

Running head: PROBLEM WITH FORMAL REASONING PARADIGMS

“At Least One” Problem with “Some” Formal Reasoning Paradigms

James R. Schmidt¹ & Valerie A. Thompson²

¹ University of Waterloo, ² University of Saskatchewan

James R. Schmidt, Department of Psychology, University of Waterloo, 200 University Ave. W.,
Waterloo, Ontario N2L 3G1. E-mail: j4schmid@watarts.uwaterloo.ca.

Abstract

In formal reasoning, the quantifier “Some” means “At least one and possibly all.” In contrast, reasoners often pragmatically interpret “Some” to mean “Some but not all” on both immediate inference and Euler circle tasks. It is still unclear whether pragmatic interpretations can explain the high rates of errors normally observed on syllogistic reasoning tasks. To answer this question, participants in the current experiments were presented with either standard quantifiers or clarified quantifiers designed to precisely articulate their logical interpretation. In Experiment 1, reasoners made significantly more logical responses and significantly fewer pragmatic responses on an immediate inference task when presented with logically-clarified as opposed to standard quantifiers. In Experiment 2, this finding was extended to a variant of the immediate inference task in which reasoners were asked to deduce what followed from premises that they were to assume were false. In Experiment 3, we used a syllogistic reasoning task and observed that logically-clarified premises reduced pragmatic and increased logical responses relative to standard ones, providing strong evidence that pragmatic responses can explain some aspects of the errors made in the syllogistic reasoning task. These findings suggest that standard quantifiers should be replaced with logically-clarified quantifiers in teaching and in future research.

“At Least One” Problem with “Some” Formal Reasoning Paradigms

Syllogistic reasoning is a widely used measure of formal, deductive reasoning. In this task, reasoners are presented with two premise statements (e.g., “Some of the As are Bs” and “All of the Bs are Cs”) and are then asked to decide whether a given conclusion statement logically follows from the premises (e.g., “Some of the As are Cs,” which is a valid conclusion). Syllogistic reasoning is used as a vehicle to investigate such disparate phenomena as belief bias (Evans, Barston, & Pollard, 1983; Klauer, Musch, & Naumer, 2000; Newstead, Pollard, Evans, & Allen, 1992; Thompson, Striener, Reikoff, Gunter, & Campbell, 2003), the role of working memory in reasoning (Capon, Handley, & Dennis, 2003; Copeland, & Radvansky, 2004; Gilhooley, Logie, & Wynn, 1999, Quayle & Ball, 2000), strategies in reasoning (Bacon, Handley, & Newstead, 2003; Bucciarelli & Johnson-Laird, 1999; Chater & Oaksford, 1999), and disruptions to reasoning performance caused by age (Fisk & Sharp, 2002; Gilinsky, & Judd, 1994) and other factors (Fisk, Montgomery, Wareing, & Murphy, 2005; Smeets & De Jong, 2005).

Although often taken as a measure of logical or analytic reasoning, it is almost certainly the case that performance on this task also encompasses a number of non-analytic processes. For example, it is well known that the believability of both the premises and the conclusion have a large impact on the inferences reasoners are willing to endorse (e.g., Evans et al., 1983; Klauer et al., 2000; Newstead et al., 1992; Thompson, 1996; Thompson et al., 2003). Even when the believability of the material is not an issue (i.e., premises and conclusions describe arbitrary or abstract relations), the reasoner must still interpret the task, the instructions, and the meaning of the quantifiers used in the problems. Poor performance, therefore, may not necessarily represent poor logical reasoning, but, instead, may reflect differences between the reasoner’s and the

experimenter's interpretation of the task requirements. For example, reasoners conflate the concepts of logical possibility with logical necessity (e.g., Evans, Handley, & Harper, 2001; Evans, Handley, Harper, & Johnson-Laird, 1999; Newstead, Thompson, & Handley, 2002), and employ heuristic, rather than logical, strategies (Chater & Oaksford, 1999). Consequently, it is difficult to make attributions about the contributions of analytic processes to performance in the absence of a well-articulated model of interpretation (Evans & Thompson, 2004; Thompson, 2000).

Thus, the goal of this paper is to investigate the how interpretation of the quantifiers used in syllogistic problems contributes to variability in reasoning performance. Surprisingly, relatively little is known about how interpretations along this dimension affect reasoning. There is also a well-developed literature documenting reasoner's interpretation of quantified premises in isolation (e.g., Begg & Harris, 1982; Déret, 1998; Evans et al., 1999; Newstead, 1989, 1995; Newstead & Griggs, 1983; Politzer, 1991; Rosenthal, 1980), but little that examines the relationship between these interpretations and reasoning on a complex task.

Of particular interest is the quantifier "Some," which is used differently in formal logic than it is used in everyday speech. In formal logic, "Some" means "At least one and possibly all." Pragmatically, however, this interpretation of the word "Some" is infelicitous, as "Some" means "Some but not all" in everyday speech. For instance, if James tells Sarah "Some of the employees are part of the union," then Sarah will infer that some employees are *not* part of the union (because *only* some of the employees are); Sarah would reasonably assume that if James had meant "All," then he would have said "All." This is termed the Gricean maxim of quantity (or informativeness; Grice, 1975/2002). Thus, when reasoners are presented with a premise such as "Some of the As are Bs," they will make the sensible conclusion that "Some *but not all* of the

As are Bs.” As a result, some syllogistic reasoning “errors” may not be errors at all. Rather, the error may reflect a Gricean, rather than a logical, interpretation of the premises.

Indeed, the evidence supports the conclusion that reasoners make pragmatic interpretations of quantified premises (e.g., Begg & Harris, 1982; Déret, 1998; Evans et al., 1999; Newstead, 1989, 1995; Newstead & Griggs, 1983; Politzer, 1991; Rosenthal, 1980). Much of the data is derived from the immediate inference task, where reasoners are presented with a single premise statement (e.g., “All of the As are Bs”) and are then asked to judge whether a given conclusion statement follows (e.g., “None of the As are Bs,” which would be false). Pragmatic interpretations are common. For instance, reasoners are 22% *more* likely to conclude that “Some of the As are Bs” follows from “Some of the As are not Bs” than from “All of the As are Bs,” even though the former is invalid and the latter is valid (Evans et al., 1999). Furthermore, although instructing reasoners on the logical interpretation of the word “Some” increases logical responses (Newstead, 1995), it does not prevent pragmatic responses (Begg & Harris, 1982; Newstead, 1989; Newstead & Griggs, 1983). Similar findings have been obtained using a different paradigm, the Euler circle task, in which the conclusions are circle diagrams instead of statements (Begg & Harris, 1982; Newstead, 1989).

Although there has been a long-standing assumption that these types of Gricean interpretations contribute to the large number of errors typically observed in syllogistic reasoning tasks, the existing data do not corroborate this hypothesis. For example, whereas Newstead and his colleagues (Newstead, 1989, 1995; Newstead & Griggs, 1999) have demonstrated that while such interpretations are common on immediate inference tasks, they appear to account for little, if any, variance on the more complex, syllogistic reasoning task. For example, when one looks at the errors reasoners make, there is little evidence to show that they are produced by a Gricean

interpretation of the premises (Newstead, 1995); indeed, such errors tend to be produced less often than expected by chance. However, Newstead (1995) found a small correlation between the tendency to endorse Gricean interpretations on an Euler circle task and the probability of endorsing a conclusion consistent with a Gricean interpretation in the syllogistic reasoning task. He also observed a small, but statistically reliable reduction (28% to 19%) in responses consistent with a Gricean interpretation when participants were given instructions about the logical meaning of “Some.” Additionally, Roberts, Newstead, and Griggs (2001) have shown that a portion of syllogistic errors are consistent with a pragmatic interpretation of the premises, assuming that they also make reversible interpretations of the premises (e.g., such that “Some of the As are not Bs” entails “Some of the Bs are not As”).

One interpretation of these data is that the expression of Gricean interpretations is but one of many factors that contribute to performance on these tasks (Roberts et al., 2001). This would make the calculation of chance rates of observations difficult. How many pragmatically-consistent errors does one need to observe to conclude that they occur a significant proportion of the time in a context where responses are multiply determined? Moreover, given that overriding the inclination to interpret “Some” as “Some but not all” is a working-memory demanding task (Feeney, Scrafton, Duckworth, & Handley, 2004), instructions to do so might have limited utility in the context of a highly working-memory demanding task, such as syllogistic reasoning.

In the current paper, we have adopted an alternative approach, namely to replace the ambiguous quantifiers with new quantifiers that precisely articulate their logical interpretation. This manipulation should not pose extra demands on working memory; thus, to the extent that “errors” on syllogistic tasks reflect pragmatic interpretations of the quantifiers, performance should improve with the logically-clarified versions.

We know of only two studies that have attempted to reduce the ambiguity of the premises by substituting alternative phrases for them (Ceraso & Provitera, 1971; Newstead & Griggs, 1999). In those studies, the quantifiers were replaced with extended sentences (e.g., the sentence “Whenever I have a square block it is blue, but I also have some blue blocks that are not square” indicates that some *but not all* blue blocks are square). It is important to note, however, that the elaborated premises did not specify the *logical* meaning of quantifiers, but instead reduced ambiguity by eliminating some possible interpretations of the quantifiers. Indeed, in some cases, the meaning of the quantifiers was altered to such an extent that the elaborated and traditional versions entailed different valid conclusions.

Nonetheless, the elaborated premises had a significant impact on the conclusions that participants endorsed, suggesting that some variability in syllogistic reasoning can be attributed to ambiguity in the interpretation of the premises. The goal of our study is to take this one step further and to clarify the logical meaning of the particular quantifiers by replacing them with statements that precisely articulate their logical meaning, namely “At least one” (similar to Geurts, 2003). The use of “At least one” quantifiers should decrease reasoners’ inclination to make pragmatic responses for two reasons. First, “At least *one*” stresses the focus on *particular* (i.e., individual) cases instead of sets of cases. Second, “At least one” pragmatically allows for the possibility “All.” Both of these characteristics are important, because a statement such as “Some of the As are Bs” logically allows for the possibility that only *one* A is a B, that *all* As are Bs, or anything in between. The only possibility ruled out is that none of the As are Bs, and “At least one” is consistent with this logical meaning. That is, “At least one” clearly asserts that there is one A that is a B and allows for any or all of the remaining As to be Bs.

The experiments that follow have two goals. The first is to demonstrate that reasoners

interpret our new logically-clarified “At least one” quantifiers as we intended them to; in Experiments 1 and 2, we tested the hypothesis that logically-clarified premises will reduce pragmatic responses and increase logical responses on two versions of the immediate inference task. The second goal was to extend this analysis to syllogistic reasoning; in Experiment 3, we predicted that logically-clarified premises would likewise facilitate logical responses on a syllogistic task. Finally, we discuss the implications of our findings for theories of syllogistic reasoning.

Experiment 1

Participants in this study completed an immediate inference task with one of three types of quantifiers. The first was the standard particular premises (“Some” and “Some...not”). The second was the logically-clarified quantifiers (“At least one” and “At least one...not”). Finally, as a control condition, we also included pragmatically-clarified quantifiers (“Some but not all” and “Some but not all...not”), which should provide an estimate of reasoners’ inclination to make pragmatic responses when it is unambiguously appropriate to do so.

We were interested in two types of responses, pragmatic and logical. Logical responses should be based on the Square of Opposition, as illustrated in Figure 1. In contrast, pragmatic interpretations should resemble the pattern depicted by the Triangle of Opposition in Figure 2. Figure 1 depicts the relationship among the four standard quantifiers: “All,” “No,” “Some,” and “Some...not.” Traditionally, each quantifier represents one of four moods that make up an orthogonal combination of universality (universal, particular) and polarity (affirmative, negative). “All” is the universal affirmative (e.g., “All of the chemists are beekeepers”), “No” is the universal negative (e.g., “None of the chemists are beekeepers”), “Some” is the particular affirmative (e.g., “Some of the chemists are beekeepers”), and “Some...not” is the particular

negative (e.g., “Some of the chemists are not beekeepers”).

Figure 1 about here

However, given that reasoners often interpret “Some” pragmatically rather than logically, Begg and Harris (1982) argued that reasoners divide logical reality into just three categories: “All,” “No,” and “Some but not all.” In this view, there are truly only three logical moods, which could be modelled as a “Triangle of Opposition” as depicted in Figure 2. The “All” and “No” moods are the same as in the Square of Opposition, but the “Some” and “Some...not” moods are replaced by a single “Some but not all” (partition) mood (where “Some” and “Some...not” are merely exemplars of the partition mood). Under this interpretation, each of the three moods (universal affirmative, universal negative, and partition) should be mutually exclusive. That is to say, no two of these moods (e.g., “All” and “No”) can be true at the same time, so the truth of one mood (e.g., “All”) implies the falsity of the remaining two (i.e., “No” and “Some but not all”).

Figure 2 about here

If the quantifier manipulation is successful in changing the frequency of pragmatic responses, then differences among the three types of quantifiers should be observed. The rate of logical responses should be highest with the logically-clarified premises, and the rate of pragmatic responses should be highest with the pragmatically-clarified premises. That is, participants given logically-clarified premises should produce responses consistent with the

Square of Opposition depicted in Figure 1, whereas pragmatically-clarified premises should elicit responses consistent with the Triangle depicted in Figure 2. Under the assumption that traditional quantifiers are ambiguous in their interpretation, we expect a mixed pattern for these.

However, as illustrated in Figures 1 and 2, the differences among the quantifiers should be selective, applying only to problems in which a pragmatic interpretation of the quantifier would lead to a different conclusion than the logical interpretation. The two types of problems for which these differences should be observed are called *subcontrary* and *subaltern* problems.

For *subcontrary* problems, reasoners are presented with a particular premise followed by the reverse-polarity conclusion (i.e., “Some of the As are Bs” followed by “Some of the As are not Bs,” or vice versa). Under a pragmatic interpretation, these conclusions follow validly. Under a logical interpretation, however, neither conclusion is necessary, because “Some” allows the possibility of “All” and “Some...not” allows the possibility of “None.” Consequently, we expected that the rates of endorsement for these conclusions to be highest for the pragmatically-clarified premises, followed by the standard and the logically-clarified premises; conversely, the pattern of logically valid responses (i.e., to indicate that these conclusions are possible) should be the reverse.

For *subaltern* problems, reasoners are presented with a particular premise followed by the same-polarity universal conclusion (i.e., “Some of the As are Bs” followed by “All of the As are Bs” or “Some of the As are not Bs” followed by “None of the As are Bs”). These conclusions are logically, but not pragmatically possible. Thus, reasoners should be more inclined to reject these inferences in the pragmatically-clarified condition than the standard condition, followed by the logically-clarified condition. Again, logical responses (i.e., to indicate that these conclusions are possible) should be the mirror image of this pattern.

For the two remaining problem types there should be no effect of the quantifier manipulation, because the logical and pragmatic interpretations of the premises lead to the same conclusion. For *identity* problems, reasoners are presented with a premise followed by the same statement as a conclusion (i.e., “Some of the As are Bs” followed by “Some of the As are Bs” or “Some of the As are not Bs” followed by “Some of the As are not Bs”). Both logically and pragmatically, these conclusions necessarily follow. For *contradictory* problems, reasoners are presented with a premise followed by a conclusion with the opposite universality and polarity (i.e., “Some of the As are Bs” followed by “None of the As are Bs” or “Some of the As are not Bs” followed by “All of the As are Bs”). Both logically and pragmatically, these conclusions are necessarily false.

In summary, for two critical problem types (subaltern and subcontrary) logical responses should be increased and pragmatic responses should be decreased in the logically-clarified condition and the reverse should be true in the pragmatically-clarified condition. For the remaining two problem types (contradictory and identity) responses should be consistent for all three quantifier types.

Method

Participants. Twenty-four participants with no background in logic completed Experiment 1. Six participants were volunteers and the remaining 18 were University of Saskatchewan undergraduates that received course credit or five dollars for participating.

Materials and Design. Each problem was presented on the page in the following form:

IF IT IS TRUE THAT:

At least one of the chemists is a beekeeper

THEN IS IT THE CASE THAT:

All of the chemists are beekeepers Y__ N__ M__

None of the chemists are beekeepers Y__ N__ M__

Some of the chemists are beekeepers Y__ N__ M__

Some of the chemists are not beekeepers Y__ N__ M__

All problems related one category of individuals to another category of individuals. Each of the 48 unique categories used was either a profession or a hobby. The particular affirmative and particular negative premises were presented using standard (“Some,” “Some...not”), logically-clarified (“At least one,” “At least one...not”), or pragmatically-clarified quantifiers (“Some but not all,” “Some but not all...not”). This resulted in six premise-types (2 polarities x 3 quantifiers). Four versions of each premise type were created using unique combinations of hobbies and professions. The conclusion statements were always the four standard moods (“All,” “No,” “Some,” and “Some...not”) and were followed by three response options: “Y” (for “yes”), “N” (for “no”), and “M” (for “maybe”).

Procedure. The 24 problems were presented in four blocks, each with one instance of each premise type. The order of the items within blocks was randomized for each group of four participants. Orthogonal to this, the four conclusion statements were presented in four counterbalanced orders. Each participant was given a booklet that began with the following instruction page:

Immediate Inference Task

IF IT IS TRUE THAT:

Either a magpie or a magpie and a crow are in the yard

THEN IS IT THE CASE THAT:

A magpie is in the yard Y__ N__ M__

A magpie is not in the yard Y__ N__ M__

A crow is in the yard Y__ N__ M__

A crow is not in the yard Y__ N__ M__

Assume that the first “premise” statement (i.e., “Either a magpie or a magpie and a crow are in the yard”) is true. You must then decide whether each of the four following (“conclusion”) statements is **NECESSARILY** true, **NECESSARILY** false, or possibly true given that the premise statement is true.

For instance, because the premise is true, with or without a crow, there must be a magpie in the yard. As such, the first conclusion statement (“A magpie is in the yard”) is necessarily true. Thus, you would check the space next to the “Y” (for “yes”).

Conversely, the second conclusion statement (“A magpie is not in the yard”) is necessarily false. Thus, you would check the space next to the “N” (for “no”).

The remaining two conclusion statements are both possible but not necessary because a crow could or could not be in the yard (e.g., the magpie could be alone in the yard). For both of these statements you should therefore check the space next to the “M” (for “maybe”).

In the problems presented below assume that all the statements refer to a gathering of people in a room. After you have finished a problem, please do not return to it (even if you realize that you have made a mistake). If you are unclear about any of these instructions, please ask the experimenter to clarify them before you continue. When you are ready to begin, flip the page.

Participants were tested alone or in small groups. No time limit was set for completing the problems. Most participants took about 15 minutes.

Results

The dependent measures for Experiment 1 were proportion logical and pragmatic responses; as the overall results for these dependent measures were always complementary, the analysis will focus only on the logical responses. The data are plotted in Figure 3.

 Figure 3 about here

A 3 (quantifier; logically-clarified, standard, pragmatically-clarified) x 4 (problem type; contradictory, subaltern, subcontrary, identity) within-group ANOVA for logical responses revealed a significant main effect for quantifier, $F(2,46) = 20.178$, $MSE = 1.115$, $p < .001$, a main

effect for problem type, $F(3,69) = 103.565$, $MSE = 8.774$, $p < .001$, and an interaction, $F(6,138) = 17.160$, $MSE = .746$, $p < .001$. Planned comparisons were then conducted.

We predicted that subaltern problems (i.e., problems for which “Some” or “Some...not” was the premise and “All” or “No,” respectively, was the conclusion) would elicit different responses in the three quantifier conditions. A logical interpretation acknowledges that, for instance, “All” is not impossible given “Some;” thus, clarifying the meaning of “Some” should facilitate this interpretation. As predicted, participants made more logical responses with “At least one” quantifiers (.65) relative to “Some” quantifiers (.36), $t(23) = 3.446$, $p = .002$, $SE_{diff} = .085$, and fewer logical responses with “Some but not all” quantifiers (.05) relative to “Some,” $t(23) = 4.541$, $p < .001$, $SE_{diff} = .068$. As expected, these findings were mirrored by a reduction in pragmatic responses for “At least one” (.34) relative to “Some” quantifiers (.64), $t(23) = 3.562$, $p = .002$, $SE_{diff} = .083$, and an increase in pragmatic responses for “Some but not all” quantifiers (.94) relative to “Some,” $t(23) = 4.460$, $p < .001$, $SE_{diff} = .068$.

We predicted that subcontrary problems (i.e., problems for which “Some” was the premise and “Some...not” the conclusion, or vice versa) would elicit different responses in the three quantifier conditions. A logical interpretation acknowledges that “Some” does not imply “Some not,” and vice versa; thus, clarifying that “Some” means “At least one” should block this interpretation. Consistent with this prediction, participants made more logical responses with “At least one” quantifiers (.51) relative to “Some” quantifiers (.31), $t(23) = 2.483$, $p = .021$, $SE_{diff} = .081$, and less with “Some but not all” quantifiers (.12) relative to “Some,” $t(23) = 3.016$, $p = .006$, $SE_{diff} = .063$. As expected, these findings were mirrored by a reduction in pragmatic responses for “At least one” (.44) relative to “Some” quantifiers (.64), $t(23) = 2.369$, $p = .027$, $SE_{diff} = .084$, and an increase in pragmatic responses for “Some but not all” quantifiers (.83)

relative to “Some,” $t(23) = 3.194, p = .004, SE_{diff} = .060$.

For contradictory problems, we expected no differences between premise conditions, given that the conclusion is impossible regardless of whether a pragmatic or logical interpretation is made (e.g., “None” cannot follow from either “Some” or “At least one”). Consistent with this prediction, no differences were found between “Some” (.94) and “At least one” quantifiers (.97), $t(23) = .917, p = .369, SE_{diff} = .035$, or between “Some” and “Some but not all” quantifiers (.93), $t(23) = .432, p = .670, SE_{diff} = .022$.

Similarly, no differences were expected for identity problems. Consistent with this prediction, there was no difference in logical responses for “Some” (.99) and “Some but not all” quantifiers (.97), $t(23) = 1.381, p = .181, SE_{diff} = .011$. However, contrary to predictions, logical responses for “At least one” quantifiers (.80) were significantly reduced relative to “Some” quantifiers, $t(23) = 3.182, p = .004, SE_{diff} = .060$. This was mirrored by a rise in “maybe” responses for “At least one” (.20) relative to “Some” (.01), $t(23) = 3.171, p = .004, SE_{diff} = .061$. This apparent anomaly is likely due to the fact that although logically-clarified versions of the premises were used, the conclusions were presented in the standard form. Whereas logically-clarified statements allow for the possibility of universal (i.e., “All”) or singular (i.e., “one”) interpretations, standard statements imply more than one but not all; as such, given “At least one,” “Some” does not *pragmatically* follow (i.e., if “All” or “one” is true). Thus, these data serve as additional evidence that participants do not interpret standard and logically-clarified statements to mean the same thing.

Discussion

Our manipulation successfully reduced the number of pragmatic responses given to particular quantifiers. Specifically, for subaltern and subcontrary problems, the rate of pragmatic

responses was reduced by 30% and 20%, respectively, with logically-clarified “At least one” quantifiers; this decrease was accompanied by concomitant increases in logical responses. Thus, reasoners interpret “At least one” to mean something more similar to the logical meaning of particular connectives than the traditional “Some.” It is worth noticing, however, that performance for the logically-clarified premises was not at ceiling, indicating that we were not 100% successful in promoting a logical interpretation.

In addition, it is clear that *some* reasoners made a logical interpretation of the standard quantifiers, given that pragmatic responses were more prevalent in the pragmatically-clarified condition. These data suggest that the standard quantifiers are ambiguous, and produce highly variable interpretations. In order to avoid the variance in performance that can be attributed to differences in interpretations, researchers might be advised to specify the intended meanings of the premises using either the pragmatically-clarified or logically-clarified premises.

Experiment 2

The goal of Experiment 2 was to replicate and extend the findings of Experiment 1 in the following ways. First, Experiment 1 used a within-group design, opening up the possibility of carry-over effects. That is, given that the various quantifiers were presented in close proximity, participants may have made a conscious effort to discriminate their meanings. Thus, we wished to rule out the possibility that the high rate of logical responses and low rate of pragmatic responses on the logically-clarified problems were artificially elevated and decremented, respectively, due to exposure to the standard problems. To accomplish this goal, Experiment 2 used a between-groups design. Pragmatically-clarified quantifiers were not presented in Experiment 2.

A second goal of Experiment 2 was to address the concern raised in the results of

Experiment 1 by providing logically-clarified conclusions as well as logically-clarified premises. That is, those participants who received “At least one” premises also received “At least one” conclusions. By so doing, we were able to provide a more appropriate identity condition for logically-clarified quantifiers. In addition, this manipulation allowed us to increase the number of conclusions predicted to differ for standard and logically-clarified premise types, in particular, for *subaltern* problems with universal premises. For these problems, participants are presented with a universal premise followed by the particular premise with the same polarity (i.e., “All of the As are Bs” followed by “Some of the As are Bs” or “None of the As are Bs” followed by “Some of the As are not Bs”). These conclusions are logically necessary, but pragmatically impossible. Thus, logical “yes” responses should be increased and pragmatic “no” responses should be decreased in the logically-clarified condition relative to the standard condition.

For the remaining new problem types, no effect of quantifier is expected, because both logical and pragmatic interpretations lead to the same response. Specifically, *contradictory* and *identity* problems with universal premises follow the same logic as contradictory and identity problems with particular premises. In *contrary* problems, participants are presented with a universal premise followed by the reverse-polarity universal premise (i.e., “All of the As are Bs” followed by “None of the As are Bs,” or vice versa). The type of particular quantifier cannot affect performance in this condition because particular quantifiers do not appear in these problems.

The final goal of Experiment 2 was to extend our analysis to situations in which the premises are assumed to be false. For instance, if it is false that “Some of the As are Bs,” then is it the case that “None of the As are Bs?” Thus, for half of the problems, participants were asked to judge what is implied by a true premise, and for the remaining half, they were asked to judge

what follows from a false premise.

As we pointed out earlier, the pragmatic interpretation outlined in Figure 2 suggests that the three moods (universal affirmative, universal negative, and partition) will behave as mutually exclusive. That is to say, no two of these moods (e.g., “All” and “No”) can be true at the same time, so the truth of one mood (e.g., “All”) implies the falsity of the remaining two (i.e., “No” and “Some but not all”). Conversely, a false premise (e.g., “All”) implies that each of the remaining two moods (i.e., “No” and “Some but not all”) are still possible. Consequently, in the false premise task the majority of a pragmatic reasoner’s responses should conform to a simple pattern, where all conclusions will be regarded as possible except the false identity conclusions (including the subcontrary conclusions, because “Some” and “Some...not” are regarded as the same). Reasoners should be more likely to adopt this pragmatic strategy as their primary strategy in the standard condition relative to the logically-clarified condition.

The pattern of responses under a logical interpretation will be different. To derive a logical interpretation, one needs to determine what follows logically from the contradiction of the premise. For example, if the premise “Some of the As are Bs” is false, then this can be converted to mean that the premise “None of the As are Bs” is true. The reasoner can then use this converted “No” premise to evaluate the conclusions (i.e., “All” and “Some” are false and “No” and “Some...not” are true). Reasoners should be more likely to adopt this strategy in the logically-clarified than the standard condition.

We also anticipated that reasoners will have difficulty in making inferences from false premises. For example, reasoners often have difficulty determining which situations contradict or are inconsistent with a set of premises (e.g., Barres & Johnson-Laird, 2003). According to these authors, reasoners must first eliminate the set of true possibilities before they can derive the false

possibilities. As this may overtax working memory, reasoners are prone to error. In the current task, we hypothesised that reasoners might take a mental short-cut, in particular, by accepting the “opposite” of the premise (i.e., the reverse polarity), where “All” and “No” are opposites and “Some” and “Some...not” are opposites (e.g., a false “All” may seem to imply a true “No”). In principle, the polarity rule could be used just as easily by pragmatic and logical reasoners.

Method

Participants. Thirty-two University of Saskatchewan undergraduates with no background in logic completed Experiment 2. Participants received course credit or five dollars for participating. None of the participants in Experiment 2 had participated in Experiment 1.

Materials and Design. The same 48 categories of professions and hobbies used in Experiment 1 were used in Experiment 2. Half of the participants were presented with standard quantifiers (“Some,” “Some...not”) and half were presented with logically-clarified quantifiers (“At least one,” “At least one...not”) for both premises *and* conclusions. All participants were also presented with universal affirmative (“All”) and universal negative (“No”) premise statements. Each problem used one of the four premises with all four conclusions. Thus, each participant was presented with four unique problems, which were randomly ordered on the pages in each of six blocks for a total of 24 problems. The conclusion statements were presented in the same four counterbalanced orders as in Experiment 1. For half of the participants, participants were to assume the premise was true for the first half of the problems and false for the second half of the problems. The other half of the participants received the false premises first and true premises second. This order manipulation was orthogonal to the counterbalancing order of conclusion statements and to the quantifier manipulation. The premise headings were changed such that the word “TRUE” or “FALSE” was bold, italicized, and underlined (e.g., “IF IT IS

TRUE THAT:”) for the true and false premise problems, respectively. As in Experiment 1, participants were to put a check next to the “Y,” “N,” or “M” for each conclusion statement.

Procedure. The procedure for Experiment 2 was identical to Experiment 1 except that there were two instruction pages. The instruction page for the true premise problems was identical to that used in Experiment 1. The instruction page for the false premise problems was only slightly modified (i.e., because different conclusions follow from true and false versions of the same premise), but was otherwise identical.

Results

The dependent measures for Experiment 2 were proportion logical and pragmatic responses; as before we will base our analysis primarily on the logical pattern. Results will be presented in three subsections: (1) overall logical responses, (2) true premise problems with particular premises (i.e., the conditions used in Experiment 1), and (3) true premise problems with universal premises, which provides an additional test of the quantifier hypothesis. Note that five participants (four in the standard condition and one in the logically-clarified condition) were removed from Experiment 2 for having three or more errors on any of the identity, contradictory, or contrary problems. Failure to respond logically to these self-evident problems suggests that these participants did not comprehend the task. We did not use a similar exclusion criterion for Experiment 1 because this would have concealed the reduction in acceptance of identity problems in the logically-clarified condition. Removing the five participants from the present experiment does not alter any of the findings reported below.

Overall Logical Responses. To broadly characterize the data, a 2 (quantifier; standard, logically-clarified) x 2 (premise truth; true, false) mixed ANOVA for proportion logical responses was conducted. These data are reported in Figure 4. As is clear in the figure, more

logical responses were given for true (.81) than false (.57) premise problems, $F(1,28) = 64.691$, $MSE = .014$, $p < .001$. As expected, the ANOVA also revealed that logical responding was higher for logically-clarified problems (.75) than standard problems (.63)., $F(1,30) = 5.906$, $MSE = .037$, $p = .021$. The interaction was non-significant, $F(1,28) = .067$, $MSE = .014$, $p = .797$, indicating that the advantage for logically-clarified quantifiers is about the same with true and false premise problems.

 Figure 4 about here

However, our suspicion that participants would have difficulty with the false premise problems was confirmed. This data was quite noisy both within and between participants, making more fine-grained analyses difficult to interpret and low in power. In particular, many participants appeared to have simply reversed the polarity of the premise, as we hypothesized. Several participants specifically edited the problems in this way in their testing books (e.g., the “not” in “Some...not” problems would be crossed out, or “None” would be crossed out and replaced with “All”). Interpretation of responses therefore becomes difficult (e.g., if a reasoner thinks that a false “At least one” premise is equivalent to a true “At least one...not” premise but otherwise reasons logically, then they will think that the logically-necessary “No” conclusion is merely possible, which is exactly the response one would expect the reasoner to make if they solved the problem pragmatically). For the rest of this section we will focus on the true premise data.

True Particular Premise Problems (Some and Some...not). The proportion logical response data for problems with a true particular premise are presented in Figure 5. A 2

(quantifier; logically-clarified, standard) x 4 (problem type; contradictory, subaltern, subcontrary, identity) ANOVA for logical responses revealed a significant main effect for quantifier, $F(1,25) = 7.672$, $MSE = .752$, $p = .010$, a main effect for problem type, $F(3,75) = 28.008$, $MSE = 1.499$, $p < .001$, and an interaction, $F(3,75) = 5.641$, $MSE = .302$, $p = .002$. Planned comparisons were then conducted.

 Figure 5 about here

As in Experiment 1, a quantifier effect was predicted for both the subaltern (Some/Some...not followed by All/ None) and subcontrary problems (Some followed by Some...not, and vice versa). The data were consistent with both predictions. For the subaltern problems, participants made more logical responses with “At least one” quantifiers (.82) relative to “Some” quantifiers (.43), $t(25) = 3.142$, $p = .004$, $SE_{diff} = .125$. This was mirrored by a reduction in pragmatic responses for “At least one” quantifiers (.17) relative to “Some” (.56), $t(25) = 3.287$, $p = .003$, $SE_{diff} = .119$. Similarly, for the subcontrary problems, there were more logical responses with “At least one” quantifiers (.66) relative to “Some” quantifiers (.36), $t(25) = 2.075$, $p = .048$, $SE_{diff} = .147$, a trend that was again mirrored by a reduction in pragmatic responses for “At least one” quantifiers (.23) relative to “Some” (.60), $t(25) = 2.580$, $p = .016$, $SE_{diff} = .141$.

Responses to the contradictory and identity problems were not expected to differ for standard and logically-clarified quantifiers. For the contradictory problems, the rates of logical responses was almost identical for “Some” (.96) and “At least one” quantifiers (.97), $t(25) = .229$, $p = .821$, $SE_{diff} = .033$; the same was true for identity problems for “Some” (1.00) and “At least one” quantifiers (.97), $t(25) = 1.667$, $p = .108$, $SE_{diff} = .020$.

True Universal Premise Problems (All and None). The proportion logical response data for problems with a true universal premise are presented in Figure 6. A 2 (quantifier; logically-clarified, standard) x 4 (problem type; contradictory, subaltern, contrary, identity) ANOVA for logical responses revealed a significant main effect for quantifier, $F(1,25) = 6.203$, $MSE = .193$, $p = .020$, a main effect for problem type, $F(3,75) = 14.449$, $MSE = .474$, $p < .001$, and an interaction, $F(3,75) = 3.643$, $MSE = .120$, $p = .016$. Planned comparisons were then conducted.

 Figure 6 about here

Logically-clarified quantifiers were expected to increase logical responses for universal subaltern problems (All/ None followed by Some/ Some...not). Consistent with this prediction, participants made significantly more logical responses with “At least one” quantifiers (.87) relative to “Some” quantifiers (.58), $t(25) = 2.086$, $p = .047$, $SE_{diff} = .136$. As expected, this was mirrored by a reduction in pragmatic responses for “At least one” quantifiers (.03) relative to “Some” (.35), $t(25) = 2.965$, $p = .007$, $SE_{diff} = .106$.

Differences between the quantifiers were not expected for the remaining problem types, and the pattern of responses was almost identical for the two premise types. The proportion of logical responses for “Some” and “At least one” quantifiers was .99 and 1.00 respectively for contrary problems, $t(25) = 1.124$, $p = .272$, $SE_{diff} = .013$; .96 and 1.00 for contradictory problems, $t(25) = 1.573$, $p = .128$, $SE_{diff} = .026$; and 1.00 and 1.00 for identity problems.

Discussion

These findings replicate and extend the findings from Experiment 1. For both false and true premises, “At least one” quantifiers reduced pragmatic responses and promoted logical

responses on an immediate inference task. Moreover, the size of the effect was substantial. For the true premise data, logical responses were between 30% and 40% higher in the logically-clarified than the standard condition for the critical problem types, resulting in an increase in overall logical responses from about 50% to about 75%. In the false premise data, there were similar increases in logical responding in the logically-clarified condition. Finally, consistent with other findings (Barres & Johnson-Laird, 2003), we observed that participants had difficulty working out the implications of false premises, many choosing to simply reverse the polarity of the premise (e.g., a false “All” implies a true “No”).

Experiment 3

The previous two studies provide evidence that pragmatic responses in an immediate inference task can be reduced with logically-clarified quantifiers. The goal of this experiment was to determine whether performance on the more complex syllogistic reasoning task can be facilitated by presenting the premises in logically-clarified format. Recall that in the syllogistic reasoning task reasoners are presented with two premise statements followed by a conclusion statement. The premises and conclusions each contain one of four quantifiers used on the immediate inference task, and are composed of three terms, called the A, B, and C terms. The B term is repeated in the premises; the conclusion joins the A and C terms as illustrated below:

All of the Beekeepers are Artists
Some of the Cyclists are Beekeepers
 Therefore, some of the Cyclists are Artists

In the current study, half the participants solved syllogisms using the logically-clarified premises and half solved syllogisms with standard premises. If pragmatic interpretations are a source of error in syllogistic reasoning task, logical performance should be increased (and pragmatic responses reduced) using logically-clarified relative to standard premises.

Method

Participants. Seventy-eight University of Saskatchewan undergraduates with no background in logic completed Experiment 3. Participants received course credit or five dollars for participating. None of the participants in Experiment 3 had participated in either of the first two studies.

Materials and Design. As in Experiments 1 and 2, the A, B, and C terms of the syllogisms were instantiated using combinations of hobbies and professions. Half of the participants were presented with standard quantifiers for particular statements (“Some,” “Some...not”) and half were presented with logically-clarified quantifiers for both the premises and conclusions (“At least one,” “At least one...not”). There are 20 syllogisms for which a pragmatic interpretation of the premises predicts a different response than a logical interpretation. Each of the ten premise pairs was accompanied by two conclusions, as illustrated below:

IF IT IS TRUE THAT:

Some of the chemists are not beekeepers

All of the beekeepers are musicians

THEN IS IT THE CASE THAT:

None of the musicians are chemists Y__ N__ M__

Some of the musicians are chemists Y__ N__ M__

Both of the provided conclusions were logically possible, but not logically necessary. That is, the conclusions were consistent with a logical interpretation of the premises, but were not necessitated by the premises. The conclusions differed, however, in terms of their status under a pragmatic interpretation; one conclusion was pragmatically impossible and the other pragmatically true. We will refer to these as “pragmatically false” and “pragmatically true” conclusions. For example, “None of the musicians are chemists” is a logical possibility, but only if the first premise is interpreted to allow for “None of the chemists are beekeepers” to be true.

On the other hand, under a pragmatic interpretation (in which some chemists must be beekeepers), the conclusion that “None of the musicians are chemists” is impossible, and the conclusion is false. Thus, participants who made a pragmatic interpretation should choose “no” and those making a logical interpretation should choose “maybe.” Similarly, under a pragmatic interpretation, the conclusion “Some of the musicians are chemists” follows necessarily from the premises.

All of the syllogisms are presented in Table I. Because the logical response to all of the critical syllogisms was “maybe,” four simple filler questions were added that had non-“maybe” answers. The filler questions are also presented in Table I. Thus, each participant was presented with fourteen unique problems, with two conclusions each. The problems were always presented in the same order with filler problems strategically placed throughout (problems 1, 4, 8, and 11). However, the order of conclusions (universal first or particular first) was randomly varied across problems and counterbalanced across participants.

 Table I about here

Procedure. The procedure for Experiment 3 was identical to the previous two studies with a few exceptions. First, part of the sample was tested in a large group (half in the standard and half in the logically-clarified condition). Second, the instruction page used a simple syllogism for explaining when “Y,” “N,” and “M” responses were appropriate (but was otherwise identical to Experiment 1). This problem is also presented in Table I.

Results

The dependent measures for Experiment 3 were proportion logical and pragmatic

responses for the 10 critical syllogisms; again, however, we will focus our analysis on the logical responses. The logical response was always “maybe.” The pragmatic response was always “yes” and “no,” respectively, for pragmatically true and pragmatically false problems. For the four filler problems, participants had similar rates of logical responses in the “At least one” (.77) and “Some” (.67) conditions, $t(28) = 1.555$, $p = .131$, $SE_{diff} = .067$. These problems will not be discussed further.

The proportion of logical responses is presented in Figure 7. A 2 (quantifier; standard, logically-clarified) x 2 (problem type; pragmatically true, pragmatically false) ANOVA for logical responses revealed a significant main effect for problem type, $F(1,76) = 9.098$, $MSE = .024$, $p = .003$, indicating that more logical responses were made to pragmatically true than pragmatically false conclusions (i.e., for the pragmatically false responses, participants frequently responded “no” rather than “maybe”). More importantly, the ANOVA also revealed a main effect for quantifier, $F(1,76) = 10.635$, $MSE = .120$, $p = .002$, indicating that overall problems with logically-clarified quantifiers were given logical responses more often than problems with standard quantifiers. There was no interaction, $F(1,76) = 2.612$, $MSE = .024$, $p = .110$. To verify that the difference between the quantifiers was observed for both problem types, planned t -tests were computed. For pragmatically false problems, participants made significantly more logical responses to logically-clarified (.44) than to standard quantifiers (.22), $t(76) = 3.742$, $p < .001$, $SE_{diff} = .006$. This was mirrored by a greater number of pragmatic responses for standard (.73) relative to logically-clarified quantifiers (.51), $t(76) = 3.652$, $p < .001$, $SE_{diff} = .006$. Similarly, for pragmatically true problems, participants made significantly more logical responses to logically-clarified (.48) than to standard (.33) quantifiers, $t(76) = 2.250$, $p = .027$, $SE_{diff} = .006$. Again, this was mirrored by a greater number of pragmatic responses for standard (.58) relative to logically-

clarified quantifiers (.40), $t(76) = 3.104$, $p = .003$, $SE_{diff} = .006$.

 Figure 7 about here

Discussion

The results of Experiment 3 showed that pragmatic responses decrease with the use of logically-clarified “At least one” premises on those syllogisms that should produce different conclusions depending on whether a logical or pragmatic interpretation was made of the particular premises. This difference was also accompanied by an increase in logical responses. As such, the present data show that pragmatic responses *do*, on their own, influence the outcome of reasoning on the syllogistic inference task, in that the influence of pragmatics could not be reduced if pragmatics were not involved in the task in the first place. Thus, where previous reports (Newstead, 1995, 2003; Roberts et al., 2001) were unclear about the role of Gricean interpretations per se (as opposed to reversible Gricean interpretations), the present results demonstrate that pragmatic responses do occur in the syllogistic reasoning task and that using logically-clarified quantifiers can reduce the inclination of reasoners to make these pragmatic responses.

General Discussion

In three experiments, we demonstrated that logically-clarified premises reduced pragmatic responses and facilitated logical responses in the immediate inference and syllogistic reasoning tasks. Moreover, the effect of our quantifier manipulation was large (e.g., as high as a 40% in Experiment 2). Most importantly, the logically-clarified premises similarly facilitated logical responses on a syllogistic reasoning task, demonstrating that at least some of the “poor”

performance on that task is attributable to the ambiguity of the premises.

There has been a long-standing debate regarding the degree to which “errors” reflect failures in logic or differences in interpretation. For example, Henle (1962) argued that there truly are no logical errors, but only misinterpretations of task premises (i.e., participants always reason logically, only sometimes they use the wrong premises). Weaker claims have been forwarded by mental logic theories (e.g., Newstead & Griggs, 1999) and verbal reasoning theory (Polk & Newell, 1995), which also emphasize the role of encoding processes on task performance (see also, Revlis, 1975). The fact that our experiments show that disambiguation of premises can substantially reduce errors can be considered support for these approaches, although it is also clear that participants also make logical errors including misunderstanding logical necessity (e.g., Evans et al., 1999) and failing to consider relevant alternatives (Newstead et al., 2002; Torrens, Thompson, & Cramer, 1999).

There are a number of non-mutually exclusive explanations for how this facilitation might occur. The first is that standard quantifiers are ambiguous in their meaning, such that some reasoners interpret them logically and others interpret them pragmatically. The use of clarified premises may remove one of these interpretations and increase the probability that reasoners adopt the desired interpretation. The findings of Experiment 1 support this interpretation in that response patterns for the standard quantifiers were intermediate to either the pragmatic or logically clarified quantifiers. Similarly, some reasoners may be aware of the ambiguity of interpretation and this uncertainty in and of itself impedes their ability to solve capacity-demanding syllogisms.

Alternatively, use of the clarified premises may change the way in which the problems are represented. For example, the logically clarified (“At least one”) version might facilitate a

representation of concrete tokens that allows the set relationships of the premises to be clearly represented. Similar explanations have been proposed to explain why presenting information in the form of frequencies (i.e., 90 out of 100) as opposed to probabilities (i.e., 90%) increases accuracy on various probabilistic and statistical reasoning problems (Evans, Handley, Perham, Over, & Thompson, 2000; Girotto & Gonzales, 2001; Sloman, Over, Slovak, & Stiebel, 2003).

Regardless of the mechanism involved, our findings have immediate practical implications for studies using syllogistic reasoning. Specifically, given the ambiguity in the meaning of the traditional quantifiers, reasoning researchers would be better served by using clarified quantifiers in place of the standard quantifiers. The choice of whether to use pragmatically- or logically-clarified premises would depend on the goal of the study. If the goal is to gain a measure of logical ability, using logically-clarified premises will eliminate an unwanted source of error variance. This is especially important when syllogistic reasoning is used as an index of logical reasoning, for example, when it is used to assess the role of working memory in reasoning (e.g., Capon et al., 2003; Copeland & Radvansky, 2004; Gilhooley et al., 1999; Quayle, & Ball, 2000) or the effects of age and other variables on reasoning (e.g., Fisk et al., 2005; Fisk, & Sharp, 2002; Gilinsky, & Judd, 1994; Smeets & De Jong, 2005). In such cases, it is not clear the extent to which variability in reasoning performance is due to logical competence or interpretive processes.

In contrast, it may be of interest to investigate the contribution of pragmatic factors to reasoning performance, and our methodology offers a means to do so. By comparing performance on pragmatically-clarified, standard, and logically-clarified versions of the premises, it is possible to gauge, for a particular participant, the contribution of interpretive factors to their performance. Conversely, if a researcher wished to study failures in reasoning, then removing as

much variability as possible from the encoding phase would allow a more accurate and unconfounded look at these errors. Thus, appropriate use of standard versus logically-clarified quantifiers highly depends on the researcher's goals.

It is important to note, however, that even with logically-clarified premises reasoning is far from perfect. There are at least two explanations for this: First, it is possible that (at least for some reasoners) the meaning of these logically-clarified statements is still not perfectly clear. Alternatively, it is possible that there are other performance factors, such as working-memory limitations and misunderstanding of logical concepts that impact performance. Further research is required to determine how many of these errors are due to misinterpretation of premises during encoding and how many are due to failures in reasoning. Further research is also warranted to discover new ways of clarifying the meaning of quantifiers in logic tasks. However, although it certainly does not seem to be the case that logically-clarified premises resolve all pragmatic misinterpretations, it is our position that our logically-clarified premises should be favoured in further research, because they reduce a potentially irrelevant source of variance and thereby produce a purer measure of logical competence.

A further and perhaps more important practical implication of the results of the present investigation applies to the teaching of logic to students. Logic textbooks represent the particular affirmative and particular negative moods with the standard "Some" and "Some...not" statements, respectively. The present results suggest a refinement of this approach. Learning of the logical relations in the Square of Opposition would be greatly facilitated with the use of logically-clarified "At least one" and "At least one...not" statements. A change of quantifier use is therefore justified for practical reasons.

A change in quantifier use in research and in teaching is also justified on theoretical

grounds. The *particular* affirmative and *particular* negative moods are supposed to refer to *particular* instances of category members. Consistent with this, “At least *one*” refers to a *particular* member and indicates that there may be more particular members (with no upward boundary, thus not excluding the universal). “*Some*,” on the other hand, is *not* consistent with the theoretical underpinnings of the particular moods, because “*Some*” refers to a *group* of individuals. This is problematic for a number of reasons. First, this seems to rule out a singular (i.e., “one” or “one...not”) interpretation, because “one” is less than “Some” and “one...not” is more than “Some...not.” Second, this also rules out a universal (i.e., “All” or “No”) interpretation, because “All” is more than “Some” and “No” is less than “Some...not.” Both the singular and the universal are supposed to be possible interpretations of particular moods. As such, logically-clarified quantifiers should be favoured over standard quantifiers for theoretical reasons given that they are better representatives of the particular moods and for practical reasons given that reasoners better understand them.

Conclusions

The results of the three studies reported here using the immediate inference and syllogistic reasoning paradigms bring into question the continued use of standard “Some” and “Some...not” particular quantifiers in research and teaching. As demonstrated here, logically-clarified “At least one” and “At least one...not” quantifiers have a meaning more pragmatically consistent with the logical meaning of the particular moods. This leads to greater logical performance by participants in immediate inference and syllogistic reasoning tasks and would likely facilitate learning of logical relations greatly in the classroom.

Author Note

James R. Schmidt and Valerie A. Thompson, Department of Psychology, University of Saskatchewan.

This research was supported by a grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) to VAT and by an NSERC Undergraduate Student Research Award (USRA) to JRS. Correspondence concerning this article should be addressed to James R. Schmidt, Department of Psychology, University of Waterloo, 200 University Ave. W., Waterloo, Ontario N2L 3G1. E-mail: j4schmid@watarts.uwaterloo.ca.

References

- Bacon, A. M., Handley, S. J., & Newstead, S. E. (2003). Individual differences in strategies for syllogistic reasoning. *Thinking & Reasoning, 9*, 133-161.
- Barres, P. E., & Johnson-Laird, P. N. (2003). On imagining what is true (and what is false). *Thinking and Reasoning, 9*, 1-42.
- Begg, I., & Harris, G. (1982). On the interpretation of syllogisms. *Journal of Verbal Learning and Verbal Behavior, 21*, 595-620.
- Bucciarelli, M., & Johnson-Laird, P. N. (1999). Strategies in syllogistic reasoning. *Cognitive Science, 23*, 247-303.
- Capon, A., Handley, S., & Dennis, I. (2003). Working memory and reasoning: An individual differences perspective. *Thinking & Reasoning, 9*, 203-244.
- Ceraso, J., & Provitera, A. (1971). Sources of error in syllogistic reasoning. *Cognitive Psychology, 2*, 400-410.
- Chater, N., Oaksford, M. (1999). The probability heuristics model of syllogistic reasoning. *Cognitive Psychology, 38*, 191-258.
- Copeland, D. E., & Radvansky, G. A. (2004). Working memory and syllogistic reasoning. *The Quarterly Journal of Experimental Psychology, 57A*, 1437-1457.
- Déret, D. (1998). Le syllogisme catégorique: Une réexamination du terme “certains” dans une perspective développementale. *Canadian Journal of Experimental Psychology, 52*, 174-182.
- Evans, J. St. B. T., Barston, J. L., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory & Cognition, 11*, 295-306.
- Evans, J. St. B. T., Handley, S. J., & Harper, C. N. J. (2001). Necessity, possibility and belief: A

- study of syllogistic reasoning. *The Quarterly Journal of Experimental Psychology*, 54A, 935-958.
- Evans, J. St. B. T., Handley, S. J., Harper, C. N. J., & Johnson-Laird, P. N. (1999). Reasoning about necessity and possibility: A test of the mental model theory of deduction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 1495-1513.
- Evans, J. St. B. T., Handley, S.J., Perham, N., Over, D.E., & Thompson, V.A. (2000). Frequency versus probability formats in statistical word problems. *Cognition*, 77, 197-213.
- Evans, J. St. B. T., & Thompson, V. A. (2004). Informal reasoning: Theory and method. *Canadian Journal of Experimental Psychology*, 58, 69-74.
- Feeney, A., Scrafton, S., Duckworth, A., & Handley, S. J. (2004). The story of *some*: Everyday pragmatic inference by children and adults. *Canadian Journal of Experimental Psychology*, 58, 121-132.
- Fisk, J. E., Montgomery, C., Wareing, M., & Murphy, P. N. (2005). Reasoning deficits in ecstasy (MDMA) polydrug users. *Psychopharmacology*, 181, 550-559.
- Fisk, J. E., & Sharp, C. (2002). Syllogistic reasoning and cognitive ageing. *The Quarterly Journal of Experimental Psychology*, 55A, 1273-1293.
- Geurts, B. (2003). Reasoning with quantifiers. *Cognition*, 86, 223-251.
- Gilhooley, K. J., Logie, R. H., & Wynn, G. (1999). Syllogistic reasoning tasks, working memory, and skill. *European Journal of Cognitive Psychology*, 11, 473-498.
- Gilinsky, A. S., & Judd, B. B. (1994). Working memory and bias in reasoning across the life span. *Psychology and Aging*, 9, 356-371.
- Giroto, V., & Gonzales, M. (2001). Solving probabilistic and statistical problems: A matter of

- information structure and question form. *Cognition*, 28, 247-276.
- Grice, H. P. (2002). Logic and conversation. In D. J. Levitin (Ed.), *Foundations of cognitive psychology: Core readings* (pp. 719-732). Cambridge, MA: MIT Press. (Reprinted from *Syntax and semantics 3: Speech acts*, pp.26-40, by P. Cole & J. Morgan, Eds., 1975, New York: Academic Press)
- Henle, M. (1962). On the relation between logic and thinking. *Psychological Review*, 69, 366-374.
- Klauer, K. C., Musch, J., & Naumer, B. (2000). On belief bias in syllogistic reasoning. *Psychological Review*, 107, 852-884.
- Newstead, S. E. (1989). Interpretational errors in syllogistic reasoning. *Journal of Memory and Language*, 28, 78-91.
- Newstead, S. E. (1995). Gricean implicatures and syllogistic reasoning. *Journal of Memory and Language*, 34, 644-664.
- Newstead, S. E. (2003). Can natural language semantics explain syllogistic reasoning? *Cognition*, 90, 193-199.
- Newstead, S. E., & Griggs, R. A. (1983). Drawing inferences from quantified statements: A study of the Square of Opposition. *Journal of Verbal Learning and Verbal Behavior*, 22, 535-546.
- Newstead, S. E., & Griggs, R. A. (1999). Premise misinterpretation and syllogistic reasoning. *The Quarterly Journal of Experimental Psychology*, 52A, 1057-1075.
- 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.
- Newstead, S. E., Thompson, V. A., & Handley, S. J. (2002). Generating alternatives: A key

- component in human reasoning? *Memory & Cognition*, 30, 129-137.
- Politzer, G. (1991). Comparison of deductive abilities across language. *Journal of Cross-Cultural Psychology*, 22, 389-402.
- Polk, T., & Newell, A. (1995). Deduction as verbal reasoning. *Psychological Review*, 102, 533-566.
- Quayle, J. D., & Ball, L. J. (2000). Working memory, metacognitive uncertainty, and belief bias in syllogistic reasoning. *The Quarterly Journal of Experimental Psychology*, 53A, 1202-1223.
- Revlis, R. (1975). Two models of syllogistic reasoning: Feature selection and conversion. *Journal of Verbal Learning and Verbal Behavior*, 14, 180-195.
- Roberts, M. J., Newstead, S. E., & Griggs, R. A. (2001). Quantifier interpretation and syllogistic reasoning. *Thinking and Reasoning*, 7, 173-204.
- Rosenthal, D. J. A. (1980). Logical quantifiers in affirmative and negative contexts. *The Journal of General Psychology*, 102, 99-106.
- Sloman, S.A., Over, D., Slovak, L., & Stibel, J.M. (2003). Frequency illusions and other fallacies. *Organizational Behavior and Human Decision Processes*, 91, 296-309.
- Smeets, G., & De Jong, P. J. (2005). Belief bias and symptoms of psychopathology in a non-clinical sample. *Cognitive Therapy and Research*, 29, 377-386.
- Thompson, V. A. (1996). Reasoning from false premises: The role of soundness in making logical deductions. *Canadian Journal of Experimental Psychology*, 50, 315-319.
- Thompson, V. A. (2000). The task-specific nature of domain-general reasoning. *Cognition*, 76, 209-268.
- Thompson, V. A., Striemer, C. L., Reikoff, R., Gunter, R. W., & Campbell, J. I. D. (2003).

Syllogistic reasoning in time: Disconfirmation disconfirmed. *Psychonomic Bulletin & Review*, 10, 184-189.

Torrens, D., Thompson, V. A., & Cramer, K. M. (1999). Individual differences and the belief bias effect: Mental models, logical necessity, and abstract reasoning. *Thinking & Reasoning*, 5, 1-28.

Table I
Syllogisms Presented in Experiment 3

Premise Pairs	Conclusions
Example Syllogism:	
Aab-Abc	Aac, Eac, Aca
Filler Syllogisms:	
Aba-Acb	Aca, Oca
Aab-Abc	Aca, Oca
Eab-Acb	Eca, Ica
Aba-Abc	Eca, Ica
Critical Syllogisms:	
Oab-Abc	Eca, Ica
Eba-Ocb	Aca, Oca
Aab-Icb	Aca, Oca
Eba-Obc	Aca, Oca
Oba-Abc	Eca, Ica
Eab-Obc	Aca, Oca
Aba-Ocb	Eca, Ica
Eab-Ocb	Aca, Oca
Aba-Obc	Eca, Ica
Iba-Abc	Aca, Oca

Note: lowercase *a* and *c* represent the terms presented in the premises and conclusions and lowercase *b* represents the term that joins *a* and *c* in the premises. A = All, E = No, I = Some, O = Some... not. For instance, the premise pairs for “Oab-Abc” are “Some of the As are not Bs” and “All of the Bs are Cs.”

Figures

Figure 1. The Square of Opposition. There are four moods in the Square: A (universal affirmative; “All of the As are Bs”), E (universal negative; “None of the As are Bs”), I (particular affirmative; “Some of the As are Bs”), and O (particular negative; “Some of the As are not Bs”).

Figure 2. The Triangle of Opposition. There are three moods in the Triangle: A (universal affirmative; “All of the As are Bs”), E (universal negative; “None of the As are Bs”), and U (partition; “Some but not all of the As are Bs,” which is assumed to be equivalent pragmatically to “Some of the As are Bs” and “Some of the As are not Bs”).

Figure 3. Proportion of logical responses and 95% confidence intervals as a function of problem and premise type in Experiment 1.

Figure 4. Proportion of logical responses and 95% confidence intervals as a function of premise truth and premise type in Experiment 2.

Figure 5. Proportion of logical responses and 95% confidence intervals for the four true particular premise problem types as a function of type of quantifier in Experiment 2.

Figure 6. Proportion of logical responses and 95% confidence intervals for the four true universal premise problem types as a function of quantifier type in Experiment 2.

Figure 7. Proportion of logical responses and 95% confidence intervals for pragmatically true and pragmatically false syllogisms as a function of premise type in Experiment 3.

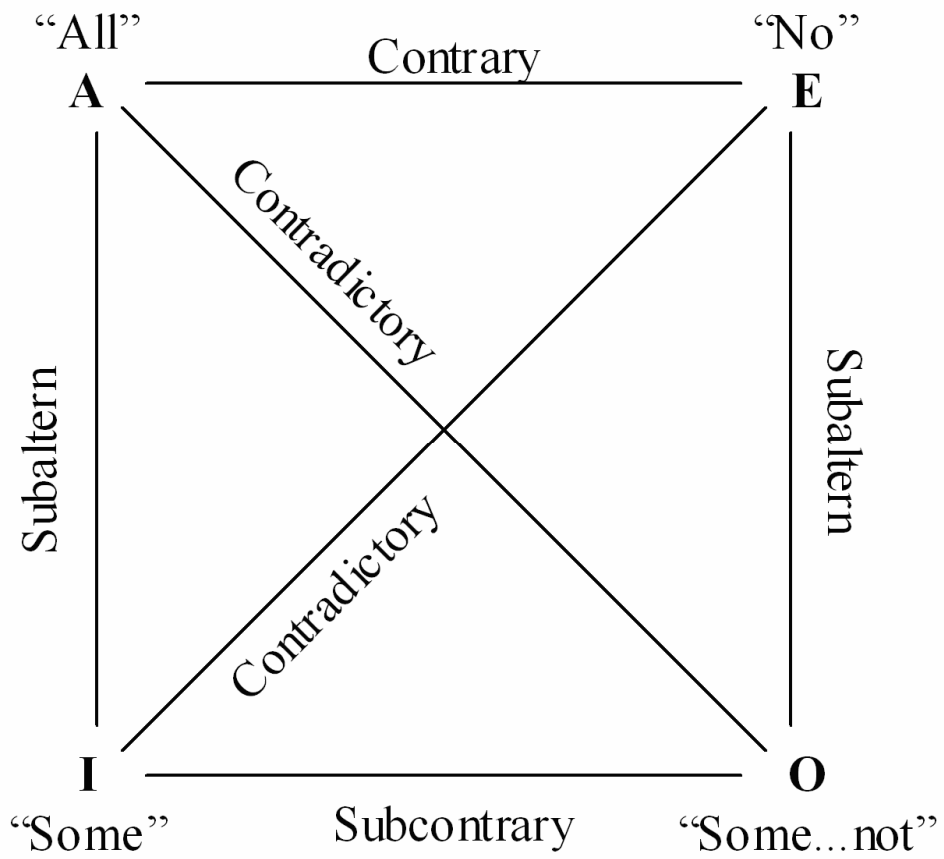


Figure 1 (C300).

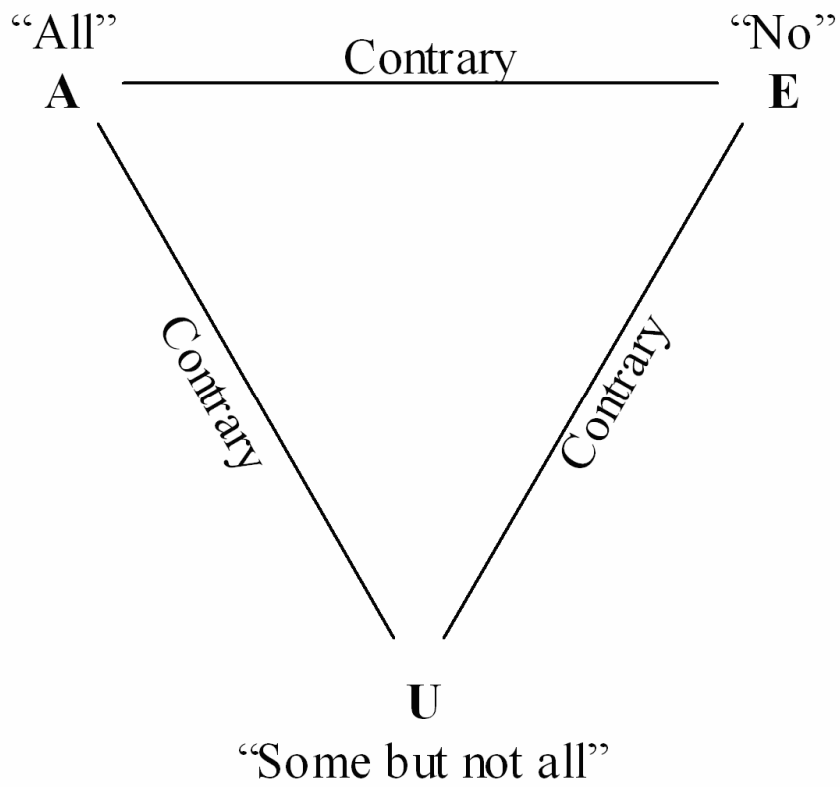


Figure 2 (C300).

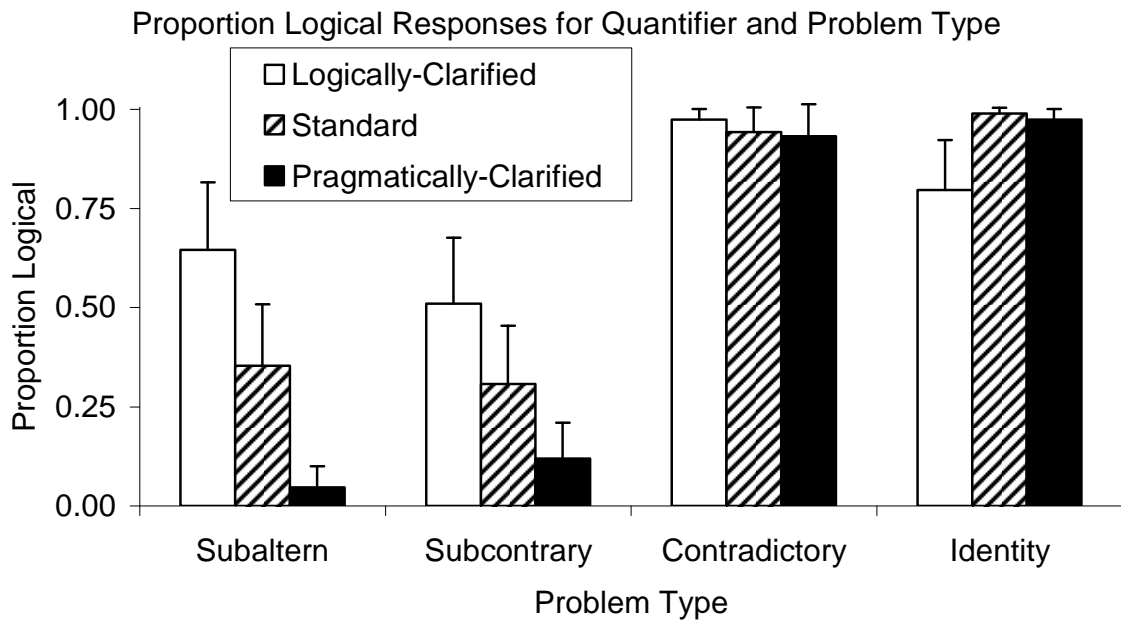


Figure 3 (C300).

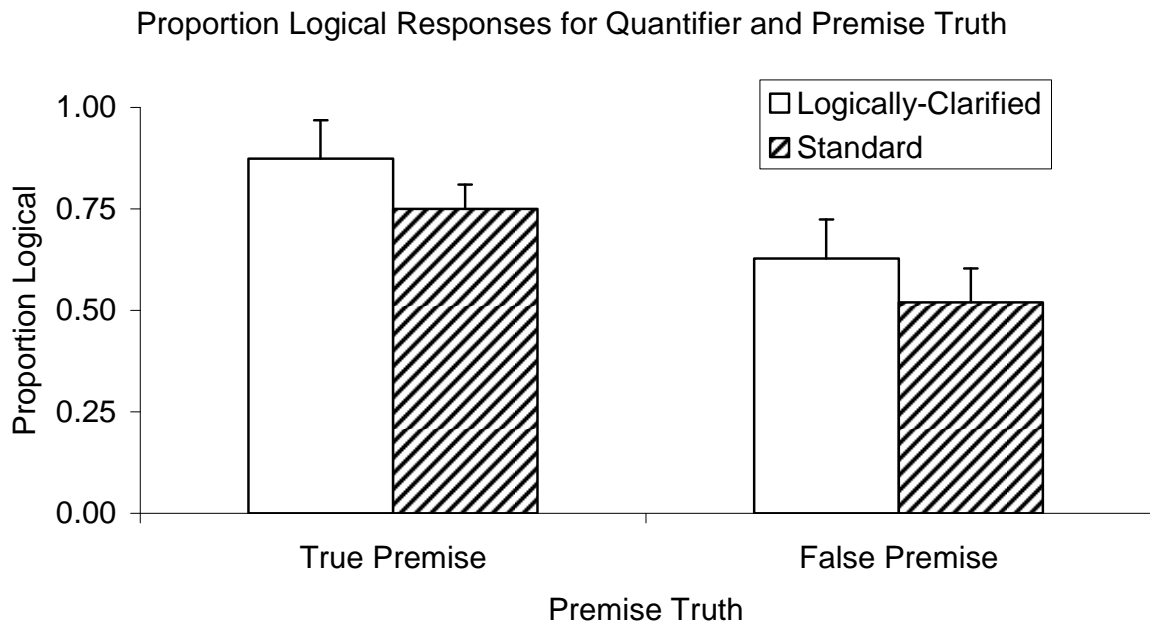


Figure 4 (C300).

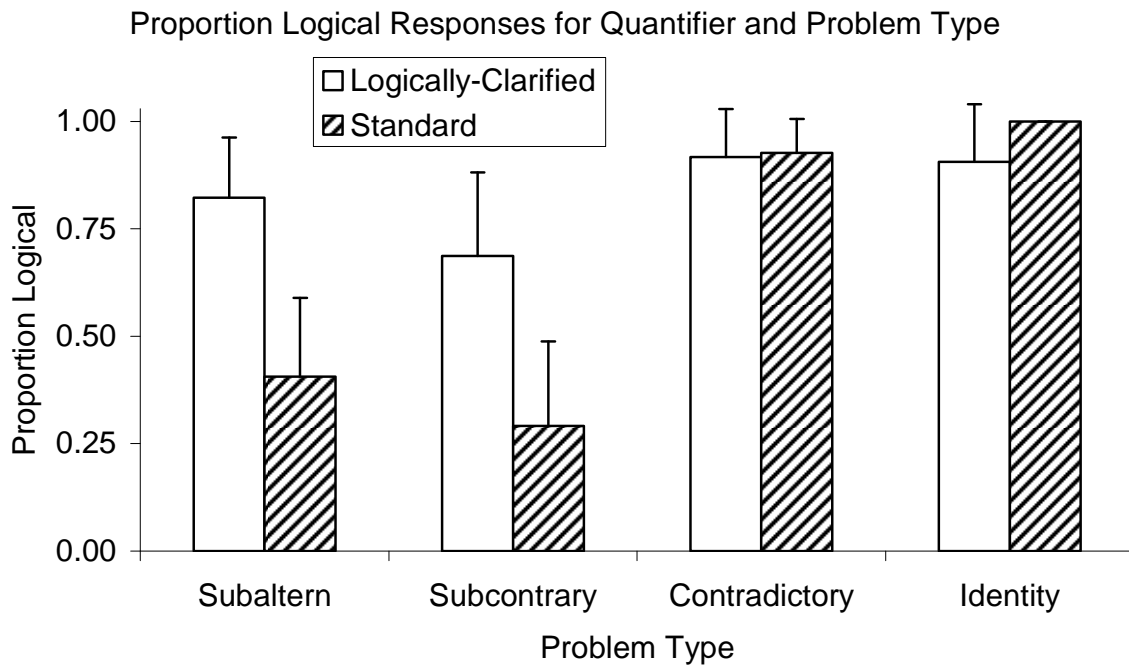


Figure 5 (C300).

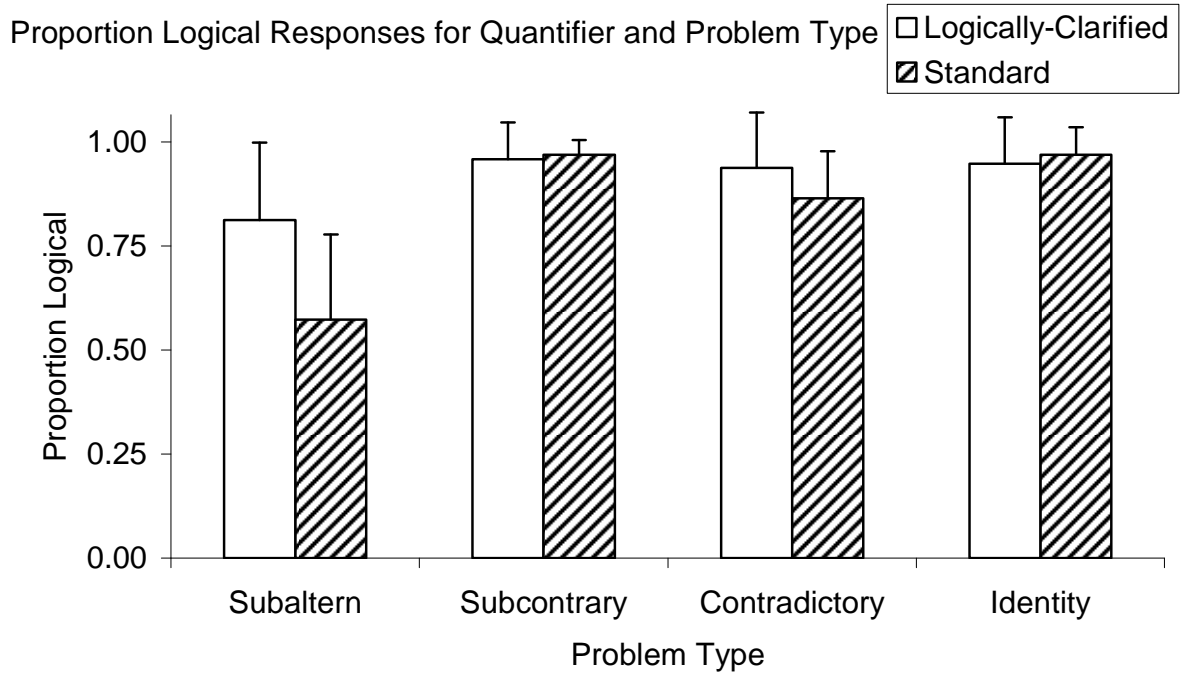


Figure 6 (C300).

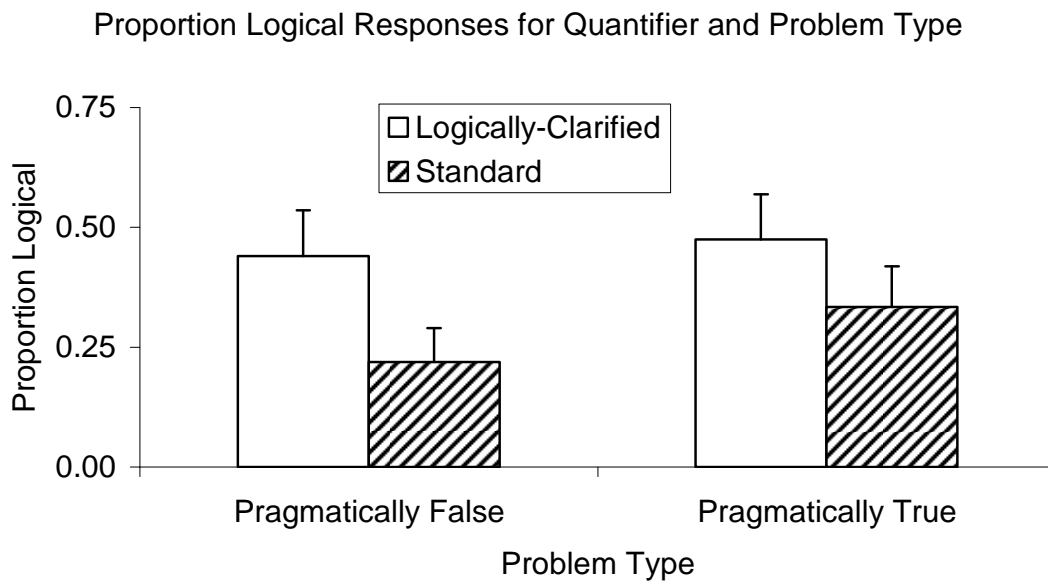


Figure 7 (C300).