Research Article

THE IMPACT OF NECESSITY AND SUFFICIENCY IN THE WASON FOUR-CARD SELECTION TASK

Woo-kyoung Ahn[1] and Loranel M. Graham[2]

[1]Yale University and [2]University of Louisville

Abstract—*Performance in the Wason card selection task is often improved given thematic content. Such content effects have been considered evidence against human rationality. We propose that the role of content lies in specifying premises underlying "if* P*, then* Q*" rules. Unlike thematic rules, abstract conditional rules do not explicitly provide material interpretation (*P *is a sufficient but not necessarily a necessary condition for* Q*), resulting in nonnormative responses. When necessity–sufficiency relations were explicated, normative responses were elicited and effects of other logically irrelevant components disappeared. The results suggest that content effects are compatible with human rationality.*

The question of whether human reasoning is based on formal logic-based, domain-general mechanisms or on pragmatic content-based, domain-specific mechanisms has served as the focus of a long-standing debate in psychology (e.g., Cheng & Holyoak, 1985; Henle, 1962; Rips, 1994). At the heart of the ongoing debate on this issue lies Wason's four-card selection task.

The Wason selection task involves presenting an "if *P,* then *Q"* rule along with cases that need to be selected to determine if the rule holds or if it has been violated. In the original abstract version of the task, the rule, "If a card has a vowel on one side, it has an even number on the other side," was presented along with four cards showing an *A* (*P*), a *G* (not-*P*), a *4* (*Q*) and a *7* (not-*Q*). According to formal logic, *P* and not-*Q* cards should be selected, because no other pairing can falsify the rule. Although the task is superficially straightforward, people's performance on it was surprisingly poor. Most people selected the *P* and *Q* cards or the *P* card alone (e.g., Wason, 1966, 1968), suggesting that people cannot reason logically.

Later studies demonstrated dramatic improvement in performance when rules involved thematic content, such as "If a person drinks beer, that person must be over 21" (e.g., Evans, 1982; Griggs, 1983). For supporters of human rationality, such improvement seemed to convey even worse news about humans' reasoning ability, because if people reasoned logically, then logically irrelevant components (i.e., content) should have no influence on performance. Hence, it was argued that theories based on formal rules cannot account for why people fail with an abstract rule but not with thematic ones (Manktelow & Over, 1987). Subsequently, the content effect tended to be explained by domain-specific mechanisms.

For instance, Cheng and Holyoak (1985) proposed that in reasoning with conditional rules, people rely on pragmatic reasoning schemas (PRSs), which are sets of generalized rules consisting of goals (e.g., desirable actions) and relationships to these goals (e.g., preconditions). Table 1 presents the permission PRS. Note that P4 in

the permission schema readily provides information that a *P* and not-*Q* combination (the correct answer for the selection task) is a rule violation. According to the PRS theory, thematic content rules facilitate performance because they invoke a PRS.

Cosmides (1989) offered another domain-specific theory, claiming that humans have evolved to be sensitive to violations of social-contract rules that take the form of "If you take a benefit, you pay a cost." Violators, or "cheaters," are those who take the benefit (*P*) without paying the cost (not-*Q*). According to Cosmides, when the content triggers this cheater-detecting algorithm, performance on the selection task is improved. Gigerenzer and Hug (1992) argued further that cases of cheating would vary with the reasoner's perspective. For instance, in the rule "If an employee works on the weekend, then that employee gets a day off during the week," the employer would consider not working on the weekend (not-*P*) and getting a day off (*Q*) as cheating, whereas the employee would consider working on the weekend (*P*) and not getting a day off (not-*Q*) as cheating. Indeed, given identical rules, their participants selected *P* and not-*Q* cards from one perspective, but not-*P* and *Q* cards from the opposite perspective. This effect of a perspective shift was presented as evidence against a strictly logic-based reasoning mechanism and against the PRS account (but see Holyoak & Cheng, 1995).

The aim of the current article is to reopen the debate on domain-specific versus domain-general mechanisms of reasoning and to account for the effect of both content and perspective in terms of formal, domain-general components. Recall that in formal logic, "if *P,* then *Q*" means that *P* is a sufficient condition for *Q* (i.e., *P* will always result in *Q*), but *P* does not have to be a necessary condition for *Q* (i.e., *Q* may occur without *P*). Selecting *P* and not-*Q* cards is the normative answer for the selection task only under this so-called material interpretation.

Note, however, that everyday usage of if-then clauses does not always conform to this interpretation. For instance, consider a person who says, "If you eat food with high fat, you will have a heart attack." This person might mean that (and the listener might understand the person to mean that) if someone eats high-fat food, he or she is more likely to get a heart attack (i.e., eating high-fat food is not sufficient for a heart attack).

Indeed, "if *P,* then *Q*" statements in everyday casual conversation span all possible combinations of necessity–sufficiency relations between *P* and *Q,* as shown in Figure 1.[1] For each relation in Figure 1, an example rule is provided in italics. To explicate the necessity–sufficiency relations, appropriate modal auxiliaries (e.g., "may") are included in the rules. In this figure, segments inside each box (e.g., for the unnecessary-but-sufficient rule, *P* & *Q,* not-*P* & *Q,* and not-*P* & not-*Q*) indicate states that should exist under the corresponding conditional rule. Below each box, next to "Violations" are states that are

Address correspondence to Woo-kyoung Ahn, Department of Psychology, Yale University, 2 Hillhouse Ave., New Haven, CT 06520; e-mail: ahn@pantheon.yale.edu.

1. We do not claim that necessary and sufficient relations are psychological primitives because there are many possible ways of implementing them.

Copyright © 1999 American Psychological Society

Necessity and Sufficiency

Table 1. *Pragmatic reasoning schema for permission*

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.

---

impossible if the rule is true or obeyed. Normative selections, indicated by pointing fingers, correspond to cards that have the potential to reveal these violations. It should be noted that normative answers in the selection task vary as a function of necessity–sufficiency relations (see also Thompson, 1994, 1995).[2]

Now, imagine participants confronted with a novel rule like the letter-vowel one. They have no way of knowing which interpretation should be imposed, and consequently they might intuitively select cards that look relevant to the task or salient in the rule (e.g., *P* and *Q* cards; see Oaksford & Chater, 1994, for a rational analysis of these choices). In contrast, when thematic rules show facilitation in the selection task, they all indicate, through experimentally provided contexts or existing background knowledge, that *P* is unnecessary-but-sufficient for *Q*. For instance, all undergraduates know that a person drinking beer must be over age 21 (*P* is sufficient for *Q*) but a person drinking Coke may be over 21 (*P* is unnecessary for *Q*). We propose that the critical determinant for the content effect in the selection task is not the fact that it triggers domain-specific reasoning schemas or a cheater-detecting algorithm, but the fact that the thematic content allows people to impose the material interpretation of a given conditional rule (see also Rips, 1994, for a similar discussion).

To test this prediction, Experiment 1 manipulated the necessity–sufficiency relations in rules novel to participants, including the notorious letter-vowel rule as well as social-contract rules. We predicted that if a conditional rule is clearly presented as an example of an unnecessary-but-sufficient rule, the traditionally normative response of *P* and not-*Q* would increase even with abstract materials (Platt & Griggs, 1995). Furthermore, other types of normative responses should be obtained from other possible formulations of the necessity–sufficiency relations, as shown in Figure 1. Experiment 2 took this approach one step further and pitted the necessity–sufficiency relations against the effect of a perspective shift.

EXPERIMENT 1

Method

Four content materials were used. Three concerned social interactions ("If a player won the last game, that person buys a round of drinks for the others"; "If a person takes action A, they first satisfy pre-

---

2. We do not imply that these four possible formulations exhaust all possible meanings of if-then statements in everyday language. For instance, a rule "If you are interested, then there is a Hitchcock movie on TV" does not warrant deductions from antecedent to consequent, and its truth value is based on that of the consequent only (Johnson-Laird & Byrne, 1991).

---

condition P"; "If a student enrolls in Jazz Music, they enroll in Rock Music"), and one was the letter-number rule. For each material, four versions were created to manipulate necessity–sufficiency relations. The unnecessary-but-sufficient, necessary-but-insufficient, and necessary-and-sufficient versions were used for the experimental conditions. The fourth version left the necessity–sufficiency relations unspecified and served as the baseline condition[3] (see Appendix A for the four versions and the cards used for the letter-number material). In creating three experimental versions, we added "might" or "must" in the consequent, and context stories were added to clarify the necessity–sufficiency relations. In total, 16 problems were developed (4 content materials × 4 versions). Out of these, 4 problems were assigned to each participant using a Latin square design to select one of each content and one of each version. Each problem was presented on a separate page of a booklet in a randomized order. Participants were told to determine which card or cards they would definitely need to turn over in order to find out whether or not the rule had been violated. Twenty-two Yale University undergraduate students, unfamiliar with the selection task, participated in the experiment.

Results and Discussion

The data were collapsed over the four content rules because no difference was found among them. Table 2 shows the mean proportions of critical responses for each condition. The most frequent responses were the predicted normative ones (indicated in boldface). Selection of *P* and not-*Q* cards was significantly higher for the unnecessary-but-sufficient condition than for others, all *p*s < .05.[4] Selection of not-*P* and *Q* cards was significantly higher for the necessary-but-insufficient condition than for the others, all *p*s < .05. Finally, selection of all four cards was significantly higher for the necessary-and-sufficient condition than for the others, all *p*s < .05. In the control condition, no particular pattern of responses was obtained.

Selections of *P* alone and of the *P*-and-*Q* combination were dramatically reduced, compared with performance in typical Wason card selection tasks.[5] The selection of the *P*-and-*Q* combination was most common in the necessary-and-sufficient relations (32% of responses), presumably because the normative answer of all four cards for this condition subsumes this response. The selections of *P* alone and of the *P*-and-*Q* combination are often considered evidence for a confirmation bias, or the lack of appreciating disconfirming evidence. However, when the necessity–sufficiency relations were explicitly specified, no such confirmation bias was observed.

Of particular interest in this study is performance on the abstract letter-number material. Over the three experimental conditions, 50.0% of responses were normative ones. This figure did not differ from the

---

3. Graham (1995, Experiment 1) included an unnecessary-and-insufficient condition using a familiar rule ("If a couple has a good time on the first date, they get married") and found normative responses (selecting no card) were most frequent.

4. All statistical analyses reported in this article are based on McNamer's test for dependent proportions (Agresti, 1990).

5. These responses were also reduced in the control condition, presumably because of the use of a within-subjects design. But the within-subjects manipulation does not seem to invalidate the main finding because the first responses were consistent with the overall pattern of the results.
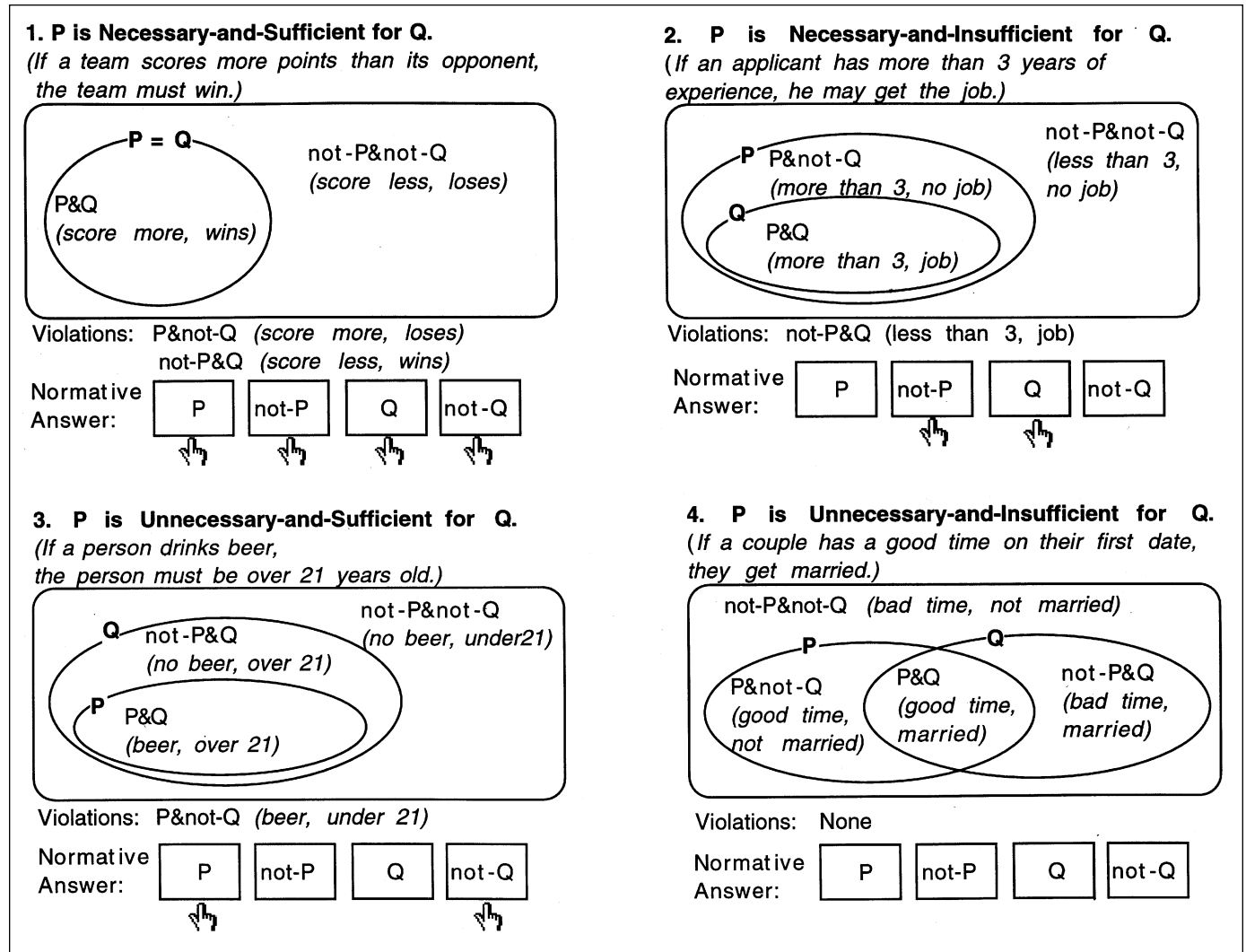
Woo-kyoung Ahn and Loranel M. Graham

1. P is Necessary-and-Sufficient for Q.
(*If a team scores more points than its opponent, the team must win.*)

P = Q
P&Q (*score more, wins*)
not-P&not-Q (*score less, loses*)

Violations: P&not-Q (*score more, loses*)
not-P&Q (*score less, wins*)

Normative Answer: | P | not-P | Q | not-Q |

2. P is Necessary-and-Insufficient for Q.
(*If an applicant has more than 3 years of experience, he may get the job.*)

P
P&not-Q (*more than 3, no job*)
Q
P&Q (*more than 3, job*)
not-P&not-Q (*less than 3, no job*)

Violations: not-P&Q (*less than 3, job*)

Normative Answer: | P | not-P | Q | not-Q |

3. P is Unnecessary-and-Sufficient for Q.
(*If a person drinks beer, the person must be over 21 years old.*)

Q
not-P&Q (*no beer, over 21*)
P
P&Q (*beer, over 21*)
not-P&not-Q (*no beer, under 21*)

Violations: P&not-Q (*beer, under 21*)

Normative Answer: | P | not-P | Q | not-Q |

4. P is Unnecessary-and-Insufficient for Q.
(*If a couple has a good time on their first date, they get married.*)

not-P&not-Q (*bad time, not married*)
P
Q
P&not-Q (*good time, not married*)
P&Q (*good time, married*)
not-P&Q (*bad time, married*)

Violations: None

Normative Answer: | P | not-P | Q | not-Q |

Fig. 1. Four possible necessity–sufficiency relations between *P* and *Q* in "if *P*, then *Q*." The sentence above each diagram is a corresponding real-life example.

Table 2. *Mean proportions of the critical selection combinations in Experiment 1*

Response	Condition			
	Unnecessary-but-sufficient	Necessary-but-insufficient	Necessary-and-sufficient	Control
P and not-*Q*	**.50**	.00	.05	.18
Not-*P* and *Q*	.05	**.59**	.00	.14
All four	.00	.05	**.55**	.18
P alone	.23	.05	.09	.23
P and *Q*	.05	.18	.32	.09

Note. Selections predicted by the necessity–sufficiency relations are in boldface.

proportion of normative responses with the social-interaction materials (57.1%).

To summarize, Experiment 1 indicated the importance of necessity–sufficiency relations in conditional reasoning. Even with identical content material, the selection of cards varied as a function of the necessity–sufficiency relations established between the antecedent and the consequent. Similar studies conducted previously (Graham & Ahn, 1996; Thompson, 1995) examined the effect of people's existing background knowledge about the necessity–sufficiency relations of familiar rules. For instance, we (Graham & Ahn, 1996) tested the rules listed as real-life examples in Figure 1 and found normative responses. In these studies, however, when the necessity–sufficiency relations varied, the rule content also varied because the rules used were familiar to participants. Hence, it was difficult to assess whether the results were due to the necessity–sufficiency relations or to the item differences. Experiment 1 is the first study to vary the necessity–sufficiency relations of rules in test items while keeping their content constant. For this reason, the results of the current study provide stronger evidence against a content-based theory.

EXPERIMENT 2

Experiment 2 attempted to explain the effect of perspective observed by Gigerenzer and Hug (1992) in terms of necessity–sufficiency relations. Consider the "day-off rule" in their study, "If an employee works on the weekend, then that employee gets a day off during the week." According to Gigerenzer and Hug, the perspective of a participant determines which cases would be considered cheating (P and not-Q from the employee's perspective, and not-P and Q from the employer's perspective), which, in turn, determines the selection of cards.

Note, however, that the day-off rule is ambiguous in its necessity–sufficiency relation. We argue that when a participant's perspective is shifted, his or her understanding of the necessity–sufficiency relation of the rule is shifted as well. Specifically, from the employee's perspective, the rule is interpreted as an unnecessary-but-sufficient rule: If an employee works on the weekend (P), that employee must get a day off during the week (Q). This interpretation results in the selection of P and not-Q on the selection task. However, an employer will interpret the same rule as a necessary-but-insufficient one: If a person works on the weekend, he or she may or may not take a day off during the week (P is insufficient for Q), but if a person takes a day off during the week, he or she must have worked on the weekend (P is necessary for Q). Thus, the responses from the employer's perspective would be selections of not working on the weekend (not-P) and taking a weekday off (Q). That is, we propose that there is nothing special about the perspectives except that each highlights one specific necessity–sufficiency interpretation geared toward a party's focus of interest.

To test this interpretation, in Experiment 2, we pitted necessity–sufficiency relations against perspectives. Gigerenzer and Hug (1992) would predict that a shift in perspective would produce a reversal of information selection, regardless of necessity–sufficiency relations, because what matters is detecting cheaters. However, we predicted that rather than the perspectives, the necessity–sufficiency relations would determine performance.

Method

Experiment 2 used four social-contract materials previously used by Gigerenzer and Hug (1992). Each material varied in two ways. First, as in Gigerenzer and Hug's study, participants were told to suppose they were a particular person, either Party A or Party B (these parties corresponded, respectively, to someone who would consider a combination of P and not-Q as cheating or someone who would consider a combination of not-P and Q as cheating). For this manipulation, we adapted context stories from Gigerenzer and Hug, retaining the cost-benefit and perspective specifications.

Second, for both versions of each story (differing in perspective), the rule was either unnecessary-but-sufficient or necessary-but-insufficient. In order to minimize the item differences, we developed the necessary-but-insufficient rules (e.g., "If an employee gets a day off during the week, then that person might have worked on the weekend") by reversing the unnecessary-but-sufficient rules ("If an employee works on the weekend, then that person must get a day off during the week") and changing "must" to "might." For this reason, P and Q cards corresponded to different events in the different necessity–sufficiency conditions (e.g., for unnecessary-but-sufficient, P corresponded to "working on the weekend," whereas for necessary-but-insufficient, P corresponded to "getting a day off"). In the context stories, a minimal number of sentences was added to specify necessity–sufficiency relations (see Appendix B).

To summarize, four forms of each social-contract situation were constructed: Party A/unnecessary-but-sufficient, Party B/unnecessary-but-sufficient, Party A/necessary-but-insufficient, and Party B/necessary-but-insufficient. In total, there were 16 problems (4 content materials × 4 forms). A Latin square design was used to assign 4 problems to each participant, 1 of each content and 1 of each form. The procedure was the same as in Experiment 1. Twenty-three Yale University undergraduate students, unfamiliar with the selection task, participated in the experiment.

Results and Discussion

Table 3 presents the mean proportions of selections of the two critical combinations: (a) P and not-Q and (b) not-P and Q. For both Party A and Party B, the selection of P and not-Q cards was significantly higher for the unnecessary-but-sufficient condition than for the

Table 3. *Mean proportions of the critical selection combinations in Experiment 2*

| | Condition | | | |
| | Unnecessary-but-sufficient | | Necessary-but-insufficient | |
Response	Party A	Party B	Party A	Party B
P and not-Q	**.70**	**.52**	.13	.17
Not-P and Q	.00	.09	**.30**	**.35**

Note. Selections predicted by the necessity–sufficiency relations are in boldface. The cheater-detecting algorithm predicts that Party A would select P and not-Q, and Party B would select not-P and Q.

necessary-but-insufficient condition, ps < .05. However, there was no effect of perspective within each necessity–sufficiency condition, ps > .10. Similarly, the selection of not-P and Q was significantly higher for the necessary-but-insufficient condition than for the unnecessary-but-sufficient condition, for both Party A and Party B, ps < .01. There was no effect of perspective within each necessity–sufficiency condition, ps > .10. No other selections accounted for more than 20% of the responses.

In sum, Experiment 2 supports necessity–sufficiency relations as an account of why a change in perspective produced changes in information selection in Gigerenzer and Hug's (1992) study. These results stand in contrast to Gigerenzer and Hug's interpretation of the effects of perspective shift, and suggest that their results were due to a shift in participants' understanding of the necessity and sufficiency of the rule, rather than to a cheater-detecting algorithm.

GENERAL DISCUSSION

Previously, performance on the Wason card selection task has been measured based on the assumption that "if P, then Q" is properly interpreted as a material conditional, or what we have termed an unnecessary-but-sufficient rule. However, in everyday conversation, the P and Q relationship is not restricted to this interpretation. Thus, we argued that the facilitative effect of content-based rules occurs because of individuals' understanding of thematic rules as unnecessary-but-sufficient, an understanding that is absent for abstract or arbitrary rules. The two studies reported here show that when the necessity–sufficiency relations are clarified, the most frequent responses correspond to normative answers, regardless of whether the rules are in social-contract forms or abstract ones, and regardless of a reasoner's perspective.

As discussed earlier, Cheng and Holyoak (1985) argued that the application of PRSs underlies performance on the selection task. Their findings to date do not necessarily conflict with the necessity–sufficiency account. As explained earlier, when a rule is presented in the form of P1 in Table 1, a permission schema is invoked and P4 indicates that P and not-Q is a violation. P1 is an example of unnecessary-but-sufficient rules, and hence both accounts make the same prediction. A rule may also be presented as one that maps onto P3. This time P corresponds to "precondition" and Q to "action." The violation would still be the case in which an action is performed (Q) without satisfying the precondition (not-P). P3 in the permission schema is an example of necessary-but-insufficient rules, and therefore both accounts predict selections of not-P and Q cards given P3. However, our results, as well as results from Graham (1995), showed that there are also necessary-and-sufficient and unnecessary-and-insufficient relations, for which other selection patterns are predicted and obtained, and these patterns of response are not yet accommodated by the PRS theory. In addition, our analysis is more parsimonious in that a single psychological mechanism can account for a variety of domains; there is no need to propose separate schemas whenever performance on the selection task is better than on the original abstract version of the task.

Yet these domain-specific approaches do make contributions to our approach. The normative answers specified in Figure 1 can be yielded only when one considers all possible relations between the presence or absence of antecedent and consequent. Relying on familiar concepts such as permission and obligation, or being sensitive to cheaters, can serve as a means to represent necessity–sufficiency relations easily. However, we argue that necessity–sufficiency relations are more fun-

damental factors than domain-specific mechanisms. Triggering a cheater-detection algorithm was not sufficient to induce normative responses in Experiment 2. In addition, we argue that PRSs are effective because they explicate necessity–sufficiency relations.

Recently, Sperber, Cara, and Girotto (1995) proposed the relevancy theory of the selection task, a theory based on cognitive effect and processing effort. On the effect side, a conditional rule (e.g., "If it is a real champagne, it is expensive") can be interpreted either as P and Q exist (e.g., there exists an expensive champagne) or as P and not-Q cannot exist (e.g., real champagne cannot be cheap). According to Sperber et al., when the rule serves as a denial of there being cases of P and not-Q, a state with P and not-Q becomes particularly relevant, resulting in success on the selection task. In their experiment, for instance, a leader of a religious sect was described as stating the rule "If a woman has a child, she has had sex," in the context of contesting the allegation that his sect's virgin girls were artificially inseminated (i.e., the leader claimed that P and not-Q cannot exist). Given such a context, selection of P and not-Q cards significantly increased. In our framework, what such contexts did was to explicate the rule as an unnecessary-but-sufficient one. On the effort side, Sperber et al. proposed that one method for reducing cognitive effort is to make P-and-not-Q easier to represent, by lexicalizing not-Q (e.g., "unmarried" instead of "not married"). Indeed, they found improvement in performance with such manipulations. However, in our Experiment 1, half of the rules used lexicalized not-Qs, but no advantage was found. In the unnecessary-but-sufficient condition, selections of P and not-Q were 45.5% for lexicalized materials and 54.5% for unlexicalized ones. The inconsistency in these findings seems to suggest that when the necessity–sufficiency relations are made explicit, processing effort does not matter.

In conclusion, what does the current study say about human rationality? Our claim is consistent with Henle's (1962) conclusion in that errors found in syllogisms do not necessarily reflect a breakdown of the deductive process, but rather reflect changes in the material from which the reasoning proceeds. We argue that particularly in the selection task, poor performance from abstract rules is to be expected because of ambiguities in their necessity–sufficiency status. In addition, we argue that relying on content should not be equated with irrationality because content specifies the necessity–sufficiency relations, which are formal components in conditional reasoning. In this way, the current study demonstrates that neither the poor performance resulting from abstract rules nor the content effect serves as valid evidence against human rationality.

REFERENCES

Agresti, A. (1990). *Categorical data analysis.* New York: Wiley.

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

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.

Evans, J.St.B.T. (1982). *The psychology of deductive reasoning.* London: Routledge & Kegan Paul.

Gigerenzer, G., & Hug, K. (1992). Domain-specific reasoning: Social contracts, cheating, and perspective change. *Cognition, 43,* 127–171.

Graham, L. (1995). *A two-component theory of conditional reasoning: Content and process in the Wason four-card selection task.* Unpublished doctoral dissertation, University of Louisville, Louisville, Kentucky.

Graham, L., & Ahn, W. (1996). Necessity and sufficiency as an account for biases in Wason's card selection task. In *Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society* (pp. 552–555). Mahwah, NJ: Erlbaum.

Griggs, R.A. (1983). The role of problem content in the selection task and in the THOG problem. In J.St.B.T. Evans (Ed.), *Thinking and reasoning: Psychological approaches* (pp. 16–43). London: Routledge & Kegan Paul.

Henle, M. (1962). On the relation between logic and thinking. *Psychological Review, 69,* 366–378.

Holyoak, K.J., & Cheng, P.W. (1995). Pragmatic reasoning with a point of view. *Thinking and Reasoning, 1,* 289–313.

Johnson-Laird, P.N., & Byrne, R.M.J. (1991). *Deduction.* Hillsdale, NJ: Erlbaum.

Manktelow, K.I., & Over, D.E. (1987). Reasoning and rationality. *Mind and Language, 2,* 199–219.

Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review, 101,* 608–631.

Platt, R.D., & Griggs, R.A. (1995). Facilitation and matching bias in the abstract selection task. *Thinking and Reasoning, 1,* 55–70.

Rips, L.J. (1994). *The psychology of proof.* Cambridge, MA: MIT Press.

Sperber, D., Cara, F., & Girotto, V. (1995). Relevance theory explains the selection task. *Cognition, 57,* 31–95.

Thompson, V.A. (1994). Interpretational factors in conditional reasoning. *Memory & Cognition, 22,* 742–758.

Thompson, V.A. (1995). Conditional reasoning: The necessary and sufficient conditions. *Canadian Journal of Experimental Psychology, 49,* 1–58.

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

Wason, P.C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology, 20,* 273–281.

APPENDIX A: RULES AND CONTEXT STORIES FOR THE LETTER-NUMBER MATERIAL OF EXPERIMENT 1

The following context stories for the letter-number material illustrate how the necessity–sufficiency relations were manipulated in Experiment 1.

- *Unnecessary-but-sufficient version:*
  The secret agency of North Korea is making secret codes for English. The first step is to line up alphabets with numbers. So, they look like the following: A = 1, B = 2, C = 3, D = 4, E = 5, F = 6, G = 7, H = 8, I = 9, J = 10, K = 11, L = 12, M = 13, N = 14, O = 15, P = 16, Q = 17, R = 18, S = 19, T = 20, U = 21, V = 22, W = 23, X = 24, Y = 25, Z = 26. They made up cards with a letter on one side and a corresponding number on the other side. In this coding scheme, all vowels have odd numbers. However, some consonants also have odd numbers. Therefore, when they made the cards, they followed the rule "If a card has a vowel on one side, then it must have an odd number on the other side."
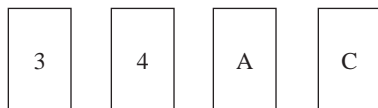 Cards:

  | A | C | 3 | 4 |

- *Necessary-but-insufficient version* (The context story was the same as the unnecessary-but-sufficient version except for the last sentence):
  Therefore, when they made the cards, they followed the rule "If a card has an odd number on one side, then it might have a vowel on the other side."
  Cards:

  | 3 | 4 | A | C |

- *Necessary-and-sufficient version:*
  The secret agency of North Korea is making secret codes for English. They first made up some cards with a letter on one side and a number on the other side. They look like the following: A = 1, B = 2, C = 4, D = 6, E = 7, F = 8, G = 10, H = 12, I = 13, J = 14, K = 16, L = 18, M = 20, N = 22, O = 23, P = 24, Q = 26, R = 28, S = 30, T = 32, U = 31, V = 32, W = 34, X = 36, Y = 38, Z = 40. In these cards, if there is a vowel on one side, then there is an odd number on the other side. Only a vowel can have an odd number on the other side. All consonants have an even number on the other side. When they made the cards, they followed the rule "If a card has a vowel on one side, then it must have an odd number on the other side."
  Cards:

  | A | C | 3 | 4 |

- *Control version:*
  The secret agency of North Korea is making secret codes for English. They made up cards with a letter on one side and a corresponding number on the other side. In making these, they followed the rule "If a card has a vowel on one side, then it has an odd number on the other side."
  Cards:

  | A | C | 3 | 4 |

APPENDIX B: SAMPLE MATERIAL USED IN EXPERIMENT 2

In this example of the materials used in Experiment 2, boldface indicates the alternative sentences used for the perspective manipulation, and italics indicate the sentences used for the necessity–sufficiency manipulation (italicized sentences in brackets were for the necessary-and-insufficient condition).

- *Day-off rule:*
  There is a sweatshirt factory. Having employees work on the weekend is a benefit for the employer because the factory can make use of its machines and be more flexible. On the other hand, working on the weekend is a cost for the employee because it takes away from home life. In this factory, if an employee works on the weekend, then that person must get a day off during the week. Employees can also get a day off if they work a certain amount of overtime at night.

  Suppose you are an employer at the factory. [Suppose you are an employee at the factory.] You are aware of the rule *"If an employee works on the weekend, then that person must get a day off during the week." ["If an employee gets a day off during the week, then that person might have worked on the weekend."]* You want to check whether or not the rule has been violated based on the four cases available. *Keep in mind that employees can also get a day off during the week if they work a certain amount of overtime at night. [Keep in mind that employees who worked on the weekend must get a day off during the week.]*