

Understanding quantified categorical expressions

RUSSELL REVLIN and VON OTTO LEIRER
University of California, Santa Barbara, California 93106

A model of categorical inference (Revlis, 1975b) claims that a conversion operation participates in the encoding of quantified, categorical expressions. As a consequence, a reasoner is said to interpret such sentences as "All A are B" in a way that permits it to also be the case that "All B are A." The present study examines this conception of encoding using a sentence-picture verification task. In two experiments, students were asked to judge whether one of five possible Euler diagrams was true or false of a categorical expression (e.g., All A are B, No A are B, Some A are B, Some A are not B). Verification errors support a three-stage verification model whose major component is access to a "meaning stack" representing the progressive analysis of categorical relations; at the top of that stack is a converted reading of the input sentence. These findings have implications for current conceptions of categorical inference and semantic retrieval.

Some of the recent models of syllogistic reasoning make strong and largely unsupported claims concerning our understanding of the categorical relations we are asked to reason about (e.g., Erickson, 1974; Johnson-Laird, 1975; Revlis, 1975b). One such model, the conversion model (Revlis, 1975b), although accurate in predicting reasoners' decisions, makes a particularly counterintuitive claim regarding our representation of quantified relations; to wit, a conversion operation is said to participate in the encoding of such relations, so that when a reasoner is told that all A are B, he/she interprets this proposition to mean that the converse is also true (all B are A). In summary, the reasoner is said to make decisions from an interpretation, which, if expressed as a sentence, would be markedly different from the sentence as actually presented and intended by the experimenter. The purpose of the present study is to test for the operation of conversion in the encoding of quantified, categorical expressions using a sentence-picture verification task rather than an extended inference task. The findings have implications for models of syllogistic reasoning, as well as for conceptions of categorization and semantic retrieval (e.g., Just, 1974; Meyer, 1970).

Conversion Operation

The presence of conversion was first noted as a psychologically relevant variable by Wilkins (1928), but she was at a loss to account for its presence. More recent findings in psycholinguistics and categorical reasoning suggest that conversion of quantified relations may be based on one or both of the following two factors:

This research was supported in part by NSF Grant BNS 78-24763 to the first author. Requests for reprints should be sent to Russell Revlin, Department of Psychology, University of California, Santa Barbara, California 93106.

(1) the ambiguity of the relational copula "is a" that appears in all standard syllogisms, and (2) elementary comprehension mechanisms for analyzing quasisymmetrical relations.

The first factor is historically the earliest explanation for conversion and empirically the most tenuous. Chapman and Chapman (1959) report that reasoners appear to encode "is a" as "is equal to" (an identity relation) rather than the logical "is included in" (proper inclusion). Chapman and Chapman claim that this interpretation is a result of reasoners' experience with algebra and geometry, in which "is a" may be suitably understood as "is equal to." One implication of this conception of conversion is that the universal affirmative, "All A are B," would be interpreted as "A equals B" (identity relation). Three studies of encoding challenge this hypothesis. Erickson, Wells, and Traub (Note 1) had students draw Venn diagrams as representations for all possible premises and premise pairs. They found that only 40% of the subjects systematically interpreted universal affirmatives as set identity. Neimark and Chapman (1975) examined the interpretation of sentences and diagrams within a developmental paradigm and found that college-aged students misinterpreted universal affirmatives in this fashion only 20% of the time. In contrast, Bucci (1978) found that 92% of children (6- to 8-year-olds who had had no experience with the special terms of algebra and geometry) make conversion errors.

A more viable explanation for conversion in our view is that it is an instance of general comprehension processes. We claim here that in the comprehension of quantified relations between distinct categories, there is an automatic operation that treats inputs as symmetric, so that on hearing, for example, "All A are B," the listener automatically entertains the notion that all

B are A. This encoding may not be stored, however, under specifiable conditions (see "Blocking," below). The conversion operation is said to apply not only to abstract, symbolic inputs (in which letters stand for categories), but to concrete ones as well (see Revlin & Leirer, 1978). As a result of this operation, the encoding for $A \rightarrow B$ relations (independent of quantifier) includes $B \rightarrow A$.

Note that this conception of conversion is substantially different from that offered by Chapman and Chapman (1959), with which it is often confused (e.g., Johnson-Laird & Steedman, 1978). Chapman and Chapman claimed that conversion applies primarily to universal-affirmative sentences, so that given "All A are B," the listener is said to understand this relation to mean that $A = B$. The admittedly counterintuitive view taken here is that subjects actually reverse the subject and predicate terms in the utterance as part of the comprehension process and that this is no mere accident (in contrast, see Newell, in press).

The justification for this conception of conversion as reflecting efforts toward symmetry is twofold: Our representation for categorical relations permits it, and conversational and communicative postulates invite it. In examining the nature of our representation for categorical expressions, Bucci (1978) notices that there is a discernible developmental sequence in the comprehension of quantified, categorical expressions. In the earliest stages, the child (tested at age 6 years) comprehends expressions such as "All the circles are yellow" to mean that everything in the universe of discourse (geometric shapes in a box) is both a circle and yellow: [(All) (circles) (yellow)]. This is an order-free relation (structure-neutral), in which one might think that circleness and yellowness were properties of all things in the universe of discourse (i.e., All things are yellow circles). This hypothesis accounts for some of the findings of Bucci (1978) and of Inhelder and Piaget (1964).

Bucci's (1978) data argue that in a second stage of development, the child (tested at age 11 years) is able to bind the quantifiers to the subject and predicate terms and derives a meaning that is closer to the logician's identity relation but that still allows conversion. Here, in glossing the question, "Are all the circles yellow?" the child understands that all circles are yellow and that all yellow objects in the set are circles, but that there may exist objects that are neither circles nor yellow.

In the third stage, adult reasoners convert the question "Are all the circles yellow?" to a form that allows "All yellow things are circles" and yet still has "available" the set-inclusion relation, "All of the circles are yellow; Some yellow things may not be circles." It is not clear from Bucci's (1978) data whether this unconverted inclusion relation is recomputed from the converted relation in response to specific queries or it exists in memory following comprehension of the

initial proposition. The conversion model makes a specific and important assumption in this regard in claiming that the inclusion meaning is present in memory even at the time when the reasoner appears to be employing a converted interpretation. This assumption will be treated below in a discussion of the hypothesized memory structure that stores the reasoner's encoding. In any case, we infer from Bucci's findings that the adult's representation of the categorical relation makes available a structure-neutral symmetric reading.

The conversion model of formal inference described here assumes that for adults, the $B \rightarrow A$ reading of an $A \rightarrow B$ statement is prepotent because of communicative presuppositions. This is especially clear in the encoding of particular negatives (e.g., Some A are not B). The principle involved is that in the act of communicating a relation, the most informative form is used. For example, when a speaker asserts a particular negative without a context (e.g., Some plants are not sale items), the listener knows that the converse of the statement is also true (Some sale items are not plants); this is so because if any universal relation between plants and sale items were appropriate (e.g., "All sale items are plants") or ("No sale items are plants"), normal conventions of communication would require their use (Grice, 1967). Consequently, the assertion of a particular negative readily leads to the acceptance of its converse.

It is a reasonable extension of this notion that the student may judiciously question the relation from the perspective of B given the sentence "All A are B." Quantified relations are at least quasymmetrical in the sense that for every quantified expression of the form "A relation₁B," there is a true expression of the form "B relation₂A." Such reversals of subject and predicate terms are the hallmark of symmetrical relations; true symmetry holds in just those cases in which relation₁ = relation₂. In the case of the four categorical propositions (All A are B, No A are B, Some A are B, Some A are not B), two cases of true symmetry occur (conversion or reversal of subject and predicate classes leave meaning unchanged): No A are B \rightarrow No B are A; Some A are B \rightarrow Some B are A. In addition, two cases of limited symmetry occur: All A are B \rightarrow Some B are A; No A are B \rightarrow Some B are not A. In all cases, there appears to be at least an intuitive basis for asserting a quasymmetrical relation between the subject and predicate classes in a syllogism. In addition, Tsai (1977) provides evidence that when confronted with ambiguous relations in a reasoning task, students frequently assume such relations are either symmetrical or both symmetrical and transitive. Such assumptions may be isomorphic to gestalt judgments of "good figure" and automatically function to transform potentially ambiguous relations into conceptual good figures (cf. Desoto & Kueth, 1958, 1959).

This notion of symmetry has become part and parcel of the explanations given to inference errors in recall

for linear orderings (Potts, 1978) and categorical orderings (Griggs, 1978). This also explains many circumlocutions used by students in avoiding universally quantified categoricals when a nonsymmetrical interpretation is intended (Thomas, Note 2).

Blocking Conversion

Conversion of quantified relations between nominal categories will not always be the encoding of choice. Conversion may be blocked in those cases in which relations are sufficiently explicit for symmetry not to be justified: For example, students are unlikely to treat a sentence of the form "All dogs are animals" as permitting the relation "All animals are dogs." We propose two criteria for the blocking of the conversion operation: (1) semantics of the relational terms and (2) distinguishability of the categories. Semantic analysis reveals that relations such as "taller than," "to the left of," "inside of," "is happier than," and "dominates" are nonreversal relations and are so treated by students in reasoning tasks (e.g., Desoto & Kuethe, 1959; Johnson-Laird, 1975). Conversion may be blocked in a manner analogous to having special components in the lexical entries for such terms that do not permit reversibility (or limitations in the procedural understanding of relational terms that do not permit the full expression of conversion).

Distinguishability of the categories refers to the saliency of the disjunction of the categories, that is, to the portion of the categories that do not participate in the stated relationship. In such relations as "All A are B," if A and B are relatively equal in size and if the portion of B that is not contained in A is small relative to the union of A and B, symmetry is permitted. If the portion of B not included in the relation is large, it becomes salient and demands a special label and consideration. For example, Revlis (1975a) has shown that a sufficient condition for symmetry interpretation (conversion) exists when reasoners are shown an expression such as "All A are B" in which at least 85% of the predicate category, B, is contained in the union of the subject and predicate categories; in contrast, a sufficient condition for blocking conversion exists when only 50% of B is contained in the union of the subject and predicate categories. While these criteria for blocking conversion are not definitive, they provide heuristics for independent predictions of when conversion is likely to occur and when it is likely to be blocked.

Conversion and its blocking would at first appear to reflect solely the subjects' knowledge of the relational terms. If this is so, it must entail deep long-term knowledge, because the action of conversion is not readily conditioned by the immediate past experience with concrete examples (e.g., Revlin, Ammerman, Petersen, & Leirer, 1978; Tsai, 1977). This implies the obligatory nature of the encoding mechanism and suggests that conversion is not simply one among many ways to dis-

ambiguate sentences, but rather that conversion should be viewed as the default encoding. This is shown most strongly for symbolic statements (e.g., All A are B), but it is clearly present in the encoding of concrete relations as well (e.g., All plants in the shop are sale items, see Revlin et al., 1978).

The foregoing has been by way of motivation for the hypothesis that a conversion operator is applied in the encoding of quantified, categorical expressions and that, among other elements, the listener stores a meaning equivalent to the converse of the input sentence. The manner of storage of this information and the subject's ability to retrieve entailed relations are described in the following section concerning the "meaning stack."

Encoding

The conversion model assumes that the encoding of categorical statements proceeds as a progressive extraction of elementary features (Revlis, 1975a). The temporal sequence involved is analogous to the construction of a stack of meanings for each categorical statement. Figure 1 shows that this entails the isolation of orthographic-acoustic features, quantifiers, subject, object, and relational terms. Most important for the present study is that the reasoner stores both the intended inclusion relation ($n - 1$ level) and the converted meanings (n th level) in his data base, with the converted proposition given priority in the "stack" of meanings because its representation is generated later than the unconverted ones. By this notion of a pushdown stack, we encompass both the reasoner's current understanding of the input sentence and the temporal sequence leading to that understanding. It should be pointed out that the k th level corresponds to the features of quantifier and polarity that have played an important role in an explanation of syllogistic reasoning errors by Woodworth and Sells (1935) (i.e., the "atmosphere effect"). They conjectured that conversion was due to an extraction of

<u>STACK LEVEL</u>	<u>SENTENCE CODE</u>
N	QUANTIFIER (PREDICATE(PREDICATE IS A SUBJECT))
N-1	QUANTIFIER (SUBJECT(SUBJECT IS A PREDICATE))
.	
.	
K	(QUANTIFIER) (SUBJECT) (PREDICATE)
K-1	(QUANTIFIER) (SUBJECT)
K-2	(QUANTIFIER)
.	
.	
K-i	PHONOLOGICAL FEATURES
.	
.	
K-j	ORTHOGRAPHIC FEATURES

Figure 1. Meaning stack of quantified categorical expressions.

these elements and that the reasoner's judgments in a syllogism task reflected just these features when the reasoner was unable to "see" the more complex relations (cf. Revlis, 1975b, for a process model embodying this explanation).

The k th through $n - 1$ levels of the meaning stack correspond also to Bucci's (1978) claim of a developmental sequence, beginning with a structure-neutral interpretation (k) and leading progressively to class inclusion ($n - 1$). Notice that the present formulation departs from Bucci's in at least two ways. First, the adult is claimed here to replicate the entire developmental sequence during the encoding of input sentences. Second, each encoding is not computed just when needed from the structure-neutral representation, but the entire stack is normally computed and stored, with each operation being applied to the immediately preceding encoding. These assumptions are motivated by constraints on Revlis' (1975a) reasoning model and are carried forward here.

The Task

The predictive value of a reasoning model that includes conversion as an essential encoding operation has already been established (e.g., Revlin & Leirer, 1978). However, direct support for the underlying conversion mechanism has not been evaluated. To this end, the present study examines the encoding of quantified, categorical expression in a sentence-picture verification paradigm. The present task is a simple one. Students are asked to say whether an Euler diagram¹ is true or false of a quantified sentence. For purposes of the present study, it is assumed that the verification process proceeds through four stages, as illustrated in Figure 2. This representation of the verification paradigm is similar to those shown by Carpenter and Just (1975), Chase and Clark (1972), and Trabasso, Rollins, and Shaughnessy (1971).

Stage 1. This stage has two parts. In the first, the sentence is encoded by creating a meaning stack of relations, as illustrated in Figure 1. In the second part, the picture is encoded in two steps: (1) isolation of the subject, predicate, and relational terms that have been specified by the sentence encoding and (2) a determination of what the relation in the picture should be, given the subject and predicate terms. Table 1 illustrates the

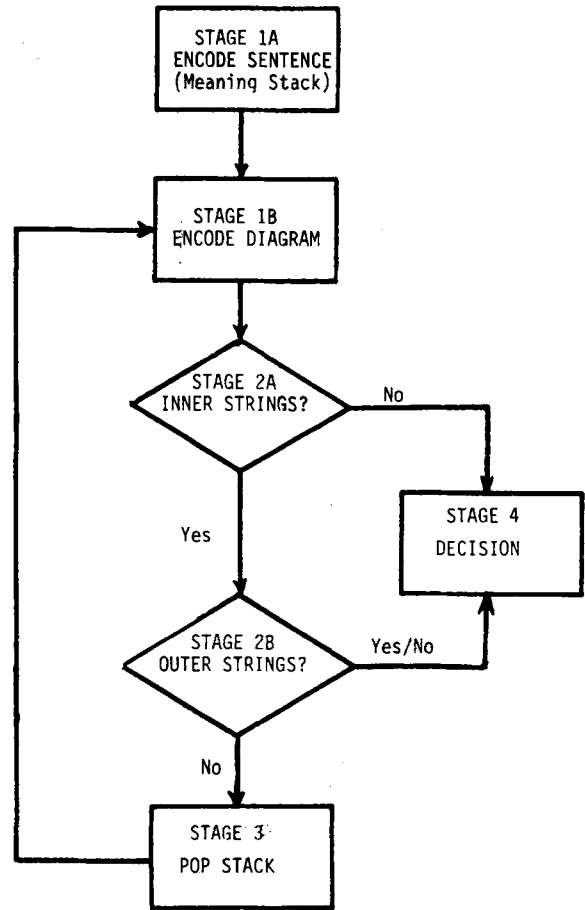


Figure 2. Sentence-picture verification model.

verification codes for each of five diagrams and the universal-affirmative sentence. It should be clear that the picture codes are not mere affirmations or denials of the sentence codes (e.g., Just, 1974) but are independently derived, depending only on an extraction of the salient components of the sentence.

Stage 2. Here, sentence and picture encodings are compared in a serial, self-terminating manner, beginning with the inner strings of the two encodings (strings are demarcated by brackets). The comparison process proceeds in two steps. In the first step, the inner strings are compared. They will be incongruous when one proposition affirms a relation and the other denies

Table 1
Verification Sequence for Universal Affirmatives

Picture	Picture Code	Stage 2A	Stage 2B	Stage 3	Stage 4
Identity	[All(B [Aff(B are A)])]	Y	Y		Y
Exclusion	[All(B [Neg(B are A)])]	N			N
Subset	[Some(B [Aff(B are A)])] [All(A [Aff(A are B)])]	Y	N Y	Y	Y
Superset	[All(B [Aff(B are A)])]	Y	Y		Y
Disjunction	[Some(B [Aff(B are A)])] [Some(A [Aff(A are B)])]	Y	N N	Y	N

Note—Based on sentence code [All(B [Aff(B are A)])]. Y = yes; N = no.

one, no matter what the quantifier or order of terms. In this case, since further processing is unnecessary, an immediate "no" response is called for, and the process enters Stage 4, the decision stage (see below). In the second step, the outer strings are compared. If the outer strings also match, the decision process enters Stage 4; otherwise, processing includes Stage 3.

Stage 3. This stage embodies the accessing of less derived encoding in the meaning stack. If a mismatch in the outer string occurs in Stage 2, the meaning stack is popped, and a decision is based on the next proposition retrieved (i.e., reentry of Stage 1B). This recursion occurs only in those cases in which both the sentence and picture encodings affirm a relation or both deny a relation.

Stage 4. If the sentence and picture codes are congruent, the subject will respond "yes"; if they are incongruent, the subject will respond "no."

It will be helpful to illustrate the process with some specific examples. The encodings and stage sequences for universal affirmatives are summarized in Table 1; in addition, Figure 3 shows the verification sequence for a universal-affirmative sentence (All A are B) and an identity diagram. The student encodes the sentence with the nth level equivalent to [All (B [Aff (B are A)])], in which this converted reading has the quantifier modifying the predicate class, B. The A and B categories are isolated in the picture, which the student encodes as [All (B [Aff (B are A)])]. The two encodings match on their inner and outer strings, and the student responds "yes." This is a logically correct judgment, although it is reached through a nonlogical, converted encoding of the sentence.

Figure 4 shows the decision sequence for an incorrect "yes" judgment. Here, the reasoner finds the sentence and picture codes congruent, although logic compels otherwise. To anticipate the final discussion somewhat, notice that the sentence and inclusion diagram pair in this figure is the troublesome false superset problem from Meyer (1970) that compelled his "category interchange" hypothesis.


STACK LEVEL	SENTENCE CODE	PICTURE CODE
N	[ALL (B[Aff(B are A)])]	[ALL(B[Aff(B are A)])]
		
N	[ALL (B[Aff(B are A)])]	→ B-A
N-1	[ALL (A[Aff(A are B)])]	
.		
.		
.		
INPUT:	ALL A ARE B	

Figure 3. Comparison of a universal-affirmative sentence with an identity diagram.

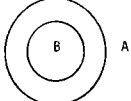
STACK LEVEL	SENTENCE CODE	PICTURE CODE
N	[ALL (B[Aff(B are A)])]	ALL (B[Aff(B are A)])
		
N	[ALL (B[Aff(B are A)])]	→ B-A
N-1	[ALL (A[Aff(A are B)])]	
.		
.		
.		
INPUT:	ALL A ARE B	

Figure 4. Comparison of a universal-affirmative sentence with a superset diagram.

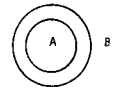
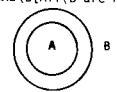
STACK LEVEL	SENTENCE CODE	PICTURE CODE
N-1	[ALL (A[Aff(A are B)])]	[ALL (A[Aff(A are B)])]
		
N-1	[ALL (A[Aff(A are B)])]	→ A-B
N	[ALL (B[Aff(B are A)])]	[SOME (B[Aff(B are A)])]
		
N	[ALL (B[Aff(B are A)])]	→ B-A
N-1	[ALL (A[Aff(A are B)])]	
.		
.		
.		
INPUT:	ALL A ARE B	

Figure 5. Comparison of a universal-affirmative sentence with a subset diagram.

Figure 5 illustrates a slightly different situation. The reasoner is shown the sentence "All A are B" and a set-inclusion diagram in which A is shown as a subset of B in the diagram. Logically, the sentence and diagram should be congruent. However, the picture code is initially formulated as [Some (B [Aff (B are A)])], which is incongruent with the sentence encoding [All (B [Aff (B are A)])].² Here, the reasoner invokes the third stage, pops his meaning stack for the sentence, and reencodes the picture accordingly. By this process, the picture and sentence will be judged congruent, and the student will once again respond correctly.

There are two major predictions from the model in terms of verification accuracy. The first of these is termed a concordance effect. As a result of conversion, the encoding and congruity-matching operations will, on some occasions, result in a decision that is concordant with a rational or logical decision. At other times, the encoding and congruity-matching operations will result in a decision that is discordant with logic. The concordance effect refers to the prediction that we

should observe high verification accuracy when the claimed operations are concordant with logic and low verification accuracy when the operations are discordant with logic. These predictions contrast with those of other models of quantified sentence verification (e.g., Just, 1974; Meyer, 1970) in which uniform accuracy across sentence-picture pairs is expected.

Related to concordance is the fact that conversion may be blocked when the semantics of the relational terms or the distinguishability of the categories make conversion pragmatically unacceptable. For example, if the sentence were "All A are included in B," conversion to "All B are included in A" would produce a proposition that lexically conflicts with the intended relation, since the including set would also have to be the included set. In such cases, the converted interpretation would not be stored. Of course, in the present example, the distinguishability of the categories would permit conversion in that case in which the included and including sets were identical (i.e., the area of non-overlap of the sets is zero). Blocking of conversion has implications for the verification task. A direct prediction is that when conversion is permitted (e.g., with an "is a" relation), the decisions for some of the sentence-picture pairs are discordant with logic and are error prone. Yet, those same sentence-picture verifications should be concordant with logic and highly accurate when conversion is blocked (when expressed with an "is included in" relation). For example, the convertible relation "All A are B" paired with the subset diagram (see Figure 5) results in a decision discordant with logic. Yet, when the relation is expressed as "All A are included in B," conversion is not permitted, and the verification should be concordant with logic.

The second area of prediction from this model concerns the effect of processing complexity on the verification judgments. It is reasonable to assume that decision accuracy should be negatively related to the number of operations that are required when the subject functions within the time constraints imposed by the present experiment. In such conditions, the subject may have insufficient time to carry out either the prescribed operations or other necessary, but as yet unspecified, operations. In the present paradigm, operational complexity is synonymous with stack popping (i.e., repeated passes through Stage 3), which may contribute a modicum of error to the reasoner's verification decisions. Reduced accuracy when Stage 3 is required will provide a tentative assessment of the model's tacit assumptions.

In addition to testing these two predictions, the study also examines additional issues concerning the students' encoding of the diagrams. Unfortunately, there is no available model of Euler diagram encoding that would permit us to independently determine what components of a picture are perceptually salient factors in determining the picture code that is adduced by the reasoners. The absence of formal descriptions of picture codes has not diminished the use of quantifier-diagram encoding

experiments (e.g., Johnson-Laird, 1969a, 1969b, 1970; but see a more restricted task of Just, 1974; Just & Carpenter, 1971). Of particular importance is the uncertainty of the perceptual features that will determine the constituent representation. For example, current descriptions do not acknowledge that students may be sensitive to the relative sizes of the sets as well as to their degrees of intersection. Which set is seen as the dominant or "subject" category may be critical for the comparison. For the sentence "All A are B," the potentially dominating, including class is mentioned last; in a sense, this violates normative subject-predicate topicality constraints (cf. Johnson-Laird, 1968a, 1968b). Since it is possible that the foregoing factors may be critical, the experiments presented here will examine the importance of intersection and set-size parameters.

EXPERIMENT 1

Method

Materials. Two groups of students ($n = 11$) were asked to verify whether an Euler diagram was true or false of a quantified sentence. One group of students verified sentences for which conversion is permitted at encoding ("is a"), and a second group verified sentences for which conversion is blocked (e.g., "is included in"). All students were shown every pairing of a sentence ("All A are B," "No A are B," "Some A are B," and "Some A are not B") and an Euler diagram twice (40 decisions: 4 sentences \times 5 diagrams \times 2 replications). On one pairing, the left-right order of letters in the diagram for the categories was A-B, and on the other pairing, B-A (the specific occasion for each pairing was randomly assigned).

The sentence-picture pairs were presented in a booklet. The sentence was centered at the top of the page, followed below by the diagram. (The A set was always outlined in blue ink, the B set, in red ink.) Each page was divided into quadrants. Each basic diagram had three possible, logically equivalent variations, one of which was placed in each of the bottom left, top left, and top right quadrants. The bottom right quadrant contained the phrase "None of the above." The variations in each diagram are illustrated in Figure 6. These variations were distributed to

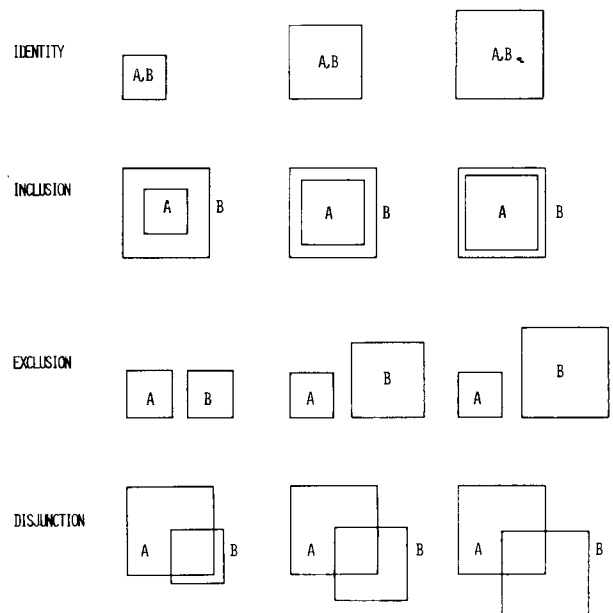


Figure 6. Euler diagrams used in Experiment 1.

the three picture quadrants by assigning the most readily discriminable diagram (the one with the largest difference in set size) to the upper left quadrant, the second most discriminable to the upper right quadrant, and the least discriminable to the bottom left quadrant. This procedure was intended to take advantage of normal reading order and to maximize verification accuracy. It also provides an opportunity to examine whether subjects' preferences for specific diagram variations would reveal regularities in subjects' expectations for relative set sizes (see below).

Procedure. The experimenter read aloud the standard definitions for each quantifier ("all" = each and every; "no" = none; "some" = at least one, possibly all; "some not" = at least one is not, possibly none are) and illustrated the task with a universal negative sentence and an exclusion diagram (a "true" sentence-picture pair that did not play a critical part in evaluating the motivating hypotheses). Students were allowed 20 sec for each verification decision, which consisted of two parts: (1) a response whether the sentence and diagram were congruent and (2) a ranking of the pictures in terms of subject's intuitions concerning the "informativeness" of the pictures vis-à-vis the sentence.

Subjects. The subjects were 44 introductory psychology students fulfilling a course requirement. They were run in groups of three or four in sessions lasting 30 min. Half of the students were randomly assigned to the experimental groups described above; the remaining students participated in a replication of the "is a" group.

Results

The accuracy scores for each group are presented in Table 2, which shows a significant concordance effect: Verifications were more accurate when conversion was claimed to be concordant with the logical prescription than when it was discordant with logic ["is a" group: $T(20) = 1.9, p < .05$]. In contrast, when the expressed relation blocks conversion in encoding ("is included in"), no significant effect was observed. Together, these data show the predicted interaction between convertibility of the relations and concordance: Verification errors increased in only those situations in which conversion produced a decision that was discordant with logic [$F(1,20) = 4.9, p < .05$].

To assess the reliability of the effect, a replication of the "is a" group was performed. Twenty-two students were asked to make verification judgments on the identical materials and under the same conditions as the "is a" group; the results are also presented in Table 2. The data show that the concordance effect was replicated. Concordant sentence-picture pairs were more accurately verified than were discordant pairs [$T(42) = 5.2, p < .01$].

There was a possible confound in the data: All of the discordant cases were ones in which the reasoner

Table 3
Percent Verification Accuracy at Two Levels of Operational Complexity

Relation	Stage 3	No Stage 3
is a	68.2	92.8
is a*	69.8	84.5
is included in	82.9	90.2

*Replication.

should logically reject the diagram as false of the sentence. It is possible, therefore, that these results more parsimoniously reflect the truth value of the verification rather than encoding per se. To examine whether the concordance effect might reflect truth-value parameters, we compared the verification accuracy of "true" and "false" sentence-picture pairs for which the encoding was concordant with the a priori (logical) truth values. This analysis showed only a modest true-false difference and no significant truth-value effect or interaction between truth value and quantifier for either the original group or the replication group. Consequently, the concordance effect cannot be readily attributed to the confounding of truth value and conversion.

The present model assumes that processing complexity adversely affects verification accuracy, where complexity is synonymous with repeated passes through Stage 3 (stack popping). The findings support this claim: Table 3 shows that verification accuracy was lower when Stage 3 was repeated than when it was not. This held for the original group ["is a": $T(20) = 3.78, p < .01$], as well as for its replication ["is a": $T(42) = 6.5, p < .01$]. Since stack popping was not required for the "is included in" relation, we expected that no processing complexity effect would be manifested [$T(20) = 1.1, p > .1$]. Again, Table 3 affirms the prediction and shows an interaction between processing complexity and sentence relation [$F(1,20) = 6.9, p < .01$]. It should be added that the extra operation required to pop the meaning stack did not confound with the concordance effect, since the discordant problems did not entail the popping operation. In fact, since popping occurred only on the concordant verifications, their true accuracy may have been higher than reported above.

Students were asked to rank order the diagrams for each verification in terms of the congruence between the picture and the sentence. It was felt that the first-choice evaluations might provide an independent basis for a hypothesized picture encoding process, since codes adduced by others (e.g., Just, 1974) are motivated primarily by fitting latency measures by following very closely the propositional structure of the sentence. Unfortunately, no systematic effect of sentence-picture pairs on preference judgments was apparent. Any effect that may have been present would have been obscured by the overwhelming tendency of students to select the top left diagram as most representative of the sentence (75.1% of the time). Since this may reflect an effect of discriminability, Experiment 2 included a necessary

Table 2
Percent Verification Accuracy: Area Fixed

Relation	Concordant †			Discordant †
	True	False	Average	
is a	86.4	77.9	82.1	68.2
is a*	84.1	78.4	81.3	52.3
is included in	90.9	91.7	91.3	93.2

*Replication. †With logic.

modification in the design to clarify any effect of picture format that may be critical for the verification task.

While the trends in the data accord well with the predictions, it is also clear that the precise levels of verification accuracy are not what would be expected: namely, that students would be perfectly accurate on concordant problems and perfectly inaccurate on discordant ones. To examine whether the disparity reflects nonhomogeneity of subject strategies, the data were analyzed in terms of individual performance functions. Three of the 11 subjects in the "is a" group made no errors. One subject made a single error on concordant problems, but none on discordant ones. Finally, seven students (approximately 64%) made many errors on the discordant problems (average accuracy = 35.7%) but made few errors on concordant problems (average accuracy = 71.4%).

The replication group shows a similar profile. Of 22 students, 2 made no errors, 1 made errors on concordant problems but not on discordant ones, and 19 students (86%) made many errors on discordant problems (average accuracy = 39.5%) and few errors on concordant ones (average accuracy = 80.9%). In both groups, the concordance effect was substantial and was contributed to by a majority of the students. The performance seems consistent with the hypothesis that verification accuracy should be relatively high for concordant problems and diminished for discordant problems. In contrast, the findings provide little indication of the picture encoding processes. Given the consistent preference for the left-most diagram in Experiment 1, it is not possible to support any picture encoding hypothesis that might make differential predictions concerning the encoding of sentences and pictures in the present paradigm.

EXPERIMENT 2

The preferences shown for Euler diagrams in Experiment 1 provide little support for the notion that students expect specific ratios for the sizes of the to-be-verified sets (over the range employed here). In addition, if the subject category is expected to be the largest, dominating set, then students should experience uniform difficulty in verifying universal-affirmative sentences against identity, exclusion, subset, and disjunction diagrams (in which the subject category was not greater than the predicate category). However, verification accuracy on these diagrams ranged from 54.5% to 100%.

An alternate conception to be considered here is that picture encoding is sensitive to the degree of shared membership (overlap) between the subject and predicate categories, that is, "distinguishability" (see introduction). Revlin and Leirer (1978) show that when such relations as "All A are B" are expressed propositionally, if at least 85% of the B category overlaps the A cate-

gory, it is sufficient for students to treat the relation as identity in a syllogistic reasoning task. To address this possibility in a picture encoding task, subjects in Experiment 2 verified Euler diagrams in which the degree of intersection was manipulated. The findings were compared with those of Experiment 1.

Method

Two groups of students ($n = 11$) were asked to determine whether a picture was true or false of a sentence, using the same procedures and relations ("is a" and "is included in") as in Experiment 1; only the nature of the diagrams was different. Our purpose was to keep the sizes of categories fixed, whenever possible, while varying the percentage of each category that contributes to the category intersection. This objective was accomplished with some difficulty: (1) Intersection area and set size are correlated in two of the five diagrams (subset and superset); (2) intersection area cannot be varied in two diagrams (identity and exclusion); (3) intersection and set size can be systematically varied only in the disjunctive diagram.

Within these constraints, set size was held constant, whenever possible, and intersection varied over three levels (25%, 50%, and 75%).³ If picture codes are sensitive to the degree of intersection, verification accuracy should reflect this on the disjunctive problems. For example, in the case of a universal-affirmative (All A are B) sentence and a disjunctive diagram, the greater the percentage of the B set contributing to the intersection with A, the more readily students will encode the diagram as "All B are A," thereby leading to an erroneous verification.

Subjects. The subjects were 44 introductory psychology students fulfilling a course requirement. Half of the students were randomly assigned to the two experimental groups ("is a" and "is included in"); the remaining students participated in a replication of the "is a" group.

Results

The processing assumptions will be assessed in two ways. First, we will consider concordance and complexity as in Experiment 1, using those sentence-picture pairs that are similar to those in Experiment 1 (all but the disjunctive ones). Second, the importance of intersection area for picture codes will be examined by comparing the verifications of the two experiments on the disjunctive sentence-picture pairs.

The verification scores are presented in Table 4, which shows that the present conditions closely replicated those of Experiment 1. Verification judgments were more accurate when the quantified relation was expressed as "is included in" rather than "is a" [$F(1,20) = 8.8, p < .01$]. This effect of relation was the result of a significant concordance effect [$F(1,20) = 12.0, p < .01$] that reduced the accuracy of verifications with "is a"

Table 4
Percent Verification Accuracy: Intersection Area Varied

Relation	Concordant †			Discor- dant †
	True	False	Average	
is a	82.9	77.3	80.1	45.5
is a*	74.6	81.8	78.2	54.5
is included in	84.1	93.2	88.7	84.1

*Replication. †With logic.

but had no appreciable effect on “is included in” [Concordance by Relation: $F(1,20) = 8.5, p < .01$]. The reliability of these findings was assessed by asking an additional 22 students to solve the identical problems as the “is a” group had solved. Those data are also presented in Table 4, which shows that the concordance effect was replicated: Concordant sentence-picture pairs were more accurately verified than were discordant ones [$T(42) = 2.69, p < .01$].

As in Experiment 1, this concordance effect was not the result of a confounding between truth value and conversion: There was no significant truth-value effect or interaction between truth value and quantifier.

The model correctly predicts the effect of processing complexity: Table 5 shows that the verifications that required Stage 3 processing were less accurate than those that could be reached without Stage 3 [“is a”: $T(20) = 2.34, p < .05$]. This was also shown for the replication groups [“is a”: $T(42) = 1.6, p = .05$]. As predicted, no effect of complexity was shown for similar verifications in the “is included in” group, since Stage 3 processing was not required of those verifications.

A comparison of the verification accuracy of the “is a” groups in Experiments 1 and 2 on the same sentence-picture pairs shows that there was no discernible difference between them; together, both groups showed an overall concordance effect [$F(1,20) = 30.9, p < .01$], with the effect slightly greater for the intersection group (Experiment 2) [Groups by Concordance: $F(1,20) = 6.3, p < .05$].

The findings of the two experiments are also comparable in terms of strategies used to make verifications. Of the 11 students in the “is a” group, only 1 was completely accurate; the remaining 10 students were more accurate on concordant than on discordant problems (accuracy = 87.9% and 40.0%, respectively). A similar pattern was shown for the replication group; of 22 students, 18 were more accurate on concordant problems than on discordant ones (accuracy = 70.8% and 50.0%, respectively).

This experiment explored the importance of intersection area for sentence-picture verifications by manipulating the overlap area in each of the disjunctive diagrams. In Experiment 1, the sizes of the sets varied, whereas the absolute intersection area was fixed. Figure 7 shows the percentage of erroneous verifications on which the subjects selected each disjunctive diagram

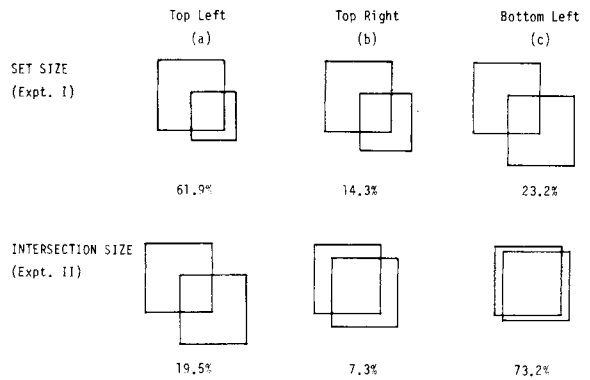


Figure 7. Percent erroneous verifications (universal-affirmative sentences) on which subjects preferred each disjunctive diagram.

as true of the universal-affirmative sentence (error rate = 44.4%). While the preference was clearly for the diagram consistent with the conversion model (i.e., in which the predicate category was maximally consumed in the intersection), we are unable to discern from these data alone whether the preferences for specific diagrams reflect (1) the left-right ordering of the diagrams (i.e., reading order), (2) the proportion of the total union of the two sets consumed in the intersection, or (3) the proportion of the predicate set in the intersection (the conversion assumption).

The disjunctive diagrams from Experiment 2 are also presented in Figure 7. They reflect the fact that the sizes of the sets are fixed, whereas the proportion of the two sets contributing to the intersection varies. The findings in Figure 7 inform us concerning two questions. First, preference judgments probably do not reflect reading order. Diagram preference was roughly invariant across position: Diagram c in Experiment 1 (the least optimal reading position) was accepted as often as its structurally equivalent form, Diagram a, in Experiment 2. The latter held what should be the optimal reading position (assuming that reading order did not vary with experiments). As a consequence, we are reasonably justified in treating the preference shown for the conversion-predicted Diagram a in Experiment 1 as reflecting subject’s encodings of the universal-affirmative sentence.

Second, the findings of Experiment 2 (Figure 7) show that when an error occurred (overall error rate = 68.2%), the most preferred diagram was the one that maximized both the total area of the union of the sets participating in the intersection (Diagram c, Experiment 2). This diagram is the one predicted to be most congruent with the converted encoding of the sentence.

Notice that a nonconverted reading might also produce such preferences, since the subject category is also maximally consumed in the intersection (i.e., Most A are B). The preferences in Experiment 1 for Diagram a (Figure 7), however, make this somewhat unlikely but still tenable. Systematic manipulations of specific alternative diagrams might permit a definitive evalua-

Table 5
Verificational Accuracy: Two Levels of Operational Complexity (Experiment 2)

Relation	Stage 3	No Stage 3
is a	69.3	83.7
is a*	62.5	75.0
is included in	84.1	89.0

*Replication.

tion of the locus of this intersection effect. However, it is not our intention here to exhaust the value of preference judgments, but rather to use such measures as indices of relevant stimulus characteristics to which picture codes might be sensitive. Standard latency measures to single sentence-picture pairs would provide a more substantial basis on which to evaluate such codes (see General Discussion).

GENERAL DISCUSSION

The present study examines a conception of sentence encoding that makes two strong claims: first, that a conversion operation participates in the encoding of quantified categorical relations (e.g., conversion model of Revlin & Leirer, 1978); second, that temporally less derived interpretations are available to the reasoner as a function of the task constraints. Using a verification paradigm, the two claims are entailed in the concordance and complexity predictions. The findings support the concordance predictions that decisions are most accurate when the hypothesized encoding operations are concordant with the logically prescribed judgments and least accurate when the hypothesized operations are discordant with the logical prescriptions. An alternative hypothesis is that the concordance effect reflects the accidental reversal of subject and predicate terms (e.g., Meyer, 1970; Newell, in press) rather than the deliberate operation of conversion at encoding. If this hypothesis were correct, concordance would be shown for relations expressed with the copula "is included in," as well as with "is a." However, the concordance effect is shown only for those relations in which conversion should be permitted ("is a") and not for those in which conversion should be blocked ("is included in"). This supports the model's contention that conversion deliberately participates in the encoding of quantified, categorical expressions.

The findings also support the complexity hypothesis that decision accuracy is negatively related to the number of operations that are required by the task constraints. In the present case, the major contributor to processing complexity is the reasoner's attempt to access less derived encodings (stack popping: repeated passes through Stage 3 of the verification model). It should be noted that the additional operations required to pop the reasoner's meaning stack do not confound with the concordance effect, since the discordant problems do not entail the popping operation.

While it is the purpose of the present study to focus solely on the distribution of verification errors, a direct prediction from the present formulation is that decision latencies as well as verification accuracy should reflect the processing assumptions made here. A partial test of these claims is provided by Just (1974), who measured decisions and decision latencies to categorical sentence-picture pairs, in which the pictures contained concrete collections of circles and squares (rather than the

present Euler diagrams). While only his universal-affirmative sets are pertinent to the present issues, the ordinal predictions hold.

A superficial summary of Just (1974) shows only a modest effect of conversion: When conversion in encoding accords with logic, the decisions are only slightly more accurate than when conversion and logic are discordant (Experiment 1, 95.4% and 92.9%, respectively; Experiment 2, 96.1% and 91.7%, respectively). However, this is far greater than it appears. The small concordance effect in Just may reflect increased processing complexity among the concordant sentence-picture pairs. Table 6 shows that error rates for verifications requiring Stage 3 are many times larger than for those verifications that do not require Stage 3 (by a factor of 11 in Experiment 1 and by a factor of 6 in Experiment 2). This effect of complexity is also observed in the decision latencies. For universal-affirmative sentences, verifications take 291 msec longer when Stage 3 functions than when it does not. In contrast, no such effect is found in the verifications of particular affirmatives, in which Stage 3 does not operate (the difference in error rates is .3%, and the latency difference is 23 msec.). In addition, the current framework correctly predicts that while concordant no-pop verifications should be far more accurate than the discordant (no-pop) ones, the decision latencies should be similar because the same processes apply. These ordinal predictions are supported by Just (1974), although it must be acknowledged that the magnitudes of the effects are smaller in Just's study than in the present study. This may be a consequence of the generally higher accuracy in Just, which reflects a greater emphasis on precision in such reaction time tasks, commensurate with an increased number of practice trials.

The importance of practice is amply shown in Experiment 3 of Just (1974). To control response probabilities, students received additional presentations of the subset diagram paired with the universal-affirmative sentence. This resulted in a reduction of both verification errors and decision latencies on these sentence-picture pairs, as well as an elimination of the processing complexity effect. However, the concordance effect was preserved: Concordant verifications were more accurate than discordant ones (3.1% error and 11.1% error, respectively).

Table 6
Percent Verifications Error (E) for Universal Affirmatives

	Concordant with Logic				Discordant with Logic	
	Stage 3		No Stage 3		No Stage 3	
	E	T	E	T	E	T
Experiment 1	18.0	2489	1.6	2120	7.1	2213
Experiment 2	8.8	1684	1.5	1474	8.3	1474

Note—T = decision times in milliseconds. Universal affirmatives are patterned after Just (1974).

Table 7
Percent Verification Errors and Set Size

	Concordant with Logic	Discordant with Logic
	Conversion Consistent	
Small Superset	5.7	13.6
Large Subset	4.0	25.9
Total	4.9	19.8
	Conversion Inconsistent	
Large Superset	2.4	5.8
Small Subset	4.6	4.2
Total	3.5	5.0

Note—Patterned after Meyer (1970).

Just's (1974) study sought to show that normal sentence-picture verification processes could account for the semantic retrieval data of Meyer (1970). Interestingly, some of Meyer's data also bear on the encoding notions discussed here. The most discrepant condition in that study occurred when students were asked to verify false universal-affirmative expressions, in which the sentence was of the form "All A are B" and the semantic relation had B as the included category and A as the including (or superset) category. Error rates on these verifications were twice that of any other conditions in the study. To account for this unpredicted finding, Meyer suggested the "category interchange" hypothesis, in which students confuse the subject and predicate terms prior to Stage I analysis (for both universal-affirmative and particular-affirmative sentences). Notice that this is quite similar to the notion of conversion, although Meyer did not specify any mechanism by which conversion (interchange) would occur.

In the superset condition, the conversion model predicts that increases in erroneous verifications will occur in those instances that maximize conversion. These will be cases in which both the subordinate and superordinate categories are of similar size (Revlis, 1975a). Meyer's (1970) paradigm permits an examination of this prediction, since it separately varies subordinate and superordinate set sizes. Table 7 presents the error rates from Meyer as a function of the relative set sizes and shows that the set-size effect, unique to the superset condition, corresponds closely to that predicted from the model.⁴ To wit, verification errors are more frequent when the sets are similar in size (either a large subordinate set or a small superordinate set) than when they are dissimilar in size (either a small subordinate set or a large superordinate set). While Meyer's model eloquently predicts these findings based on complex analyses of exemplars and attributes of the subject and predicate sets, it has been instructive to explore the generality of the present model for semantic as well as diagram verifications.

Developmental Sequence

Neimark and Chapman (1975) also employed a verification task using quantified sentences and Euler dia-

grams (all diagrams were present on each trial). The purpose of the study was to examine verification errors developmentally. The findings of that study can be interpreted in terms of the conversion model. With respect to universal-affirmative propositions, errors of the youngest children (7th-graders) include selecting the alternatives that are correct for the reversed form of the proposition, so that these children treat "All A are B" as if it were "All B are A." In addition, when children select too many alternatives, they tend to accept two inclusion diagrams: This is equivalent to accepting both the original inclusion relation and its converse.

One consequence of the approach taken here is that the development of categorical inference and categorization rules may reflect general linguistic development (that permits the comprehension of quantified, categorical relations) as much as it may entail a specific developmental sequence in the construction of logical inference and categorization rules. Such possibilities suggest that research with independent task environments, such as the present study, may provide important convergent validation of our theories of formal inference and human rational competence.

REFERENCE NOTES

1. Erickson, J., Wells, G. L., & Traub, B. H. *Tests of a model of formal syllogistic reasoning*. Paper presented at the annual meeting of the Psychonomic Society, Boston, 1974.
2. Thomas, J. *Quantifiers and question-asking* (Research Rep. RC 5866). Yorktown Heights, N.Y.: IBM Thomas J. Watson Research Center, 1976.
3. Revlin, R., & Tomlinson-Brown, S. *Euler's representation for quantified, categorical expressions* (Tech. Rep. 79-4). Santa Barbara: University of California, 1979.

REFERENCES

- BUCCI, W. The interpretation of universal affirmative propositions. *Cognition*, 1978, 6, 55-77.
- CARPENTER, P. A., & JUST, M. A. Sentence comprehension: A psycholinguistic processing model of verification. *Psychological Review*, 1975, 82, 45-73.
- CHAPMAN, L. J., & CHAPMAN, J. P. Atmosphere effect re-examined. *Journal of Experimental Psychology*, 1959, 58, 220-226.
- CHASE, W., & CLARK, H. H. Mental operations in the comparison of sentences and pictures. In L. Gregg (Ed.), *Cognition in learning and memory*. New York: Wiley, 1972.
- DESOTO, C., & KUETHE, J. Perception of mathematical properties of interpersonal relations. *Perceptual and Motor Skills*, 1958, 8, 279-286.
- DESOTO, C., & KUETHE, J. Subjective probabilities of interpersonal relationships. *Journal of Abnormal and Social Psychology*, 1959, 59, 290-294.
- ERICKSON, J. A set analysis theory of behavior in formal syllogistic reasoning tasks. In R. L. Solso (Ed.), *Theories of cognitive psychology: The Loyola symposium*. Hillsdale, N.J.: Erlbaum, 1974.
- GLASS, A., HOLYOAK, K., & SANTA, J. *Cognition*. Menlo Park: Addison-Wesley, 1979.
- GRICE, H. O. *The logic of conversation: William James lectures*. Cambridge, Mass: Harvard University Press, 1967.
- GRIGGS, R. Drawing inferences from set inclusion information

- given in text. In R. Revlin & R. Mayer (Eds.), *Human reasoning*. Washington, D.C: Winston, 1978.
- INHEDER, B., & PIAGET, J. *The early growth of logic in the child: Classification and seriation*. New York: Norton, 1964.
- JOHNSON-LAIRD, P. N. The choice of the passive voice in a communicative task. *British Journal of Psychology*, 1968, **59**, 7-15. (a)
- JOHNSON-LAIRD, P. N. Interpretation of the passive voice. *Quarterly Journal of Experimental Psychology*, 1968, **20**, 69-73. (b)
- JOHNSON-LAIRD, P. N. On understanding logically complex sentences. *Quarterly Journal of Experimental Psychology*, 1969, **21**, 1-13. (a)
- JOHNSON-LAIRD, P. N. Reasoning with ambiguous sentences. *British Journal of Psychology*, 1969, **60**, 17-23. (b)
- JOHNSON-LAIRD, P. N. The interpretation of quantified sentences. In F. D'Arcais & W. J. M. Levelt (Eds.), *Advances in psycholinguistics*. New York: American Elsevier, 1970.
- JOHNSON-LAIRD, P. N. Models of deduction. In R. Falmagne (Ed.), *Reasoning: Representation and process in children and adults*. Hillsdale, N.J: Erlbaum, 1975.
- JOHNSON-LAIRD, P. N., & STEEDMAN, M. The psychology of syllogism. *Cognitive Psychology*, 1978, **10**, 64-99.
- JUST, M. A. Comprehending quantified sentences: The relation between sentence-picture and semantic memory verification. *Cognitive Psychology*, 1974, **6**, 216-236.
- JUST, M. A., & CARPENTER, P. A. Comprehension of negation with quantification. *Journal of Verbal Learning and Verbal Behavior*, 1971, **10**, 244-253.
- MEYER, D. On the representation and retrieval of stored semantic information. *Cognitive Psychology*, 1970, **1**, 242-299.
- NEIMARK, E., & CHAPMAN, R. Development of the comprehension of logical quantifiers. In R. Falmagne (Ed.), *Reasoning: Representation and process in children and adults*. Hillsdale, N.J: Erlbaum, 1975.
- NEWELL, A. Reasoning, problem solving and decision processes: The problem space as a fundamental category. In R. Nickerson (Ed.), *Attention and performance VIII*. Hillsdale, N.J: Erlbaum, in press.
- POTTS, G. The role of inference in memory for real and artificial information. In R. Revlin & R. Mayer (Eds.), *Human reasoning*. Washington, D.C: Winston, 1978.
- REVLIN, R., AMMERMAN, K., PETERSEN, K., & LEIRER, V. Category relations and syllogistic reasoning. *Journal of Educational Psychology*, 1978, **70**, 613-625.
- REVLIN, R., & LEIRER, V. The effects of personal biases on syllogistic reasoning: Rational decisions from personalized representation. In R. Revlin & R. Mayer (Eds.), *Human reasoning*. Washington, D.C: Winston, 1978.
- REVLIS, R. Formal syllogistic reasoning: Logical decisions from a complex data base. In R. Falmagne (Ed.), *Reasoning: Representation and process in children and adults*. Hillsdale, N.J: Erlbaum, 1975. (a)
- REVLIS, R. Two models of syllogistic reasoning: Feature selection and conversion. *Journal of Verbal Learning and Verbal Behavior*, 1975, **14**, 180-195. (b)
- TRABASSO, T., ROLLINS, H., & SHAUGHNESSY, E. Storage and verification stages in processing concepts. *Cognitive Psychology*, 1971, **2**, 239-289.
- TSAL, Y. Symmetry and transitivity assumptions about nonspecified logical relations. *Quarterly Journal of Experimental Psychology*, 1977, **29**, 677-684.
- WILKINS, M. C. The effect of changed material on ability to do formal syllogistic reasoning. *Archives of Psychology*, 1928, **102**, 83.
- WOODWORTH, R. S., & SELLS, S. B. An atmosphere effect in formal syllogistic reasoning. *Journal of Experimental Psychology*, 1935, **18**, 451-460.

NOTES

1. We deliberately use Euler diagrams in contrast to Venn diagrams, since the latter presuppose the existence of a third, unspecified category denoted by the set of elements not included in the union of the subject and predicate classes of the categorical relation. The presence of additional sets has implications for the convertibility of some sentences (see Revlin & Tomlinson-Brown, Note 3, and the "availability strategy" of Glass, Holyoak, & Santa, 1979).

2. Note that since the picture code depends upon the sentence subject, predicate, and polarity (affirmative or negative), the initial code for the subset diagram will be expressed by the affirmative relation "Some B are A" when the sentence is affirmative (e.g., "All A are B" and "Some A are B") and by the negative relation, "Some B are not A" otherwise.

3. Greater than 75% overlap was not possible, since preliminary data showed that at greater overlap, students could not physically distinguish between the line drawings that represented the sets.

4. Predictions concerning other conditions cannot be assessed, since Meyer (1970) varied the frequency of sentence-relation pairs to hold response probability constant.

(Received for publication February 26, 1980;
revision accepted April 9, 1980.)