Perspective Shifts on the Selection Task: Reasoning or Relevance?

Jonathan St.B. T. Evans and John Clibbens, University of Plymouth, UK

Commentary on “Pragmatic Reasoning With a Point of View” by Keith J. Holyoak and Patricia W. Cheng

The single most investigated problem in the psychology of reasoning is the Wason selection task (see Evans, Newstead, & Byrne, 1993, Chapter 4 for a detailed review). Nevertheless, there is room for considerable doubt as to whether subjects’ choices on this task actually reflect a process of reasoning at all, as we shall argue here (see also Evans, in press). Selection task studies can broadly be divided into those using abstract or arbitrary problem content and those using thematic content. At one time it appeared that subjects were very poor at solving the abstract task but quite good on the thematic task, but we now know this to be a double simplification. First, there are manipulations that cause subjects to choose largely correctly on the abstract task (see Platt & Griggs, 1993, 1995), and second, the pragmatic influences introduced by realistic content can sometimes trigger non-normative selection patterns.

The Wason selection task is described by Holyoak and Cheng (this issue) and will not be repeated in detail here. Suffice it to say that subjects are asked to test a rule of the form if \( p \) then \( q \) by deciding whether or not to turn over each of four cards displaying the values \( p, not-p, q, \) and \( not-q \). If the facing side of the card provides information about the antecedent of the rule, then the hidden side will provide information about the consequent, and vice versa. The logically correct choices are the cards \( p \) and \( not-q \) which could provide evidence that a card violates the rule. This choice is rarely observed on the abstract task where arbitrary relations such as “If there is a vowel on one side of the card then there is a consonant on the other side of the card” are used. It has been known for many years, however, that certain kinds of thematic phrasings of the rule, together with an appropriate context, will produce this logically correct pattern.

One of the most important new findings in this literature in recent years is that the pattern of selections observed can be radically altered by instructions that manipulate the perspective of the subject in terms of which character they identify with in a scenario. For example, given the statement:

If you tidy your room you may go out to play
together with an appropriate scenario involving a working mother and untidy child, the subject cued with the perspective of the child will tend to select the $p$ (tidied room) and $not-q$ (was not allowed out to play) cards, whereas when cued with the perspective of the mother, the subject will choose instead the $not-p$ (did not tidy room) and $q$ (went out to play) cards (Manktelow & Over, 1992). Holyoak and Cheng (this issue) discuss such findings in some detail and provide an account in terms of their pragmatic reasoning schema (PRS) theory (Cheng & Holyoak, 1985). This theory proposes that people have access to abstract, but pragmatically sensitive, schemas, which contain embedded production rules. The schema account of the perspective shift finding, elaborated in detail by Holyoak and Cheng (this issue) is essentially that the former pattern arises due to elicitation of a permission schema for the child, and the latter due to elicitation of an obligation schema for the mother, effectively reinterpreting the rule as "If you go out to play then you must tidy your room first". We provide detailed discussion of this example later.

We find the schema theory ingenious, to fit quite well with a range of findings on the thematic selection task, and to be more convincing in general than rival accounts in terms of social contract theory (Cosmides, 1989). However, we do feel obliged to criticise the theory on grounds of parsimony. First, the theory is incomplete as an account of the selection task in that it has nothing to say about the causes of card choices on the abstract form of the task. Thus separate and additional proposals are needed to account for the abstract task data. Second, Holyoak and Cheng's account of the thematic selection task appears to be over-elaborate and assumes more than is necessary to account for the data. We agree with them that choices on the thematic selection task are influenced by pragmatic factors and that this causes problems for theories of reasoning involving content-independent mechanisms of reasoning. However, there are in general many kinds of pragmatic influence that could not be described in terms of schemas.

The central requirement of a theory of pragmatics is to account for the way in which the interpretation of a communicative act is influenced by the specific context in which it is processed. It is clear that such a theory is necessary in view of the fact that the same utterance can have an unlimited number of distinct interpretations, depending on such features of the communicative context as the identity of the speaker and the audience, their shared assumptions, the physical surroundings, and the immediately preceding discourse (Sperber & Wilson, 1986). Pragmatics thus has to account for the way in which information derived from the act of communication itself (including both its content and the signalled intention of communicating) is combined with information derived from other sources, including sense perception and information held in memory (both immediate and long-term). Where previous experience has resulted in the development of knowledge schemas in particular domains, and where this fact is known to both speaker and audience, such schemas form part of the cognitive
context in which an utterance is interpreted. However, a complete pragmatic theory, as indicated earlier, will need to provide an integrated account of the use of schemas and all other information sources. In the case of the selection task, we have evidence of pragmatic effects that cannot be explained by invoking schema-based production rules, in particular the tendency for content and context to inhibit the “matching bias” effect, discussed later.

THE HEURISTIC-ANALYTIC THEORY AND EVIDENCE FOR RELEVANCE EFFECTS

The heuristic-analytic theory of Evans (1984, 1989) proposes that we reason with highly selected representations of the problem information together with relevant information retrieved from memory. This representation is formed by preconscious heuristic processes which are rapid and parallel in nature. Recently, Evans has stressed the notion of perceived relevance and suggested connections with implicit cognitive systems of a possibly connectionist nature (see, for example, Evans, 1995). The analytic reasoning applied to selected information was originally unspecified, although connections have recently been explored with the mental model theory (Johnson-Laird & Byrne, 1991). The basic idea is that relevance is equivalent to explicit representation in a mental model (see Evans, 1991, 1993, 1995). However, in the case of the Wason selection task—both abstract and thematic—we are yet to be convinced that any process of reasoning affects the selections made. It appears that subjects simply choose the cards that are preconsciously cued as relevant. Such conscious reasoning as does occur serves only to rationalise the choices already made (see Evans, 1995, in press).

Let us briefly summarise the evidence for this claim with the abstract selection task before focusing on the thematic task. The first problem is that logical performance on the abstract task is disproportionately low compared with other conditional reasoning tasks. Selection of the not-q card is typically less than 10% with the abstract affirmative conditional. By contrast, Evans et al. (1993, Table 2.4) show that the Modus Tollens (given not-q, infer not-p) inference with similar conditional rules is made by between 41% and 81% (median 63%) in studies reported in the literature. Another logically equivalent measure is the identification of p and not-q as falsifying cases of the conditional in the truth table task which the great majority of subjects succeed in identifying with affirmative rules (see, for example, Evans & Newstead, 1977). The fact that subjects understand the logical principles underlying the choice of the not-q card, but nevertheless ignore it, makes sense in the two stage heuristic-analytic theory. If a card is not seen as relevant at the heuristic stage then analytic reasoning is never applied to it. The reason that the card is not seen as relevant on the abstract task is due to “matching bias”. Matching bias is a tendency to choose cards that are explicitly named in the rule, which means, for example,
that the form if $p$ then not-$q$ is associated with most correct selections—the $p$ and $q$ cards in this case (see Evans, 1989, and in press, for detailed discussion of matching and relevance). The evidence shows that when the relevance of the false consequent is enhanced by manipulation of negative components in the rules, then its rate of selection increases dramatically.

A possible view is that subjects make relevance judgements on the abstract selection task, but somehow start to use reasoning on the thematic selection task—such as by application of pragmatic reasoning schemas. Certainly, subjects are affected quite strikingly by the presence of thematic content. For example, Griggs and Cox (1983) and Evans (1995) have shown that introduction of unhelpful pragmatic content, which does not facilitate correct selections, nevertheless entirely removes the phenomenon of matching bias that is so reliably observed with abstract content. However, this seems to us to be a clear example of a pragmatic influence that cannot be explained by a schema theory of the kind Holyoak and Cheng propose. Moreover, some recent evidence suggests that subjects do not switch to a different process of reasoning with the thematic task but rather that relevance effects of a different kind prevail.

Evans (in press) reports some studies using a new methodology. A computerised version of the Wason selection task was used, in which subjects select cards by moving a mouse pointer on to the card and then clicking the mouse. However, subjects were also instructed to point at any card they were thinking of selecting, but not to click until they were sure. The program logged cumulative pointing time, described as card inspection times, prior to any card being clicked on. Now, if relevance is the sole determinant of card choices, then we would expect subjects to select the cards to which they attend, i.e. inspection times should be much larger for selected than non-selected cards. On the other hand, if subjects are using analytic reasoning, considering each card and deciding whether to select or reject it, then no such correlation should exist.

The first experiment reported by Evans (in press) used four selection tasks: two abstract and two thematic. One of the abstract rules used a negative consequent which is known to facilitate false consequent selections due to matching bias. One thematic rule was known to be a weak and the other a strong facilitator of correct choices. The card selection patterns were as expected, but the main interest lies with the inspection-time findings. The results were massively as predicted by the relevance theory. First, the cards that were most often selected were inspected for much longer periods. A more impressive result, however, was obtained by comparing inspection times within each card between the subjects who selected the card and those who rejected it. The subjects choosing any given card spent far longer inspecting it (see Fig. 1).

Although the pattern was observed throughout, it was least marked on the strong facilitator rule. This raised a question as to whether some different process might occur with thematic rules. A second study was therefore devised in which four new rules, all thematic, were involved. Of particular relevance
FIG. 1. Mean inspection times for selected and non-selected cards on each rule in the study by Evans (in press) Experiment 1 (n = 30).

Here is the fact that two of the rules were taken from recent studies of perspective shifts, one designed to produce the p & not-q pattern and the other the not-p & q pattern. Again the card selection patterns were broadly in line with expectations, showing that the pointing method did not change the selection process. In this study, once again, a marked correlation was observed between selection rates and inspection times, whether examined between or within cards (see Fig. 2).

Let us consider the impact of these findings on theories of the thematic selection task. What has been shown is (a) that subjects attend almost exclusively to the cards they choose on the abstract selection task, and (b) that the same applies to the thematic selection task, so that whatever factors cause a shift in responding (including perspective) also cause a corresponding shift in attention. The inspection-time findings are supported by independent evidence from verbal protocol analyses which also suggest that subjects think only about the cards they select (Beattie & Baron, 1988; Evans, 1995, Experiment 6). Thus it appears that the subject is from the outset thinking about the cards she or he will choose. During this time it appears from the evidence of protocol analysis
that the subject is thinking about the consequences of selecting the card, but that this analysis effectively serves only to rationalise a choice already made. Non-selected cards are not eliminated by reasoning, however—it appears that they are simply not thought about at all.

PRAGMATIC RELEVANCE VERSUS PRAGMATIC SCHEMAS AS AN ACCOUNT OF PERSPECTIVE SHIFTS

What implications does this evidence carry for the interpretation of perspective shifts in the thematic selection task? We agree with Holyoak and Cheng that a pragmatic, non-logical mechanism is responsible. However, the proposal of full-blooded schemas containing embedded production rules seems to us to be over-elaborate, and unnecessary to account for the data. As discussed earlier, the mechanism is also too specific to form the basis of a general theory of pragmatic influences that are so pervasive in discourse comprehension and reasoning. With regard to the specific evidence on the selection task, there is no direct, process-tracing evidence that people reason by use of production rules. Indeed, we will argue that such evidence of this type that does exist—referred to earlier as
providing evidence of relevance effects—is difficult to reconcile with the
proposals of the PRS theory.

Let us look at the proposal of production rules in particular. With a
permission schema the proposed rules are:

P1 If the action is to be taken then the precondition must be satisfied
P2 If the action is not to be taken then the precondition need not be satisfied
P3 If the precondition is satisfied, then the action may be taken
P4 If the precondition is not satisfied, then the action must not be taken

For example with the Manktelow and Over (1992) rule:

If you tidy your room you may go out to play

with the child perspective this schema is instantiated as follows:

precondition = tidy room
action = go out to play

Holyoak and Cheng propose that the instantiated rules guide reasoning
and thus mediate the card choices. For example, as only P1 and P4 contain the modal
must, then only the cards to which they apply (p and not-q) will need to be
selected. Hence subjects will correctly choose the cards "tidied room" and "did
not go out to play". The child would rightly feel cheated if he tidied his room
and then was not let out.

The explanation of the perspective shift is that when the mother’s perspective
is given, the obligation schema is evoked instead which in its general form has
the rules:

O1 If the precondition is satisfied then the action must be taken
O2 If the precondition is not satisfied then the action need not be taken
O3 If the action is to be taken then the precondition may have been satisfied
O4 If the action is not to be taken then the precondition must not have been
satisfied.

This is instantiated with values:

precondition = go out to play
action = tidy room

and the rule is reinterpreted as "If you go out to play you must tidy up your
room first." Now the rules with must are O1 and O4 which apply to the not-p
("did not tidy room") and q ("went out to play") cards, and these are the ones
subjects tend to select with this perspective.

How does this schema-based reasoning with instantiated production rules
work? We can only presume the following: the subject takes each of the four
rules in turn and attempts to find a rule to apply to it (alternatively the subject regards each card in turn and looks for a rule that applies to it). The point about production rules is that you match their conditions and then see what actions are required. When a rule matches, say P3 to the q card “went out to play” you then look at the consequent of the rule. In this case it says that the action may be taken, but, as it need not be, you decide that it is not necessary to turn over the card—and so on. The evidence for this theory is entirely in the form of its predictions about the response to different types of problems such as that offered in the new experiment of Holyoak and Cheng (this issue). Surely it must be possible to test more directly for the procedures specified by use of process-tracing methods such as latency measures and verbal protocol analysis. Herein lies the problem.

It seems to us that PRS theory makes sense only if we assume that card selections reflect a process of reasoning that involves considering all four cards and making a decision to turn or not turn each depending on the rule. However, this is precisely what the evidence from card inspection times suggests does not happen (Evans, in press). From the outset subjects focus on the cards to be selected and do not appear to spend any significant time thinking about the cards they will reject. If subjects were applying production rules we might, if anything, expect the opposite trend. On selected cards the clear affirmative actions might be quickly appreciated whereas realising that “may” meant you need not turn the card would take longer to work out. In any event, we are aware of no proposals in the current formulation of PRS theory that predict rapid elimination of non-selected cards accompanied by significant scrutiny of selected cards.

It seems indeed that rapid preconscious processes cue the pragmatic relevance of the cards, and it must be at this level that the perspective shift is occurring. At this time we do not have a theory of the mechanism by which this occurs, although we presume that it may have a network or connectionist nature. A problem for our account, however, is to explain why subjects spend some considerable time thinking about cards that they are apparently predetermined to choose. We assume that they are applying analytic processes to justify or rationalise a choice already made and protocol analyses are consistent with this. However, we should note a somewhat different interpretation in a recent and alternatively devised relevance theory of the selection task by Sperber, Cara and Girotto (in press) which builds on the successful relevance theory of discourse by Sperber and Wilson (1986). They reject the heuristic-analytic distinction and argue that relevance guides all thinking. They suppose in general that subjects attempt to maximise relevance while minimising cognitive effort. On the selection task subjects rapidly focus on inferable consequences that appear relevant in the context and spend time checking that the card in focus is indeed a potential instance of the inferred features.
In summary, our challenge to Holyoak and Cheng is twofold: first can they provide us with any direct evidence of a process of reasoning which looks like systematic application of production rules, and second, how can they explain the evidence for relevance effects, and in particular selective attention to the cards to be chosen?

REFERENCES


The Taming of Content:  
Some Thoughts About Domains and Modules  
Gerd Gigerenzer, Max Planck Institute for Psychological Research, Munich

Commentary on “Pragmatic Reasoning With a Point of View” by Keith J. Holyoak and Patricia W. Cheng

Peter Wason invented the Selection Task in 1966. Thirty years and many, many experiments later, two results are evident for me. First, the view that sound reasoning can be reduced to propositional logic (or first order logic) is myopic. Human thought operates in more dimensions than entailment and contradiction (Strawson, 1952). We need to work out how the mind infers the meaning of a conditional statement from the content of the Ps and Qs and the social context in which the statement is uttered, rather than exclaiming “Cognitive illusion! Hurrah! Error!” whenever human reasoning cannot be reduced to propositional logic. Second, the hope that these “errors” or their flip-side, the “facilitation” of logical reasoning, would be the royal road to discovering the laws of human reasoning did not materialise. This hope was fueled by the (misleading) analogy with visual illusions (Gigerenzer, 1991). What were seen as “errors” were attributed to deeper cognitive deficits such as confirmation biases, to crude heuristics such as availability, or simply to “the subjects’ incorrigible conviction that they are right when they are, in fact, wrong” (Wason, 1983, p.356). Unfortunately, this programme of research has brought little progress on the theoretical front.

In the last decade, a few adventurous researchers freed themselves from the straitjacket of propositional logic and looked at dimensions of reasoning beyond entailment and contradiction. The content rather than the logical structure of the conditional statement moved into the foreground. As early as 1972, Wason and Johnson-Laird pointed out (p.245) that, contrary to their expectation, “content is crucial” to reasoning and that “any general theory of human reasoning must include an important semantic component”. But neither they nor the others who studied human reasoning at that time found a way to include semantics in a theory of reasoning. Instead, content remained but a decorative element in reasoning problems—which either might “confuse” subjects or “facilitate” their logical reasoning. In the 1980s, Patricia Cheng (Cheng & Holyoak, 1985; Cheng, Holyoak, Nisbett, & Oliver, 1986) and Leda Cosmides (1989; Cosmides & Tooby, 1989, 1992) dared to introduce the content of the Ps and Qs into their theories. Cheng and her collaborators postulated a set of pragmatic reasoning schemas, such as permission and obligation schemas. Cosmides and Tooby postulated Darwinian algorithms, with social contracts and threats as examples. Content, the terra incognita where no established researcher on reasoning dared or cared to venture, became a legitimate topic of study. Cheng and her
colleagues made a significant first proposal: permission and obligation schemas, each defined by four production rules. This was an important move away from propositional logic, but the production rules still resembled the four rules of the truth table, with Ps and Qs replaced by "actions" and "preconditions". Cosmides and Tooby made a bolder and theoretically richer leap, connecting information search with pragmatic goals such as cheater detection, with cost-benefit analyses, and with the broader evolutionary theory of reciprocal altruism. These researchers set the stage for the discovery of a genuinely pragmatic dimension of reasoning unknown to logic: a person's perspective (Gigerenzer & Hug, 1992; Light, Girotto, & Legrenzi, 1990; Manktelow & Over, 1991).

The semantic and pragmatic approach to reasoning, however, is still in its infancy: researchers often continue to focus on the "logically correct answer", to report their results in terms of "logical facilitation", or to criticise a competing pragmatic theory by saying that its semantic conditionals are simply logical biconditionals (e.g. Politzer & Nguyen-Xuan, 1992). Too much energy has been spent in the debate on asserting who is wrong, and too little on actually developing the theories. But whether a theory is "right" or "wrong" cannot be decided unless it is well specified in the first place. For instance, one group now needs to work out exactly what the Darwinian algorithms are, and another group needs to put pragmatic contents—a person's point of view, their goals, their cost-benefit computations among others—explicitly into the production rules that define permission and obligation schemes. Unless this theoretical specification is accomplished, one can only make plausible arguments that one's theory could be consistent with the perspective effect. For instance, Holyoak and Cheng's account, in this issue, of the perspective effect along the lines of Politzer and Nguyen-Xuan (1992) is plausible, but the pragmatic reasoning schema theory is not richly specified enough to allow for a deduction of this effect. The challenge is not to slip content and perspective effects in through the back door, but to devise richer theories that allow semantic and pragmatic dimensions in through the front door.

So, how can we make progress in the study of reasoning? How can we improve on the first ideas suggested by Cheng and Holyoak, and Cosmides and Tooby? I will not join the battle over who is right. Instead, I will try to work out some of the unresolved theoretical issues and suggest how these could be solved.

BEYOND "LOGICAL FACILITATION": WHERE TO GO FROM HERE?

First we need to get clear what we were studying when we discovered perspective effects. What I saw as the issue was reasoning about conditional statements in natural languages—not deductive or logical reasoning. If one wanted to study logical reasoning, one would have to instruct subjects: "Forget the content and just treat this problem as an exercise in propositional logic." But
this is not what has been done in the last 30 years. Given the way this research has been conducted, the question to ask is: How does a mind understand a natural language conditional, and what cognitive processes and behaviours are controlled by that understanding?

I start with two assumptions. First, understanding a conditional statement is impossible without analysing its semantic content and the pragmatic context in which the statement is uttered. Or, as Fillenbaum (1977) put it, “Mind your Ps and Qs”. The second assumption is that minds reason about conditional statements neither at the level of the particular nor at the abstract level of propositional logic, but at an intermediate level. This intermediate level of abstraction retains enough of semantics and pragmatics but at the same time discards enough particulars to allow for both rich and fast cognitive processes and behaviour. What I call domain-specific reasoning is carried out at this intermediate level of abstraction.

Here is an example. Suppose you say to me: “If you touch me, I’ll kill you”. I infer that if I don’t touch you, you won’t kill me. Thus, from “if P then Q”, I infer “if not-P then not-Q”. This is logically invalid, but it may save my life. Threats, warnings, social contracts, bribes, and many other classes of conditional statement do not follow propositional logic. They have their own “natural logics”. Note the plural. The term “natural logics” refers to how people reason with or in natural language (Fillenbaum, 1976, 1977). For instance, linguists study paraphrases and “invited inferences” (as opposed to truth-preserving logical inferences). A conditional threat can be paraphrased with “P and Q” as well as with “not-P or Q”. You could as well have said “Touch me and I kill you” or “Don’t touch me or I kill you”. These paraphrases bypass propositional logic in that the English “or” and “and” do not map onto the logical OR and AND. Now take a social contract: “If you mow my lawn, I’ll give you $50”. The social contract can be paraphrased by “P and Q”, but not by “not-P or Q”, as was the case with the threat. “Don’t mow my lawn or I’ll give you $50” is not an acceptable paraphrase. These examples illustrate that invited inferences and paraphrases systematically vary between domains, such as conditional threats, warnings, bribes, and social contracts (Fillenbaum, 1976, 1986).

It is not just paraphrases that are specific to domains. What would a research programme that studied domain-specific reasoning look like?

MENTAL MODULES: A RESEARCH PROGRAMME

Here I propose a two-step programme for studying reasoning about conditional statements (and beyond). The first step is to model the mapping algorithm that recognises a particular statement as falling into the range of a domain; the second step is to work out the subsequent activity of the module: how it deploys attention, makes inferences, and executes other processes.
1. Search for the Mapping Algorithm

Here is the first problem to be solved: There is a potentially unlimited number of conditional statements $C_i$ (in English or any other language) and a smaller number of domains $D_j$. What is the algorithm that people use to map $C_i$ onto $D_j$? Note that we have simplified the problem by only considering linguistic information. Once we have solved this problem, we may go on to tackle the issue of how nonverbal information—such as facial cues that signal a threat (Ekman, 1982) or motion cues signalling a self-propelled actor (Premack, 1990)—is used in parallel with verbal information.

Our problem is like solving one equation with two unknowns. We know the $C_i$ but not the mapping algorithm or the domains. One strategy is to assume a value for one of the two unknowns and try to solve for the other. Let us assume a few domains $D_j$ and solve for the algorithm. I suggest proceeding in this way because we have several converging proposals for candidate domains, but I do not know of any specified proposals for mapping algorithms.

What are plausible candidates for domains? With respect to conditional statements, linguists have distinguished inducements, such as conditional promises and bribes, and deterrents, such as conditional threats and warnings (Fillenbaum, 1977, 1986). Evolutionary psychologists focusing on important adaptive problems in the history of humans have proposed strikingly similar candidates; social contracts and social exchange, precautions, and threats (Cosmides & Tooby, 1992). Cognitive psychologists have postulated permissions and obligations (Cheng & Holyoak, 1985) as candidate domains, which include social exchange but also other forms of social regulations and deontic reasoning (Over & Manktelow, 1993). No complete list of domains is known, nor is one needed; it is sufficient to start analysing a few.

If we provisionally commit ourselves to a group of domains then we can ask: What is the mapping algorithm? That is, how does a mind infer that a particular semantic conditional is a threat and not a social contract or something else? Here is an idea. I conjecture that only a small number of dimensions are needed to make fairly reliable classifications of individual English conditionals into domains. For instance, all statements of the kind "if $P$ then $Q$" in which you issue a threat to me seem to have in common that (i) $P$ is an action of mine, (ii) $Q$ is an action of yours, (iii) both actions are possible in the future, (iv) $P$ is first and $Q$ second, and (v) $Q$ has negative consequences for me. This is most likely not the fullest possible characterisation of a conditional threat, but with a few hundred examples of conditional threats we could find the dimensions I have overlooked. Now consider the social contract: "If you mow my lawn, I'll give you $50". The values on the dimensions (i), (ii), and (iii) are as characteristic of a social contract as of a threat and would not distinguish between them. The strict temporal order in (iv) is not characteristic of a social contract because it
may or may not hold: you may pay me in advance or after the work. The
decisive difference is in feature (v). If your future action Q is a benefit for me,
then I can infer that the statement is a social contract rather than a conditional
threat. As a last example, imagine you issue a warning to me: “If you give him
your finger, he’ll take your hand”. Warnings share with threats the values (i),
(iii), (iv), and (v), but differ on (ii). Q is typically not an action of yourself but
rather an action of a third party (as in the last example) or a negative
consequence for me (“If you ski in this snow, you may break a leg”). These are
some candidate dimensions from which a richer account of pragmatic reasoning
schemas might be constructed.

If this view is correct, the algorithm for mapping conditional statements C_i
onto domains D_j could be modelled analogously to a key, where the teeth are
the values on the dimensions. Each threat, each social contract, and so on has
the same characteristic profile, and the mapping algorithm uses the dimensions
on which they differ to discriminate between them. Of course, there will be
some residual uncertainty, which may be further reduced by non-verbal
information. The challenge is to examine hundreds of conditional statements to
improve on the preliminary sketch I have given here. Working models of the
mapping algorithm can subsequently be tested in the form of computer programs
together with a knowledge component, e.g. for defining positive and negative
consequences. My intuition is that the mapping algorithm uses no more than a
half-dozen content dimensions of the Ps and Qs to infer a domain.

2. The Control Structure of Cognitive Modules

Once the mapping algorithm has inferred a domain (e.g. “threat!”), a “cognitive
module” that controls the processes necessary for coping with this type of
situation is activated. The advantages of cognitive modules that respond to the
social and physical environment at the intermediate level of a domain (rather
than at the level of the particular or of logical abstraction) are the following.
First, attention can be focused. For instance, if the situation is identified as a
social contract, then attention can be directed to information that could reveal
that one is being cheated (Cosmides, 1989; Gigerenzer & Hug, 1992). If,
however, a threat is identified, information that could reveal being cheated needs
no attention, but information that can reveal being bluffned or double-crossed
does. Thus, modules can help to reduce one fundamental problem of induction:
What to observe? (Popper, 1959). Second, inferences can be made more
efficiently. Modules that contain semantic and pragmatic structures enable the
organism to react quickly and to reduce the problem of computational explosion
when making inferences. For instance, John Garcia (e.g. Garcia & Koelling,
1966; Garcia y Robertson & Garcia, 1985) showed that when the taste of
flavoured water was paired with experimentally induced nausea, rats could learn
in just one trial to avoid the flavoured water, even if the nausea occurred two
hours later. In contrast, when the taste of flavoured water was repeatedly paired with electric shock, rats had great difficulty learning to avoid the water. Thus, rats seem to have a built-in specific mechanism for food avoidance that enables them to develop some associations rapidly but not others. Note that such specific mechanisms are often phrased in terms of "constraints" on inference or learning. The more appropriate statement seems to me that these mechanisms enable, not constrain, inference. Semantic relations built into mechanisms can enable what "unconstrained", that is, content-independent, algorithms could not do in the first place because combinatorial explosion might paralyse any system that is truly domain-general (Cosmides & Tooby, 1994). Third, learning could hardly proceed without domain-specific devices—an argument that Chomsky made against Skinner's view of language learning. Garcia's experiments are further examples in support of this argument.

The second problem that needs to be solved, then, is how to model the control structure of a module? Here are some thoughts.

In order to design the structure of a module, we need to define more clearly the notion of a module. (In the literature, almost everything has been suggested as a domain to which some module is attached; see Hirschfeld & Gelman, 1994). Fodor (1983) distinguished "horizontal" from "vertical" faculties. Current curricula and textbooks typically organise cognitive psychology according to the doctrine of horizontal faculty psychology, assuming that the mind is divided up into general-purpose processes: memory, attention, thinking, judgment, perception, volition, and so forth. All memories are in the same place; they may depend on time and rehearsal, but not on the content of the memory. Similarly, in this view all thoughts, judgments, and so forth are of one kind across content domains. For the "vertical" view of faculties, however, there are no such things as memory, attention, thinking, judgment, perception, volition, and so forth. Instead there are domain-specific capacities, each with different mechanisms in which the horizontal faculties are like parts of an engine. The modules we are searching for are domain-specific and are thus vertically, not horizontally, organised.

For Fodor there exist about six input systems, one for each of the five senses and one more for language. He envisions modules as functionally more specific than these six systems, including, for instance, modules designed for colour perception, analysis of shape, recognition of faces, and recognition of voices. Fodor's focus on input systems makes him believe that modules are informationally encapsulated.

Where I depart from Fodor is when he restricts modules to informationally encapsulated input systems, arguing that central cognitive processes such as thinking tend to be domain-general. The mind in Fodor's Modularity of mind (1983) is decidedly anti-modular at its centre. The stronger modularity thesis is that central cognitive processes like thoughts are modular to some important degree, too. Evidence for this stronger thesis stems from studies of cognition and
development (e.g. Barkow, Cosmides, & Tooby, 1992; Goldstein & Weber, in press; Leslie, 1994; Premack & Premack, 1994; Sperber, 1994; Todd & Miller, 1993) and from selective impairments following brain damage (e.g. Caramazzo, Hillis, Leek, & Miozzo, 1994). This descriptive evidence is complemented by the normative argument that content-independent formulations of principles of rational reasoning and decision making, such as consistency, are essentially confused, in the sense that there is no way to determine whether choice is consistent or rational without referring to something external to choice behaviour, such as a person’s motives and values (e.g. Gigerenzer, in press; Sen, 1993).

Furthermore, if we think of modules as hierarchically organised, then modules can activate and inhibit one another and the notion of encapsulation is misleading. Take for instance a sequential, tree-like arrangement of modules and mapping algorithms. Assume you are out in the woods at night, it is windy, and you notice at some distance a large, dark, moving object. The postulated mapping algorithm in your brain would analyse the motion pattern to classify the object as either “self-propelled” (animal or human) or not self-propelled (plant or physical object), to use David Premack’s terms. If the algorithm infers that there is a self-propelled object, a module for unrecognised self-propelled objects may be activated, which initiates physiological, emotional, and behavioral reactions that alert the organism and direct attention to information that can reveal whether it is human or animal. A second, more specialised mapping algorithm may infer from shape and motion information that the object is human. This inference may in turn activate a module lower in the hierarchy that initiates reactions appropriate for unidentified humans (warnings, threats, and the like) and guides attention to information that could reveal whether that human is friend or enemy (or predator or prey, in the case of an animal), and so on. Note that such hierarchically organised, specialised modules can act quickly, as only a few branches of the hierarchical tree are travelled. The organism thereby avoids combinatorial explosion. For instance, if the first mapping algorithm had inferred that the object is not self-propelled, then any information that could reveal whether it is a human or animal, and subsequently, friend or enemy, or predator or prey, would not need to be searched for.

There need not be a one-to-one correspondence between domain and module. The domain of a module may shift over evolutionary time. Sperber (1994) tries to capture this phenomenon with his distinction between the “proper” and the “actual” domain of a module. The proper domain of a module designed for social contracts may once have been exclusively the small hunter-gatherer group in which trust had to be established, forgiveness granted, and repeat-cheaters expelled. The “actual” domain today, in contrast, may range from the task of managing a large company to board games like Diplomacy.
THE TAMING OF CONTENT

In his seminal book, *The taming of chance* (1990), Ian Hacking described how chance was tamed by statistical laws in the sense that chance became the very stuff of the fundamental processes of nature and society. Statistical mechanics, quantum theory, and evolutionary theory epitomise that revolution. Chance was no longer the essence of the lawless and unpredictable. In theories of reasoning, content has played a role similar to the one chance once played in theories of nature: something to ignore or banish. The taming of content by a new class of theories of reasoning can fundamentally change our understanding of the mind, just as the probabilistic revolution once changed our understanding of nature and society (Gigerenzer et al., 1989).

But we are only beginning to build theories that model the role content plays in reasoning: domains and modules mark only a first step. Cognition is still studied by most psychologists in terms of what Fodor calls horizontal faculties: deductive thinking, probabilistic reasoning, problem solving, and so on. And the preferred models are propositional logic, probability theory, or variants thereof that ignore the content and pragmatics of thought. In this article, in contrast, I have sketched a two-step programme for studying modular thought. I have not provided an answer to the questions I proposed to study, but have at least tried to define these questions and to outline possible research strategies. Let the work begin.

ACKNOWLEDGEMENTS

I am grateful to Valerie Chase and Lorraine Daston for many helpful comments on an earlier version of this paper.

REFERENCES


Contextual Factors in Deontic Reasoning

Vittorio Girotto, University of Trieste and CREPCO, Aix-en-Provence

Commentary on “Pragmatic Reasoning With a Point of View” by Keith J. Holyoak and Patricia W. Cheng

The discovery of perspective effects is one of the most interesting findings in the recent literature on the selection task. Several investigators have found that people’s selections can vary as a function of the specific perspective that they are asked to take (cf. Gigerenzer & Hug, 1992; Light, Girotto, & Legrenzi, 1990; Manktelow & Over, 1991; Politzer & Nguyen-Xuan, 1992). Consider a conditional contractual regulation of the form:

If you perform action A1 for me, I’ll perform action A2 for you

in which the two actions increase the positive utility of the two target actors. Given the task of checking for the possible violation of the rule from the hearer’s perspective, people will tend to select the “action A1 performed” (p) and “action A2 not performed” (not-q) cases. In other words, (using the terminology proposed by Holyoak & Cheng, this issue; henceforth H&C), they will tend to select the cases corresponding to the “duty of the hearer towards the speaker fulfilled” and the “right of the hearer against the speaker acquired”. By contrast, if the same rule is checked from the speaker’s point of view, people will

1For an analysis of the entire set of contract proposals, including those expressing negative utilities, cf Legrenzi, Politzer, & Girotto (in press).
tend to select the “action A1 not performed” (i.e. “duty of the hearer towards the speaker not fulfilled”, \( \text{not-p} \)) and “action A2 performed” (i.e. “right of the hearer against the speaker not acquired”, \( q \)) cases. According to pragmatics schemas theory (H&C; Politzer & Nguyen-Xuan, 1992), the two indicated perspectives activate different schemas (obligation and permission, respectively), which determine the selection of the different cases.

The results reported by H&C provide further evidence of perspective effects by showing that the selections elicited by a typical permission rule can also be elicited by a rule that in isolation would be ambiguous between a permission and an obligation, provided that a disambiguating context cues the permission schema. H&C also provide a convincing defence of pragmatic schema theory, which seems to offer a better interpretation of perspective shifts and related phenomena than rival theories of deontic reasoning on the selection task.

However, a specific aspect of perspective effects is not discussed in the H&C paper (nor in most studies on this topic), despite its potential relevance. If one considers the percentages of selection patterns reported in the studies in which perspective phenomena have been investigated by using similar or identical deontic problems (see Table 1), a rather puzzling picture emerges. A qualitative analysis of the available data indicates that selection patterns do not replicate well both in pairs formed by an original study and a replication, and in pairs of independently conducted studies using similar rules. In particular, a striking

<table>
<thead>
<tr>
<th>Rule and Study</th>
<th>Perspective and Predicted Pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Speaker (( \text{not-p} &amp; q ))</td>
</tr>
<tr>
<td><strong>Shop offer</strong></td>
<td></td>
</tr>
<tr>
<td>Manktelow &amp; Over (1991/Expt. 3)</td>
<td>85</td>
</tr>
<tr>
<td>Politzer &amp; Nguyen-Xuan (1992)</td>
<td>17</td>
</tr>
<tr>
<td><strong>Mother promise</strong></td>
<td></td>
</tr>
<tr>
<td>Manktelow &amp; Over (1991/Expt. 1)</td>
<td>68</td>
</tr>
<tr>
<td>Manktelow &amp; Over (1991/Expt. 2)</td>
<td>86</td>
</tr>
<tr>
<td><strong>Day-off rule</strong></td>
<td></td>
</tr>
<tr>
<td>Gigerenzer &amp; Hug (1992)</td>
<td>61</td>
</tr>
<tr>
<td>Holyoak &amp; Cheng (this issue)</td>
<td>57</td>
</tr>
<tr>
<td>Liberman &amp; Klar (in press)</td>
<td>58</td>
</tr>
<tr>
<td><strong>Subsidy rule</strong></td>
<td></td>
</tr>
<tr>
<td>Gigerenzer &amp; Hug (1992)</td>
<td>59</td>
</tr>
<tr>
<td>Holyoak &amp; Cheng (this issue)</td>
<td>41</td>
</tr>
</tbody>
</table>
difference emerges by comparing the results obtained by Manktelow and Over (1991/Expt 3) with those reported by Politzer and Nguyen-Xuan (1992). Despite the use of a virtually identical rule:

If you spend more than £100, you may take a free gift

as opposed to:

If you spend more than 10,000 Francs, the salesman must stick on the back of the receipt a gift voucher for a gold bracelet

and the identical indication of a specific perspective (in both studies, seller-speaker vs consumer-hearer), only Manktelow and Over’s problem elicited reasonably high rates of the predicted selection patterns in the two perspective conditions (85% of not-p & q vs 62% of p & not-q). Politzer and Nguyen-Xuan, in fact, obtained very low rates of the expected performances in both conditions (17% vs 35%, respectively; Klar & Liberman [personal communication] have replicated Politzer & Nguyen-Xuan’s results by using their “shop-offer” problem). This difference in performance in similar conditions is not limited to independent studies. Manktelow and Over’s “mother promise”:

If you tidy up your room, you may go out to play

elicited the response that was predicted to be dominant in only one of the two conditions in which subjects had to assume the “mother-speaker” perspective.

Finally, in the H&C paper, the rates of the expected performances on the “day-off” and “subsidy” problems are generally lower than those originally obtained by Gigerenzer and Hug (1992). More importantly, for the “employee-hearer” condition of the “day-off” problem the predicted dominant combination (p & not-q) was produced by only 23% of the subjects (vs 75% in Gigerenzer & Hug’s study, and 67% in Liberman & Klar, in press).

It might be argued that it is useless to try to explain these differences, which essentially emerge from the comparison of unrelated studies. Moreover, it might be argued that most deontic reasoning problems produce replicable effects. However, theories that are primarily devoted to the explanation of content effects in reasoning should offer at least some post hoc interpretations. Actually, the pragmatic schema theory can explain the indicated results by stressing the role of contextual variables. Pragmatic schemas are by definition (cf H&C) “context sensitive”, so that it is possible to explain these results by taking into account the contextual constraints on the activation of the schemas. For example, Girotto and Light (1992) have attributed the differences in response in the two versions of the “shop-offer” problem to the different representation of the potential violators offered by the two scenarios. In Politzer and Nguyen-Xuan’s scenario, the speaker (a famous Paris jeweller) had no apparent motive for violating the
offer, and the hearer (an anonymous, but probably rich customer) had no apparent opportunity (and probably no motive either) to do so. By contrast, Manktelow and Over’s scenario depicted both parties as potential violators, with clear motives and opportunity (the speaker was a shop manager in a British industrial town during a depression and the hearer was a customer from the same area who was suspected of taking more than he/she was entitled to).

The mixed results obtained by Manktelow and Over for the “mother promise” problem can be explained along the same lines. Unless the scenario explicitly indicates that the “mother has already been unfair” with her children, it is difficult to represent her as a potential violator who decides not to give the promised reward to a deserving promisee, and consequently to check for the relevant cases \((p \& \neg q)\). Of course, the explicit indication of a previous violating behaviour by the child is not necessary when subjects are cued in the mother’s perspective. In sum, when the scenario does not make the violations of a rule from a particular perspective intrinsically plausible, the activation of the appropriate schema and the consequent selection of the relevant cases are unlikely to occur.

This account presents a twofold difficulty. First of all, it does not cover all the existing data. In particular, the differences in responses for the hearer-employee condition of the “day-off” problem (cf Table 1) do not appear interpretable along the explanation proposed, given that the scenario used by H&C seems to be identical to the ones used by Gigerenzer and Hug (1992) and by Liberman and Klar (in press). Thus, one can only link them to possible cross-cultural differences in experiencing the checked violations (are companies’ frauds against employees more serious, or are these frauds just more common in Germany, or in Israel, than in the US?). Second, an interpretation based on contextual constraints upon the representation of the potential violators makes problematic the derivation of risky predictions from the available theories of deontic reasoning. All existing theories can provide an explanation of deontic reasoning performance based on the crucial role of contextual factors. The only exception is represented by the original version of the social contract theory, according to which perceiving a rule as a social contract is a sufficient factor for eliciting robust and replicable content effects (Cosmides, 1989).

The modified version of social contract theory proposed by Gigerenzer and Hug (1992) could provide a context-based interpretation. This theory (like the original one) appears to be limited in its scope when compared to pragmatic schema theory. In particular, it cannot account for deontic reasoning phenomena outside those related to the subclass of contractual deontic rules (cf the H&C paper). However, it might explain the failures to replicate the main findings on perspective change; in some versions of the problems (as in those used by Politzer & Nguyen-Xuan, 1992), subjects could have difficulty in taking a specific perspective due to insufficient contextual information. This extension of the theory is legitimated by Gigerenzer and Hug’s acknowledgement of “the
pragmatics of who reasons from what perspective to what end" (1992, p.166) as a sufficient factor for eliciting robust effects on deontic reasoning.

Recent accounts of selection performance based on theories of reasoning not uniquely devoted to the deontic domain, seem to offer (at least implicitly) a similar interpretation. In their explanation of content effects, Oaksford and Chater (1994) made reference only to “type of rule” (obligation vs. permission) and perspective that the subject must take. However, it is possible to specify their model (according to which card selections reflect an attempt by subjects to maximise expected utilities), by assuming that potential instances of unfairness will not be uncovered in contexts where unfairness is not likely to occur, as is the case, for instance, in Manktelow and Over’s (1991) first version of the “mother promise” problem.

Similarly, Rips (1994) argues that “focus on the obligatory nature” of the rule (i.e. the possibility of reading a conditional as implying OBLIGATORY [P/Q]), and instructions regarding the party who might be breaking it, explain deontic selection performances. However, in his explanation contextual variables are also assumed to affect selections (“content effect seems to break down when the cover story fails to guide subjects to [the OBLIGATORY (P/Q)] interpretation of the conditional...” (Rips, 1994, p.331).

Mental model theory (Johnson-Laird & Byrne, 1991) attributes to content and context the possibility of representing the falsifying cases (in particular, for deontic rules, the representation of violations) and the selection of the relevant cases.

Finally, according to Sperber, Cara, and Girotto (in press) success in the selection task depends on the interpretation of the conditional rule as a denial of the existence of the counterexamples to the rule (i.e. not exist [p & not-q]). Some rule-context pairs can cause subjects to make this interpretation. In particular, deontic rules that are interpreted as forbidding the occurrence of p & not-q cases yield good performance. When the deontic rule is a contract, it is possible to produce perspective shifts (i.e. change the representation of the forbidden cases), by instructing subjects to adopt a specific point of view (speaker vs hearer) and by providing sufficient information about the potential violations. This last proposal has the merit of generality: a linguistic pragmatic analysis of the task allows prediction of selection patterns in both deontic and non-deontic versions. Sperber et al. (in press) were able to show that a non-deontic version of the task produces successful performance when p & not-q cases are at least as easy to represent as p & q cases, and when the context indicates that knowing whether there are p & not-q cases would have greater effects than knowing whether there are p & q cases.

Good performance with non-deontic problems was also obtained by Love and Kessler (1995) who proposed a similar interpretation. Contexts that (1) indicate “how and why” to expect p and not-q cases; and (2) focus on the consequences of their occurrences, yield high percentages of p & not-q answers, irrespective
of the presence of a deontic content. In a similar vein, Liberman & Klar (in press) have shown that factors related to task interpretation explain the results obtained with contractual deontic rules, in particular perspective effects.

The moral of this brief analysis of the literature is that pragmatic schema theory (and a fortiori social contract theory) is not only necessarily restricted to the explanation of some knowledge domains (such as the deontic one), but it also needs to provide an explanation for a corpus of data on the deontic domain, which do not replicate well. The theory can provide such an explanation on the basis of an assumption about the role of contextual–interpretational factors. The apparently cross-theoretical nature of this assumption makes the schema view less original.

Deontic reasoning is undoubtedly domain-specific. However, as shown by the present analysis, it is questionable whether such a kind of reasoning can be validly investigated with a task generally requiring complex texts that are likely to introduce several confounding variables. An effort to devise new experimental tools would therefore seem to be necessary.

ACKNOWLEDGEMENTS

I would like to thank Dan Sperber, Jonathan Evans and Yechiel Klar for their comments on a previous version of this paper.

REFERENCES


Patricia Cheng and Keith Holyoak have played a central role in revealing the mysteries of the selection task (Wason, 1966). They argue that the mind contains "pragmatic reasoning schemas"; that is, principles such as:

If the precondition is not satisfied then the action must not be taken.

and they have shown that when a selection task is likely to map onto the appropriate schema, subjects are likely to make a correct selection (Cheng & Holyoak, 1985). Likewise, Holyoak and Cheng (this issue) argue that the phenomena of "point of view" can be explained by combining the schemas governing permission and obligation. Readers new to the selection task may be puzzled by what it is, and so we will begin with a brief history (see also Evans, Newstead & Byrne, 1993).

The pioneer of formalism, Jean Piaget, wrote that if individuals have to verify whether x implies y, then they "will look in this case to see whether or not there is a counterexample x and non-y" (Beth & Piaget, 1966, p.181). The selection task confronts this beautiful hypothesis with an ugly fact. The task itself consists in asking subjects to choose which of four cards should be turned over in order to find out whether a conditional assertion is true or false. In a typical version of the task, the four cards have a number on one side and a letter on the other side, which the subjects know, and the cards are laid out on a table with A, B, 2, and 3 uppermost. The conditional to be tested is:

If there is an A on one side of a card, then there is a 2 on the other side.

Intelligent adults tend to select the A card, or the A and the 2 card, and they fail to select the 3 card. In other words, given a rule of the form: if x then y, they select x, but conspicuously fail to select not-y. It is as though they have still to learn what counts as a counterexample to a conditional. However, when they are asked to generate a falsifying instance of a conditional, they usually construct...
cases of x and not-y (see e.g. Oaksford & Stenning, 1992). A striking acceleration in intellectual development is accomplished merely by changing the content of the selection task. Robust effects occur with deontic conditionals concerning what is permissible (Cosmides, 1989; Griggs & Cox, 1982). Thus, as Griggs and Cox showed, the conditional:

If a person is drinking beer then the person must be over 18

tends to elicit the selection of the correct potential counterexamples (x: the card corresponding to a beer drinker, and not-y: the card corresponding to someone younger than 18). Cheng and Holyoak (1985) advanced their pragmatic theory in order to explain this sort of phenomenon. They argued that the conditional maps onto the pragmatic schema:

If the precondition is not satisfied (e.g. person is not over 18 years) then the action (e.g. drinking beer) must not be taken and it, in turn, elicits the correct selection of cards.

Manktelow and Over (1991) carried out the selection task with the deontic conditional:

If you tidy your room then you may go out to play

and the subjects' selections depended on whose point of view they were asked to take. The mother who laid down the law is concerned that her child does not cheat, and the subjects who had to take her point of view tended to select the cards:

- did not tidy (not-x) went out to play (y)

The child is concerned that the mother does not renege on the deal, and the subjects who had to take the child's point of view tended to select the cards:

- tidied the room (x) did not go out to play (not y).

Similar effects have been demonstrated by Politzer and Nguyen-Xuan (1992), who also showed that subjects with a neutral point of view tended to select all four cards.

The question is: what accounts for the effects of point of view? Our plan in what follows is, first, to analyse Holyoak and Cheng's answer; second, to describe the theory of mental models; and, third, to argue that it gives a more comprehensive answer.
The central claim of Holyoak and Cheng’s theory is that individuals have knowledge of the following sorts of principles governing permissions:

If the precondition is satisfied, then the action may be taken

and the following sorts of principles governing obligations:

If the precondition is satisfied, then the action must be taken.

These principles are known as “pragmatic reasoning schemas”. Performance in the selection task, Holyoak and Cheng write (this issue, p. 291), “will be facilitated . . . when the stated rule has content that evokes a schema, and the correspondence between the stated rule and the schema rules is such that the latter map onto rules of standard logic”. They then explain the effects of point of view in terms of mapping the conditional to schemas concerning the relevant individual’s rights and others’ duties towards this individual. Thus, from the mother’s point of view, the conditional about tidying the room maps onto the permission schema given earlier for the child; whereas from the child’s point of view, the conditional maps onto the obligation schema given earlier for the mother. How these mappings are made is not yet specified by the theory.

The theory of pragmatic reasoning schemas is ingenious, it has provoked interesting research, and it may be true. However, we do have three misgivings about it. First, it has a narrow purview. It has little psychological justification from outside the selection task. A major task for Holyoak and Cheng is to show how the theory applies to other sorts of reasoning and thinking. Second, the pragmatic reasoning schemas are stated using the modal auxiliaries “may” and “must”. These verbs are systematically ambiguous, referring either to what is possible (or necessary) epistemically, for example:

It may rain tomorrow

or to what is possible (or necessary) deontically, for example:

You may smoke.

This ambiguity shows that the schemas are high-level rather than foundational. Holyoak and Cheng point out that rights and duties are interdefinable, but there is a more fundamental relation familiar to logicians:
possible  =  not necessarily not the case  
necessary  =  not possibly not the case

This relation echoes the interdefinability of the existential and universal quantifiers:

\[ \text{some}  =  \text{not all are not} \]
\[ \text{all}  =  \text{not some are not} \]

A natural interpretation of these relations is that at the heart of epistemic and deontic modalities are at least two distinct sets of possible states of affairs ("possible worlds")—those that are epistemically possible, and those that are deontically possible. If an event may happen, then it occurs in at least some member of the relevant set (epistemic or deontic); if an event must happen, then it occurs in all members of the set. Individuals know that certain actions are possible, that certain actions are permissible, and that certain actions create obligations. They can envisage such actions by building mental models of them. This knowledge and their knowledge of the meaning of "may" and "must", and other modal terms, provides the foundation for the principles that they can consciously articulate, e.g. "if you carry out certain acts, then you are morally committed to carrying out others". We take pragmatic reasoning schemas to be a systematic statement of such high-level principles, and our point is that they are not the deepest level of epistemic and deontic knowledge. A foundation is necessary along the lines that we will sketch out in the next section (see also Johnson-Laird, 1978).

Third, although people have knowledge that can be captured in pragmatic reasoning schemas, we have yet to be convinced that it is this knowledge in this format that is responsible for insight into the selection task. There is an alternative possibility to which we now turn.

THE THEORY OF MENTAL MODELS

The phenomena of the selection task cry out for Occam's razor. In our view, the theory of mental models provides such a parsimonious account (see e.g. Johnson-Laird & Byrne, 1991), and it goes beyond the data to make novel predictions. We will derive these predictions presently, but first we will outline the model theory.

The theory postulates that reasoning—deductive or inductive—is a process in which reasoners first represent the truth conditions of premises, and then use this representation together with their knowledge to construct mental models of the relevant situations. These models may take the form of visual images, but their critical feature is their structure. Thus, an exclusive disjunction:

There is a king or else there is an ace, but not both
calls for two alternative models (one for each possibility), which we represent
in the following diagram:

\[
\begin{array}{c}
k \\
a
\end{array}
\]

where \(k\) denotes a model of the king and \(a\) denotes a model of the ace. The
representation of explicit information is kept to a minimum so as not to overload
working memory. These models are partially *implicit* because they do not make
explicit that an ace does not occur in the first model and a king does not occur
in the second model. Reasoners need to make a mental "footnote" that the first
model exhausts the hands in which a king occurs and the second model exhausts
the hands in which the ace occurs. (We have sometimes used square brackets to
represent such footnotes in the diagrams, see Johnson-Laird & Byrne, 1991, but
will forgo the practice here.) These footnotes, provided they are remembered,
can be used to make models wholly explicit if necessary. Likewise, a
conditional, such as:

If there is a king in the hand, then there is an ace

is initially represented by two models:

\[
\begin{array}{c}
k \\
a
\end{array}
\]

\[
\ldots
\]

Individuals grasp that both cards may be in the hand, but defer a detailed
representation of the case where there is not a king in the hand. The ellipsis
accordingly signifies a wholly *implicit* model, i.e. one that has no explicit
content. Individuals need to make a mental footnote that the king is exhaustively
represented in the explicit model, i.e. it cannot occur in the hands represented
by the implicit model. Hence, the footnote can subsequently be used to infer that
the antecedent does not occur in the implicit model:

\[
\begin{array}{c}
k \\
a
\end{array}
\]

\[
\neg k
\]

where \(\neg\) represents negation. As aces are not exhaustively represented in the
initial models, they may, or may not, occur in the hands represented by the
implicit model:

\[
\begin{array}{c}
k \\
a
\end{array}
\]

\[
\neg k
\]

\[
\neg k \quad \neg a
\]
In this way, the initial models can be fleshed out so that they are completely explicit.

In general, if a conclusion holds in all the models of some factual premises, it describes a necessary conclusion; if it holds in most of the models, it describes a conclusion that is probable; and if it holds in at least one model, it describes a conclusion that is possible. Similarly, if a conclusion about an action holds in all the models of some deontic premises, it describes an obligation; if it holds in most of the models, it describes an action that is less than obligatory but more than permissible; and if it holds in at least one model, it describes an action that is permissible. These principles provide a foundation for factual, probabilistic, and deontic inferences.

People who have no training in formal logic appear to reason on the basis of mental models: they take longer and make more errors for inferences that call for multiple models than for inferences that call for only one model (see e.g. Johnson-Laird & Byrne, 1991; Johnson-Laird, Byrne, & Schaeken, 1992). They are also susceptible to “illusory” inferences predicted by the theory (see Johnson-Laird & Savary, 1995). For example, given the following premises:

Only one of the following assertions is true:
If there is a king in the hand, then there is an ace.
If there is a queen in the hand, then there is an ace.

they infer that an ace is more likely to be in the hand than a king—a conclusion that follows from constructing models of the overall exclusive disjunction using implicit models:

\[ k \quad a \]
\[ q \quad a \]

In fact, it is impossible for there to be an ace in the hand. If only one of the assertions is true, the other assertion must be false, and so either there is a king without an ace, or there is a queen without an ace. (If the premises are interpreted as bi-conditionals, the assertion is a tautology—and again one should not conclude that the ace is more probable.) There are a number of such illusions using different connectives, and they cannot be predicted by theories using valid formal rules of inference, because such rules cannot yield systematically invalid conclusions.

Holyoak and Cheng refer to the model theory as using “content-free model-theoretic procedures” (this issue, p.292) and as based on “content-independent reasoning procedures” (this issue, p.304, footnote 3). In fact, as befits a theory that was originally devised to explain the comprehension of discourse (see
Johnson-Laird, 1983), the model theory is neither content-free nor uses content-independent procedures. Content affects at least three distinct processes. First, it influences the initial interpretation of premises (see e.g. Byrne, 1989). Thus, as Johnson-Laird and Byrne (1991, p.46) point out, it will affect whether a conditional is interpreted as a one-way implication or a bi-conditional. Second, it affects the search for counterexamples to putative conclusions. If a conclusion is believable, reasoners are likely to accept it, but if it is unbelievable then, as Oakhill, Johnson-Laird, and Garnham (1989) have demonstrated, reasoners are likely to search more assiduously for counterexamples. Third, content affects the process of fleshing out implicit models (Johnson-Laird & Byrne, 1991, p.73). In fairness to Holyoak and Cheng, it is true that the model theory gives no detailed account of how knowledge is triggered during the process of comprehension. Nor, it might be added, does any other theory, including the pragmatic schemas theory, which does not explain how expressions in natural language trigger the pragmatic schemas. In short, the model theory assumes that semantic content and general knowledge play a critical role in reasoning—in interpreting premises, in fleshing out their interpretations, and in influencing the search for alternative models.

MENTAL MODELS, THE SELECTION TASK, AND POINT OF VIEW

In the selection task, subjects have a tendency not to construct counterexamples but to reason instead on the basis of their models of the conditional. A conditional, such as:

If there is an 'A', then there is a '2'

yields the initial models:

A 2

Subjects will select a card if it bears on the truth value of the conditional, and so they should select A alone or A and 2, depending on whether their mental footnotes correspond to a one-way or bi-conditional interpretation. These are the predominant selections for such neutral conditionals. The theory goes beyond the phenomenon to make a further prediction: there should be a correlation between the interpretation of the conditional (as a one-way or bi-conditional) and the pattern of selections. This correlation has been recently confirmed by Francesco Cara and Stefana Broadbent (personal communication).

Subjects should get the selection task right when they correctly construct counterexamples to the conditional. In an earlier analysis (Johnson-Laird & Byrne, 1991, p.80), we wrote:
An insightful performance may further depend on an explicit representation of what is not possible, i.e. the real impossibility given the rule [if A then 2] of:

\[ A \rightarrow 2. \]

This assumption appears to be correct. It is necessary to construct a model representing a counterexample to the conditional:

Impossible: \( A \rightarrow 2. \)

A deontic conditional also has counterexamples, although they do not render the conditional false but rather are violations of the principle it expresses. The key prediction is accordingly:

Any experimental manipulation that leads reasoners to flesh out their models of the conditional, and, in particular, to construct an explicit model of an appropriate counterexample, should enhance performance in the selection task. Such manipulations include the use of a procedure or a content likely to make counterexamples to the conditional salient, either by triggering specific memories or by eliciting a model in which such violations are represented explicitly. The model theory therefore accounts for the enhanced performance with realistic materials, including those of a deontic sort.

The key prediction goes beyond deontic effects and applies to any materials whatsoever (Johnson-Laird & Byrne, 1991, pp.80-81). It has recently been corroborated by four independent studies.

First, David Green and Rodney Larking (1995) asked subjects to construct counterexamples to conditionals that were not deontic, and showed that this manipulation enhanced performance of the selection task.

Second, Daniel Sperber, Francesco Cara, and Vittorio Girotto (in press) used a more indirect procedure with several sorts of materials that were not deontic. For example, they told their subjects that a certain machine generated cards according to the rule:

If a card has an A on one side, then it has a 2 on the other side.

The machine went wrong and ceased to obey the rule, but it has been repaired, and the subject have to check that the job has been done properly. They are thus likely to represent the machine’s potential error explicitly: \( A \rightarrow 2. \) The subjects in this condition and other similar ones were more successful in the selection task.

Third, Roberta Love and Claudius Kessler (1995) have independently obtained similar results. For example, they used the conditional rule:
If there are Xow then there must be a force field where the Xow are strange crystal-like living organisms who depend for their existence on a force field. In a context that suggested the possibility of counterexamples—mutant Xows who can survive without a force field—the subjects carried out the selection task more accurately than in a control condition that did not suggest such counterexamples.

Fourth, Nira Liberman and Yechiel Klar (in press) have shown experimentally that apparent effects of a schema for “checking for cheaters” are better explained in terms of subjects’ perception of appropriate counterexamples and of the relevance of looking for counterexamples.

The effects of point of view appear to arise from those conditional assertions that elicit a bi-conditional interpretation (see e.g. Fillenbaum, 1977). For example, the conditional:

If you tidy your room then you can go out to play

is usually taken to imply:

If you don’t tidy your room then you cannot go out to play.

When the models for such a bi-conditional are fleshed out in a fully explicit way, they are as follows:

\[
\begin{array}{ccc}
\text{t} & \text{p} \\
\neg \text{t} & \neg \text{p} \\
\end{array}
\]

where t denotes tidying your room and p denotes going out to play. The particular cards that individuals choose in a selection task based on a bi-conditional will depend on what sort of counterexample they represent in explicit models. There are two potential counterexamples:

\[
\begin{array}{ccc}
\text{t} & \neg \text{p} \\
\neg \text{t} & \text{p} \\
\end{array}
\]

Hence, a proper test of the conditional calls for selecting all four cards (as a neutral point of view elicits from subjects, see Politzer & Nguyen-Xuan, 1992). If subjects construct an explicit model of the first counterexample, they will select the t and not-p cards; if they construct an explicit model of the second counterexample, they will select the not-t and p cards. What happens in Manktelow and Over’s (1991) experiment, according to our account, is that subjects who interpret the conditional from the child’s point of view construct the models:
from which, by negation, they derive the counterexample:

\[ \neg t \rightarrow \neg p \]

In contrast, subjects who interpret the conditional from the mother's point of view construct the models:

\[ \neg t \rightarrow \neg p \]

from which, by negation, they derive the counterexample:

\[ \neg t \quad p \]

In both cases, the subjects select the cards corresponding to their respective counterexamples. The model theory accordingly explains the phenomena of point of view in the selection task: when deontic rules are equivalent to bi-conditionals, instructions can lead to explicit models of one or other (or both) of the counterexamples.

The theory goes beyond the known results to make a further prediction. With a factual conditional that strongly suggests a bi-conditional interpretation, such as:

If the Greeks disarmed then the Turks disarmed

the same phenomena should occur as with the deontic conditionals. When counterexamples are made salient, subjects with a neutral point of view should tend to select all four cards. The Greek point of view, however, makes salient the counterexample:

Greeks disarmed \( \rightarrow \) Turks disarmed

whereas the Turkish point of view makes salient the counterexample:

\( \rightarrow \) Greeks disarmed \quad Turks disarmed

Such a result would show that a deontic content is not essential for the subject's point of view to influence their performance in the selection task.

CONCLUSION

The model theory makes sense of the phenomena of the selection task: individuals will select those cards corresponding to a counterexample only if
they construct its explicit model. In the deontic domains explored by psychologists, the premises are open to a bi-conditional interpretation, and accordingly have two distinct sorts of counterexamples. The particular selections made by subjects will depend on which of these counterexamples they represent in an explicit model, and their point of view will affect which counterexamples they construct. There is no need to invoke pragmatic reasoning schemas in order to explain the phenomena. Moreover, pragmatic reasoning schemas cannot explain the results of the four studies using materials that were not deontic in which insight into the task was enhanced by making counterexamples more salient.

Our argument does not rule out the existence of pragmatic reasoning schemas, and indeed they may be used in the selection task. If theoretical entities should not be multiplied unnecessarily, however, then there is a need both for further empirical findings to bolster Holyoak and Cheng’s theory and for an extension of the theory to deal with domains other than the selection task. The onus is not on pragmatic schemas alone, however. What is missing from the model theory (and the pragmatic theory) is a detailed explanation of how conditionals and other sorts of assertion elicit relevant general knowledge.

ACKNOWLEDGEMENT

This research was carried out, in part, with support from the James S. McDonnell Foundation.

REFERENCES


Two and Three Stage Models of Deontic Reasoning

Mike Oaksford, University of Warwick, UK, and Nick Chater, University of Oxford, UK

Commentary on “Pragmatic Reasoning With a Point of View” by Keith J. Holyoak and Patricia W. Cheng

Holyoak and Cheng (this issue; henceforth “H & C”) provide a computational-level analysis (Marr, 1982) of deontic reasoning that corrects a probably too simplistic view of deontic rules (Cosmides, 1989) by introducing important ideas from jurisprudence. This analysis addresses the frequently cited criticism of pragmatic reasoning schema (PRS) theory that it does not account for the selection of the not-p and q cards in some versions of the thematic selection task. H & C suggest that people possess two PRSs, one the original permission schema from Cheng and Holyoak (1985) and an obligation schema derived from Politzer and Nguyen-Xuan (1992). They observe that these schemas are interdefinable because the antecedents and consequents of the rules that make them up involve rights and duties that are complementary. Rights and duties implicitly introduce two individuals who have different roles—e.g. employer and employee. By focusing on their rights, one of the individuals may interpret a rule as a permission whereas the other may interpret it as an obligation.
Different rules mediate schema access—P3 for the permission schema and O1 for the obligation schema. For a permission schema access via P3 leads to clause order inversion of the P1 rule, turning an \( \text{if } p \text{ then } q \) rule into an \( \text{if } q \text{ then } p \) rule. Using the rules in the permission schema with this clausal inversion yields the \( \neg p, q \) selection pattern.

H & C also note that their approach is consistent with subjective utility approaches to deontic reasoning (Kirby, 1994a; Manktelow & Over, 1991; Oaksford & Chater, 1994). They point to an important distinction, which we focus on in this commentary, between the factors that lead people to take a particular perspective and the effects of taking that perspective. Consistent with subjective utility approaches H & C argue that taking a particular perspective affects the utilities people assign to the various possible outcomes determined by a deontic rule. We find H & C’s computational-level proposals for what gets computed in taking perspectives both compelling and novel. However, H & C also propose an algorithmic-level account of how people make their actual card selections using pragmatic reasoning schemas (PRS) with which we take issue in this commentary.

THREE-STAGE MODEL

H & C’s account suggests a three-stage model of the processes involved in deontic reasoning:

(i) **Perspective Assignment**: Determining which perspective to adopt.
(ii) **Deontic Inference**: From the perspective determined in (i) decide which deterministic inferences to make.
(iii) **Inference Withdrawal**: Dependent on the probabilities and utilities assigned in (i), withdraw some inferences made in (ii).

So for example, take H & C’s day-off rule:

If an employee works on the weekend, then that person gets a day off during the week \( \text{ (1) } \)

Employers interpret (1) as:

If an employee works on the weekend, then that person **may** take a day off during the week \( \text{ (2) } \)

because they focus on their rights in this contractual arrangement. (1) also leads to an assignment of high utility to cases of abuse where an employee takes a day off without having worked at the weekend, i.e. the \( \neg p, q \) case. (2) leads people to access the permission schema via P3 which creates deterministic permission rules with appropriately inverted antecedent and consequent clauses. These
COMMENTARIES
deterministic rules yield the now familiar not-p, q card selections. H & C suggest that the degree to which subjects draw these deterministic inferences will depend on their expected utility. So, if the expected utility of using the rule to guide card-turning behaviour is low then subjects will tend not to turn the card. The expected utility of turning, for example, the not-p may be low for two reasons. First, it may be very unlikely that someone who doesn’t work at the weekend takes a day off. Second, it may be that although the benefits of detecting violators are high, the costs of false alarms are not small enough for subjects to discount them. For example, it may be important in a company where there is a strong union not to offend innocent workers and hence to avoid cases other than not p, q. According to H & C’s PRS account, low expected utility of turning the not-p would have to lead to the inference based on P4—that this card must be turned over—being withdrawn.

TWO-STAGE MODEL
In this commentary we argue that the three-stage model implied by H & C (this issue, see also Stevenson & Over, 1995) is unparsimonious and that a more parsimonious two-stage model is consistent with Oaksford and Chater’s (1994) formal computational-level analysis of the deontic selection task:

(i) **Perspective Assignment:** Determining which perspective to adopt.
(ii) **Maximise Expected Utility:** Calculate expected utilities for choices based on background knowledge and (i).

Reasoners determine the case to assign positive utility according to H & C’s account (i), but then base card selection purely on the expected utility calculation. Using H & C’s example (1) we now show how Oaksford and Chater’s (1994) computational-level analysis can account for card selection after perspective assignment. We label the antecedent of (I) the condition (con) and the consequent the action (act).

Oaksford and Chater (1994) use two contingency tables to describe peoples’ expectations about the probabilities of the various instances. The first represents the contingencies on the assumption that people obey the rule (Table 1a) and the second represents the contingencies on the assumption that people ignore the rule (Table 1b), i.e. they behave as if the condition and action are independent events. People may take more than these two attitudes towards a regulation, e.g. they may assume that there are “negative” people who always do the exact opposite of a regulation. However, Oaksford and Chater (1994) assume that 1(b) is a sufficiently general “catch-all”. The parameters of the model are P(M_1), the probability that people are obeying the rule, hence P(M_D), the probability that people are ignoring the rule, = 1 - P(M_1); P(con), the probability of the condition; and P(act), the probability of the action.
TABLE 1
Contingency Tables for $M_0$ and $M_1$

<table>
<thead>
<tr>
<th></th>
<th>$act$</th>
<th>$\overline{act}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$con$</td>
<td>$a$</td>
<td>0</td>
</tr>
<tr>
<td>$\overline{con}$</td>
<td>$(1-a)b$</td>
<td>$(1-a)(1-b)$</td>
</tr>
</tbody>
</table>

1(a) shows the table of probabilities appropriate for when the rule is obeyed $M_0$.
1(b) shows the equivalent table for when the rule is ignored $M_1$.

A corresponds to the probability of $con$, $P(con)$, and $b$ corresponds to the probability of $act$ in the absence of $con$, $P(act|\overline{con})$.

TABLE 2
Utilities of Card Combinations for the Employer and Employee Perspectives

<table>
<thead>
<tr>
<th>Employer</th>
<th>$act$</th>
<th>$\overline{act}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$con$</td>
<td>-0.1</td>
<td>-0.1</td>
</tr>
<tr>
<td>$\overline{con}$</td>
<td>5-0.1</td>
<td>-0.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Employee</th>
<th>$act$</th>
<th>$\overline{act}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$con$</td>
<td>-0.1</td>
<td>5-0.1</td>
</tr>
<tr>
<td>$\overline{con}$</td>
<td>-0.1</td>
<td>-0.1</td>
</tr>
</tbody>
</table>

Oaksford and Chater (1994) assign utilities to the various instances in the contingency tables in Table 1. Table 2 shows these utilities for H & C's example (1). We assigned a small negative utility (-0.1) to every combination of cards, because of the assumption of a fixed cost for turning any card. For the employer's perspective, we assigned a large positive utility (+5) to finding cases where the action occurs but the condition is not satisfied. For the employee's perspective we assigned a large positive utility (+5) to finding cases where the condition is satisfied but the action is not performed. The numerical values are arbitrary—all that is important is that the positive utility is large in comparison to the cost for turning over a card.

Expected utilities ($EU$) can then be calculated for each card using equations (3) to (7):

$$EU(con) = P(act|con)U(con,act) + P(\overline{act}|con)U(con,\overline{act})$$  \hfill (3)

$$EU(\overline{con}) = P(\overline{act}|con)U(\overline{con},act) + P(\overline{act}|\overline{con})U(\overline{con},\overline{act})$$  \hfill (4)

$$EU(act) = P(con|act)U(con,act) + P(con|\overline{act})U(con,\overline{act})$$  \hfill (5)

$$EU(\overline{act}) = P(con|\overline{act})U(con,\overline{act}) + P(\overline{con}|\overline{act})U(\overline{con},\overline{act})$$  \hfill (6)
Where the conditional probabilities \( P(x|y) \) are the expected values calculated with respect to the two contingency tables:

\[
P(x|y) = P(x|y, M_f)P(M_f) + P(x|y, M_p)P(M_p)
\]  

(7)

In equations (3)-(6) the expected utility of each card is calculated as the weighted sum of the utilities of each possible outcome given the visible face of the card. The weights are the probabilities of each outcome.

Oaksford and Chater (1994) fixed \( P(M_f) = 0.5 \) on the reasonable assumption that subjects are uncertain whether people are obeying or disobeying the rule. They then calculated expected utilities for each card by averaging over all pairs of values for \( P(\text{con}) \) and \( P(\text{act}) \) in the range 0.1 to 0.9 at 0.1 intervals. We show these values in Table 3 for the employer’s and the employee’s perspective, assuming a permission rule of the form if con, then may act. Table 3 reveals the same behaviour as that observed by H & C.

H & C argue that for their employee-O1 and employer-P3 conditions subjects interpret the ambiguous task rule as a permission, i.e. if con, then may act. Consistently assigning \( p \) to con and \( q \) to act, Table 3 reveals that the following card selections maximise expected utility. In the employer-P3 condition, subjects should select the \( \text{not}-p \) and \( q \) cards, and in the employee-O1 condition, subjects should select the \( p \) and \( \text{not}-q \) cards. In the employer-O1 condition subjects are presented with the rule: An employee must have worked on the weekend if the person takes a day off during the week. This rule is equivalent to the obligation rule, if the person takes a day off during the week, they must have worked at the weekend. Notice that this reverses the clausal order, i.e. this is an if act (q), then con(p) rule. But maximising expected utility still involves turning the \( \text{not}-p \) and \( q \) cards because of the clausal inversion from the permission to the obligation rule. In sum, if subjects are maximising expected utility then Oaksford and Chater’s (1994) model predicts just the results H & C.

### TABLE 3

<table>
<thead>
<tr>
<th>Card Face</th>
<th>(i) Employer</th>
<th>(ii) Employee</th>
</tr>
</thead>
<tbody>
<tr>
<td>act</td>
<td>+1.20</td>
<td>-0.10</td>
</tr>
<tr>
<td>act</td>
<td>-0.10</td>
<td>+2.31</td>
</tr>
<tr>
<td>con</td>
<td>-0.10</td>
<td>+2.23</td>
</tr>
<tr>
<td>con</td>
<td>+1.03</td>
<td>-0.10</td>
</tr>
</tbody>
</table>

Average expected utilities for each card face (action, not-action, condition, not-condition) (i) the employer’s perspective, and (ii) the employee’s perspective.
found in their experiment. Thus it appears that both the two- and the three-stage models can explain H & C's results. General criteria of parsimony therefore suggest that we should prefer the two-stage model.

There is also data where expected utility is clearly doing all the explanatory work. Kirby (1994a; see also Over & Evans, 1994; Kirby, 1994b) explicitly varied the probabilities and utilities in the deontic selection task. He used an obligation rule and the equivalent of an employer's perspective. Kirby found that when the cost of a false alarm increased fewer subjects selected the not-\(q\) card and as the cost decreased more subjects selected the not-\(q\) card. Moreover, selection of this card decreased as the probability of disobeying the rule decreased. Oaksford and Chater (1994) show that this behaviour is consistent with their optimal data selection model. Importantly the expected utility approach can capture much more of the variation in the data on the assumption that the proportion of subjects choosing a card directly reflects its expected utility. In sum, we can explain all the relevant data without invoking explicit rule-based accounts of deontic inference, i.e. without invoking the second stage (ii) of the three-stage model. Thus although we agree that H & C's computational-level account of perspective assignment provides a better and more comprehensive account than alternatives, we believe that the algorithmic theory based on PRSs about how this affects deontic inference is not needed to explain the data.

Several issues arise from this discussion. First, an account of perspective shifts is largely independent of the PRS account of deontic inference. Although both the two- and three-stage models of deontic reasoning require an account of perspective assignment, they use different processes from then on. Therefore we need not tie theories of perspective assignment to any particular theory of reasoning, be it mental logics, mental models, heuristics, or pragmatic reasoning schemas.

Second, although there have been informal (Manktelow & Over, 1991) approaches suggesting that subjective utility may play an important role in deontic reasoning, recent formal accounts (Kirby, 1994a; Oaksford & Chater, 1994) are crucial to demonstrating the viability of this approach. Oaksford and Chater's formal model is the most comprehensive account, explaining a wide range of data concerning the non-independence of card selections (Pollard, 1985), the negations paradigm (e.g. Evans & Lynch, 1973), the therapy experiments (e.g. Wason, 1969), the reduced array selection task (Johnson-Laird & Wason, 1970), work on so-called fictional outcomes (Kirby, 1994a) and deontic versions of the selection task, including perspective and rule-type manipulations (e.g. Cheng & Holyoak, 1985), and the manipulation of probabilities and utilities in deontic tasks (Kirby, 1994a). It remains to be seen if other proposals such as epistemic utility (Manktelow & Over, 1991; Over & Evans, 1994) can be formalised and shown to provide a better account of the data on deontic reasoning tasks and other versions of Wason's (1966, 1968) selection task.
Third, the three-stage model shares some common characteristics with some proposals of Stevenson and Over (1995). In the context of the conditional inference paradigm, they argue that deductive reasoning and probabilistic reasoning are complementary. They consider the following example:

If John goes fishing, he will have a fish supper.  
If John catches a fish he will have a fish supper.  
John will not have a fish supper.

Presenting (8) and (10) together (without 9) yields high levels of *modus tollens* inferences to “John did not go fishing” (Byrne, 1989). Adding (9) however, suppresses the number of these inferences that people draw. Stevenson and Over (1995) argue that this is because having “extended our beliefs by performing an instance of *modus tollens*... we [may] get more information... leading us to doubt the major premise...[and so]...we could be led to express doubt about the conclusion.” In other words, probability judgements may lead to inferences being withdrawn.

However, the data (Byrne, 1989; Cummins, Lubart, Alksnis, & Rist, 1991; Stevenson & Over, 1995) seems perfectly compatible with a purely probabilistic model. Subjects simply select the conclusion that has the highest conditional probability given the minor premise. We can represent the rule as a probability model as in Table 1(a) and 1(b). So for example, from 1(a) the conditional probability of *p* given *not-q* is zero, whereas the conditional probability of *not-p* given *not-q* is 1.\(^3\) However, if we allow that the rule is uncertain, i.e. the conditional probability of *q* given *p* is less than 1 (i.e. the *p, not-q* cell is non-zero), then these probabilities will vary. In sum, it seems highly likely that we can construct a purely probabilistic model of these data. In consequence, models that include a deductive component and a probability component are unparsimonious—it could be probabilities all the way down!

To summarise, we have argued that two-stage models that do not include a rule-governed or deductive stage of deontic inference provide more parsimonious accounts of deontic reasoning than three-stage models. We agree that an account of perspective assignment is a necessary component of deontic reasoning and that H & C have by far the best account to date of how to achieve this. However, we doubt that H & C’s PRS account of deontic inference is necessary to account for the data on deontic reasoning performance.

NOTES

1For the general case of deontic reasoning, see Oaksford and Chater (1994), pp.621–625.
2\(P(\text{act})\) can be calculated from \(P(M_1), P(\text{con}), \text{and } P(\text{act|con})\) (Oaksford & Chater, 1994, Equation 8).
3From Table 1(a) the conditional probability of *p* given *not-q* = \(\frac{0}{(1-a)(1-b)}\) and the conditional probability of *not-p* given *not-q* = \(\frac{(1-a)(1-b)}{(1-a)(1-b)}\)
REFERENCES


On the Relationship between Pragmatic Schemas and Mental Logic

David P. O’Brien, City University of New York, USA

Commentary on “Pragmatic Reasoning With a Point of View” by Keith J. Holyoak and Patricia W. Cheng

Holyoak and Cheng (this issue, p.304) noted that “content effects of the sort observed in research on reasoning about regulations clearly lie beyond the scope of existing psychological models of reasoning based on variants of formal logic (e.g. Braine & O’Brien, 1991). In general, proponents of the logic-based
approach have conceded that domain-general natural inference procedures must be supplemented by pragmatic extensions of the sort proposed by PRS theory.  I agree that the sorts of inferences described by PRS theory should be viewed as supplementary to those described by mental-logic (ML) theory—PRS theory alone is sufficient to account for only a fraction of the sorts of inferences and judgments made with and about conditionals. PRS theory and ML theory accomplish different things. PRS theory provides an account for certain content effects that are not explained by ML theory alone. Note, however, that many sorts of things other than rights and duties, or permissions and obligations, are expressed with the particle if and with other logic particles in natural languages, and ML theory provides an account of inferences that are available for logic particles across differing sorts of content—something that PRS theory cannot do, and indeed is not trying to do.

Holyoak and Cheng (this issue, p.290) also noted that “research on reasoning about regulations has played a central role in the development of theories of conditional reasoning,” and that deontic concepts underlie these regulations. Without doubt they are right that much recent research has centred on deontic conditionals; several intriguing effects of content on performance on tasks derived from Wason’s selection task have been explained through reference to deontic regulations. The query raised here concerns how central a role such research should play—how much of the ordinary usage of conditionals concerns deontic regulations? I know of no standard frequency counts of the various sorts of conditionals, but to get one estimate of how common deontic conditionals are, I counted the number of occurrences of the word if in the New York Times on the day I started writing this commentary (Saturday, 3 December 1994). I found 83 instances. Of these, only five could possibly be classified as conveying a permission or an obligation, and only one of these five seems clearly to be of the sort described by PRS theory: “If you earned more than $150,000 last year, you have to pay 110 percent of last year’s amount to be safe.” The other four do not quite as obviously fit the sorts of production rules described by PRS theory for permissions or obligations, although they reasonably might evoke such interpretations, e.g. “If there is no sale, there are no taxes” might evoke O2. The remaining 78 of the 83 occurrences of if seem to have nothing at all to do with deontic regulations. (I shall return to these later.) At least on this frequency count, deontic conditionals appear to be of a relatively rare sort, and thus do not seem to be the sort on which to base a general theory of conditional reasoning (or a fortiori a theory of deductive reasoning generally).

What does ML theory propose for conditionals? In general, our approach has three parts: A set of content-general inference schemas, a reasoning programme that applies the schemas in a line of reasoning, and an independently motivated set of pragmatic principles that are used in interpreting linguistic input and in making additional inferences. The theory contains two principal schemas for
conditionals: *modus ponens* (MP) and a schema for conditional proof (CP), which are discussed in detail in Braine and O’Brien (1991). Described simply, MP holds that given two propositions of the form *if p then q* and *p*, one can infer *q*. CP holds that to derive or infer *if p then ..., first suppose p; for any proposition *q* that follows from the supposition *p* taken together with other information assumed, one may assert *if p then q*. (CP in ML theory differs from the conditional-proof schemas found in standard-logic systems in that an argument under a supposition can be made on extralogical inferences, and in that no proposition can be used in a derivation leading to a conditional under a supposition that is inconsistent with that supposition.) That MP inferences are made routinely even by children is well established (see review by Braine & Rumain, 1983; for discussions of why the occasional failure to select the *P* card in the selection task does not impeach the routine manner in which people make MP inferences, see O’Brien, 1993, 1995). Less attention has been paid in the literature to CP, although CP reveals more about the meaning of conditionals than does MP, in that CP reveals that the basic meaning of *if* is found in supposition: An *if* clause poses a supposition, and a *then* clause asserts what follows under that supposition.

The reasoning programme includes a direct-reasoning routine (DRR) and some more sophisticated strategies. The DRR is claimed to be applied automatically and available universally, whereas the strategies are assumed to be effortful and to require acquisition, and large individual differences are assumed in the availability of the strategies. The DRR is quite limited and describes the minimal ways in which the schemas can be applied. For MP the DRR simply requires that the two requisite propositions be considered together for the inference to be made. For CP the DRR provides only that when an *if* statement is to be evaluated, the antecedent of that conditional is added to the premise set and its consequent treated as a conclusion to be evaluated. Clearly, however, even children show evidence of understanding conditionals in ways that go well beyond what CP and the DRR provide, e.g. *if* statements that require suppositions that would not be forthcoming on the DRR alone are made spontaneously and routinely by children (e.g. Bowerman, 1986; Reilly, 1986).

Although 78 of the 83 occurrences of *if* from the *New York Times* that were mentioned earlier are not captured by PRS theory, they can be evaluated by CP, i.e. by treating the antecedent as an additional premise and evaluating the consequent, for example: "If Dow acquires majority ownership in Buna, it will become the fourth largest chemical company in Germany", "If Hampton gets 58 yards a game, he will break 1,000 yards for the fourth straight year", "She faces one year in prison if convicted", "If something is stolen, I can have its photo on national television in two minutes." Some of the *if* occurrences in the *New York Times* were special usages: *even if*, *as if*, *what if*, and *if* as an apparent disjunction; these also are interpretable from the perspective of CP in which *if* indicates supposition. "I don’t know if my job will be eliminated", for example,
might appear to be using if to convey whether, but it concerns whether a supposition will turn out to be factual (as in “I suppose I will lose my job, but I don’t know”). Even if indicates that even with supposition of an alternative, the consequent still follows, e.g. “Even if cutting taxes is your goal, the variable annuity is a good choice.” As if expresses an equality with, or similarity to, what would follow under a counterfactual supposition, e.g. “He was received much as if he had been a member of their own family.” What if expresses a request to pursue what follows under a supposition, e.g. “What if the value of securities plunges in the meantime?”. ML theory captures the suppositionality that is common across these conditionals of various content, whereas PRS theory does not.

Although the basic part of ML theory for conditionals, i.e. MP, CP, and the DRR, allows us to understand what all of these different sorts of conditionals have in common, it does not account for all of the sorts of conditional-reasoning inferences people make. The basic part of mental logic is intended to describe only what is universally available. Additional sorts of inferences are available both on the strategies and on pragmatic processes; some of the inferences made on pragmatic processes are logically sound and others are not.

Just because sound inferences such as modus tollens (MT) or contraposition are not universally available does not mean that they are not commonly available in certain circumstances. Scholnick and Wing (1991), for example, reported that children often refute an adult conditional by citing a counterexample—the adult’s conditional is taken as a supposition that leads to an inference that is contradicted by the counterexample; children evidently understand that this falsifies the supposition. Given that this is a sort of reductio ad absurdum argument, why do adults often fail to make MT inferences in laboratory tasks? The difficulty would seem to be in the strategic step of setting up a supposition in the service of finding a contradiction. Indeed, it is the sophistication of such strategic reasoning that is impressive when Sherlock Holmes realises that the dog would have barked if the murderer had been a stranger, so the murderer could not have been a stranger. Among the production rules described by PRS theory are two that correspond to a specialised application of MT for certain sorts of universally quantified deontic regulations (P4 and O4 for linguistic input mapped onto the permission or obligation schemas). If one views the PRS schemas as supplementary to those of ML theory, one arrives at the following scenario. Because MT is not a basic schema, such inferences will not be made reliably across all sorts of content. Some people on some occasions will make the appropriate judgment from a reductio argument, and some people on some occasions will make the judgment on pragmatically invited biconditional-like interpretations. When a conditional evokes P4 or O4, as PRS proposes for permissions and obligations, the appropriate response will be common. This scenario seems to me to correspond to the published data. (The extent to which some non-deontic conditionals also might lead reliably to MT inferences is yet to be discovered.)
Not all inferences drawn from conditionals are valid entailments. People often judge \( \text{if } p \text{ then } q \text{ and } q \text{ to entail } p \), and \( \text{if } p \text{ then } q \text{ and not } p \text{ to entail not } q \). Although these classical fallacies are not valid in standard logic, such inferences often are pragmatically sanctioned, e.g. someone who is told that they will be paid if they do some work will expect to be paid only if they do the work. (Not that they would feel cheated if they were paid anyway; they might think, though, that they were getting away with something.) Among the production rules described by PRS theory are some that block these otherwise pragmatically invited fallacies (P2, P3, O2, and O3). If one again views the PRS schemas as supplementary to those of ML theory, one arrives at the following scenario. When faced with a conditional and either the denial of its antecedent or the affirmation of its consequent, nothing in the basic part of ML theory applies. People then might reasonably conclude that nothing follows; for a variety of pragmatic reasons, however, people can be invited to accept the classical fallacies (see O'Brien et al., 1989, for a discussion of a bias against indeterminate responses). When a conditional evokes the PRS production rules, the pragmatically invited fallacies can be blocked. Again, this scenario seems to me to correspond to the published data. For MT and the classical fallacies, PRS theory thus describes one set of circumstances (for deontic regulations) in which appropriate conditional-reasoning inferences that are not available on the basic part of mental logic can be made. In principle, no contradiction exists between the sorts of inferences described in each of the two schema theories.

What do the perspective shifts of the sort described by Manktelow and Over (1991), Politzer and Nguyen-Xuan (1992), Gigerenzer and Hug (1992), and Holyoak and Cheng (this issue) have to say about the claims of ML theory? PRS theory and ML theory both propose that inference schemas are applied not to the surface-structure forms of linguistic input, but to propositions as interpreted. In principle, this is not incompatible with what Braine and O'Brien (1991) proposed: People prefer semantically plausible interpretations over those based on syntactic cues alone. What PRS theory provides is a schema structure onto which surface structures that express deontic regulations can be mapped.

I turn now to a difference in the claims of the two approaches. Cheng and Holyoak (1985), Cheng, Holyoak, Nisbett, and Oliver (1986), and Holland, Holyoak, Nisbett, and Thagard (1986) argued that the PRS schemas are acquired inductively. Content-general inference schemas of the sort described in ML theory would be abstracted from the earlier content-specific ones. Cheng et al. (1986, p.318) wrote, however, of a “negative conclusion about the prevalence of a natural logic based on syntactic rules,” and Holland et al. claimed (1986, p.282) “that this level of abstraction in conditional reasoning is seldom attained.” More recent articles by advocates of PRS theory have not expressed such negative assessments of the ML schemas (e.g. Holyoak & Cheng, this issue; Kroger, Cheng, & Holyoak, 1993), but they have not addressed further
what the developmental relationship is between the content-general and counter-
specific sorts of schemas.

The inductive-acquisition claim of PRS theory seems to call for a description
of development in which children begin to use and understand if statements
through the accumulation of content-specific pragmatic schemas, first acquiring
one, then adding a second, and so forth. Abstraction of features that are
common across these schemas would occur relatively late in development, if at
all; content-general schemas such as MP and CP would come in late and be
rather rare. From the perspective of ML theory, however, MP and CP provide
a basic understanding of conditionals, and should be available early; content-
specific schemas that provide additional knowledge about the relations between
antecedent and consequent would be added through experience.

The latter scenario is more sensible. Both adults (as shown in the description
of if statements from the New York Times) and young children (Bowerman, 1986;
Reilly, 1986; Scholnick & Wing, 1991) use if across a much wider range of
content than would be supported by the sorts of schemas described by PRS
theory. The sheer number of additional content-dependent schemas that would
be required makes the inductive-acquisition approach of PRS theory
implausible. Further, the evidence reviewed by Braine and O’Brien (1991)
shows that MP and CP are available across a wide variety of sorts of content
early in development. An additional reason follows from the observation by
Fodor (1975) and Macnamara (1986) that in order to acquire propositional
knowledge about the world it is necessary to represent it, and this requires a
format for representation. PRS theory has not addressed why deontic regulations
are represented as conditionals, whereas ML theory describes the logical
primitives that allow knowledge about deontic regulations to be acquired in the
form of conditionals.

Some of the details of how PRS theory works still need to be worked out. For
example, Holyoak and Cheng claim that from an employer’s perspective “If an
employee works on the weekend, then that person gets a day off during the
week” will be mapped onto with P3 as “If an employee works on the weekend,
then that person may take a day off during the week.” Once this representation
is made, the rule can then be “recast in the form of Rule P1”. (this issue, p.297)
How this recasting is done is left unspecified, and I remain unconvinced that
someone necessarily would infer “If an employee takes a day off during the
week, then that person must have worked on the weekend” from the P3
representation. Earlier (this issue, pp.294–295) we are told that at the deontic level
the rules may be “pragmatically biconditional”. Is this how one gets from P3 to
P1? Or does any one of the production rule always evoke all others? Or perhaps
any one of the permission production rules will be recast as P1? At any rate,
clearly I do not yet see how the theory operates as one moves about within or
across the production rules of the two schemas.
An additional question concerns the use of the assumption that people focus on their own rights and on the duties of others. Although this may often be true (and may be the default value), I find it easy to imagine situations where the opposite is the case. For example, someone who is afraid of being caught cheating (taking something without permission, or failing to meet an obligation) might be concerned more with their own duties and the rights of others. A worker who is afraid of losing their job when they violate a company regulation might respond very differently from a person who is afraid of missing out on a benefit. This suggests to me that the mapping from surface structure to PRS production rule is governed not simply by a focus on one's own rights and the duties of others, but—dare we say it—by the situation's utilities. Note that a utility would provide knowledge about a regulation's topic, but would not by itself provide the schematic structure onto which the regulation can be mapped. Because the assumption of focusing on one's own rights and on the duties of others is not part of PRS theory per se, this required use of utilities to explain which perspective is taken is not damaging to PRS theory. What it suggests, I believe, is an interdigitation of PRS with utility theory—a direction that may prove more fruitful than the recent attempt to place utilities within a mental-models framework (e.g. Manktelow & Over, 1995).

REFERENCES


---

Perspectives, Preferences, and Probabilities

D.E. Over, University of Sunderland, UK, and
K.I. Manktelow, University of Wolverhampton, UK

Commentary on “Pragmatic Reasoning With a Point of View” by
Keith J. Holyoak and Patricia W. Cheng

It is now a neat 10 years since the publication of a seminal paper in the psychology of reasoning, “Pragmatic reasoning schemas” by Cheng and Holyoak (1985). They were the first to recognise that deontic contexts were instrumental for inducing “correct” performance on Wason's selection task, to test this idea explicitly, and to provide a special account of this performance. Since 1985, a large body of empirical and theoretical work has accumulated, issuing from several groups of researchers around the world. We shall give a brief review of some of our own ideas (given more fully in Manktelow & Over, 1995).

PREFERENCES AND PERSPECTIVES

Our approach to deontic selection tasks, and deontic thought in general, is decision theoretic. The deontic thought we are concerned with is a kind of practical reasoning. It is aimed at deciding which actions to perform to achieve goals, and not at acquiring true beliefs about the world. Our two basic concepts are thus subjective preferences, from which can be technically derived a notion
of utility or value, and subjective probability judgements. We take people to have preferences between certain possible states of affairs, and to be able to make probability judgements about them. If you, a smoker, prefer to be healthy, and think that you are more likely to attain this state by giving up smoking, you will by extension prefer giving up smoking to carrying on with it. Giving up smoking will be your goal, and all things being equal, you will then assert something like, “I should give up smoking.”

We do sometimes make deontic statements just for our own benefit, as when we are struggling with weakness of the will, and finding it hard to give up a habit like smoking. (See Over & Manktelow, 1993, on weakness of the will and deontic reasoning.) Most commonly, however, a social context has to be taken into account to understand deontic statements fully. Consider how often these statements are made by one party with the aim of affecting the behaviour of another (cf Cheng & Holyoak, 1995, p.398). We have called the party uttering such a rule the agent, and the party whose behaviour is its target the actor (Manktelow & Over, 1991). Each party must hold, and be mutually assumed to hold, preferences among the states of affairs described by the deontic rule and the context for its utterance to be effective. To take our well-worn example:

(1) If you tidy your room then you may go out to play.

Assume that this is uttered by a mother in the attempt to influence her young son. In the standard case, the statement would not be made if the mother (the agent) did not prefer the room to be tidy rather than untidy, and believe that her son (the actor) preferred going out to staying in. It is just possible that the son would not accept the statement. He might think that he ought to stay in, even if he tidies his room, as he could prefer to get his homework out of the way early. There may be real children like this somewhere, but let us suppose that the son in our example is not one of them.

Such statements and their contingencies can be straightforwardly mapped onto the selection task: cards are constructed with “room tidied” (p) or “room not tidied” (not-p) on one side and “went out” (q) or “did not go out” (not-q) on the other.

Now we can predict a range of responses in subjects cued to the perspectives and their attendant preferences. Subjects will seek to uncover possible costs to their perspective—this could help them to take corrective action or to avoid the costs in the future. The agent will prefer situations in which p is made true and be particularly sensitive to situations in which p is not made true (her goal in this example is a tidy room); the actor will prefer situations in which q is made true and be particularly sensitive to situations in which q is not made true (his goal is to go out to play). There are not two, as often presumed, but six ways in which, depending on the actions of one party or the other, an outcome arises that represents a cost to either (in the sense that a party’s goal is unattained):
1. Agent sees that P is true but does not allow Q to be made true;
2. Agent sees that P is not true but still allows Q to be made true;
3. Actor makes P true but does not make Q true;
4. Actor does not make P true but still makes Q true;
5. Agent sees that P is not true and does not allow Q to be made true;
6. Actor does not make P true and does not make Q true.

Selection task predictions in each case would be as follows. 1: p, not-q; 2: not-p, q; 3: p, not-q; 4: not-p, q; 5: not-p, not-q; 6: not-p, not-q. Cases 1 and 4 are those that have been commonly confirmed by other researchers, including Holyoak and Cheng in their current paper; Cases 1, 2, 3, and 4 were confirmed by Manktelow and Over (1991) using an example about a shop; Cases 5 and 6 (where neither party’s goal is attained) are covered in Manktelow and Over (1995) but have not, as far as we know, been tested for.

It is not true to state, as do Holyoak and Cheng, that they or other schema theorists such as Politzer and Nguyen-Xuan (1992; see Holyoak & Cheng, this issue, pp.293–294) have either predicted or observed this full range of cases. Holyoak and Cheng’s current account, based on the relation of rights and duties (this issue, pp.296–297), does not accommodate cases 2 and 3 (2: mother sees room is not tidy but still allows child out to play; 3: child tidies room but does not go out to play), but we have shown that people are as sensitive to these cases as they are to the classic “cheating” cases 1 and 4.

**PERSPECTIVES AND PROBABILITIES**

Probabilities are the third crucial factor (as first pointed out at the end of Manktelow & Over, 1991). We have argued before that the study of deontic reasoning can help to link the traditionally disparate research areas of reasoning and decision making. (See Evans, Over, & Manktelow, 1993, for this argument and a more general one on the need to integrate these areas.) Combining basic preferences, and so subjective utility or value, with subjective probability judgements produces subjective expected utility. It should be possible to demonstrate such a combination of factors in deontic reasoning.

Empirical work reporting the effects of probability manipulations on responses to deontic selection tasks has just started to appear, and Holyoak and Cheng mention an important example in their final section. Kirby (1994) used a modified version of Griggs and Cox’s (1982) drinking-age task, in which the target rule was:

(2) If a person is drinking beer then the person must be over 21 years of age.

Subjects were more likely to choose a false consequent card referring to a drinker aged 19 than one referring to a 4-year-old, presumably because of a
judgement that small children would be unlikely to be drinking beer. In our own work, we have used an adaptation of Cheng and Holyoak's (1985) immigration task (Manktelow, Sutherland, & Over, 1995; Over, Manktelow, & Sutherland, 1994) to decrease selections of a true antecedent card. The rule was:

(3) If a person has ENTERING on one side of their immigration card then they must have CHOLERA on the reverse side.

The violating case here would be one with ENTERING but without “cholera”. Using multiple instances of p, not-p, q, and not-q, we were able to introduce a probabilistic factor by labelling the cards with countries of origin; either tropical, where cholera was said to be common, or European. In one condition, fewer subjects (57%) chose European ENTERING cards compared to tropical ENTERING cards (98%). It is just as costly to allow someone suffering from cholera to enter the country from Europe as it is from the tropics, but the expected cost of the former is less than that of the latter.

DEONTIC REASONING, DUTIES, AND RIGHTS

Sentences in deontic reasoning characteristically contain special modal terms, such as the deontic “must” and “may”, and have a different logical form from other modal statements, let alone from extensional ones. Compare:

(4) If it is raining then the barometric pressure must have fallen.

(5) If John is to go out to play then he must tidy his room.

The consequent of (4) logically implies that the barometric pressure has fallen. If we find that this pressure has not actually fallen, we may infer the negation of the consequent of (4), and then use modus tollens to infer that it is not raining. In contrast, the consequent of (5), with its deontic “must”, does not logically imply that anyone tidies the room. People can violate whatever obligations they have, and if that occurs in this case, we cannot logically infer the negation of the consequent of (5). Still less could we use modus tollens to infer that John does not go out to play.

A standard, indicative selection task could be based on (4), while a deontic one could be based on (5). In spite of their recognition that the ability to engage in deontic reasoning is a special one, Holyoak and Cheng would claim that (4) and (5) are “formally-equivalent rules”. But a deontic conditional is not formally the same as a non-deontic one, as we have illustrated. They claim later in their paper that, after learning that the consequent of a deontic conditional like (5) has been violated, one can perform modus tollens, but again our example shows that this is not so. At a deep level, both indicative and deontic selection tasks can be seen as decision problems (Kirby, 1994; Over & Evans, 1994). However, an
understanding of this similarity should be built on an appreciation of the
difference between the two cases, and particularly on the logical difference
between indicative and deontic conditionals.

Moreover, we have to be clear about this difference if we are to get a good
account of the nature of the schemas postulated by Holyoak and Cheng. As we
have pointed out before, these schemas themselves contain the deontic terms
“must” and “may” (Manktelow & Over, 1991, 1995). Their presence means that
the conditionals in the schemas are unlike the kind of production rules in
systems like ACT-R and SOAR. We would like to hear much more about these
conditionals and the schema system in which they are to operate. What other
rules does this system have? To what extent does it resemble a deontic logic,
or incorporate a very limited deontic logic within it?

The reasons Holyoak and Cheng give for rejecting the idea that there is a
mental deontic logic are not decisive. One reason they state seems to depend on
their view that modus tollens is needed in deontic logic to account for the results
in deontic selection tasks. But we have already illustrated that modus tollens
cannot be used in a deontic selection task. In general, standard deontic logics tell
us nothing about violations, but only about theorems, characteristically
containing the deontic “must” or “may”, that can be deduced from some given
axioms. This is, in fact, a better reason for holding that something more is
needed for ordinary deontic reasoning than a deontic logic. Yet another reason
Holyoak and Cheng give for this conclusion is that the deontic S5 logic will not
allow obligations to conflict with each other. Against this, it could be pointed
out that some weaker logics, such as minimal deontic logic, will allow
conflicting obligations (Chellas, 1980, 6.5).

Certainly, a mental deontic logic could not be all there is to our understanding
and production of deontic discourse and deontic reasoning. As we have just said,
deontic logics, at least as so far formulated, are not about violations, and as we
have seen, preference judgements are needed before deontic tasks can be
properly understood. (See Manktelow & Over, 1992, for an experiment that
supports this.) We have also argued that preference judgements are needed to
explain why people assert and accept deontic statements. Rules in deontic logic
are unlike those in a natural deduction formulation of extensional logic: the
deontic rules are not of any real help in introducing deontic terms into our
discourse. For example, perhaps the least controversial rule in deontic logic is
that we may infer that a proposition is obligatory if we have found that it is a
logical truth. This is of somewhat limited value in ordinary affairs.

On the other hand, there must be further mental rules containing the deontic
“must” and “may” in what Holyoak and Cheng propose. For example, if the
precondition is not satisfied, P4 in their “permission” schema allows us to infer
that the action must not be taken, and from that, we can intuitively infer that it
is not the case that the action may be taken. But Holyoak and Cheng give us no
further rule for making this intuitively correct inference. To justify further
inferences like this, it looks as though their schema system would indeed have to incorporate some limited deontic logic, or at least some principles that tell us how their “must” and “may” are related to each other. Some of these principles may in turn have to depend on whether some conflict in obligations has arisen.

Another point about which we are unsure is whether Holyoak and Cheng’s schema system accounts for the non-monotonic nature of ordinary deontic reasoning (Manktelow & Over, 1995; Over & Manktelow, 1993). A child may accept (1) or (5) and infer that he ought to tidy the room as he wants to go out. But if the house suddenly catches fire, no doubt he will go out without tidying up. For us, this means that he can understand his mother’s preferences between states of affairs, and can then act consistently with these preferences, as well as his own interests in this case, rather than conforming rigidly to any rule she explicitly gives him. Evidence of a non-monotonic aspect in ordinary deontic reasoning comes from the experiment we described earlier, in which true antecedent selections were decreased.

Our view has always been that a variety of representations will probably be necessary to account for human reasoning (e.g. Manktelow & Over, 1990, 1995). Part of our understanding of deontic discourse could well be embodied in schemas and even in some really basic principles of some deontic logic. Much deontic reasoning could be described informally as schematic: people often seem to know quickly and easily what they must or may do in more or less routine situations. Sometimes people prefer one state of affairs to another just because of what they infer from legal, moral, or other rules that we have been taught.

We do hold that preferences between possible states of affairs are represented in mental models when people are trying to decide which possible rules to assert or accept in the first place, or to act on in some new or unexpected situation. A full account of deontic reasoning must include mental models, and we would strongly agree with what Holyoak and Cheng say at the end of their paper on the need to combine work on deontic schemas with that on models of utility. But they present our view as if we consider the acceptability of a deontic statement only from the perspective of an authority. We have, however, called this perspective that of the agent and distinguished it from the perspective of the actor, who can have different preferences. Agents make deontic statements in light of their own preferences, but actors may not accept these agents as legitimate authorities, or as ones who really have power over them. Perhaps the son in our example will conclude that he may go out, without tidying his room, because he thinks his mother is unlikely to punish him.

Finally, we acknowledge another important step that Holyoak and Cheng have taken, by focusing on duties and rights, and by their use of the work of Hohfeld. They seem to speak as if these concepts are always expressed in deontic conditionals, but we would ask if they really do hold this. Consider their example:
(6) If an employee works on the weekend, then that person gets a day off during the week.

In Holyoak and Cheng's analysis, the antecedent of (6) apparently refers to a duty that the employees have towards their employer. The employees may have this duty—it could be a requirement of their contracts, for example, that they work at the weekend in exchange for the right to a day off during the week. But (6) could also be an offer or an inducement to employees who have no duty to work at this time, but who can do so if they see that it is in their interest.

Holyoak and Cheng claim that there are some counterexamples to our account of obligation and permission. Their example of a parking regulation raises an important point: such rules are not usually laid down on their own. The Los Angeles civic authorities presumably think that what the populace really should do is to obey all their regulations; these are presented as a unit among which one is not allowed to pick and choose. If the populace did this, the city would save so much money that it would have no need of fines. (The serious problem of choosing which combinations of actions one ought to perform is discussed in Manktelow & Over, 1995.) A more cynical view would be that the parking regulation is the deontic equivalent of a lie. The authorities prefer to proclaim it for their own selfish purposes, but they do not themselves endorse its content.

In Holyoak and Cheng's other example, the government could prefer that women be free to choose for themselves whether to have an abortion, and that the total number of abortions so chosen be small. These are consistent preferences, and no counterexamples to what we say about permission. Again, another way to look at this case would make one more critical. The government may actually be acting to restrict the free choice of some women by its funding policy. It could minimise the number of abortions while increasing free choice, in the true sense, by making birth control more widely available. Of course, any agents or actors may have different goals, and so different deontic judgements, from that which they profess.

We ourselves would stress that we have not tried to present an account of duties and rights, and still less of their relation to obligations and permissions, which is subtle. For example, your right to free speech implies that you have a right to make many false statements; even so you think that you should not do this. It is even wrong to say that the government has given you permission to do it. Another topic that should be explored more fully by psychologists is the relation between legal, moral, and prudential obligations and permissions, and the topic we have mainly been concerned with: what it means to decide which actions one should or may perform, all relevant things considered. Holyoak and Cheng have shown that psychology has something to learn about deontic reasoning from legal theory, as it does, we would add, from moral philosophy and political theory. In return, these other fields, with their lack of an
experimental base, should take a serious interest in what Holyoak and Cheng have done.

REFERENCES


