

Markers of social group membership as probabilistic cues in reasoning tasks

Gary L. Brase

University of Sunderland, UK

Reasoning about social groups and their associated markers was investigated as a particular case of human reasoning about cue–category relationships. Assertions that reasoning involving cues (group markers) and associated categories (social groups) elicits specific probabilistic assumptions are supported by the results of three experiments. This phenomenon remains intact across the use of categorical syllogisms (Experiment 1), conditional syllogisms (Experiment 2), and the use of social groups that vary in their perceived cohesiveness, or entitativity (Experiment 3). Implications are discussed for various theories of reasoning, and additional aspects of social group/coalitional reasoning are also discussed.

‘Serpent!’ screamed the pigeon.

‘I’m *not* a serpent!’ said Alice indignantly. ‘Let me alone!’

‘...No, no! You’re a serpent; and there’s no use denying it. I suppose you’ll be telling me next that you never tasted an egg.’

‘I *have* tasted eggs, certainly,’ said Alice, who was a very truthful child; ‘but little girls eat eggs quite as much as serpents do, you know.’

‘I don’t believe it,’ said the pigeon; ‘but if they do, why, then, they’re a kind of serpent, that’s all I can say.’

Alice’s Adventures in Wonderland, by Lewis Carroll (1971/1865, pp. 45–47)

Here Lewis Carroll recounts a peculiar form of “pigeon logic”:

All serpents eat eggs.

All little girls eat eggs.

Therefore, all little girls are serpents.

Correspondence should be addressed to Gary L. Brase, Sunderland Business School, Department of Psychology, St. Peter’s campus, University of Sunderland, Sunderland, SR6 0DD, UK. Email: gary.brase@sunderland.ac.uk

This article is based in part on a dissertation submitted in partial fulfillment of the requirement for a doctorate at the University of California, Santa Barbara. The author would like to extend thanks to Leda Cosmides, John Tooby, Daphne Bugental, Denise Cummins, Laurence Fiddick, David Over, Russ Revlin, and several anonymous reviewers for helpful comments, and to Sandra Brase for advice and support.

As obvious as the pigeon's confusion may be to us right now, psychological research has shown that human beings, when faced with similar situations, are often no more logical.

Consider the following set of premises:

All freshmen wear green caps;

All these fellows wear green caps;

When Wilkins (1928) gave these premises to college undergraduates, a majority of the subjects generated invalid inferences (e.g., that "All these fellows are freshmen"). We are faced with the uncomfortable indication that humans are no more logical than Lewis Carroll's pigeon. In fact, Carroll's pigeon and Wilkins' subjects are apparently showing the same phenomenon in reasoning, which is called a *content effect*. Briefly, a content effect occurs when someone uses knowledge that is associated with the content of an argument (e.g., eggs, serpents, freshmen, and caps) to reach their answer, rather than relying only on the logical form of the argument as one is expected to do in a logical reasoning task.

The existence of content effects has been used by domain-general theories of reasoning to accommodate many of the extra-logical inferences that people make (e.g., Evans, Barston, & Pollard, 1983; Evans, Over, & Manktelow, 1993; Johnson-Laird & Byrne, 1991). Content effects may happen for any domain in which a person has acquired relevant knowledge, with the specific effect being either facilitation or suppression of the logical inferences. In some examples, this seems to be straightforward and sufficient as an explanation. For example:

Bill Clinton has gotten sexual favours from an employee

People in positions of power get sexual favours from employees

Therefore, Bill Clinton has been in a position of power

In this example the content tends to compel people to accept the conclusion even though it does not logically follow from the premises. The reason the conclusion appears to be valid is because, in the real world, it is a true statement of fact. That is, specific and relevant knowledge about the content drives the effect. The freshmen/green caps example from Wilkins (1928) appears to be the same sort of effect—however, it is importantly different in that the content is not about actual, real-life people. "Freshmen" is a conceptual category of people; a social group. Furthermore, it is not actually the case (so far as I can tell) that freshmen have ever all worn green caps. So if in fact the Wilkins result is a content effect, it is a content effect that relies on content knowledge at a relatively abstract conceptual level (the level of social group members being identified by a particular shared property). A crucial question emerges: At what point do content effects that involve conceptual categories of events in the world (e.g., freshmen), become no

longer a mere content effect and graduate to an inference procedure specific to that conceptual domain?

There is no consensus as to how or why a collection of content effects can graduate to being a domain-specific inference procedure. The recent history of human reasoning theories has seen several theorists attempt to make this jump with varying degrees of success. For instance, the social contract theory of Cosmides and Tooby (1989, 1992) claims the following origins:

When we began this research in 1983, the literature on the Wason selection task was full of reports of a wide variety of content effects, and there was no satisfying theory or empirical generalization that could account for these effects. When we categorized these content effects according to whether they conformed to social contracts, a striking pattern emerged. Robust and replicable content effects were found only for rules that related terms that are recognizable as benefits and cost/requirements in the format of a standard social contract... No thematic rule that was not a social contract had ever produced a content effect that was both robust and replicable. (Cosmides & Tooby, 1992, p. 183)

To the idea of domain-specific reasoning procedures for social contracts, Cosmides, Tooby, and colleagues have proposed additional domain-specific reasoning procedures for dealing with social threats and hazards (Cosmides & Tooby, 1989; Fiddick, Cosmides, & Tooby, 1995, 1996; Rutherford, Tooby, & Cosmides, 1996). At this point, these theories are hotly debated, but they are clearly treated as theories about specific reasoning procedures rather than simply "content effects".

Are there other areas in which content effects may actually be indicative of specific inference procedures that are not yet understood? Cosmides and Tooby (1992; Tooby & Cosmides, 1988), and the results of items like the Wilkins example, suggest that there may be inference procedures that are invoked in social group situations. The field of social psychology supports the general idea that dealing with social groups involves complex cognitive processes. The present paper is concerned with just one aspect within this area: the relationship between social groups and group markers.

GROUPS AND GROUP MARKERS

Social psychologists have studied the dynamics of social groups for decades, revealing a complex web of psychological phenomena. Some the most general conclusions from these efforts are that people very readily categorise other people into groups (Hirschfeld, 1996), perceive people in terms of their group identities (Hamilton & Sherman, 1996), and treat people in those groups in particular ways based on their group-identities (Mullen, Brown, & Smith, 1992; Sherif, et al., 1961; Tajfel, 1970, 1982). Furthermore, it is actually fairly difficult to prevent these group-oriented perceptions and behaviours, suggesting that the

psychological processes underlying them are pervasive and easily invoked. Many of these coalitional behaviours, as they occur in laboratory-formed groups on the initiation of psychologists, can appear to be dysfunctional (as when members of two groups created in a completely and explicitly random way show biases in favour of their fellow group members). However, these behaviours that may appear strange in a laboratory, seem in fact to be well designed for enhancing the status of coalitions in natural situations. Studies using naturally occurring coalitions—e.g., castes in India (Bohra, 1979); national ethnicities (Gallos, Callan, & Parslow, 1982; Sande, Goethals, Ferrari, & Worth, 1989); religious groups (Taylor & Jaggi, 1974); and racial groups (Greenberg & Rosenfield, 1979)—tend to show coalitional behaviours as considerably more functional than they appear in laboratory results. One key to this paradox is that groups in the real world are almost never randomly determined; groups are instead more typically formed around ideas, interests, or features common to the group members. In the real world, therefore, phenomena such as in-group bias are sensible (even if not particularly admirable) behaviours.

Another example of these apparently deviant behaviours in relation to groups comes from research on the *representativeness heuristic*, a classic example of which is that a woman, “Linda”, with a very politically liberal background is judged more likely to be (a) a feminist and a bank teller than (b) just a bank teller (Tversky & Kahneman, 1983). Judging (a) to be more probable than (b) violates the normative rules for the conjunction of probabilities, but this judgement can be partially understood by appealing to the notion that social groups (e.g., feminists) are structured around the ideas, interests, and features of their members (on the representativeness heuristics, see also Gigerenzer, 1991). Memberships in social groups appear to powerfully indicate particular personal characteristics of those members (and vice versa).

Within the wide variety of groups in the world, some groups are more cohesive (or “groupish”) than others. How people perceive this social group cohesiveness varies accordingly, and Donald Campbell (1958; also see Hamilton & Sherman, 1996) labelled this characteristic that groups have in greater or lesser amounts *entitativity* (i.e., being like or having the properties of a real entity). Some groups are more entitative than others. Campbell envisaged perceived entitativity as something that varied on a continuum, and I agree, but for purposes of research I will consider *strong coalitions* (groups with very high entitativity), *weak coalitions* (groups with lower entitativity), and *grouping* (groups with little or no entitativity). I also propose that a major underlying factor that determines perceived entitativity is the degree to which all the group members share common objectives, recognise this conflux of interests, and behave as a cooperative unit in the furtherance of their objectives. The difference between these three groups can be seen in a quote from Campbell’s original description (1958, p. 18):

A band of Gypsies is empirically harder, more solid, more sharply bound than the ladies aid society, and the high school basketball team... falls somewhere in between.

Entitativity, as hypothesised here, is fundamentally about shared mental states, although it may be inferred from cooperative and coordinated behaviours. Strong coalitions are therefore groups in which there are a number of shared objectives, all mutually recognised as such, and the group members engage in planned and coordinated behaviours in the pursuit of those objectives. Weak coalitions are groups in which there are one or a few shared objectives, and the behaviours of group members (taken as a whole) are not as planned and coordinated with each other. Groupings are groups in which there are no shared objectives, or any shared objectives are not mutually recognised, or any shared objectives do not lead to planned and coordinated behaviours.

Interestingly, groups with a high degree of entitativity are generally perceived as having some degree of “psychic reality”, a phenomenon that actually supports the current conception of entitativity. Mental states can be and are attributed to coalitions in a way that is relatively independent of the mental states of the individuals constituting a coalition (e.g., statements such as “The French want open borders” and “The Republicans dislike the Democrats” are accepted as sensible statements about the world). In other words, coalitions—as a whole—are attributed with beliefs and desires. This makes sense if understanding coalitions involves understanding the idea of multiple individuals acting in a concerted fashion due to shared mental representations of a common objective. The attribution of the common beliefs to the group, rather than to each individual within the group, is an extremely efficient and usually accurate shortcut. For the purposes of reasoning about coalitions, this property allows us plausibly to treat coalitions as unitary items that can be assigned labels, traits, and other properties in much the same way that such properties are assigned to individuals (for similar ideas regarding the construction of “institutional facts”, see Searle, 1995). Furthermore, if there are reasoning procedures specific to the domain of social groups, they should be elicited more effectively in situations that involve a strong coalition.

Grant for the moment that coalitional groups indicate a shared set of mental representations (beliefs, goals, intentions, etc.). As mental representations cannot be directly observed, group membership is not immediately obvious to an observer. Physical markers of these beliefs and goals are often used to effectively communicate these memberships and implied beliefs. These *coalitional markers* are physical representations of internal mental states (group membership and the entailing objectives and beliefs). Ribbons pinned to lapels, fraternity pins, police badges, gang tattoos, and secret handshakes are all examples of external, physical markers of coalition memberships. Coalitional membership (which is a state of

mind) is, in fact, something that must be inferred from physical and behavioural cues. While some of the cues used by others to infer memberships may be incidental (e.g., track marks on a heroin addict's arms), *intentionally* produced cues of coalitional existence are more properly and consensually considered coalitional markers¹. Coalitional markers, do in fact, occur universally (LeVine & Campbell, 1972) and evidence indicates that they have existed as an aspect of human societies for at least as long as recorded history (Davis, 1985; Marcus, 1993). The purposes of coalitional markers appear to be recognising, cooperating, and coordinating within the group, and/or warning and intimidating those outside the group. Thus, coalitional markers, as used in this research, are cues—intentionally produced—to communicate information about coalitional affiliations to others.

Effective (i.e., “good”) coalitional markers appear to have certain probabilistic properties that are important: markers are items that have a very low (i.e., at or near zero) natural base rate of occurrence in the world. This zero base-rate property produces another curious feature: although markers are intentionally produced they are never designed to also serve a functional or utilitarian use. This (along with the intentional production properties of coalitional markers) creates very high category validity—where M stands for marker and G stands for group member, $p(M|\sim G)$ is minimised and $p(M|G)$ is maximised. Thus, for example, the properties of having eyes, or even having blue eyes, should be difficult to construe as markers (they have large, non-restricted reference classes, they are not intentionally created, and they are very functional)². On the other hand, having pierced eyebrows, tattooed cheeks, or other ritual mutilations should be readily suspected as coalitional markers. In summary, markers are proposed to have two critical properties: A very high categorisation validity and intentional production. A cue that is associated with a coalition should be perceived as a marker the more it is both (a) not likely to occur in the general environment (i.e., the base rate is zero) and (b) clearly intentional (i.e., not accidental, a byproduct of some other process, or an uncontrollable event).

Such a situation, in terms of the hypothesised probabilities of group markers in the general world and as cues of social group membership, leads to a model of how probabilistic inferences could be made from the presence of a marker to a

¹This uses the folk concept of intentionality (Malle & Knobe, 1997), which predicates beliefs and desires (i.e., the marker can in a sense actually be considered as a physical proxy for the common objectives underlying the coalition).

²As one reviewer noted, the use of blue eyes as a group marker was famously used in a study of prejudice and discrimination (Peters, 1987). This line of research has, indeed, shown the proclivity of even young children to form powerful social groups, but it should also be noted that the original study actually used strips of cloth that fastened at the neck as collars on the children to more effectively distinguish blue-eyed children from brown-eyed children. Apparently, eye colour by itself was not believed to be adequate as a group marker.

person's group membership. If, when one identifies a group (G) and something as an associated group marker (M; for example, a piece of red ribbon as a marker for AIDS awareness), one makes two default assumptions: (1) that the category validity, $p(M|G) = x$, which is some positive value and (2) that the error rate, $p(M|\sim G) = 0$, then one should draw inferences about group membership based on the presence or absence of the marker (as well as the reverse). In such a case, we can use Bayes' theorem to see that the base rates cancel out of the equation and the posterior probability, $p(G|M)$, which is the same as the cue validity³ (CV_M), is equal to 1:

$$p(G|M) = \frac{p(G)p(M|G)}{p(G)p(M|G) + p(\sim G)p(M|\sim G)}$$

$$p(G|M) = \frac{P(G) * x}{p(G) * x + p(\sim G) * 0}$$

$$p(G|M) = p(G) * x / p(G) * x$$

$$p(G|M) = 1$$

With complete cue validity, two inferences are allowed—from cue (group Marker) to category (Group membership), and from not-category (non-Group member) to no-cue (no Marker). Importantly, these inferences should be allowable, based on cue validity, regardless of the formal logic structure within which they exist. Furthermore, if the category validity is taken to be equal to one, i.e., $p(M|G) = x = 1$, such a situation would lead people to make biconditional conclusions from the cue to the category (and vice versa), even if such conclusions are deductively unwarranted (for similar ideas, which will be dealt with in the discussion section, see Kirby, 1994; and Chater & Oaksford, 1999).

GENERAL PREDICTIONS

The following experiments all involve reasoning about social groups and associated coalitional markers within reasoning tasks. The general idea is to

³Cue validity is the probability that a cue occurs in instances of a target category, divided by the probability that the cue occurs regardless of category membership (Beach, 1964; Reed, 1972). The cue validity (CV_M) can be mathematically expressed as:

$$CV_M = \frac{p(M|G)}{p(M|G) + p(M|\sim G)}$$

evaluate the effects of social groups of varying types and group markers of varying quality. Specifically, are group markers utilised in people's inferences as to group memberships, and are these inferences sensitive to the quality of both the group markers and the quality of the entitative state of the groups? Is so, once a marker and a group have been intercontingently associated (as a cue and the relevant category), certain decision rules can be expected to follow from that association. To the extent that a reasoner makes the assumptions that $p(M|G) > 0$ and $p(M|\sim G) = 0$, this licenses certain inferences (given a person with the marker, that person is inferred to be a member of the coalition, and given that a person is not a member of the coalition, that person is inferred not to have the marker). If a reasoner further assumes that $p(M|G) = 1$, then biconditional inferences are allowed. Expressed in terms of first order logic, [(“If G then M”) & (“If $\sim G$ then $\sim M$ ”)] licenses a biconditional inference and, therefore, logically implies “If M then G”. In other words, when one makes these two assumptions (of category validity and cue validity), then one is warranted in inferring “If M then G” from “If G then M”. Thus, when told “All group members have the marker” (and the aforementioned assumptions are made), people should presume that “All the people with marker are in the group”. Similarly, in so far as “All people with the marker are in the group” leads people to make these cue and category validity assumptions, then they should also infer that “All group members have the marker”.

In the following experiment, subjects were given booklets of syllogism completion items that contained statements about social groups. Some items associated social groups with cues that could be suitable as group markers. Other items associated social groups with cues that were less suitable as group markers (i.e., had higher base rates and/or were not intentional). As many previous studies have focused on a different reasoning task (the Wason selection task), it may be necessary to point out that the present studies use a syllogism completion task because the considerations of social groups and group markers dealt with here relate to affirmation situations (using markers to identify group members). The selection task, on the other hand, is used primarily in situations of violation and falsification.

EXPERIMENT 1

The first experiment uses categorical syllogism completion items in two different figures (for a review of syllogism structures, see Mayer, 1983). One figure allows a deductively valid conclusion, and the other figure does not allow a deductively valid conclusion but often leads to inferences by people nonetheless.

Participants

A total of 53 undergraduate students (15 males and 38 females) at a southeastern American university participated in this study for extra credit in psychology courses. The average age of the participants was 24.7 years old.

Materials and Procedure

Each participant was given a booklet containing an instruction sheet and 17 test items. Each test item contained two initial premises and then six possible conclusions, from which participants were asked to select the conclusion they believed to be correct (i.e., complete the syllogism). All participants selected just one of the conclusions for each item. 11 of the 17 test items were distractor items, involving social groups but with various different types of premises and conclusions. The remaining six test items (see Table 1) were taken from a $2 \times 2 \times 3$ collection of premise pairs with different figures (2 forms), different content words (2 types), and different marker item base rates (3 base rates). One of the two different figures permits an informative, logical conclusion:

- A. All people with [cue X] are in [social group Y]
[Person Z] has [cue X]
 Therefore, [Person Z] is in [social group Y]

The function of these syllogisms in the present experiment was: (1) as a control to make sure that participants were faithfully performing the experiment; and (2) as a baseline of performance to compare against syllogisms with the same semantic content but the second figure, which has no informative, logical conclusion:

- B. All people in [social group Y] have [cue X]
[Person Z] has [cue X]

This arrangement is usually called an “undistributed middle term” (the position of the item common to both premises, “cue X”, is such that the two premises cannot be deductively combined to form a conclusion), and a conclusion derived from these two premises is called an undistributed middle term (UMT) error.

The premise contents were chosen to be ambiguous, while always containing information about a social group and a cue associated (by the premise) with that social group. The purpose of using this type of content was to avoid, as much as possible, any content effects due to specific knowledge of actual social groups in the real world. The social groups used as contents for these experimental syllogisms were selected to be close to neutral with respect to positive or negative connotations, as determined in a pre-test of connotative values for a large number of social groups. The marker cues given as the common term to both premises

TABLE 1
Experiment 1, stimuli

	<i>Version 1 (eyes content)</i>	<i>Version 2 (ears content)</i>
Bad quality cues		
<i>Control</i>	All people who have eyes are Fishermen Mark has eyes	All people who have ears are Nannies Larry has ears
<i>Undist. middle term</i>	All people who are Fishermen have eyes Mark has eyes	All people who are Nannies have ears Larry has ears
Low quality cues		
<i>Control</i>	All people who have bloodshot eyes are Acrobats Adam has bloodshot eyes	All people who have attached earlobes are Bankers Ralph has attached earlobes
<i>Undist. middle term</i>	All people who are Acrobats have bloodshot eyes Adam has bloodshot eyes	All people who are Bankers have attached earlobes Ralph has attached earlobes
High quality cues		
<i>Control</i>	All people who have a serpent tattoo are Bricklayers Brian has a serpent tattoo	All people with a "P" shaved above their ears are Plumbers Steven has a "P" shaved above his ear
<i>Undist. middle term</i>	All people who are Bricklayers have a serpent tattoo Brian has a serpent tattoo	All people who are Plumbers have a "P" shaved above their ears Steven has a "P" shaved above his ear

Organized by marker item base rates (Bad quality, Low quality, and High quality), content words (Version 1 and Version 2), and syllogism figure (control and undistributed middle term) Distractor items are omitted.

(e.g., serpent tattoo) were chosen to systematically vary in their quality as group markers, from "bad" (obligate physical traits with high base rates and no intentionality, e.g., eyes and ears), to "low quality" (possible physical traits with medium base rates and questionable intentionality, e.g., bloodshot eyes and attached earlobes), to "high quality" (very low base rate and clear intentionality, e.g., serpent tattoo and "P" shaved above the ear).

Each participant was presented, interspersed with the distractor items, with the two different contents in opposite logical forms (i.e., half the participants saw content version 1 in the control figure and content version 2 in the UMT figure. The other participants saw the content version 1 in the UMT figure and content version 2 in the control figure). Because each participant received in the booklet some version of the control syllogisms figure, and the opposite version of the

UMT syllogism figure, no participants saw two items with repeated contents and every participant saw exactly half of all the stimuli used in this study. Following the two premises, several possible conclusions were given, and participants were asked to indicate which conclusion they believed to be correct. An example of a complete item is:

- All Bricklayers have a serpent tattoo
 Brian has a serpent tattoo
- (A) Therefore, all Bricklayers are Brians
 - (B) Therefore, Brian is a Bricklayer
 - (C) Therefore, some Bricklayers are Brians
 - (D) Therefore, some Brians are Bricklayers
 - (E) Therefore, Brian is not a Bricklayer
 - (F) Cannot reach a conclusion

Participants were run in small groups ranging in size from 2–15 persons. There was no time limit and all subjects finished in less than 15 minutes.

Results and discussion

Studies have always found very good performance in reasoning about the items that have the control figure arrangement. Because some of the premises in these syllogisms obviously violate real-world knowledge (e.g., “All people with eyes are Fishermen”), high percentages of participants drawing inferences from these syllogisms will indicate that participants are actually attempting to reason from the premises rather than relying just on any real-world knowledge and relations. Of more interest are the results from the UMT syllogisms. It is predicted that probabilistic information entailed in cue–category relationships can be used as relevant grist by whatever inductive machinery is invoked by this task. From this perspective, a baseline rate of undistributed middle term (UMT) errors should occur when the premise contents involve social groups and people with “bad” quality coalitional markers, and progressively more of these UMT errors should occur as the quality of the marker cues improves—i.e., “high quality” coalitional markers (items with clear intentionality and a low base rate) will evoke the most UMT errors.

The results of Experiment 1 are shown in Table 2, expressed as the proportions of different responses for each item (responses to the distractor items are not reported). Because the responses for the two different contents were quite similar (difference of proportion tests revealed no significant differences between analogous pairs), these two sets of results were collapsed to simplify the results, as shown in Figure 1. For the premises with informative, logical conclusions, the results showed a high number of correct inferences drawn across all types of meaningful contents (error rates: .23 for bad markers, .15 for low

quality markers, and .11 for high quality markers). As mentioned before, some of the premises with bad quality marker contents violated real-world knowledge, and there was a marginally significant effect of this in the error rates (.23 vs .15: $z = 1.05$, $p = .15$, $h = .20$; and .23 vs .11: $z = 1.64$, $p = .052$, $h = .32$; see Cohen, 1988, regarding effect size estimates). The fact that most participants (77% and upwards) drew correct logical inferences from these unbelievable premises indicates that the participants were largely, but not entirely, uninfluenced by real-world considerations of believability.

The number of UMT errors that could be attributed to a biconditional interpretation (responses of “All A are C”, “Some C are A”, and “Some A are C”) was lowest for the bad quality marker contents (.38), higher for the low quality marker contents (.55), and highest for the high quality marker contents (.77).

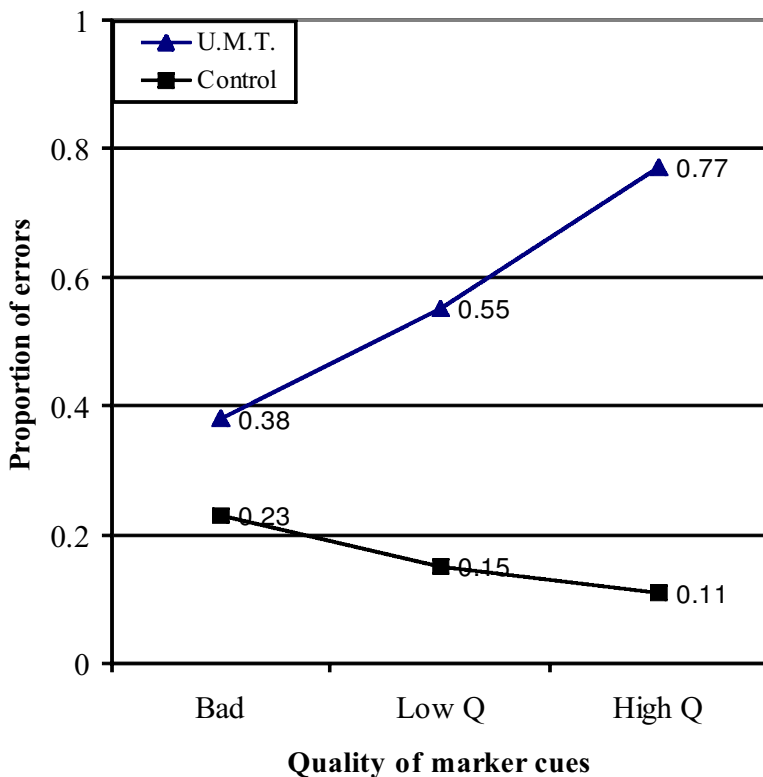


Figure 1. Experiment 1 results: Proportions of incorrect responses for the control premise sets. (Responses other than UA (All A-C)) and proportion of undistributed middle term errors that could be attributed to a biconditional interpretation (All A-C, Some A-C, Some C-A) with the data collapsed across the two different types of content.

TABLE 2
Experiment 1, responses

	<i>All C-A</i>	<i>All A-C</i>	<i>Some C-A</i>	<i>Some A-C</i>	<i>No A-C</i>	<i>None</i>
<i>Premise sets with informative, formal logic conclusions (All B-C, All A-B)</i>						
<i>Bad markers</i>	.037	.778	.037	–	–	.148
<i>Low quality</i>	.037	.852	–	.037	–	.074
<i>High quality</i>	.037	.926	–	–	–	.037
<i>Bad markers</i>	–	.769	–	.038	–	.192
<i>Low quality</i>	–	.846	–	.038	–	.115
<i>High quality</i>	.038	.846	.038	.038	–	.038
<i>Premise sets with undistributed middle terms (All C-B, All A-B)</i>						
<i>Bad markers</i>	.038	.346	.038	.077	–	.500
<i>Low quality</i>	–	.538	.077	.038	–	.346
<i>High quality</i>	–	.538	.038	.154	–	.269
<i>Bad markers</i>	.074	.296	–	–	–	.630
<i>Low quality</i>	.074	.370	–	.074	–	.481
<i>High quality</i>	.037	.778	–	.037	–	.148

Proportions of responses for control premise sets (with informative, logical conclusion) and for undistributed middle term (UMT) premise sets in Experiment 1 (within each half of the table, each row shows data for a specific content and arrangement and each column shows the proportion of participants selecting a particular conclusion). Logically correct responses are in italics, and modal responses are in bold.

With the responses dummy coded as either logically correct (0, for “Cannot reach a conclusion”) or logically incorrect (1, for any other response), a repeated measures ANOVA of the three types of premise contents with undistributed middle term arrangements showed a significant difference existed between the levels, $F(2, 104) = 22.20, p < .001$. Repeated measures difference of proportions tests showed that the incremental differences between bad, low quality, and high quality marker premise contents were small to medium effects (.38 vs .55: $z = 1.75, p = .04, h = .34$; and .55 vs .77: $z = 2.39, p = .009, h = .47$), with the total difference between bad and high quality markers being a moderate to large effect (.38 vs .77: $z = 4.06, p < .001, h = .81$).

These results strongly suggest a role of cue validity within the context of deductive reasoning tasks. The gradual change in responses appears to indicate that the underlying cause of these changes is the actual cue validity of the syllogism contents, and this in turn implies that coalitional markers, at least in this respect, are utilised in a manner that is sensitive to cue–category relationships and probabilistic cue validity.

EXPERIMENT 2

The results of Experiment 1 indicate that contents involving social groups and coalitional markers influence categorical reasoning. This second experiment investigated if similar contents have an analogous effect on reasoning with conditional premises. The same predictions apply; reasoning formats—categorical premises, conditional premises, and so on—are simply tasks that psychologists use in order to get controlled glimpses at how the mind works.

The conceptual design of this experiment generally parallels that of Experiment 1, but using conditional premises rather than categorical premises and using a much larger collection of items. When given the information “If a person is in the group, then they have the marker” (i.e., category validity = 1), participants should infer that group members have the marker and that those without the marker are not in the group (which happen to be the logically correct answers). Participants should also make the remaining inferences (non-group members do not have the marker and those with the marker are group members) to the extent that the cue validity compels them to do so, i.e., higher cue validity should lead to more of these inferences, because the assumption that the error rate, $p(M|\sim G) = 0$ becomes less problematic. On the other hand, when given the information “If a person has the marker, then they are in the group” (i.e., cue validity = 1) participants should feel licensed to infer that those with the marker are group members and that non-group members do not have the marker (which now happen to be the logically correct answers) much more frequently than in the former condition, and also still in proportion to the quality of the marker as a cue. Furthermore, if subjects make the default assumption of complete category validity (because all the categories are clearly social groups and there are no indications of category validity problems), they will therefore conclude, that group members have the marker and that those without the marker are not in the group.

In summary, there are three embedded patterns that should emerge: (1) inferences based on category validity should be high overall, but the logical status of these inferences will change depending on the order of the initial information given in the conditional rule (“if group, then marker” or “if marker, then group”); (2) inferences based on cue validity should increase as the markers used become more suitable as social group markers (low base rates and clear intentionality); and (3) inferences based on cue validity should be higher when the conditional information is “switched” (“If a person has the marker than they are in the group”), as compared to the number of inferences based on cue validity when the conditional information is in the “standard” order (“If a person is in the group, then they have the marker”).

Participants

A total of 36 undergraduate students (7 males and 29 females) at a southeastern American university participated in this study for extra credit in psychology courses. The average age of the participants was 23.1 years old.

Materials and Procedure

Participants were given a booklet of conditional syllogism completion items that all had contents dealing with social groups. Some premises associated social groups with traits that would be suitable as coalitional markers, while other premises associated social groups with traits that were not as suitable as coalitional markers (i.e., had high base rates or were not intentionally produced). A catalogue of 48 conditional premise pairs was constructed. These pairs varied along three dimensions: Type of marker (3 levels: bad quality, low quality, and high quality), order of terms in the initial (conditional) premise (2 levels: standard and switched), and term in the second (indicative) premise (4 levels: group member, non-group member, with marker, no marker). This yielded 24 types of premise pairs. Each type of premise pair was developed twice for the catalogue, using different specific content words, for a total of 48 syllogisms. The same neutral social groups used in Experiment 1 were used here, but clarified with the phrase “member of the [group]” and followed by a synonym for a group (e.g., league, club, union, alliance, etc.). This was done to make the social groups more clearly coalitions rather than mere aggregations of individuals. No distractor items were used in this experiment.

The three types of cue contents were: “bad” markers (medium base rate, low intentionality traits: blue eyes & type A blood), “low quality” markers (medium base rate, unclear intentionality traits: black pants & white shirt), and “high quality” markers (low base rates and clear intentionality: diamond tattoo & silver badge). To contrast the influences of cue–category relationships and logical form, the conditional premise had two different ordering of terms:

Standard: If [Group], then [Marker]

e.g. “If a person is a member of the Bricklayers league, then he has blue eyes”

Switched: If [Marker], then [Group]

e.g. “If a person has blue eyes, then he is a member of the Bricklayers league”

The four possible second premises for a conditional syllogism (given a conditional premise “If p , then q ”: p , not- p , q , and not- q) were all used in this experiment. For the above (switched) example, the second (indicative) premises were:

the items in the catalogue). Participants were told to assume that the first two premises were true, and that each conclusion should be based on what necessarily follows from the premises. Participants were run in small groups ranging in size from 2–15 persons. There was no time limit and all participants finished within 15 minutes.

Results and discussion

As in Experiment 1, the responses of each participant to the two different sets of content words were quite similar, so these two sets were collapsed to simplify the results. Three issues will be addressed in this section: (1) Are the logical categories of the made inferences based on category validity reflexive to the ordering of the conditional premise (i.e., MP and MT when given a standard order conditional and AC and DA when given a switched order conditional)? (2) As the cues used as markers become more suitable as social group markers (with low base rates and clear intentionality) do inferences based on cue-validity increase? (3) Overall, are more cue-validity-based inferences made when the conditional premise information is in the switched order, as compared to the standard order? To evaluate these predictions, the frequencies with which different conclusions were drawn by participants, organised by the types of content in the syllogisms and the order of items in the conditional premises, are shown in Figures 2 and 3.

Effects of category validity. When the conditional premises are in the standard order (Figure 2), the aggregated frequencies of inferences follow a pattern similar to that prescribed by formal logic (i.e., MP and MT are by far the most common inferences). However, when conditional premises are in the switched order (Figure 3), the DA and AC inferences are the most frequent. In short, the number of people drawing these inferences, which are all based on category validity, is very high and very consistent. The very high number of (DA and AC) inferences in the switched conditional condition is somewhat surprising, but there are clear indications that these inferences are driven by the same forces as the category-validity-based inferences (MP and MT) in the standard conditional condition. Differences of proportion tests show that, when these inferences are compared based on their status relative to category validity (i.e., AC “switched” vs MP “standard” and DA “switched” vs MT “standard”), they are not significantly different.

Effects of cue validity. Two predictions were made regarding the effects of cue validity. The first was that better quality group markers would promote more inferences based on cue validity. The second prediction was that more of these associated inferences would be made when the conditional premise information was in the “switched” order. To assess both of these predictions, the overall

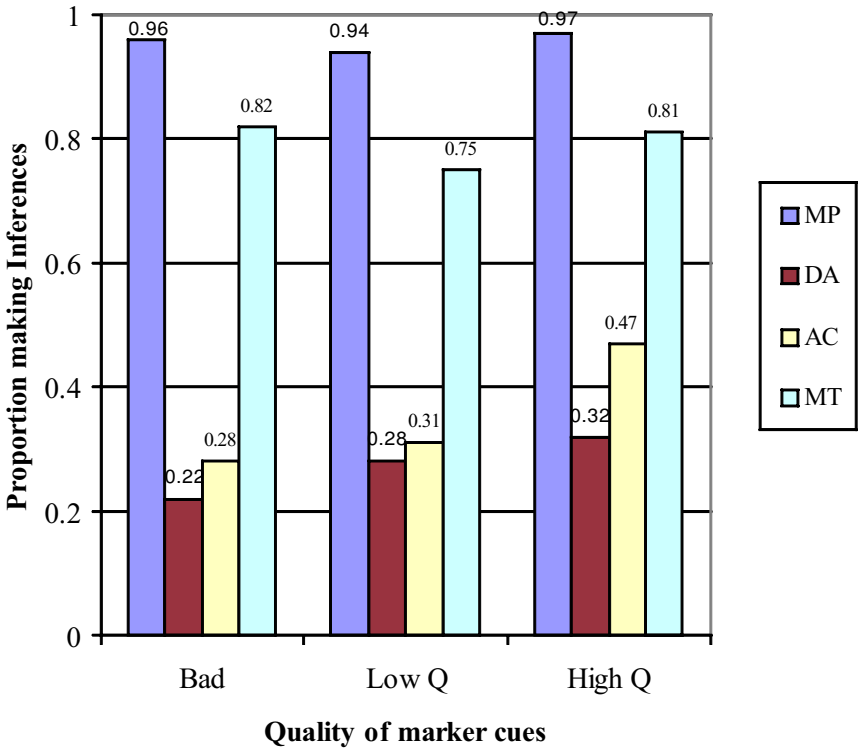
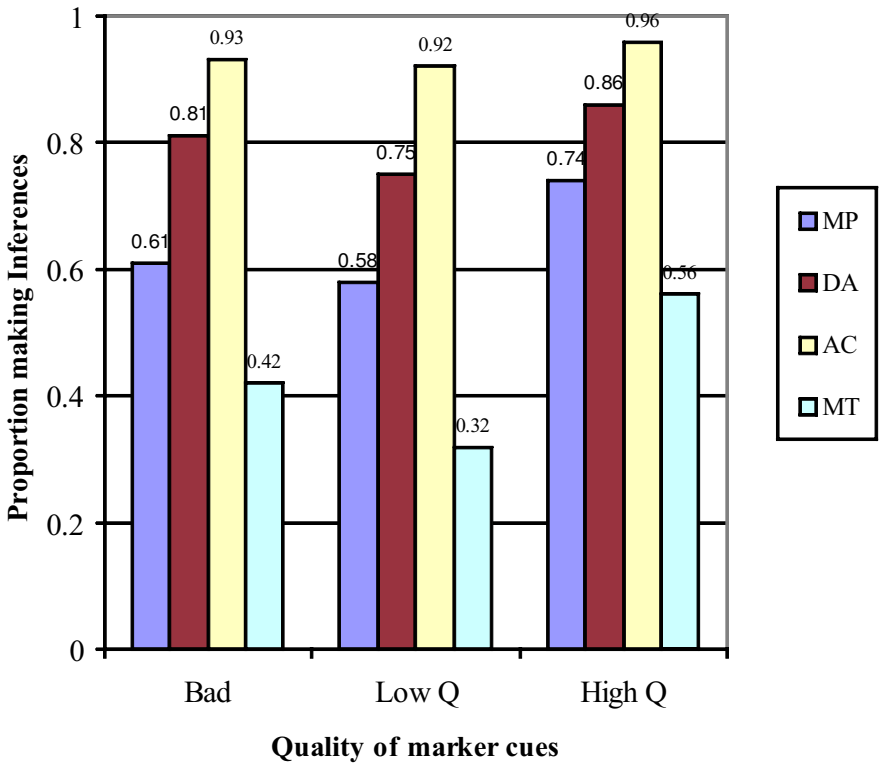


Figure 2. Experiment 2 results when the initial conditional premise is in the standard order (“If in the group, then one has the marker”). Aggregate proportions of inferences (MP = modus ponens, DA = denial of the antecedent, AC = affirmation of the consequent, MT = modus tollens) for different types of marker cues.

numbers of cue-validity-based inferences of each participant (DA and AC in the standard conditional premise condition, and MP and MT in the switched conditional premise condition) were counted. As each participant actually answered each marker type/conditional type combination twice (with the two different specific contents), a total of four possible inferences of this type could be made by each person, for each marker/conditional combination. This resulted in one “inferential strength” score for each participant, for each marker/conditional combination (with a range of 0–4). The means of these scores show that more inferences overall were drawn with “high quality” markers (1.58 in standard order, 2.58 in switched order) than with the other two types of contents (Low quality markers: 1.16 in standard order, 1.81 in switched order; Bad markers: 1.00 in standard order, 2.06 in switched order). A 2×3 repeated measures ANOVA was run on these scores, yielding a significant main effect for the type of marker in the syllogisms, $F(2, 70) = 10.51, p < .001, \eta = .23$, as well



Figures 3. Experiment 2 results when the initial conditional premise is in the switched order (“If one has the marker, then they are in the group”). Aggregate proportions of inferences (MP = modus ponens, DA = denial of the antecedent, AC = affirmation of the consequent, MT = modus tollens) for different types of marker cues.

as a significant main effect for the order of the terms (standard/switched) in the conditional premise, $F(1, 35) = 18.99, p = .001, \eta = .35$, but no interaction, $F(2, 70) = 1.58, p = .213$. Subsequent directional t -tests revealed that bad and low quality markers led to inferences significantly less often than high quality marker cues, with small to medium effect sizes. For standard order conditional premises: high quality vs low quality markers, $t(35) = 2.12, p = .02, d = .32$; high quality vs bad markers, $t(35) = 3.17, p = .001, d = .45$; low quality vs bad markers, $t(35) = 0.95, p = .18, d = .14$; for switched order conditional premises: high quality vs low quality markers, $t(35) = 1.68, p = .05, d = .35$; high quality vs bad markers, $t(35) = 1.14, p = .13, d = .22$; low quality vs bad markers, $t(35) = -0.74, p = .23, d = .17$.

Individual patterns of responses. Another approach to these results is to look at the patterns of responses made by each person. Because each participant

received all four possible second (indicative) premises for any particular content, the patterns of inferences drawn could be discerned for each participant⁴. As suggested by the previous measure of inferential strength, we should expect that the items with “high quality” marker cues will produce completely biconditional interpretation more often than other contents. That is, the drawing of all four inferences (MP, DA, AC, and MT) should be most common for contents that include “high quality” marker cues, as opposed to other contents (“Bad” and “low quality” marker cues).

The responses of each participant were coded by the types of markers and the different conditional premise orders, and participants’ responses were averaged across the two versions of premise sets with the same type of content and same conditional premise order (for instance, high quality marker traits with the standard [group-marker] conditional premise). The percentages of participants who produced each type of pattern, for patterns in which at least 15% of participants in any condition produced that pattern, are shown in Table 3. This criterion of 15% response rate eliminates all but five patterns of responses.

The pattern of all four inferences (MP, DA, AC, and MT) was made more frequently with “high quality” markers than with other types of marker cues (standard order: .08 vs .18: $z = 1.78$ $p = .04$, $h = .30$; .10 vs .18: $z = 1.38$ $p = .08$, $h = .23$; and switched: .24 vs .39: $z = 1.94$ $p = .03$, $h = .33$; .31 vs .39: $z = 1.01$ $p = .16$, $h = .17$). Although only half of these comparisons show significant differences (with effect sizes in the “small” range; between .17 and .33), they are all in the predicted relationships. There also appears to be a very pronounced tendency for participants to produce a specific pattern of MP and MT inferences

TABLE 3
Experiment 2, responses

	<i>Standard Premise Order</i> (Group-Marker)			<i>Switched Premise Order</i> (Marker-Group)		
	<i>Bad</i>	<i>Low Q</i>	<i>High Q</i>	<i>Bad</i>	<i>Low Q</i>	<i>High Q</i>
DA, AC	–	–	–	.18	.25	.17
MP, DA, AC	.04	.03	.06	.19	.19	.22
MP, MT	.51	.39	.38	–	–	.03
MP, AC, MT	.11	.11	.17	.03	.03	.07
<i>MP, AC, DA, MT</i>	.08	.10	.17	.31	.24	.39

Proportion of participants producing selected response patterns in Experiment 2, presented in terms of the conditional premises given and the type of contents used (modal responses are in bold and percentages of participants making all four inferences are in italics).

⁴This generates results that appear much like those obtainable with the Wason selection task, but are reporting the inferences made, rather than those inferences checked for violations of the conditional rule.

for syllogisms that have a standard order conditional premise, and a reversed pattern of inferences (DA and AC) for syllogisms that have a switched order conditional premise. This pattern—and specifically the reversed nature of the pattern across the two types of premises—indicates that category validity is a primary factor driving these inferences made about social groups, even when couched within a formal logic task.

EXPERIMENT 3

The previous two experiments indicate that, contrary to traditional theories of reasoning (i.e., mental models and mental rules theories), reasoning is significantly guided by the existence and nature of cue–category relationships expressed in the premises. In short, the number and types of inferences made by participants are better predicted by probabilistic models of human reasoning, as opposed to deductive competence models. However, these experiments do not distinguish very clearly between some alternative ways of understanding these effects. One interpretation is that there are inferential mechanisms designed to deal with social groups and markers of group membership. This would be a relatively narrow domain of application, and would fit well with the evolutionary model of reasoning (Tooby & Cosmides, 1988, 1989). On the other hand, the results of Experiments 1 and 2 can also be explained by more general inferential processes that deal with broader domains. For instance, Cummins and colleagues (Cummins, 1995; Cummins, Lubart, Alsknis, & Rist, 1991) suggest a domain of causal relationships, Evans, et al. (1993) suggest a process in terms of subjective expected utility, Kirby (1994) describes a process in the language of signal detection theory, and Oaksford and Chater (e.g., 1994) apply a Bayesian “rational analysis” to human reasoning. Within all these accounts, reasoning tends to involve elements of goal-directed, pragmatic judgements that are based on how well a cue predicts a category (and vice versa), although each account sees the conceptual domain somewhat differently.

Experimental distinction between all of these approaches is difficult to achieve, not only because there are several different views, but because they actually involve interconnected ideas and overlapping domains. For example, a narrowly domain-specific mechanism for reasoning about coalitions would have to entail, at the algorithmic level, some process such as signal detection, Bayesian analysis, or calculations of expected utilities. This final experiment is an initial attempt to discriminate between some of these models, using the phenomenon of entitativity. The entitativity of social groups clearly should be a factor in reasoning about coalitional markers if such reasoning is governed by a narrow domain-specific ability for coalitional reasoning. Entitativity appears to be a phenomenon fairly specific to social groups, reflecting the shared objectives and mental states of group members within a coalition. In other words, entitativity appears to be a part of any evolved coalitional psychology. If coalitional markers

are primarily cues of this “strong” form of coalitions, then non-entitative groups should be less able to support inferences of membership based on marker cues. On the other hand, group entitativity should be less relevant from a more general process model; perceptions of entitativity should not affect the general category–cue relationship that exists.

Participants

A total of 80 undergraduate students (16 males and 64 females) at a southeastern American university participated in this study for extra credit in psychology courses. The average age of the participants was 22.4 years old.

Materials and procedure

Experiment 3 returns to the use of categorical syllogisms and a portion of the stimuli used in Experiment 1. Each participant was given a booklet containing an instruction sheet and 17 test items, with each test item consisting of two initial premises and then six possible conclusions from which participants were asked to select the conclusion they believed to be correct. All participants selected just one of the conclusions for each item. As in Experiment 1, 11 test items were distractor items, involving social groups but with various different types of premises and conclusions. The remaining six test items (see Table 4) were all of the same logical figure: undistributed middle term (UMT) premises. The content of these items was systematically varied in three ways: the quality of the cues were varied (3 levels) using the same sets of contents (2 contents) as in Experiment 1 (i.e., eyes / bloodshot eyes / serpent tattoo & ears / attached earlobes / “P” shaved above ear). The nature of the social groups was also varied (2 types) as being either “weak” in entitativity or “strong” in entitativity. The relative level of entitativity of the groups was derived from previous research (Lickel, et al., 2000) on perceptions of group entitativity.

Each participant was presented, interspersed with the distractor syllogisms, with the two different contents with opposite levels of entitativity (i.e., half the participants saw content version 1 associated with a weak entitativity group and content version 2 associated with a strong entitativity group; the other participants saw the reverse pattern). Thus, no participants saw two syllogisms that repeated contents and every participant saw exactly half of all the test stimuli used in this study. Following the two premises, several possible conclusions were given, and participants were asked to indicate which conclusion they believed to be correct. A complete example of a test item (weak entitativity/high quality cue) is:

All people who are in line at the bank have a serpent tattoo

Brian has a serpent tattoo

(A) Therefore, all people who are in line at the bank are Brians

- (B) Therefore, Brian is in line at the bank
 (C) Therefore, some people who are in line at the bank are Brians
 (D) Therefore, some Brians are in line at the bank
 (E) Therefore, Brian is not in line at the bank
 (F) Cannot reach a conclusion

Participants were run in small groups ranging in size from 2–15 persons. There was no time limit and all subjects completed the booklet in less than 15 minutes.

Results

The results of Experiment 3 are shown in Table 5 as the proportions of participants making each response in each condition, and in Figure 4 as the proportions of deductively invalid biconditional inferences for each item. As in Experiment 1, there is a significant increase in the number of biconditional UMT errors committed as the quality of the marker cue improves, $F(2, 158) = 19.349$, $p < .001$, $\eta = .20$; all pairwise comparisons were significant ($p < .05$) in the predicted directions with the exception of low vs high quality cues with strong entitativity groups. Neither the main effect for level of entitativity, $F(1, 76) = 1.80$, $p = .183$, $\eta = .02$, nor the interaction of these two factors, $F(2, 158) = 1.19$, $p = .309$, $\eta = .02$, were statistically significant. Figure 4 suggests that, although strongly entitative groups may have some very slight effects on cue validity at intermediate levels and strong entitativity groups tend to elicit more UMT errors than weak entitativity groups overall, the results as a whole do not show that group entitativity was a significant factor in reasoning about social groups. In summary, the results of Experiment 3 again support a probabilistic process of

TABLE 5
Experiment 3, responses

	<i>All C-A</i>	<i>All A-C</i>	<i>Some C-A</i>	<i>Some A-C</i>	<i>No A-C</i>	<i>None</i>
<i>Weak Entitativity</i>						
<i>Bad Cues</i>	0.013	0.288	0.025	–	–	0.675
<i>Low Q cue</i>	0.013	0.375	0.025	0.013	–	0.575
<i>High Q cues</i>	–	0.475	0.013	0.025	0.013	0.475
<i>Strong Entitativity</i>						
<i>Bad Cues</i>	0.025	0.250	0.038	0.038	–	0.650
<i>Low Q cues</i>	0.013	0.438	0.025	0.050	–	0.475
<i>High Q cues</i>	–	0.500	0.025	–	–	0.475

Proportions of responses for undistributed middle term (UMT) premise sets in Experiment 3 (collapsed across the two types of content words). Each row shows data for a specific content of marker type (Bad, Low quality, and High quality) and group (Weak and Strong entitativity). Each column shows the proportion of participants selecting a particular conclusion). Logically correct responses are in italics, and modal responses are in bold.

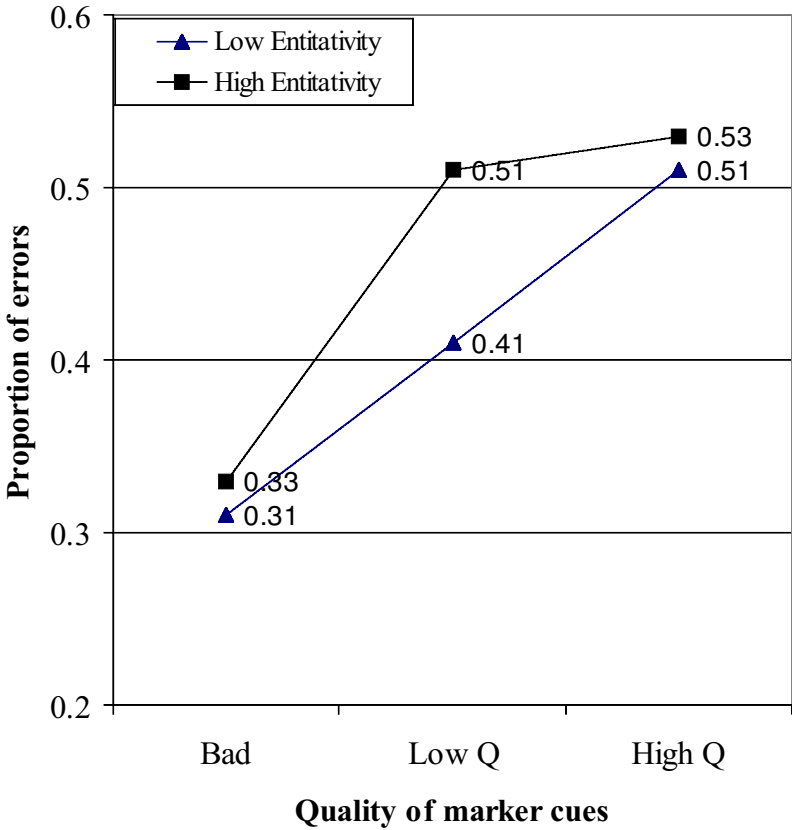


Figure 4. Experiment 3 results: Proportions of undistributed middle term errors that could be attributed to a biconditional interpretation (All A-C, Some A-C, Some C-A) with the data collapsed across the two different types of content.

using cue validities within contexts that have traditionally been considered deductive reasoning tasks. The results also fail to provide evidence for the existence of a narrow domain-specific ability for probabilistic reasoning just about coalitions and coalitional markers.

GENERAL DISCUSSION

The experiments reported here all support the thesis that reasoning situations involving social groups and associated marker cues tend to elicit reasoning not based primarily in deductive logic but rather founded in probabilistic inference. This finding holds across both categorical reasoning tasks (Experiments 1 and 3) and conditional reasoning tasks (Experiment 2). This

finding also holds across different types of social groups (Experiment 3). This section discusses the implications of these findings for various views of human reasoning, as well as for the social phenomena from which the reasoning content was drawn.

Implications for Deductive Reasoning Theories

In their basic forms the mental models and mental rules theories of human reasoning, both of which are based on a normative model of deductive logic, are unable to account for these data. Both of these theories, however, have developed ways of accounting for content-related reasoning effects. Can these accessory explanations adequately explain the present findings in a meaningful way?

The mental models approach of Johnson-Laird and colleagues (e.g., Johnson-Laird, 1983, 1985; Johnson-Laird & Bara, 1984; Johnson-Laird & Byrne, 1991; Johnson-Laird & Steedman, 1978) proposes that people form an internal representation of the reasoning problem (a model), generate a tentative conclusion, and then search for alternative possible models that would falsify the original representation. As there is some uncertainty in these representations (that is, premises can be interpreted in more than one way), there is room for people to generate alternative models for a single syllogism. The theory of mental models is a Popperian hypothesis falsification strategy (Popper, 1959) with an implicit goal of establishing the truth (though not necessarily via formal logic), assumed to be hampered by a human inability to always perform more complex computations. What initial model is generated can be influenced by semantic knowledge, and how exhaustively the initial model is evaluated is determined by cognitive limitations (or limited effort). A number of researchers find the idea of mental models intuitively appealing, but it has also been noted that this system lacks a great deal of the computational specificity desired for a scientifically testable and falsifiable theory (Rips, 1986).

Recently the mental models approach has been explicitly exported to the topic of probabilistic reasoning (Johnson-Laird et al., 1999). One could argue that this extension of the mental models approach could account for the present results. However, the ambiguity inherent in the mental models approach becomes even more obvious in this context: the mental models account of probabilistic reasoning rests on three fundamental principles: truth, equiprobability, and proportionality. The equiprobability principle is most important here; it states that “each model represents an equiprobable alternative unless individuals have knowledge or beliefs to the contrary, in which case they will assign different probabilities to different models” (this is also a prerequisite for the proportionality principle; Johnson-Laird et al., 1999, p. 68). This equipotentiality principle works unproblematically with tasks that involve choosing marbles from an urn or cards out of hats, but in the real world (where normal human reasoning occurs) it is difficult to conceive of reasoning contents that involve no prior

“knowledge or beliefs” about probability (Brase, in press). In a very real sense, then, all the actual inferential “work” is done by routine violations of this equipotentiality principle (in the form of prior knowledge and beliefs). In other words, this is new terminology for “content effects” as an explanatory label. Although it is certainly possible for one to simply dismiss all these inferences as content effects, it seems more sensible to propose that explanatory theories and empirical research are needed to explore why these specific “contents” have these very particular “effects”.

Traditional viewpoints usually given as the alternative to mental models theory are the mental rules theories. Mental rules theories utilise syntactically based rule systems to describe how people reason. These rules may or may not be the rules of deductive logic, but they approximate deduction and attempt to serve the same purpose—finding the formal truth in situations. Deviations in human reasoning from the results expected by deductive logic are considered the anomalies in need of explanations. Thus, the algorithms postulated in mental rules theories, while formal-logical in nature, also include additional assumptions, for example see the algorithms proposed by Rips’ (1994) reasoning model, PSYCOP. These modifications and addenda to formal logic can be described as reflections of two thoughts: (1) that the human mind is in some way too limited to perform certain logical functions (Chapman & Chapman, 1959; Erickson, 1974, 1978; Revlin & Leirer, 1978; Revlis, 1975; Rips, 1994; Sells, 1936; Woodworth & Sells, 1935); and (2) that the mind contains accessory processes or constraints that interfere with logical processing or augment the assumed limited processing capabilities (e.g., Braine, 1978; Braine & O’Brien, 1991; Evans, et al., 1983). Either way, the situation is similar to that encountered within the mental models theory; the ability to explain the present results, as well as other “content effects”, does not lie with the core (deductive logic) ideas of the theory but with the accessories, for example, of causality and permission parameters in Rips’ model (1994) and pragmatic parameters in Brain and O’Brien’s model (1991).

What, then, of deductive logic based reasoning theories? People obviously have some ability to use deductive processes, and even the most stringent advocates of narrowly domain-specific reasoning abilities have said that, “We’re not arguing that there are no general rules” (John Tooby, in Bock & Cardew, 1997, p. 159). The use of general deductive rules, however, seem to be neither the favoured nor the default strategy for most people (for a similar conclusion regarding normative models of probability, see Gigerenzer, Todd, & the ABC Research Group, 1999).

The present studies notably used very general instructions to the participants (e.g., “Circle the letter corresponding to the conclusion you believe is correct”), and using instructions that more strictly (sometimes even explicitly) indicate that deductive reasoning is what the experimenter wants generally lead to more

deductively correct responses (Evans & Over, 1996; Newstead, Pollard, Evans, & Allen, 1992; Stevenson & Over, 1995; Vadeboncoeur & Markovits, 1999). However, that the context demands of reasoning situations can change people's responses should not obscure the fact that people quite easily manage to provide responses across all these different context situations. This means that there must be inferential processes (engaged by generic or no instructions) that operate in the absence of the engagement of any deductive abilities. This situation resembles what have been called *dual process* theories (Evans & Over, 1996; see, particularly, Stanovich, 1999; Stanovich & West, 2000): with one process being a more controlled, effortful, and analytic process of deductive reasoning, and the other a more automatic, implicit, and less cognitively demanding process (or set of processes). The later is probably more commonly used. Life, unlike psychology reasoning tasks, has no instructions, and so although we may have some deductive reasoning competence it is a more limited part of ordinary human reasoning than syllogism task results tend to suggest. Further research would be useful in understanding how and why these two proposed processes interact and are selectively utilised.

Other theories of human reasoning

The more automatic, implicit side of this dual process model may very well be more important for understanding normal human reasoning in the real world, but it is ironically less well researched and understood. The notion that this system may involve the use of probabilistic information has opened up a diverse array of theories regarding the nature of this system. There are very general models of this system, which tend to encompass all or nearly all probabilistic situations using Bayesian reasoning, signal detection theory, and models of subjected expected utility (Chater & Oaksford, 1999; Evans et al., 1993; Kirby, 1994; Oaksford & Chater, 1994; Oaksford, Chater, & Grainger, 1999). There are more narrow models of this process that focus on large classes of events (e.g., causality: Cummins, 1995; Cummins, et al., 1991; Fairley, Manktelow, & Over, 1999; Manktelow, Fairley, Kilpatrick, & Over, 2000; see also Gigerenzer et al., 1999). Finally, there are very narrow models of this process—so called “domain-specific” models—that focus on smaller classes of events (e.g., pragmatic reasoning schemas: Cheng & Holyoak, 1985; Cheng, Holyoak, Nisbett & Oliver, 1986; Holyoak & Cheng, 1995; Kroger, Cheng, & Holyoak, 1993; and evolved reasoning algorithms: Cosmides, 1989; Cosmides & Tooby, 1989, 1992, see also Cummins, 1996). There seems to be some broad recognition that, if human reasoning is not a unitary process, it is vitally important to determine the appropriate joints at which to carve human reasoning. Not only is this a fundamental issue in understanding how reasoning works, it is also a process of setting the research agenda for the future. Selecting one of these models as correct (or more likely) is beyond the scope of this paper, but I would suggest that

the use of more realistic contents, such as coalitions and coalitional markers, can be a helpful technique in this process.

On the Nature of Coalitions & Markers

The principle of parsimony states that, “All else being equal, the simpler explanation is usually the correct one.” While it can be debated whether all the probabilistic models of reasoning are empirically and theoretically of equal status, one sensible place to start would be to test the power of more narrowly applicable models against a default background of the more general probabilistic models. Research questions can thus be framed in terms of how plausible the general models are if they need to accommodate specific judgements about coalitions, markers, entitativity, and so on (i.e., are the models not equal in their predictive abilities?). Already there are a number social and situational reasoning phenomena that might usefully be exploited for this purpose.

A robust finding from Experiment 2 was that inferences (besides being more common overall with “high quality” marker cues) follow a distinct pattern. By far the most common inferences made are: (1) group membership based on the presence of a marker, and (2) lack of a marker based on non-membership in the associated group. Taken together this pattern suggests that coalitional markers may often elicit a perception of deontic entitlement; that group memberships enable but do not require the use of a coalitional marker. One direction in which to further explore this finding is to see if it generalises to all types of category–cue relationships (i.e., using non-social contents in probabilistic reasoning). Additionally, this finding also highlights a feature of markers that was not systematically explored here but may be of some significance: different markers have different characteristic properties.

Consider the similarities and differences between coalitional markers and human language. Both markers and language involve the use of a symbolic, shared representational system that enables an efficient and compact mode of transmission for complicated ideas. Coalitional markers differ from language, however, in that they allow for cryptic transmissions of information. Many coalitional markers that involve material items (e.g., badges, pins, identification cards) or stereotyped behaviours (e.g., special handshakes, gang signs) can be selectively displayed to or hidden from particular individuals, and can allow communication without sound or (in the case of behaviours) physical traces. Other coalitional markers seem to be designed specifically to prevent such spatial–temporal flexibility (e.g., tattoos, ritualistic scarring). These variations in markers may tell us something about the natures of the associated coalitions. Generally, transient markers appear to be associated more with *sub rosa* groups and individual members in unstable/unclear situations of power. Permanent markers appear to be associated with more stable, established groups and with threat/intimidation functions of markers. However, no systematic studies appear

to have been done on this issue. The permanence or transience of group markers may be useful in discriminating between general probabilistic models of reasoning, models of reasoning that focus on the domain of causal relationships (and the existence of alternative and disabling causes⁵; Byrne, 1989; Cummins, 1995; Cummins et al., 1991), and models of reasoning specifically about social groups and group markers.

Another thing that is important to note is that coalitional membership is a psychological property; one cannot directly perceive membership, but must infer its existence from other sources. In other words, the identification of coalitions and coalition members may be psychologically real, but it is not directly tangible. Similarly, one's own state of being a member of a coalition is fundamentally a mental state that cannot be observed directly by another person. Coalitional markers act as bridges that display, in an observable way, mental states. There are a number of issues regarding the nature of coalitional markers, but one of the more pressing is the issue of intentionality; must markers be produced intentionally? A number of examples of what appear to be markers are not, in fact, produced intentionally: skin colour, ethnic features, physical disabilities, and scars from various types of surgical, cultural, or drug-related activities. These examples are all aspects of the environment that can be used as cues to various things (including the beliefs of the cue possessors). But are these cues, in fact, markers? To clarify this situation it is important to distinguish between the person who produces a cue or marker and the person who perceives it. Pure coalitional markers are intentionally produced; they are an external, physical creations designed specifically to be a reflection of a group membership. Other, non-intentionally produced cues are sometimes treated as coalitional markers by observers. So, for instance, a person may be refused a job based on old heroin track-marks on his or her arms or the colour of his or her skin, because the employer perceives those cues as indicators of undesirable traits. The irony is that this type of scenario creates a self-fulfilling prophecy, in which the potential employee comes to see himself or herself as part of a particular group because of the membership cue chronically used by others (Kenrick, Neuberg, & Cialdini, 1999).

The history of humanity holds many examples that suggest that people have a difficult time *not* attributing intentionality and entitativity when encountering individuals with rare aspects of their appearance. Individuals (and even items or locations) have been routinely held to be affiliated with evil or divine forces based on unusual features of their appearance (as with unusual warts, freckles, and birthmarks being “witchmarks” in the Salem witch trials). Even social

⁵Within the terminology of reasoning about causal relationships, the inference blocking effects of alternative causes and disabling causes can be understood as additional pieces of information that reduce category and cue validity, respectively. Perhaps of equal interest is the naïve physics implied here; causality can often be inferred unless contraindicated.

groups that are usually considered very low in perceived entitativity (e.g., people standing at a bus stop) can be seen as a more cohesive group *because* they all have some common and rare feature that appears to be a group maker. On the one hand, this indicates that it may be quite difficult to experimentally dissociate very low base rate cues from perceptions of intentionality and entitativity. On the other hand, this also suggests that there are powerful and socially important inference processes at work in this domain (see also recent work by Malle & Knobe, 1997, on peoples' intuitive understanding of intentionality).

This process can also help to explain some otherwise puzzling social phenomena of "marker violations," in which: (a) an individual with a particular (immutable) feature is criticised for not holding the beliefs and objectives of a social group for which that feature is alleged to be a coalitional marker (e.g., someone with dark skin not being truly "black"), and (b) an individual who clearly *lacks* a particular (immutable) feature is criticised for holding the beliefs and objectives of a social group for which that feature is alleged to be a coalitional marker (e.g., a "white" person "acting black"). Finally, another curious phenomenon is the adoption of coalition markers by non-group members for fashion reasons. These "posers" have recently been most noticeable in relation to street gangs. Baggy clothes, clothes of certain colours, and systematically placed sports-team icons have all been used as street gang markers and then lost their "marker value" as they became general fashions. Understanding how and why these markers are adopted by non-members as fashion statements can have important theoretical and practical implications.

In conclusion, in intergroup relations there appears to be a rich source of ideas and examples of what has been called "practical reasoning" (as distinct from any abstract, logical reasoning abilities). At present there are a number of theoretical accounts of how such practical reasoning occurs, and a new testing ground for these accounts may prove useful.

Manuscript received 14 July 2000

Revised manuscript received 8 March 2001

REFERENCES

- Beach, L.R. (1964). Cue probabilism and inference behavior. *Psychological Monographs*, 78, 1–20.
- Bock, G.R., & Cardew, G. (1997). *Characterizing human psychological adaptations, Symposium held at the Ciba Foundation on 29–31 October, 1996*. Chichester, UK: John Wiley & Sons.
- Bohra, K.A. (1979). Social perception of ingroup and outgroup members in India. *Journal of Social Psychology*, 108, 9–12.
- Braine, M.D.S. (1978). On the relation between the natural logic of reasoning and standard logic. *Psychological Review*, 85, 1–21.
- Braine, M.D.S., & O'Brien, D.P. (1991). A theory of if: A lexical entry, reasoning program, and pragmatic principles. *Psychological Review*, 98, 182–203.

- Brase, G.L. (in press). Theoretical Note: Ecological and Evolutionary Validity: Comments on Johnson-Laird, Legrenzi, Girotto, Legrenzi, & Caverni's (1999) Mental Model Theory of Extensional Reasoning. *Psychological Review*.
- Byrne, R.M. (1989). Suppressing valid inferences with conditionals. *Cognition*, 31, 61–83.
- Campbell, D.T. (1958). Common fate, similarity, and other indices of the status of aggregates of persons as social entities. *Behavioral Science*, 3, 14–25.
- Carroll, L. [pseud.] (1971/1865). *Alice's Adventures in Wonderland*. London: Oxford University Press.
- Chapman, L.J., & Chapman, J.P. (1959). Atmosphere effect reexamined. *Journal of Experimental Psychology*, 58, 220–226.
- Chater, N., & Oaksford, M. (1999). The probability heuristics model of syllogistic reasoning. *Cognitive Psychology*, 38, 191–258.
- Cheng, P., & Holyoak, K. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, 17, 391–416.
- Cheng, P.W., Holyoak, K.J., Nisbett, R.E., & Oliver, L.M. (1986). Pragmatic versus syntactic approaches to training deductive reasoning. *Cognitive Psychology*, 18, 293–328.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd Edn.). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition*, 31, 187–276.
- Cosmides, L., & Tooby, J. (1989). Evolutionary psychology and the generation of culture: II. Case study: A computational theory of social exchange. *Ethology & Sociobiology*, 10, 51–97.
- Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J.H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: evolutionary psychology and the generation of culture* (pp. 163–228). Oxford: Oxford University Press.
- Cummins, D.D. (1995). Naive theories and causal deduction. *Memory & Cognition*, 23(5), 646–658.
- Cummins, D.D. (1996). Dominance hierarchies and the evolution of human reasoning. *Minds & Machines*, 6(4), 463–480.
- Cummins, D.D., Lubart, T., Alsknis, O., & Rist, R. (1991). Conditional reasoning and causation. *Memory & Cognition*, 19(3), 274–282.
- Davis, D.D. (1985). Hereditary emblems: Material culture in the context of social change. *Journal of Anthropological Archaeology*, 4, 149–176.
- Erickson J.R. (1974). A set analysis theory of behavior in formal syllogistic reasoning tasks. In R.L. Solso (Ed.), *Theories of cognitive psychology: The Loyola symposium*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Erickson J.R. (1978). Research on syllogistic reasoning. In R. Revlin & R.E. Mayer (Eds.), *Human reasoning*. New York: Wiley/Winston.
- Evans, J.St.B.T., Barston, J., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory & Cognition*, 11, 295–306.
- Evans, J.St.B.T., & Over, D.E. (1996). *Rationality and reasoning*. Hove, UK: Psychology Press.
- Evans, J.St.B.T., Over, D.E., & Manktelow, K.I. (1993). Reasoning, decision making, and rationality. *Cognition*, 49, 165–187.
- Fairley, N., Manktelow, K., & Over, D. (1999). Necessity, sufficiency, and perspective effects in causal conditional reasoning. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 52A, 771–790.
- Fiddick, L., Cosmides, L., & Tooby, J. (1995, June). *Does the mind distinguish between social contracts and precautions?* Paper presented at the meeting of the Human Behavior and Evolution Society, Santa Barbara, CA.

- Fiddick, L., Cosmides, L., & Tooby, J. (1996, June). *By accident, or by design?* Paper presented at the meeting of the Human Behavior and Evolution Society, Chicago, IL.
- Gallois, C., Callan, V.J., & Parslow, L.A. (1982). Evaluations of four ethnic groups: Level of ethnocentrism, favourability, and social distance. *Australian Journal of Psychology*, *34*, 369–374.
- Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond “heuristics and biases”. *European Review of Social Psychology*, *2*, 83–115.
- Gigerenzer, G., Todd, P.M., & the ABC Research Group (1999). *Simple heuristics that make us smart*. Oxford: Oxford University Press.
- Greenberg, J., & Rosenfield, D. (1979). Whites’ ethnocentrism and their attributions for the behavior of Blacks: A motivational bias. *Journal of Personality*, *47*, 643–657.
- Hamilton, D.L., & Sherman, S.J. (1996). Perceiving persons and groups. *Psychological Review*, *103*, 336–355.
- Hirschfeld, L.A. (1996). *Race in the making: cognition, culture, and the child’s construction of human kinds*. Cambridge, MA: MIT Press.
- Holyoak, K.J., & Cheng, P.W. (1995). Pragmatic reasoning about human voluntary action: Evidence from Wason’s selection task. In S.E. Newstead & J.St.B.T. Evans (Eds.), *Perspectives on thinking and reasoning: Essays in honour of Peter Wason* (pp. 67–89). Hove, UK: Lawrence Erlbaum Associates Ltd.
- Johnson-Laird, P.N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.
- Johnson-Laird, P.N. (1985). Deductive reasoning ability. In R.J. Sternberg (Ed.), *Human abilities: An information processing approach*. New York: Freeman.
- Johnson-Laird, P.N., & Bara, B. (1984). Syllogistic inference. *Cognition*, *16*, 1–61.
- Johnson-Laird, P.N., & Byrne, R.M.J. (1991). *Deduction*. Hove, UK: Lawrence Erlbaum Associates Ltd.
- Johnson-Laird, P.N., Legrenzi, P., Girotto, V., Legrenzi, M.S., & Caverni, J-P. (1999). Naive probability: A mental model theory of extensional reasoning. *Psychological Review*, *106*, 62–88.
- Johnson-Laird, P.N., & Steedman, M. (Eds.) (1978). The psychology of syllogisms. *Cognitive Psychology*, *10*, 64–99.
- Kenrick, D.T., Neuberg, S.L., & Cialdini, R.B. (1999). *Social psychology: Unraveling the mystery*. Boston, MA: Allyn & Bacon.
- Kirby, K.N. (1994). Probabilities and utilities of fictional outcomes in Wason’s four-card selection task. *Cognition*, *51*, 1–28.
- Kroger, J.K., Cheng, P.W., & Holyoak, K.J. (1993). Evoking the permission schema: The impact of explicit negation and a violation-checking context. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *46A*, 615–635.
- LeVine, R.A., & Campbell, D.T. (1972). *Ethnocentrism: Theories of conflict, ethnic attitudes, and group behavior*. New York: John Wiley & Sons.
- Lickel, B., Hamilton, D.L., Wierzchowska, G., Lewis, A., Sherman, S.J., & Uhles, A.N. (2000). Varieties of groups and the perception of group entitativity. *Journal of Personality & Social Psychology*, *78*, 223–246.
- Malle, B.F., & Knobe, J. (1997). The folk concept of intentionality. *Journal of Experimental Social Psychology*, *33*, 101–121.
- Marcus, M.I. (1993). Incorporating the body: Adornment, gender, and social identity in ancient Iran. *Cambridge Archaeological Journal*, *3*, 157–178.
- Manktelow, K.I., Fairley, N., Kilpatrick, S.G., & Over, D.E. (2000). Pragmatics and strategies for practical reasoning. In W. Schaeken & G. DeVooght, (Eds.), *Deductive reasoning and strategies* (pp. 111–130). Mahwah, NJ: Lawrence Erlbaum Associates Inc.

- Mayer, R.E. (1983). *Thinking, problem solving, cognition*. New York: W.H. Freeman.
- Mullen, B., Brown, R., & Smith, C. (1992). Ingroup bias as a function of salience, relevance, and status: An integration. *European Journal of Social Psychology*, 22, 103–122.
- Newstead, S.E., Pollard, P., Evans, J.St.B.T., & Allen, J.L. (1992). The source of belief bias effects in syllogistic reasoning. *Cognition*, 45, 257–284.
- Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review*, 101, 608–631.
- Oaksford, M.R., Chater, N., & Grainger, B. (1999). Probabilistic effects in data selection. *Thinking and Reasoning*, 5, 193–243.
- Peters, W. (1987). *A class divided: Then and now*. New Haven, CT: Yale University Press.
- Popper, K.R. (1959). *The logic of scientific discovery*. New York: Harper & Row.
- Reed, S.K. (1972). Pattern recognition and categorization. *Cognitive Psychology*, 3, 382–407.
- Revlín, R., & Leirer, V.O. (1978). The effect of personal biases on syllogistic reasoning: Rational decisions from personalized representations. In R. Revlín & R.E. Mayer (Eds.), *Human reasoning*. New York: Wiley/Winston.
- Revlis, R. (1975). Two models of syllogistic reasoning: Feature selection and conversion. *Journal of Verbal Learning and Verbal Behavior*, 14, 180–195.
- Rips, L.J. (1986). Mental muddles. In M. Brand & R.M. Harnish (Eds.), *The representation of knowledge and belief. Arizona colloquium in cognition* (pp. 258–286). Tucson, AZ: University of Arizona Press.
- Rips, L.J. (1994). *The psychology of proof: Deductive reasoning in human thinking*. Cambridge, MA: The MIT Press.
- Rutherford, M.D., Cosmides, L., & Tooby, J. (1996, June). *Adaptive sex differences in reasoning about self defense*. Paper presented at the meeting of the Human Behavior and Evolution Society, Chicago, IL.
- Sande, G.N., Goethals, G.R., Ferrari, L., & Worth, L.T. (1989). Value-guided attributions: Maintaining the moral self-image and the diabolical enemy-image. *Journal of Social Issues*, 45, 91–118.
- Searle, J.R. (1995). *The construction of social reality*. London: Penguin Books.
- Sells, S.B. (1936). The atmosphere effect: An experimental study of reasoning. *Archives of Psychology*, No. 200.
- Sherif, M., Harvey, O.J., White, B.J., Hood, W.R., & Sherif, C.W. (1961). *Intergroup conflict and cooperation: The Robbers Cave experiment*. Normal, OK: University Book Exchange.
- Stanovich, K.E. (1999). *Who is rational? Studies of individual differences in reasoning*. Mahwah, NJ: Lawrence Erlbaum Associates Inc.
- Stanovich, K.E., & West, R.F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, 23, 645–726.
- Stevenson, R.J., & Over, D.E. (1995). Deduction from uncertain premises. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 48A, 613–643.
- Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American*, 223, 96–102.
- Tajfel, H. (1982). Social psychology of intergroup relations. *Annual Review of Psychology*, 33, 1–39.
- Taylor, D.M., & Jaggi, V. (1974). Ethnocentrism and causal attribution in a south Indian context. *Journal of Cross-Cultural Psychology*, 5, 162–171.
- Tooby, J., & Cosmides, L. (1988). The evolution of war and its cognitive foundations. *Institute for Evolutionary Studies Technical Report 88-1*.
- Tooby, J., & Cosmides, L., (August, 1989). *Are there different kinds of cooperation and separate Darwinian algorithms for each?* Paper presented at the meeting of the Human Behavior and Evolution Society, Evanston, IL.

- Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, *90*, 293–315.
- Vadeboncoeur, I., & Markovits, H. (1999). The effect of instructions and information retrieval on accepting the premises in a conditional reasoning task. *Thinking and Reasoning*, *5*, 97–113.
- Wilkins, M.C. (1928). The effect of changed material on ability to do formal syllogistic reasoning. *Archives of Psychology*, vol. 102.
- Woodworth, R.S., & Sells, S.B. (1935). An atmosphere effect in formal syllogistic reasoning. *Journal of Experimental Psychology*, *18*, 451–560.