>7</td>
<td>11</td>
<td>11</td>
<td>13</td>
<td>Cost NOT Paid</td>
<td>22</td>
<td>17</td>
</tr>
</tbody>
</table>

When cards are sorted according to their logical category, all
problems replicate nicely over Experiments 1 and 2, except the social contract problems. When sorted according to logical category, selection frequencies for the U-STD-SC and U-SWC-SC are radically at variance with one another. When sorted according to social contract category, however, their profiles are almost identical. This indicates that for unfamiliar social contract problems, a social contract categorization scheme captures dimensions that are psychologically real for subjects, whereas a logical categorization scheme does not.

How well do SC algorithms operate in novel, versus familiar, social exchanges?

If SC algorithms are, in part, frame-builders, as proposed, one would expect them to operate in novel social exchanges, as well as in familiar ones. In fact, the 75% "falsification" rate on the U-STD-SC, an unfamiliar social contract, is equivalent to that usually found for the Drinking Age Problem, a highly familiar standard social contact (Cox & Griggs, 1982; Griggs & Cox, 1982; Griggs & Cox, 1983; see Chapter 2). This cannot be accounted for by differences in subject populations: 78% of a similar group of 23 Harvard undergraduates "falsified" on the DAP (see Experiment 6-A below). Thus, the percent of Harvard undergraduates choosing 'P & not-Q' on a STD-SC was the same, regardless of whether the social contract was very familiar or completely unfamiliar (78% v. 75%; Z = 0.26, n.s.).

The same was true of the unfamiliar switched social contract. The percentage of subjects choosing 'not-P & Q' on the U-SWC-SC did not differ significantly from the percentage choosing 'P & not-Q' for the familiar DAP (67% v. 78%; Z = 0.88,
n.s.). The hypothesis that unfamiliar SC problems generate fewer SC answers than familiar SC problems is not supported even if one uses all three problems (F-STD-SC, U-STD-SC, U-SWC-SC) in one test (78% v. 75%, 67%, L = +2,-1,-1, $F = 0.43$, n.s.). The data of Experiments 1 and 2 suggest that SC algorithms work just as well in novel social exchanges as in familiar ones -- just as frame-builders should.

Availability Assessed

Does availability have any effect at all on familiar problems?

Although AV cannot explain the precisely patterned differences in performance among unfamiliar social contracts and all other problems, availability does appear to have had a minor, somewhat erratic effect on familiar descriptive problems (see Figure 6.7).

In the literature, a standard test of the efficacy of availability is to compare an F-D to an AP. The standard "now you see it, now you don't" result of such experiments (see Chapter 2) is mirrored in Experiments 1 and 2. The difference in percentage of falsifying responses between the F-D and AP is significant in Experiment 2, but not in Experiment 1. However, the F-D does fare significantly better than the U-D in both experiments. Furthermore, the set of contrasts for Experiment 1 which assumes that the F-D outstrips both AP and U-D problems (L=+2,-1,-1) is significant ($F = 5.97$, p < .025).

These results indicate that availability can give familiar non-SC problems an advantage of 21 to 38 points over unfamiliar
non-SC problems in producing falsifying responses (average advantage = 30.5). However, using percent falsifying responses to U-Ds and APs as a baseline for comparison, the availability advantage appears to be much less important than the social contract effect. The average advantage in producing SC responses that social contract status gives to an unfamiliar problem is about 1.8 times the size of the availability advantage. And, as Critical Tests 2 and 4 of Figure 6.7 show, more SC responses were elicited by unfamiliar social contract problems than falsifying responses by familiar non-SC problems that had an availability advantage.

Summary, Experiments 1 and 2.

Unfamiliar though they were, social contract problems reliably elicited social contract answers, even when these were radically at variance with formal logic. Furthermore, non-SC problems (U-D, AP, F-D) did not show this distinctive pattern of variation. Availability alone can neither predict nor explain the results of these experiments. In addition to the social contract effect, there also appears to have been a marginal effect of availability on F-D problems.

Experiment 3

In Experiments 1 and 2, the unfamiliar social contracts used were expressed as laws of one's social group. I did this because the rules used in the literature on the Wason selection task were invariably expressed as laws. Moreover, using a social contract law let me use the exact same rule in U-SC and U-D problems.
However, social contract algorithms should work just as well with conditionals that express a private social exchange between just two individuals, rules like:

"If you do X for me, then I'll do Y for you,"
or, equivalently,

"If I do Y for you, then you do X for me."

To test this, I conducted two experiments (Experiments 3 and 4) that were identical to Experiments 1 and 2, except the unfamiliar social contracts used expressed an exchange between two individuals rather than a social law.

The predictions for Experiments 3 and 4 are identical to the predictions for Experiments 1 and 2. Hence, they provide an opportunity to replicate the results of the six critical tests for choosing between social contract theory and availability theory.

Subjects.

Twenty-four undergraduates from Harvard University participated in Experiment 3; they were paid volunteers, recruited by advertisement (11 females, 13 males; mean age: 20.0 years (no data on age of 3 subjects)).

Materials and Procedures.

The procedure was identical to that described for Experiment 1. The materials were also identical, with one exception: the U-STD-SC expressed a private exchange rather than a social law. The 'A' version rule was: "If you get a tattoo on your face, then I'll give you cassava root." The 'B' version rule was: "If you give me your ostrich eggshell, then I'll give you duiker meat." In both cases, the "deal" was offered by the person identified in the story as the potential cheater. Thus, in terms of the value
system of the potential cheater, the SC structure of both problems was: "If B(me) then C(me)" -- a STD-SC. The SC answer to such a problem is 'P & not-Q'. The U-STD-SC problems used in Experiment 3 are shown in Figure 6.8.

Note that both social contract stories include a time delay between when the potential cheater receives his benefit and when he must cough up C(cheater) -- the benefit to the other person. In most Pleistocene exchanges reciprocation was delayed, not simultaneous (see Chapter 5). Cheating is far easier when reciprocation must occur after a benefit has been received, and subjects should be more likely to suspect someone of intending to cheat in such delayed benefit transactions. In a simultaneous, face-to-face exchange, if you see that the other person has come prepared to defect, you simply withhold what he or she wants. Subjects can be expected to assume that such intercontingent behavior will occur in face-to-face exchanges, unless they are given information to the contrary.

If subjects made this assumption, what would happen to performance on an SC Wason selection task with no time delay? Subjects would fail to choose the "cost NOT paid" card (U-STD-SC: not-Q; U-SWC-SC: not-P). This card indicates that the potential cheater had, indeed, come prepared to cheat -- that he had NOT paid the cost. The subject would assume that upon seeing this, the honest party in the interaction would simply withhold the item that the potential cheater had wanted (B(cheater)). No exchange would have taken place, and therefore no cheating. Subjects would therefore choose only the "benefit accepted" card: 'P' alone on a U-STD-SC, 'Q' alone on a U-SWC-SC.
morning. Did Big Kiku get away with cheating any of these four men? Indicate only those card(s) you definitely need to turn over to see if Big Kiku has broken his word to any of these men.

A. : Big Kiku : got the tattoo : gave him : nothing
B. : Big Kiku : got the tattoo : gave him : nothing
C. : Big Kiku : no tattoo : gave him : cassava root
D. : Big Kiku : no tattoo : gave him : cassava root

The Namka are a hunter-gatherer people who live in small bands in the deserts of southwest Africa. You are an anthropologist interested in whether members of different Namka bands can trust each other.

Bo is a crafty old Namka man in the band you are studying. He is always accidentally breaking his ostrich eggshell and would like to ‘stockpile’ some — the Namka use ostrich eggshells as canteens because they are light and hold lots of water. He sees his opportunity when four men from a neighboring band stumble into camp one morning.

The four men have been on a long and unsuccessful hunting expedition. They are hungry, and they want to be able to bring meat back to their families. Bo approaches each man privately and offers him the following deal:

"If you give me your ostrich eggshell, then I'll give you duiker meat."

Bo explains that his wife is skinning the duikers today, and they won't be ready until tomorrow. However, he will need the eggshell by this evening for his son, who is leaving tonight on a week long hunting expedition. Each man accepts Bo's offer, and agrees to meet him alone in a secluded spot tomorrow to consummate the deal.

You find this deal interesting, because you happen to know that Bo, who is rather unscrupulous character to begin with, has very little duiker meat and a large family to feed. It is perfectly possible that he will cheat some of these men. You decide to "spy" on Bo and see.

The cards below have information about the four deals Bo made with these four men. What happened in one deal had no effect on the outcome of any other deal. Each card represents one man. One side of a card tells whether or not the man gave his ostrich eggshell to Bo that evening, and the other side of the card tells whether or not Bo gave that man duiker meat the next day.

Did Bo get away with cheating any of these four men? Indicate only those card(s) you definitely need to turn over to see if Bo has broken his word to any of these men.

A. : He gave : his ostrich eggshell to Bo : nothing
B. : Bo gave him : nothing
C. : He gave : Bo nothing : duiker meat
D. : Bo gave him : nothing
Such sophisticated reasoning about intercontingent behavior should come quickly and easily to subjects. In fact, I ran a few pilots in which I had forgotten to include a time delay. After doing the tasks, a number of subjects spontaneously told me they had assumed the intercontingent scenario sketched above. Their card choices on both standard and switched social contract problems were consistent with their claim.

Hence, the time delay of reciprocal altruism is an essential element in a story when the rule expresses a private exchange: it allows the potential cheater to seize the benefit before he is expected to pay the cost. The honest person then has no options.

Because the social contract rules in Experiments 1 and 2 were expressed as social laws, the U-SC rules could be identical to their corresponding U-D rules. In experiments 3 and 4, the terms used are similar to those used in the corresponding U-D rules, but, because the U-SC rules express private deals rather than laws, they could not be identical.

U-STD-SC: "If you get a tattoo on your face, then I'll give you cassava root."
U-D: "If a man has a tattoo on his face, then he eats cassava root."
U-STD-SC: "If you give me your ostrich eggshell, then I'll give you duiker meat."
U-D: "If you have found an ostrich eggshell, then you eat duiker meat."

Results.

The percent of subjects choosing 'P & not-Q' for each problem closely matches the social contract predictions shown in Table 6.1. No one chose 'not-P & Q'; this is consistent with both AV and SC. The results are remarkably similar to those for
Experiment 1.

Table 6.7

<table>
<thead>
<tr>
<th></th>
<th>P not-Q</th>
<th>not-P</th>
</tr>
</thead>
<tbody>
<tr>
<td>U-STD-SC:</td>
<td>71</td>
<td>0</td>
</tr>
<tr>
<td>U-D:</td>
<td>25</td>
<td>0</td>
</tr>
<tr>
<td>AP:</td>
<td>29</td>
<td>0</td>
</tr>
<tr>
<td>F-D:</td>
<td>38</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 6.7 shows the percent of subjects who chose either 'P & not-Q' or 'not-P & Q'. Residual responses: Of the 7 responses to the U-STD-SC that were not full SC answers, 5 were half correct, "sins of omission": 5 'P' responses (omitted not-Q). No one answered 'not-Q' (omitting P) on either the U-STD-SC or the F-D, so counting rare residuals neither increases nor decreases the magnitude of the difference between these two problems.

Critical Tests

Six critical tests pitting hypothesis AV against hypothesis SC were used in analyzing the data from Experiments 1 and 2. The same six critical tests can be carried out on the data from Experiments 3 and 4.

Critical Test 1: Does an unfamiliar standard social contract elicit the predicted SC response, 'P & not-Q'?

Percentage 'P & not-Q' responses:

<table>
<thead>
<tr>
<th>Social Contract</th>
<th>Availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prediction:</td>
<td>Prediction:</td>
</tr>
<tr>
<td>U-STD-SC v. U-D.</td>
<td>U-STD-SC &gt; U-D</td>
</tr>
<tr>
<td>high</td>
<td>low</td>
</tr>
<tr>
<td>low</td>
<td>low</td>
</tr>
</tbody>
</table>

71% > 25%
There is a highly significant 46 point discrepancy in 'P & not-Q' responses between the U-STD-SC and the U-D (71% v. 25%: $F_{1,23} = 19.46, p < .001, r = .68$). This discrepancy is predicted only by SC; AV predicts a low proportion of falsifying responses on all unfamiliar problems, whether they are social contracts or not. U-STD-SC also produces a significant "content effect" when measured against the AP (71% v. 29%: $F_{1,23} = 16.43, p < .001, r = .65$). The U-D and AP both elicited the same low levels of falsifying responses (25% v. 29%: $F_{1,23} = 0.19$, n.s.).

Critical Test 2: Are there more SC responses to an unfamiliar standard social contract than falsifying responses to a familiar descriptive problem?

Percentage 'P & not-Q' responses:

<table>
<thead>
<tr>
<th>Social Contract Prediction:</th>
<th>Availability Prediction:</th>
</tr>
</thead>
<tbody>
<tr>
<td>U-STD-SC v. F-D</td>
<td>U-STD-SC &gt; F-D</td>
</tr>
<tr>
<td>high</td>
<td>low</td>
</tr>
<tr>
<td>low</td>
<td>mid-low</td>
</tr>
</tbody>
</table>

71% > 38%

The advantage that SC status gives an unfamiliar problem is larger than the advantage availability gives a familiar descriptive problem. The U-STD-SC elicited significantly more "falsifying" (SC) responses than the F-D (71% v. 38%: $F_{1,23} = 6.57, p < .025, r = .47$). AV predicts an inequality in the opposite direction.

This supports the claim that SC algorithms are a major determinant of responses to problems involving social exchange.

Experiment 4

This experiment is identical to Experiment 2, except the
switched social contract rules used express a private exchange rather than a social law. AV and SC predictions are the same as for Experiment 2.

Subjects.

Twenty-four undergraduates from Harvard University participated in Experiment 4; they were paid volunteers, recruited by advertisement (11 females, 13 males; mean age: 19.4 years (no data on one subject's age)).

Materials and Procedures.

The procedure was identical to that for Experiment 3. The materials were also identical with one exception: the unfamiliar rules were "switched." Thus, the U-SWC-SC rules were: "If I give you cassava root, then you must get a tattoo on your face" ('A' version), and "If I give you duiker meat, then you must give me your ostrich eggshell" ('B' version). The U-Ds were the same as in Experiment 2.

Results.

The results are shown in Table 6.8; they match the social contract predictions of Table 6.3.

Critical Test 3: Does an unfamiliar switched social contract elicit the predicted SC response, 'not-P & Q'?

Percentage 'not-P & Q' responses:

<table>
<thead>
<tr>
<th>Social Contract Prediction:</th>
<th>Availability Prediction:</th>
</tr>
</thead>
<tbody>
<tr>
<td>U-SWC-SC v. U-D,AP,F-D.</td>
<td>U-SWC-SC &gt; U-D,AP,F-D</td>
</tr>
<tr>
<td>high</td>
<td>very low</td>
</tr>
<tr>
<td>75%</td>
<td>&gt; 0%, 0%, 0%</td>
</tr>
<tr>
<td>U-SWC-SC = U-D,AP,F-D</td>
<td>very low</td>
</tr>
</tbody>
</table>

227
Table 6.8 Experiment 4: Percent of subjects choosing 'P & not-Q' or 'not-P & Q' for each problem (n=24)

<table>
<thead>
<tr>
<th></th>
<th>P</th>
<th>not-P</th>
<th>Q</th>
</tr>
</thead>
<tbody>
<tr>
<td>U-SWC-SC:</td>
<td>0</td>
<td></td>
<td>75</td>
</tr>
<tr>
<td>U-D:</td>
<td>25</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>AP:</td>
<td>33</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>F-D:</td>
<td>58</td>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>

Table 6.8 shows the percent of subjects who chose either 'P & not-Q' or 'not-P & Q'. Residual responses: Of the 6 responses to the U-SWC-SC that were not full SC answers, 4 were half correct, "sins of omission": 4 'Q' responses (omitted not-P). Counting rare residuals for both the U-SWC-SC ('Q') and the F-D ('not-Q') increases the magnitude of the difference between them (92% v. 63%).

The large and significant 75 point difference between U-SWC-SC and all other problems is predicted only by SC (75% v. 0%, 0%, 0%, L = +3,-1,-1,-1: F = 207.01, p << .001, r = .87). 'Not-P & Q' was not chosen on any other problem in both Experiments 3 and 4. AV can neither predict nor explain this result.

Critical Test 4: Are there more SC responses to an unfamiliar switched social contract than falsifying responses to a familiar descriptive problem?

Percentage 'not-P & Q' responses to U-SWC-SC, Percentage 'P & not-Q' responses to F-D:

Social Contract Prediction: Availability Prediction:
U-SWC-SC v. F-D. U-SWC-SC > F-D high
              U-SWC-SC < F-D low
              high
              very low mid-low

75% > 58%

There were more SC responses to the unfamiliar U-SWC-SC than
falsifying responses to the familiar F-D, just as SC predicts (75% v. 58%; \( F = 1.64, \text{n.s.} \)). Although the difference is not significant, AV predicts an inequality in the opposite direction. When rare residuals are counted for both problems (U-SWC-SC: 'Q'; F-D: 'not-Q'), the difference is magnified, and is significant (92% v. 63%; \( F = 6.75, p < .025, r = .48 \)). Like the results of Critical Test 2, this supports the contention that SC algorithms are a major determinant of responses to problems involving social exchange.

**Experiment 3 versus Experiment 4**

Critical Test 5: Is the correct SC response to a standard social contract ('P & not-Q') very rare for a switched social contract?

AV predicts that the U-STD-SC and the U-SWC-SC should both elicit low levels of 'P & not-Q' responses because they are both unfamiliar. In contrast, SC predicts that 'P & not-Q', the dominant, SC response to the U-STD-SC, should be very rare on the U-SWC-SC.

**Percentage 'P & not-Q' responses:**

<table>
<thead>
<tr>
<th>Social Contract Prediction</th>
<th>Availability Prediction:</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>very low</td>
</tr>
<tr>
<td>low</td>
<td>low</td>
</tr>
</tbody>
</table>

71% >> 0%

Only SC predicts the large and significant 71 point discrepancy in 'P & not-Q' responses between the U-STD-SC and the U-SWC-SC (71% v. 0%; \( Z = 5.14, p < .00000025, \phi = .74 \)). The SC
prediction that the dominant, SC response to the U-STD-SC will be very rare on the U-SWC-SC was also borne out: no one gave the STD-SC answer, 'P & not-Q', in response to the U-SWC-SC.

Critical Test 6: Is the correct SC response to a switched social contract ('not-P & Q') very rare for a standard social contract?

SC predicts that the correct SC answer to a SWC-SC, 'not-P & Q', will be very rare for a STD-SC. In contrast, AV predicts that the percentage of subjects choosing 'not-P & Q' on the U-STD-SC and the U-SWC-SC will be about equal, and very low.

Percentage 'not-P & Q' responses:

<table>
<thead>
<tr>
<th>Social Contract Prediction:</th>
<th>Availability Prediction:</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>very low</td>
</tr>
</tbody>
</table>
| 75% >> 0%                  | very low                 | very low

The large and significant 75 point discrepancy in 'not-P & Q' responses between U-SWC-SC and U-STD-SC problems is predicted only by SC (75% v. 0%: Z = 5.37, p < .0000001, phi = .77).

Furthermore, the dominant, SC response to the U-SWC-SC was indeed very rare on the U-STD-SC, just as SC predicts: no one gave the SWC-SC answer, 'not-P & Q', in response to the U-STD-SC.

Social Contract Tests

As before, if the social contract view is correct, certain relations should be manifest in the data, above and beyond those addressed by the critical tests. For convenience, the data from Tables 6.7 and 6.8 are combined in Table 6.9.
Table 6.9 Percent of subjects choosing 'P & not-Q' or 'not-P & Q' for each problem

<table>
<thead>
<tr>
<th>Experiment 3 (n=24)</th>
<th>Experiment 4 (n=24)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>not-P</td>
</tr>
<tr>
<td>not-Q</td>
<td>Q</td>
</tr>
<tr>
<td>71*</td>
<td>0</td>
</tr>
<tr>
<td>U-D:</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>0</td>
</tr>
<tr>
<td>AP:</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>0</td>
</tr>
<tr>
<td>F-D:</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>0</td>
</tr>
</tbody>
</table>

*predicted SC response to social contract problems.

Are the logically distinct SC answers to standard and switched SC problems produced by the same algorithms?

The proportions of SC answers to the U-STD-SC ('P & not-Q') and U-SWC-SC ('not-P & Q') are not significantly different (71% v. 75%: Z = 0.32, n.s.), just as one would expect if these two logically distinct responses were the product of the same SC algorithms. Using percent falsifying answers to the U-D (same rule as U-SC) as a baseline for comparison, the relative advantage SC status gave in producing SC answers is very similar for both SC problems: 46 points between U-STD-SC and its U-D, 50 points between U-SWC-SC and its U-D.

For these personal exchange problems, residual responses did not split (e.g., some 'P', some 'not-Q' for the U-STD-SC) as they did for the law problems of Experiments 1 and 2; the only "sins of omission" in Experiments 3 and 4 involved choosing the card that represents what the honest person did. There were 5 'P'
alone responses for the U-STD-SC and 4 'Q' alone responses to the U-SWC-SC. This is interesting, because only Experiments 3 and 4 admit the possibility of intercontingent behavior; these are the responses one would expect if a subject read through quickly, not noticing that there is a time delay between when the potential cheater gets his benefit and when he is expected to honor his end of the deal. The numbers involved are too small to firmly attribute this pattern to the power of intercontingent reasoning, but it is an area that deserves future research.

Table 6.10 shows the frequencies with which individual cards were selected in Experiments 3 and 4.

<table>
<thead>
<tr>
<th>Logical Category</th>
<th>D-D</th>
<th>AP</th>
<th>F-D</th>
<th>U-SC</th>
<th>Social Contract Category:</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>22  21</td>
<td>23  23</td>
<td>24  23</td>
<td>23  1</td>
<td>Benefit Accepted</td>
</tr>
<tr>
<td>not-P</td>
<td>4   5</td>
<td>4   5</td>
<td>0   1</td>
<td>2   19</td>
<td>Benefit NOT Accepted</td>
</tr>
<tr>
<td>Q</td>
<td>11  11</td>
<td>7   9</td>
<td>4   5</td>
<td>0   23</td>
<td>Cost Paid</td>
</tr>
<tr>
<td>not-Q</td>
<td>12  13</td>
<td>10  14</td>
<td>10  18</td>
<td>17  2</td>
<td>Cost NOT Paid</td>
</tr>
</tbody>
</table>

Just as before, the two social contract problems replicate when sorted by social contract category, but not when sorted by logical category, as the other problems do. The social contract categorization scheme captures dimensions that are psychologically real for the subjects in Experiments 3 and 4, just as it did in Experiments 1 and 2.

How well do SC algorithms operate in novel, versus familiar, social exchanges?

A frame-builder structures novel experiences along evolutionarily relevant dimensions that it is keyed to pick up. If SC algorithms function as frame-builders, then they should operate well in novel social exchanges, like those represented in
the unfamiliar social contract problems. The percentage of SC responses elicited by the U-STD-SC -- 71% -- does not differ significantly from the 78% elicited by the familiar DAP with a similar group of Harvard students (71% v. 78%: \( Z = 0.58, \text{n.s.} \); see Experiment 6-A). Neither does the percentage of logically distinct SC responses elicited by the U-SWC-SC (75% v. 78%: \( Z = 0.26, \text{n.s.} \)). The hypothesis that familiar SC problems elicit more SC answers than unfamiliar ones is not supported by this data, even when one combines all three problems (F-STD-SC, U-STD-SC, U-SWC-SC) into one test (78% v. 71%: 75%, \( L = +2,-1,-1, F = 0.23, \text{n.s.} \)). Not only are SC algorithms keyed to abstract SC dimensions from novel situations, but they accomplish this as efficiently in novel situations as they do in familiar ones.

Availability assessed

Does availability have any effect at all on familiar problems?

The F-D did not elicit significantly more falsifying responses to either the AP or the U-D in Experiment 3 (F-D v. AP: 38% v. 29%, \( F = 1.31, \text{n.s.} \); F-D v. U-D: 38% v. 25%, \( F = 1.30, \text{n.s.} \); F-D v. AP, U-D: \( L = +2,-1,-1, F = 1.38, \text{n.s.} \)). However, in Experiment 4 the F-D did elicit more falsifying responses than both the AP and U-D (F-D v. AP: 58% v. 33%, \( F = 5.31, p < .05, r = .43 \); F-D v. U-D: 58% v. 25%, \( F = 11.50, p < 1,23 \) these two experiments than it was in Experiments 1 and 2.

In Experiments 3 and 4, availability gave familiar non-SC problems an advantage of 9 to 33 points over unfamiliar non-SC problems in producing falsifying responses (average advantage = 233
20). However, social contract status gave unfamiliar problems an average advantage of 45 points in producing SC responses -- this means the social contract advantage was 2.25 times the size of the availability advantage. Compared to the social contract effect, the effect of availability was even smaller in Experiments 3 and 4 than it was in Experiments 1 and 2.

By considering the results of all four experiments \((n=96)\), we can estimate the relative sizes of the availability and social contract effects. Measured against the U-D problems, the unfamiliar social contracts yield an effect size \((r)\) of .70, whereas the familiar descriptive problems yield an effect size of .47. Overall, then, the social contract effect is 1.49 times the size of the availability effect -- about half again as large.

These results can shed light on why performance with the transportation problem has been so erratic in the literature. When the transportation problem is measured against the AP -- the standard test for availability in the literature -- the effect size \((n=96)\) for availability is only \(r = .41\). Twenty-four \((df 23)\) is a common sample size in the literature. Assuming the true effect size is .41, a sample size of 24 would yield an \(F\) of \(1.23 \times 4.65\) -- barely over the \(p < .05\) cutoff of 4.28. With just a little sample variation, one would sometimes see an effect, sometimes not.

**Summary, Experiments 3 and 4.**

These experiments replicated the results of Experiments 1 and 2: as before, the six critical tests supported hypothesis SC over AV, and all the other social contract predictions were borne
out. Availability was shown, once again, to have a minor but erratic effect on familiar descriptive problems.

Position Effects, Experiments 1 through 4

Experiments 1 through 4 controlled for problem position: each experiment had four problems, and every problem occurred in each position the same number of times. Therefore these experiments can be used to see if there were any position effects: if the probability of a subject solving a Wason selection task correctly increases with the number of problems that subject has already been exposed to. This will be useful to know in analyzing some data from Experiment 6.

Experiments 1 through 4 do not allow a pure test of position,* however, because the possibility of transfer from a successfully solved social contract problem is confounded with position — the later a problem occurs in the booklet, the more likely it is to have followed the social contract problem. Although no one in the literature has yet found transfer from a successfully solved Wason selection task to an abstract problem (Johnson-Laird, Legrenzi, & Legrenzi, 1972; Griggs & Cox, 1982), there is one report of transfer from a social contract problem to a "semi" social contract problem, the apparel-color problem (Cox & Griggs, 1982; see Chapter 5). Success on a social contract problem could give subjects insight into the structure of the F-D transportation problem, especially in my subject population. After all, my subject population is an unusual one: the median SAT scores of Harvard students are among the highest — if not the highest — in the U.S. SATs test, in part, an applicants'
ability to discern the structure of a problem, regardless of its content; high scores indicate skilled problem solvers. Consonant with this, for my subject population, the falsification rate on the abstract problems of Experiments 1-4 was unusually high: instead of the 4-10% rate found for most studies (Wason, 1983), an average of 25% of Harvard students falsified on the abstract problems. First solving a U-SC correctly, combined with the familiarity of the transportation problem, might provide enough insight into the structure of these problems to improve performance on the transportation problem.

Table 6.11 shows the number of problems correctly solved as a function of position for Experiments 1 through 4. The hypothesis that there is a linear effect of position was tested using the contrasts \( L = -3, -1, +1, +3 \).

<table>
<thead>
<tr>
<th>Position:</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (n=24)</td>
<td>7</td>
<td>9</td>
<td>11</td>
<td>13</td>
</tr>
<tr>
<td>2 (n=24)</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>3 (n=24)</td>
<td>9</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>4 (n=24)</td>
<td>8</td>
<td>12</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>1-4 (n=96)</td>
<td>32</td>
<td>38</td>
<td>42</td>
<td>47</td>
</tr>
<tr>
<td>2-4 (n=72)</td>
<td>25</td>
<td>29</td>
<td>31</td>
<td>34</td>
</tr>
</tbody>
</table>

\( F = 3.92, \text{n.s.} \)
\( F = 1.02, \text{n.s.} \)
\( F = 0.10, \text{n.s.} \)
\( F = 3.12, \text{n.s.} \)
\( F = 6.60, p < .025, r = .15 \)
\( F = 3.18, \text{n.s.} \)

** "Correct" is defined as 'not-P & Q' for U-SWC-SC problems, 'P & not-Q' for all other problems.

* It is not clear to me that a pure test of order is possible -- all methods I have considered seem riddled with confounds.
None of these experiments showed any significant position effect. When the results of all four experiments are combined into one large experiment (n=96) there is a significant effect of order. But remember: position is confounded with transfer in these experiments, if there is evidence of transfer. Experiment 1 showed significant transfer from a successfully solved U-SC to the P-D; none of the other experiments did (Exp 1: $G_{adj} = 6.34$, $p < .01$; Exp 2: $G_{adj} = 0.17$, n.s.; Exp 3: $G_{adj} = 0.00$, n.s., Exp 4: $G_{adj} = 1.56$, n.s.; Sokal & Rohlf, 1969, p.591). When Experiment 1 is excluded from the analysis such that the three experiments showing no evidence of transfer (2-4) are analyzed together, there is no significant order effect -- and, for an effect size of $r = .15$, 72 subjects are sufficient to produce a significant F. Thus, it seems fair to conclude that the order effect found for the combined 1-4 data was due to the transfer effect in Experiment 1, not to an effect of problem position. Taken together, these results indicate that there is no effect of position alone on performance.

Experiment 5

For standard social contracts, the correct SC response happens to also be the logically correct response, 'P & not-Q'. Experiments 1 and 3 showed that STD-SC status facilitates "logical falsification" for unfamiliar rules set using concrete -- if obscure -- terms. But can STD-SC status also facilitate "logical falsification" for an abstract problem?

Experiment 5 compared performance on two abstract, prescriptive rules: one a standard social contract rule, one not
Both were set in a realistic context -- they were bureaucratic rules involving an immigration office. The story surrounding one rule defined its propositions in cost/benefit terms that made it an abstract standard social contract (A-STD-SC). The story surrounding the other rule gave each proposition a clear meaning, but these meanings did not represent costs or benefits; hence, that problem was an ordinary abstract problem (AP). AV predicts that the percentage of 'P & not-Q' responses will be equally low on both problems; SC predicts a high percentage for the A-STD-SC and a low percentage for the AP.

Subjects.

Forty-two undergraduates from Harvard University participated in Experiment 5; they were volunteers from an introductory science course. Eighteen were in the group given the AP (13 females, 5 males), 24 in the group given the A-STD-SC (16 females, 8 males).

Materials and Procedure.

Although each test booklet contained five problems, only the abstract problems are relevant to Experiment 5. The fourth problem in every booklet was an abstract problem -- either the AP or the A-STD-SC. The other problems are shown in Figure 6.9. The first and third problems were food problems (one "salad bar" problem and one "Mexican food" problem per booklet), the second problem was a semi-social contract problem similar to Cox & Griggs' (1982) apparel color problem (see Chapter 2), and the fifth problem was a pilot threat problem (see Appendix C). Experiment 5 had a between-groups design.

Experiment 5 was part of a large battery of tests given at
Most of your friends find Mexican food too spicy. That's about all you knew about Mexican food when the Association of Mexican-American Restauranters (AMAR) first hired you to conduct market research on Americans' food preferences. So you started reading up on the subject. In the course of your reading, you came across a previously done report on the eating habits of Americans in Mexican restaurants which said:

"If a person eats hot chili peppers, then he will drink a cold beer."

You decide to look into this yourself. The cards below have information about four people in a Mexican restaurant in Boston. Each card represents one person. One side of a card tells what a person ate, and the other side of the card tells what that person drank.

Indicate only those card(s) you definitely need to turn over to see if any of these people violate this rule.

A. .................. B. ..................  
| hot tea          | cold beer          |
|                 |                  |

C. .................. D. ..................  
| hot chili        | broccoli          |
| peppers          |                  |

When the Consortium of Salad-Bar Owners & Operators (CSO&O) first hired you to conduct market research on Americans' food preferences, all you knew about salad bars was that your friends invariably start at the wrong end of them -- where the bacon bits & dressing are instead of where the lettuce is. So you started reading up on the subject. In the course of your reading, you came across a previously done report on the eating habits of Americans in salad bars which said:

"If a person eats lettuce, then he will drink water."

You decide to look into this yourself. The cards below have information about four people in a salad bar in Cambridge. Each card represents one person. One side of a card tells what a person ate, and the other side of the card tells what that person drank.

Indicate only those card(s) you definitely need to turn over to see if any of these people violate this rule.

A. .................. B. ..................  
| water            | cheese           |
|                 |                  |

C. .................. D. ..................  
| coffee           | lettuce          |

You are an anthropologist studying the Kaluame people, a Polynesian culture found only on Maku Island in the Pacific. The Kaluame adopted you into their culture and gave you the job of enforcing their societal laws. One of their social rules is:

"If a person is wearing a floral print shirt, then he must be over 20 years old."

The cards below have information about four Kaluame people sitting in a temporary camp; they don't seem to know one another. Each card represents one person. One side of a card tells what kind of a shirt a person is wearing and the other side tells that person's age.

Indicate only those card(s) you definitely need to turn over to see if any of these Kaluame people are violating this rule.

A. .................. B. ..................  
| solid black shirt | floral print shirt |
|                  |                   |

C. .................. D. ..................  
| 42 years old     | 17 years old      |
| shirt            |                   |
the beginning of a class. Upon entering the class, each subject took a sealed test booklet from a stack containing booklets from the entire battery. The order of booklets in the stack was random. The first page of each booklet was an instruction sheet. The instructions were the same as for Experiments 1 through 4, except there was an additional sentence telling subjects that if a problem had more than one question associated with it, to answer the questions in the order they appear. Subjects had the opportunity to read the instructions to themselves; in addition, the instructions were read aloud to the entire class.

The two abstract problems are shown in Figure 6.10. Both rules express an alphanumeric code governing immigration documents (AP: "If a person has a 'D' rating, then his documents must be marked code '3'"; A-STD-SC: "If you fill out document 'D', then Ed will rate you code '3'"). To make sure subjects paid attention to what the letters and numbers stood for in the story, the selection task question was preceded by questions about each term. For the AP, these questions asked what each term meant; for the A-STD-SC they asked how the desires of characters in the story mapped onto each term.

The relation expressed in the AP was a sensible one -- if one translates the letters and numbers into the propositions they represent, it reads:
"If an applicant has submitted an incomplete medical record, then that applicant must see a health inspector."

I wanted both abstract rules to be set in a similar context -- thus both are bureaucratic rules involving immigration, and both are set in an immigration office. However, this constraint
You work for the national office of the U.S. Immigration Service, and it is your job to weed out corruption among immigration officers in the branch offices.

Four upstanding Mexican citizens want to immigrate to the U.S. Each needs to be rated code '3' by U.S. immigration in order to stay here; code '*7*' means instant deportation. Central Americans must file document D with immigration; Europeans must file document P. Each Mexican calls the Texas immigration office and talks to Betty, a very nice secretary who works for an immigration officer named Ed. Betty sends each a standard packet containing documents 'A' through 'H'.

And, in accordance with standard immigration law, Betty explains the law as follows:

"If you fill out document 'D', then Ed will rate you code '3'."

Ed is the very bigoted immigration officer who is in charge of processing these forms. You know that Ed hates Mexicans; he has made no secret of the fact that he does not want them coming into this country. You suspect that he will try to cheat some or all of these people out of their right to immigrate.

Questions:

1. Each Mexican wants to be rated code number ___.
2. It is in the interest of each Mexican to fill out document ___.
3. The Mexicans would not need to fill out document ___.
4. Personally, Ed would like to rate each Mexican code number ___.
5. The cards below have information about the documents of the four Mexicans who applied for immigration through Ed's office. Each card represents one person. One side of a card tells a person's letter rating and the other side of the card tells what number code rating Ed gave that person. Indicate only those card(s) you definitely need to turn over to see if the documents of any of these people violate the alphanumeric rule.

Part of your new clerical job at the local immigration office is to make sure that the documents of people applying to immigrate into the U.S. are processed according to the proper alphanumeric rules relating letter ratings to number codes. There's a whole alphabet of letter ratings, and as many different number codes. For example, a 'D' rating indicates that the applicant has submitted an incomplete medical record, an 'E' rating means the applicant has dependent relatives; code '*3*' means the applicant must see a health inspector, code '*7*' means that the applicant must undergo a security check for possible criminal activities. The documents are already known to have the correct letter rating. Your job is to make sure the documents conform to the following alphanumeric rule:

"If a person has a 'D' rating, then his documents must be marked code '3'"

Questions:

1. Which code means the applicant must see the health inspector? ___
2. Which rating means the applicant's medical record is incomplete? ___
3. Which rating means the applicant has dependent relatives? ___
4. Which code means the applicant must undergo a security check? ___
5. You suspect the absent-minded secretary you replaced did not categorize the documents correctly. The cards below have information about the documents of four people who are applying for immigration. Each card represents one person. One side of a card tells a person's letter rating and the other side of the card tells that person's number code. Indicate only those card(s) you definitely need to turn over to see if the documents of any of these people violate the alphanumeric rule.

\[
\begin{array}{c|c|c}
\text{A} & \text{D} & \text{7} \\
\text{B} & \text{3} & \text{1} \\
\text{C} & \text{F} & \text{1} \\
\text{D} & \text{3} & \text{1} \\
\end{array}
\]
made it difficult to construct a "full strength" social contract problem. By their nature, bureaucratic rules express "institutional desires", not the personal desires of the bureaucrat charged with enforcing them. What is a STD-SC from the institution's "point of view" may not be a proper social contract of any kind from the point of view of the bureaucrat administering it. This can be seen by translating the rule into cost/benefit terms from the point of view of the immigrants, the State, and Ed, the bureaucrat and potential cheater.

The opportunity to immigrate (be rated "code '3'") is considered a rationed benefit in this country; having to filling out papers is usually considered a cost/requirement. From the point of view of the immigrants' value system, the A-STD-SC rule translates to:

"If P then Q
"If you pay the cost, then you are entitled to the benefit."

Ed, the potential cheater, is the bureaucrat charged with enforcing this rule. Cheating an immigrant would clearly involve not giving him the benefit (not-Q) he is entitled to when he has paid the cost (P). Thus the correct SC answer is 'P & not-Q', making this problem a STD-SC. (Remember: a STD-SC only has the form "If B(X) then C(X)" when translated into the value system of the potential cheater. In this problem Ed is the potential cheater, not the immigrants.)

How does this rule translate from the point of view Ed, the potential cheater? A bureaucrat is charged with representing the desires of the institution that made the rule, not his own: the bureaucrat is supposed to behave as if his interests and the
interests of his employer coincide. This rule translates into a conventional STD-SC only if you view Ed as the instrument of the State -- an interpretation discouraged by the text. Subjects living in this country are presumably aware that the State -- Ed's employer -- views immigration as costly, but is willing to allow it when the person's character and finances are likely to provide benefits that offset the cost -- benefits assured, presumably, by filling out document D. If Ed, the potential cheater, is seen merely as the instrument of the State, then the rule is a proper STD-SC:

"If P then Q"  
"If we receive the benefit, then we must pay the cost."

Translating the rule into Ed's personal value system is more difficult; because Ed did not make the rule, it does not necessarily express his values. He clearly agrees that the second term is a cost -- Ed doesn't want Mexicans entering the country. However, it is not clear that a completed "document 'D'" benefits Ed in any way. From Ed's point of view, the rule probably translates to:

"If P then Q"  
"If 0(Ed) then C(Ed)."

Moreover, we know from the story that Ed prefers not-Q (not letting Mexicans in) to Q (letting them in). Hence, in terms of cheating the immigrants, Ed has both motive and opportunity.

But note that he is not cheating them in the natural selection sense of the word -- he is not absconding with a benefit without paying the required cost. Ed is not benefited by
the immigrants having filled out document D*; by rating the immigrants code 7, Ed is avoiding the imposition of a cost, but not absconding with a benefit. From the immigrants' point of view, his doing this does cheat them in the natural selection sense of the word -- they have paid the cost, but not received the benefit that this entitles them to. Thus, in the natural selection sense, the immigrants have been cheated but Ed has not cheated them. This assymetry results from the fact that Ed did not himself make this contract with the Mexicans, so it does not express his own desires.

In fact, given Ed's bigotry, he probably does not see the rule as a social contract between himself and the immigrants, and thinks of it thus:

"If a Mexican fills out document 'F', then I get to rate him code '7'."
"If 0(Ed) then B(Ed)"

From the Mexican's point of view, this rule would be:

"If 0(Mexican) then 0(Mexican)"

-- not a social contract at all!

So, the rule is slightly confusing from a social contract point of view. If one identifies with the immigrants, then the conditions that constitute cheating are clear; if one identifies with Ed, they are not so clear. For this reason, performance may be weaker on this particular A-STD-SC than on the "full strength" U-STD-SCs used in Experiments 1 and 3.

* Nor is the State. Because B(State) resides in the character and finances of the potential immigrants, B(State) can only be collected if the Mexicans are allowed to immigrate. Thus, in the natural selection sense, the State, like Ed, cannot cheat the Mexicans -- it can, however, violate the rule.
Results.

Critical Test 7: Does an abstract standard social contract elicit the predicted SC response, 'P & not-Q'?

Percentage 'P & not-Q' responses:

<table>
<thead>
<tr>
<th>Social Contract Prediction:</th>
<th>Availability Prediction:</th>
</tr>
</thead>
<tbody>
<tr>
<td>high low</td>
<td>low low</td>
</tr>
<tr>
<td>58% &gt; 22%</td>
<td></td>
</tr>
</tbody>
</table>

The A-STD-SC elicited more "falsifying" responses than the AP, just as social contract theory predicts it should (z = 2.34, p < .01, phi = .36). Hypothesis AV does not predict and cannot explain this result. However, the size of this effect is somewhat smaller than those found in Experiments 1 and 3 for the U-STD-SC v. AP comparisons -- about 57% the size (.36 for A-STD-SC v. AP, versus .61 and .65 for U-STD-SC v. AP in Experiments 1 and 3). There are at least three explanations for this smaller effect size:

1. Abstract symbols are more difficult to process than words, no matter how unfamiliar those words are. If this were true, then in Experiments 1-4, falsification rates would have been lower for AP problems than for U-D problems; they were not. Therefore, this explanation is rather unlikely.

2. Random variation in population sampling -- I have no way of assessing the merit of this explanation.

3. The A-STD-SC was not a "full-strength" social contract problem whereas the U-STD-SCs of Experiments 1 and 3 were (see discussion above).

I have no data that would allow me to choose between the second and third explanations.
Summary, Experiment 5.

An abstract standard social contract elicited a "content
effect", as result predicted only by social contract theory. The
effect was somewhat smaller than the social contract effects
found in Experiments 1 through 4 for the unfamiliar problems.

Experiment 6

So far, unfamiliar and abstract social contract problems
have all elicited high levels of predicted SC responses: these
were the theoretically crucial problems for establishing a social
contract effect and choosing between hypotheses SC and AV.

In Experiment 6, familiar STD-SC problems are pitted against
abstract problems and familiar descriptive problems. Social
contract theory predicts that a familiar STD-SC problem (F-STD-SC),
like its unfamiliar analog, will elicit high levels of "logical
falsification". Availability predicts middling performance on
both F-D and F-STD-SC problems, unless, before seeing the results
of the experiment, the individual availability theorist can
concoct reasons why one problem or the other should prompt more
counter-examples; so far, no principled way of making such
judgments has been proposed (see Chapter 3).

The literature results reported in Chapter 2 indicate that
familiar STD-SCs elicit replicable and robust social contract
effects; however, given the elusivity of other content effects on
the Wason selection task, it seemed prudent to test some F-STD-SCs
on the same subject population that I used for the critical tests
reported above.* Moreover, this allowed me to compare

* F-SWC-SCs cannot be used; experience with the more usual STD-SC form
creates too many confounds (see Chapter 2: "Deformed Social Contracts").
performance on familiar STD-SCs to performance on unfamiliar ones, to see if familiarity produces an effect over and above the social contract effect (see above: "Social Contract Tests").

Experiment 6-A

The F-STD-SC used in this experiment was a Drinking Age Problem (DAP; see Chapter 2, Figure 4.1). The percentage of 'P & not-Q' responses elicited by the DAP was compared to that elicited by a transportation problem (F-D) and a prescriptive AP (shown in Figure 6.5).

Subjects.

Twenty-three undergraduates from Harvard University participated in Experiment 6-A (16 females, 7 males); they were volunteers from an introductory science course.

Materials and Procedure.

Experiment 6-A was part of the battery of tests described in Experiment 5. Each subject solved five problems, which appeared in the following order: AP, F-D, F-STD-SC (DAP), semi-SC (see Figure 6.9), and threat (see Appendix B). Only the first three problems are relevant to this analysis.

As discussed in the section on position effects, results indicate that falsification rates are not enhanced by increasing serial position; in fact, Cox & Griggs (1982) present evidence suggesting that first solving an AP incorrectly depresses falsification rates on a following DAP. If anything, then, having the F-STD-SC in the third position should narrow the difference between it and the AP. Moreover, having it follow the F-D eliminates the possibility of transfer from a correctly
solved F-STD-SC to the F-D.

Throughout Experiment 6, availability predictions are based on Griggs & Cox's (1982) memory-cueing version of availability theory. Predictions based on Pollard's (1982) differential availability would fare much worse, because, for very familiar STD-SCs like the DAP (Experiment 6-A) and the restaurant problem (Experiment 6-B), falsifying instances, though available, should be less available than confirming instances (see Chapter 3). Other brands of availability put forth are not sufficiently well-specified to make any strong predictions.

So as not to confuse issues, the social contract predictions for F-D problems will assume no effect of availability, as before; however, social contract theory is silent on whether availability exercises an independent effect on such problems.

Results.

The results are pictured in Table 6.12.

<table>
<thead>
<tr>
<th>Table 6.12</th>
<th>Experiment 6-A: Percent of subjects choosing 'P &amp; not-Q' for each problem (n=23)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP:</td>
<td>30%</td>
</tr>
<tr>
<td>F-D:</td>
<td>57%</td>
</tr>
<tr>
<td>F-STD-SC:</td>
<td>78%</td>
</tr>
</tbody>
</table>

The tests that follow are not critical tests because the two hypotheses, SC and AV do not make radically different predictions about these problems (although the predictions of social contract theory are more tightly specified than those of availability theory). Rather, they are offered to show that the results are
consistent with social contract theory, and that social contract status produces reliable, replicable effects.

Does a familiar standard social contract elicit the predicted SC response, 'P & not-Q'?

Percentage 'P & not-Q' responses:

<table>
<thead>
<tr>
<th>Social Contract Prediction:</th>
<th>Availability Prediction:</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>high-mid low</td>
</tr>
<tr>
<td>78%</td>
<td>78% &gt; 30%</td>
</tr>
</tbody>
</table>

Experiment 6-A replicates the many experiments in the literature showing a robust content effect for the DAP. This F-STD-SC elicited a high percentage of SC responses -- 78%. Moreover, the proportion of 'P & not-Q' responses elicited by the F-STD-SC is significantly higher than that elicited by the AP (78% v. 30%: F = 21.08, p << .001, r = .69). This result is not predicted by Pollard's differential availability theory (despite his claims -- see Chapter 3), but it is consistent with Griggs & Cox's memory-cueing version of availability theory.

Are there more SC responses to a familiar standard social contract problem than falsifying responses to a familiar descriptive problem?

Percentage 'P & not-Q' responses:

<table>
<thead>
<tr>
<th>Social Contract Prediction:</th>
<th>Availability Prediction:</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-STD-SC v. F-D F-STD-SC &gt; F-D</td>
<td>F-STD-SC ≥ F-D</td>
</tr>
<tr>
<td>high</td>
<td>high-mid mid-low</td>
</tr>
<tr>
<td>78%</td>
<td>78% &gt; 57%</td>
</tr>
</tbody>
</table>

As social contract theory predicts, the F-STD-SC problem
elicited significantly more 'P & not-Q' responses than the F-D (78% v. 57%: F = 6.11, p < .025, r = .47). This result is also consistent with Griggs & Cox's availability theory, assuming that more subjects had experienced counter-instances to the DAP than to the transportation problem. It is not consistent with Pollard's differential availability.

Availability assessed: Does availability have any effect at all on familiar problems?

<table>
<thead>
<tr>
<th>Effect:</th>
<th>No effect:</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-D v. AP</td>
<td>F-D &gt; AP</td>
</tr>
<tr>
<td>57% &gt; 30%</td>
<td>F-D = AP</td>
</tr>
</tbody>
</table>

The F-D transportation problem elicited significantly more falsifying responses than the AP (57% v. 30%: F = 7.76, p < .025) Experiments 1 through 4. As before, the availability effect is smaller than the social contract effect.

Experiment 6-B

Instead of the DAP, Experiment 6-B uses a F-STD-SC that involves discovering who has been cheating on the bill at a restaurant; other than that, it is identical to Experiment 6-A. The situation -- one person contributing less than her share when a group of friends eats out -- is probably familiar to college students.

This restaurant problem involves a simultaneous exchange: SC effects on the Wason selection task can be somewhat weaker for simultaneous, face-to-face exchanges, as they imply the possibility of intercontingent behavior that would prevent
cheating (see Chapter 5 and the discussion in Experiment 3). Unfortunately, I was not thinking about this when I created the problem. Nonetheless, it should elicit a social contract effect.

Subjects.

Twenty-four undergraduates from Harvard University participated in Experiment 6-B (17 females, 6 males, no data on sex of one subject); they were volunteers from an introductory science course.

Materials and Procedure.

The restaurant problem is shown in Figure 6.11. The other materials and the procedure were identical to those described for Experiment 6-A.

Results.

The results are pictured in Table 6.13.

Table 6.13  Experiment 6-A: Percent of subjects choosing 'P & not-Q' for each problem (n=23)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>AP:</td>
<td>12%</td>
</tr>
<tr>
<td>F-D:</td>
<td>29%</td>
</tr>
<tr>
<td>F-STD-SC:</td>
<td>62%</td>
</tr>
</tbody>
</table>

Does a familiar standard social contract elicit the predicted SC response, 'P & not-Q'?

Percentage 'P & not-Q' responses:

<table>
<thead>
<tr>
<th>Social Contract Prediction:</th>
<th>Availability Prediction:</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-STD-SC v. AP</td>
<td>F-STD-SC &gt; AP</td>
</tr>
<tr>
<td>high</td>
<td>high</td>
</tr>
<tr>
<td>62%</td>
<td>12%</td>
</tr>
</tbody>
</table>

249
Every Wednesday night you go out to dinner with the same group of four friends from work. You just got your first credit card, and you are trying to build up a good credit rating. So every time the five of you go out, you pay the check with your credit card, and they reimburse you on the spot with cash.

The problem is, the last few times you all have gone out, you've ended up paying more than your share of the bill. One of your friends has been consistently cheating you. This hurts your feelings, because you are the least well off of your friends -- you feel taken advantage of. So this Wednesday night you decide to figure out who is cheating you.

At the restaurant, some of your friends order hamburgers, and others splurge on lobster. When the check comes you announce the following rule:

"If you ordered lobster, then you must put in $20."

The cards below have information about your four friends. Each card represents one person. One side of a card tells what a person ordered and the other side of the card tells how much money that person kicked in.

Indicate only those card(s) you definitely need to turn over to see if any of your friends have broken the rule you announced.

A. ..................  
: $10  
: ..................  

B. ..................  
: $20  
: ..................  

C. ..................  
: lobster  
: ..................  

D. ..................  
: hamburger  
: ..................  


The F-STD-SC elicited the predicted social contract effect (62% v. 12%: \( F = 23.00, p << .001, r = .71 \)). The problem's simultaneity does not appear to have diminished the size of the social contract effect -- .71 for the restaurant problem, .69 for the DAP.

It is difficult to say whether or not this result is also consistent with availability theory. "Although the situation should qualify as familiar, so is eating while drinking, yet many theorists argue (post-hoc) that the food problem should not elicit a content effect (see Chapter 3). Whether AV predicts such a strong content effect for the restaurant problem depends on whether the particular availability theorist believes being shorted at a restaurant is a sufficiently common experience.

This result is important precisely because the content effect on the Wason selection task has been so elusive and unpredictable. The restaurant problem is a brand new problem, never tried before. It was designed according to the specifications of social contract theory and it elicited the predicted social contract effect. It is a new addition to the arsenal of successful social contract rules.

Are there more SC responses to a familiar standard social contract problem than falsifying responses to a familiar descriptive problem?

Percentage 'P & not-Q' responses:

<table>
<thead>
<tr>
<th>Social Contract Prediction:</th>
<th>Availability Prediction:</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-STD-SC v. F-D</td>
<td>F-STD-SC &gt; F-D</td>
</tr>
<tr>
<td>high</td>
<td>low</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>F-STD-SC</th>
<th>F-D</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>low</td>
</tr>
</tbody>
</table>

62% > 29%

250
According to social contract theory, the restaurant problem should elicit more 'P & not-Q' responses than the F-D transportation problem; this prediction was borne out (62% v. 29%: \( F = 8.36, p < .01, r = .52 \)). Whether this result is also predicted by Griggs & Cox's memory-cueing depends on whether one predicts that more subjects have been shorted by a friend than have gone to Boston or Arlington by the form of transportation not mentioned in the rule -- I have no intuitions on this score. However, as it is likely that most subjects have had more experiences with their friends paying their fare share than not, this result contradicts Pollard's differential availability theory.

**Availability assessed: Does availability have any effect at all on familiar problems?**

<table>
<thead>
<tr>
<th>Effect:</th>
<th>No effect:</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-D v. AP</td>
<td>F-D &gt; AP</td>
</tr>
<tr>
<td>F-D = AP</td>
<td>29% &gt; 12%</td>
</tr>
</tbody>
</table>

Availability appears to have produced a small but significant content effect for the F-D (29% v. 12%: \( F = 4.60, p < .05, r = .41 \)). However, the social contract effect is larger, just as it has been in all previous experiments.

Taking Experiments 6-A and 6-B together and using the AP as a baseline, the size of the social contract effect is \( r = .70 \), and the size of the availability effect is \( r = .46 \). In other words, familiar social contract problems produce a social contract effect that is 1.52 times as big as the availability effect. This replicates the 50% advantage that social contract
status gave to **unfamiliar** problems in Experiments 1 through 4.

Experiments 6-A and 6-B versus Experiments 1-4: Familiar versus Unfamiliar Social Contract Problems

Social contract status has a profound effect on the percentage of SC answers elicited by a rule, but familiarity does not appear to further enhance this effect. The percent of SC responses to the unfamiliar SC problems in Experiments 1 through 4 does not differ significantly from those for the familiar SC problems of 6-A and 6-B; in fact, the unfamiliar problems elicited slightly more SC responses than the familiar ones (U-SC = 72%, F-SC = 66%; Z = 0.72, n.s.). Moreover, the effect sizes for the U-SC v. AP comparison are comparable to those for the F-SC v. AP comparison (U-SC v. AP, Experiments 1-4: .61, .69, .65, sizes for the U-STD-SC v. F-D and F-STD-SC v. F-D comparisons are also comparable (U-STD-SC v. F-D, Experiments 1 and 3: .43, .47; F-STD-SC v. F-D, Experiments 6-A and 6-B: .47, .52). Either performance is so high for unfamiliar SC problems that there is a ceiling effect, or SC algorithms operate just as smoothly in unfamiliar situations as they do in familiar ones.

Experiment 6-C

The results so far have shown that the social contract effect is so large that the percentage of SC responses usually outstrips the percentage of falsifying responses to familiar descriptive problems. The transportation problem was the F-D problem used in all the previous experiments, because this was the non-SC problem that had been most successful in eliciting
content effects in the existing literature. To show that there is nothing special about the transportation problem, in this section performance on the DAP and restaurant problems -- the two F-STD-SC problems -- is compared to performance on an array of other familiar descriptive problems. The large battery of tests described in Experiment 5 allowed this comparison.

Subjects.

In addition to the 47 subjects who solved the DAP and restaurant problems of 6-A and 6-B, 115 Harvard undergraduates participated in Experiment 6-C (71 females, 42 males, no data on sex of two subjects); they were volunteers from an introductory science course.

Materials and Procedure.

All subjects solved five problems, but the only problems relevant to this analysis are those that fell in the third position, the same position as the F-STD-SC problems of 6-A and 6-B. The F-D problem for one group (n=23) was a rule relating a person's appearance in a particular kind of attire to his age; for another group (n=49) the F-D rule related arthritis to age -- these problems are shown in Figure 6.12. A third group (n=43) had one of the two food problems described in Experiment 5 as an F-D. The first two groups had an AP and transportation problem in the first two positions; in the third group, those positions were occupied by the other food problem and the semi-SC problem (Figure 6.9). Unlike the SC v. F-D comparisons of the previous experiments, this was a between-groups design, so the tests are not as powerful.
You are a physician interested in disease in other cultures. You are studying the Kaluame people, a Polynesian culture found only on Maku Island in the Pacific. In medical school, you learned the following:

"If a person has arthritis, then he must be over 20 years old."

The cards below have information about four Kaluame people in a hospital. Each card represents one person. One side of a card tells what disease a person has and the other side of the card tells that person's age.

Indicate only those card(s) you definitely need to turn over to see if any of these Kaluame people violate this rule.

A. 15 years old
B. 58 years old
C. arthritis
D. tonsilitis

You are a fashion designer interested in the dress of other cultures. You are studying the dress habits of the Kaluame people, a Polynesian culture found only on Maku Island in the Pacific. In design school, you studied various sorts of fabrics and learned the following about floral prints:

"If a person looks fat in a floral print shirt, then he must be over 20 years old."

The cards below have information about four Kaluame people sitting around a campfire. All four are wearing floral print shirts. Each card represents one person. One side of a card tells how heavy a person looks and the other side of the card tells that person's age.

Indicate only those card(s) you definitely need to turn over to see if any of these Kaluame people violate this rule.

A. 15 years old
B. 35 years old
C. looks fat
D. looks thin
Results.

Table 6.14 displays the results. Both social contract problems elicited more 'P & not-Q' responses than each of the familiar descriptive problems. In four out of six tests this difference was significant. In two tests, those using the arthritis problem as an F-D, the difference is in the right direction but not significant. Note, however, that the percent of subjects who falsified on the arthritis problem is not outside the range found for the transportation problem with this subject population (arthritis: 57%; transportation: 58% (Exp 4), 50% (Exp 2), 57% (Exp 6-A)).

<table>
<thead>
<tr>
<th>Problem</th>
<th>F-D</th>
<th>versus</th>
<th>F-STD-SC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drinking-Age</td>
<td>78% (n=23)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Appearance:</td>
<td>41% (n=49)</td>
<td>Z = 2.97, p &lt; .0025, phi = .35</td>
<td></td>
</tr>
<tr>
<td>Food:</td>
<td>35% (n=43)</td>
<td>Z = 3.97, p &lt; .00005, phi = .49</td>
<td></td>
</tr>
<tr>
<td>Arthritis:</td>
<td>57% (n=23)</td>
<td>Z = 1.57, n.s.</td>
<td></td>
</tr>
<tr>
<td>Restaurant Problem</td>
<td>62% (n=24)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Appearance:</td>
<td>41% (n=49)</td>
<td>Z = 1.74, p &lt; .05, phi = .20</td>
<td></td>
</tr>
<tr>
<td>Food:</td>
<td>35% (n=43)</td>
<td>Z = 2.18, p &lt; .025, phi = .27</td>
<td></td>
</tr>
<tr>
<td>Arthritis:</td>
<td>57% (n=23)</td>
<td>Z = 0.42, n.s.</td>
<td></td>
</tr>
</tbody>
</table>
Summary, Experiment 6.

Two familiar standard social contracts, including one that had never been tested before, elicited the predicted social contract effects. Not only did they elicit more "falsifying" responses than abstract problems, but they also elicited more than did a wide variety of familiar-descriptive problems.