Notre Dame Journal of Formal Logic
Volume 35, Number 3, Summer 1994

The Irrelevance of Distribution for the Syllogism

WALLACE A. MURPHREE

Abstract While accepting that distribution is a coherent notion, I argue that it is nevertheless irrelevant to the working of the syllogism. Instead, I propose: (i) that a term's being distributed or undistributed in a proposition is its capacity to be replaced in a truth-preserving substitution with a narrower or a wider term; (ii) that which capacity the term has is determined by whether it occurs as the predicate of a negative or of an affirmative statement of the proposition; and (iii) that it is only the term's occurrence as the predicate of a negative or an affirmative statement—rather than its distribution value—that is relevant to syllogistic entailment.

1 Introduction Peter Geach's criticisms of distribution (see his [3], [4], and [5]) prompted its defenders to clarify the concept (see Englebretsen [2], Sommers [10], and Toms [11]); then, since the resulting notion was plainly coherent, its relevance to the syllogism seems simply to have been assumed.

However, I propose that although the revised concept is coherent and is, indeed, important in some areas in logic, it is only indirectly related to the workings of the syllogism. Accordingly, I propose that the practice of framing rules of syllogistic validity in terms of it is specious and misleading.

The paper proceeds as follows: first, I offer my version of the revised concept of distribution and undistribution, which is suggested by the formulation advanced by Stephen Barker ([1], p. 43), and then I argue that this concept is grounded in the quality, rather than the quantity, of the statement; next I propose a sketch of "the actual workings" of valid syllogisms, and then I consider how conformity to the distribution rules serves to effect such "workings." I conclude that distribution is not directly related to this situation; instead, I propose that when it is correlated to certain features of it, it is because the ground of distribution, rather than distribution itself, is actually related to it.

Received July 15, 1992; revised August 16, 1993

2 The Nature of Distribution The revised theory (as I offer it) accounts for distribution by appealing to the "scope" of terms, in the sense that relative to "rectangle," "parallelogram" has a wider scope, and hence is a "superterm," while "square" has a narrower scope, and hence is a "subterm." Now, a term is distributed in a proposition if and only if: (i) that proposition logically implies every other proposition formed by replacing that term with one of its subterms; *and* (ii) that implication holds *by virtue of* the term's being replaced with one of it subterms; and a term is undistributed in a proposition if and only if: (i) that proposition logically implies every other proposition formed by replacing that term with one of its superterms; *and* (ii) that implication holds *by virtue of* the term's being replaced with one of its superterms. Accordingly, letting "X>Y" indicate that X is a subterm of Y (and Y is a superterm of X), we can suppose progressive subterms of G (superterms of A), and subterms of S (superterms of M) to be as follows:

Subterms of G (Superterms of A)	Subterms of S (Superterms of M)
A>B>C>D>E>F>G	M>N>O>P>Q>R>S

Now, using D and P as the terms of the propositions, we can see the distribution patterns as follows:

D A P = (A>B>C>D) A (P>Q>R>S),
D E P = (A>B>C>D) E (M>N>O>P),
D I P = (D>E>F>G) I (P>Q>R>S), and
D O P = (D>E>F>G) O (M>N>O>P).

That is, given that all D are P, then it follows that all A, B, and C, as well as D, are P since D is distributed and A, B, and C are subterms of D; and not only does it follow that they are all P, but it also follows that they are all Q, R, and S as well, since P is undistributed and Q, R, and S are superterms of P. And the same holds, *mutatis mutandis*, for the other three propositions.

A term's distribution value is a necessary feature of it; this can be seen more clearly, perhaps, when the revised theory is put into the notation of the first order predicate logic, as follows: a term is distributed in a proposition if and only if: (i) that proposition logically implies every other proposition formed by *conjoining* some other term to it; *and* (ii) that implication holds *by virtue of* the term's being replaced with a conjunction; and a term is undistributed in a proposition if and only if: (i) that proposition logically implies every other proposition formed by *disjoining* some other term to it *and*; (ii) that implication holds *by virtue of* the term's being replaced with a disjunction. Hence for predicates D, P, H, and J:

(A) $(x)(\mathrm{D}x \supset \mathrm{P}x)$ logically implies $(x)((\mathrm{D}x \cdot \mathrm{H}x) \supset (\mathrm{P}x \vee \mathrm{J}x))$,
(E) $(x)(\mathrm{D}x \supset\sim \mathrm{P}x)$ logically implies $(x)((\mathrm{D}x \cdot \mathrm{H}x) \supset\sim (\mathrm{P}x \cdot \mathrm{J}x))$,
(I) $(\exists x)(\mathrm{D}x \cdot \mathrm{P}x)$ logically implies $(\exists x)((\mathrm{D}x \vee \mathrm{H}x) \cdot (\mathrm{P}x \vee \mathrm{J}x))$, and
(O) $(\exists x)(\mathrm{D}x \cdot \sim \mathrm{P}x)$ logically implies $(\exists x)((\mathrm{D}x \vee \mathrm{H}x) \cdot \sim (\mathrm{P}x \cdot \mathrm{J}x))$.

But these two approaches to the question of distribution are the same, since the conjunction of a new term to an original one serves to pick out a subterm of that original term, as the conjunction of "equilateral" to "rectangle" yields "square," which is a subterm of "rectangle"; and, likewise, the disjunction of a new term to an original one results in a superterm of that original term, as the disjunction of "rhomboid" to "rectangle" yields "parallelogram," which is a superterm of "rectangle."

However, two caveats are needed with this latter formulation. The first is that if the original term is itself a conjunction, then a superterm is formed by dropping one of the conjuncts, as "rectangle" is a superterm of "equilateral rectangle"; and if the original term is itself a disjunction, then a subterm is formed by dropping one of its disjuncts, as "rectangle" is a subterm of "rhomboid or rectangle." But this does not constitute an exception to the above formulation since the dropping of the conjunct, C, from $(\mathrm{C} \cdot \mathrm{D})$ is logically equivalent to the disjunction of $(\sim \mathrm{C} \cdot \mathrm{D})$ to it; and the dropping of a disjunct, D, from $(\mathrm{C} \vee \mathrm{D})$ is logically equivalent to the conjunction of $(\mathrm{C} \vee \sim \mathrm{D})$ to it. Hence, the superterm, "rectangle," can be formed either by dropping "equilateral" from the subterm, "equilateral rectangle," or by disjoining "nonequilateral rectangle" to it, since "equilateral rectangle or nonequilateral rectangle" is equivalent to "rectangle." And the same is the case, *mutatis mutandis*, in the formation of subterms from complex superterms.

The other caveat is that disjunction and conjunction fail to form superterms and subterms when the adjoined terms are such "maybe-not/not-quite" modifiers such as "alleged," "so-called," or "imperfect." For example, although "equilateral rectangle" is a subterm of "rectangle," "imperfect rectangle" is not, since an imperfect rectangle is no rectangle at all (see Katz and Martinich [7]).

The second clause in the definitions above require that the logical implication hold by virtue of subterm and superterm replacement, rather than by virtue of anything else. Accordingly, the predicate term in each top sentence of the three sets below is not distributed even though each logically implies the bottom sentence in its set, and each bottom sentence is formed by replacing the predicate term of the top (blue), with its subterm (blue squares).

a. Some noncircular circles are blue = $(\exists x)((\sim \mathrm{C}x \cdot \mathrm{C}x) \cdot \mathrm{B}x)$
b. Some noncircular circles are blue squares = $(\exists x)((\sim \mathrm{C}x \cdot \mathrm{C}x) \cdot (\mathrm{B}x \cdot \mathrm{S}x))$

a. All blue squares are blue = $(x)((\mathrm{B}x \cdot \mathrm{S}x) \supset \mathrm{B}x)$
b. All blue squares are blue squares = $(x)((\mathrm{B}x \cdot \mathrm{S}x) \supset (\mathrm{B}x \cdot \mathrm{S}x))$

a. All squares are blue = $(x)(\mathrm{S}x \supset \mathrm{B}x)$
b. All squares are blue squares = $(x)(\mathrm{S}x \supset (\mathrm{B}x \cdot \mathrm{S}x))$

Again, these predicates terms are not distributed since the logical implication does not hold for any of them by virtue of the subterm replacement. Rather, in the first set the implication holds by virtue of the fact that the top proposition is necessarily false and, as such, it logically implies every proposition; in the second set the implication holds by virtue of the fact that the bottom proposition is necessarily true and, as such, it is logically implied by every proposition; and in the final set the implication holds by virtue of the fact that the two sentences are logically equivalent to each other.

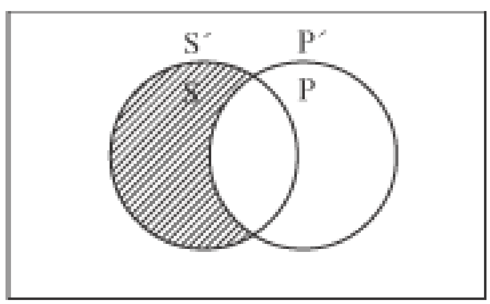
And, in general, when an implication holds by reason other than that of superterm or subterm replacement, the implication is irrelevant to the distribution value of the term.

Although a term may neither be distributed nor undistributed (as is “squares” in “Most squares are blue”), a term cannot both be distributed and undistributed at the same time. So the second clause of the definition is necessary to prevent “blue,” which is clearly undistributed in the cases above, from being ruled as both distributed and undistributed, as it would be if the determination were made by the appeal to the first clause alone.

3 The Grounds of Distribution According to this view, distribution (undistribution) is not a more basic property that somehow *makes* a term replaceable by its subterms (superterms) without risking the loss of the truth of the proposition in which it occurs; rather the distribution value of a term *is* its replaceability capacity.

So given that distribution and undistribution do not give a term its replaceability capacity, the question arises as to what confers this capacity on it. What are the grounds of distribution and undistribution?

I propose generally that distribution value is grounded in quality, rather than in quantity; however, some additional considerations are required for the specific formulation of this thesis. As a point of departure, a “proposition” will be distinguished from its various “expressions,” or “statements.” A “proposition,” I shall say, is what is depicted on a Venn diagram (having four logical spaces) when an area is shaded or an asterisk is put in an area, while an “expression” is a way of reading what it is that is on the diagram. Hence, S A P, S E P′, P′E S, and P′A S′ are four different (but equivalent) “expressions” of the one and self-same “proposition” that is displayed on the Venn diagram as:



S A P
S E P′
P′ E S
P′ A S′

That is, the Venn diagram depicts one proposition, which can be “asserted” by “stating” or (“expressing”) one of its “statements” (“expressions”). (Incidentally, the diagram cannot depict any one of the statements without depicting them all.)

Accordingly, relative to any proposition there are four terms, viz., S, S′, P, and P′, and the distribution value of each term remains constant throughout the various statements of the proposition. In fact, a term is always (d)istributed in a proposition just in case its complement is (u)ndistributed (see Williamson [12], p. 733), as is shown below:

Equivalent Statements of the Four Basic Propositions							*Distribution Values*			
							S	S′	P	P′
1. S A P	=	S E P′	=	P′ E S	=	P′ A S′	d	u	u	d
2. S E P	=	S A P′	=	P A S′	=	P E S	d	u	d	u
3. S I P	=	S O P′	=	P O S′	=	P I S	u	d	u	d
4. S O P	=	S I P′	=	P′I S	=	P′O S′	u	d	d	u

Now the conventional introduction to the doctrine of distribution proceeds by noting that universal statements distribute their subject terms while negative statements distribute their predicates. (Alternatively it might have been worded that particular statements undistribute their subject terms while affirmative statements undistribute their predicates.) At any rate, the conventional concept is that a term is distributed in a proposition when and only when reference is made to its full denotation (see Englebretsen [2] and Sommers [10]). Then the explanation proposes that the subject of a universal makes such reference because it is universally quantified; furthermore, the account of the distribution of negative predicates continues by alleging that such terms are somehow implicitly also used in a universal sense. And, certainly the subjects of universals are distributed since, for example,

> All rectangles are blue, and
> No rectangles are blue,

logically imply

> All squares are blue, and
> No squares are blue,

respectively. But, clearly, it cannot be that they are distributed because they are universally quantified since the very same distribution obtains for the subject terms in the exceptive propositions below where the quantification is not universal. That is,

> At least all but six rectangles are blue, and
> At most none but six rectangles are blue,

logically imply

> At least all but six squares are blue, and
> At most none but six squares are blue,

respectively, although in neither case is anything said about all rectangles. (See Murphree [9] for a systematic treatment of such exceptive proposition.)

The alternative that presents itself—and the one here proposed—is that it is because of its occurrence as the predicate of a negative statement that a term of a categorical proposition is distributed (and because of its occurrence as the predicate of an affirmative statement that a term is undistributed). As such, it is because S is the predicate of P′ E S, rather than because it is the subject of S A P, that it is distributed relative to the A proposition; and the same holds, it is proposed, for each distributed and undistributed term:

Distributed Predicate of Negative Statements			*Undistributed Predicate of Affirmative Statements*		
S A P	=	S E **P**′	S A P	=	S A **P**
	=	P′ E **S**		=	P′ A **S**′
S E P	=	S E **P**	S E P	=	S A **P**′
	=	P E **S**		=	P A **S**′
S I P	=	S O **P**′	S I P	=	S I **P**
	=	P O **S**′		=	P I **S**
S O P	=	S O **P**′	S O P	=	S I **P**′
	=	P′ O **S**′		=	P′ I **S**

Furthermore, this proposal holds as well for exceptive propositions since, for example,

(affirmative) At least all but six rectangles are blue, and
(negative) At most none but six rectangles are blue,

logically imply, respectively,

At least all but six rectangles are colored (superterm), and
At most none but six rectangles are light blue (subterm).

Moreover, the subject of these propositions, which was shown above to be distributed, only occurs in the predicate position of equivalent statements that are negative, as the immediate inferences below display.

1.	(affirmative)	At least all but six rectangles are blue (original)
2.	(negative)	At most none but six rectangles are nonblue (from 1. by obversion)
3.	**(negative)**	At most none but six nonblue things are **rectangles** (from 2. by conversion)
4.	(affirmative)	At least all but six nonblue things are nonrectangles (from 3. by obversion, or from 1. by contraposition)

1.	(negative)	At most none but six rectangles are blue (original)
2.	**(negative)**	At most none but six blue things are **rectangles** (from 1. by conversion)
3.	(affirmative)	At least all but six rectangles are nonblue (from 1. by obversion)
4.	(affirmative)	At least all but six blue things are nonrectangles (from 2. by obversion, or from 3. by contraposition)

And this, I submit, supports the contention that the ground of distribution is the quality of the statement in which the term occurs as predicate, rather than the quantity of the proposition in which it occurs as the subject.

Another case that supports the contention is provided by proportional quantifiers, such as “most,” “three-eighths of all,” or “75% of all.” In such propositions the subject terms have no distribution value whatever, although their predicate terms do. That is, neither

Most rectangles are blue

nor

Most rectangles are not blue

implies any proposition formed by replacing "rectangles" with "parallelograms" or with "squares," as each would if it had a distribution value. However, the predicate is undistributed in the former and distributed in the latter since they imply, respectively,

Most rectangles are colored (superterm)

and

Most rectangles are not light blue (subterm).

But now, if distribution were a matter of quantification it would seem that these subject terms would have a distribution value (of undistributed); however, if distribution be a matter of the quality of the statement in which the term occurs as predicate, as I propose, then there is a good explanation as to why they have none: it is because there is no equivalent statement of either original claim in which "rectangles" occurs as the predicate. That is,

Most S's are P (nonP)

cannot be validly converted to

Most P's (nonP's) are S.

And, the same holds for the subject terms of all other propositions having proportional quantifiers.

Now, granting that distribution is grounded in quality, the further question remains as to how this is so. Why is it that being the predicate of an affirmative statement makes a term replaceable by one of its superterms while being the predicate of a negative statement makes a term replaceable by one of its subterms?

In response to this question it is instructive to note that a term of a single-term statement is undistributed whenever the statement is affirmative, and is distributed whenever it is negative. That is, S is undistributed in Sa, $(\exists x)Sx$, and $(x)Sx$, since

Sa implies	$(Sa \vee Ta)$,	but it does not imply	$(Sa \cdot Ta)$,
$(\exists x)Sx$ implies	$(\exists x)(Sx \vee Tx)$,	but it does not imply	$(\exists x)(Sx \cdot Tx)$, and
$(x)Sx$ implies	$(x)(Sx \vee Tx)$,	but it does not imply	$(x)(Sx \cdot Tx)$;

and S is distributed in $\sim Sa$, $\sim (\exists x)Sx$, and $\sim (x)Sx$, since

$\sim Sa$ implies	$\sim(Sa \cdot Ta)$,	but it does not imply	$\sim(Sa \vee Ta)$,
$\sim (\exists x)Sx$ implies	$\sim (\exists x)(Sx \cdot Tx)$,	but it does not imply	$\sim (\exists x)(Sx \vee Tx)$, and
$\sim (x)Sx$ implies	$\sim (x)(Sx \cdot Tx)$,	but it does not imply	$\sim (x)(Sx \vee Tx)$.

Now, being the predicate term, T, of a categorical statement and being a term, T, in a single-term statement are similar in that both types of claims either affirm or deny that something is T. The categorical statements affirm or deny that at least one thing, or that all things, that are S are T, while the single-term statements affirm or deny that at least one thing (identified or unidentified), or that all things, are T.

Accordingly, I propose that the general principle of undistribution and distribution is:

> For any term, T, and superterm T*, whatever is affirmed to be T is thereby also affirmed to be T*; and for any term T, and subterm T**, whatever is denied to be T is thereby also denied to be T**.

I claim that this principle is intuitively obvious; furthermore, I claim that it, unlike the conventional approach, accounts for why the distribution values of the terms in propositions having exceptive and proportional quantifiers are what they are.

4 The Workings of the Syllogism A review of the nature of syllogistic entailment will be useful in preparation for the question of the relevance of distribution to the syllogism. It is well-known that any valid syllogism can be "reduced" either to AAA or AII, figure one. But this may be put more clearly by noting simply that every valid syllogism "has" either an AAA or AII, figure one, version. That is, since each proposition has four different, but equivalent, "statements" (or "expressions"), each argument composed of three propositions can be "stated" (or "expressed") by sixty-four different combinations of statements. That is to say, each argument has sixty-four different "versions." Of these sixty-four, some are "standard-form versions," in that they contain only three terms, while others contain four, five, or six; and of the standard form versions, at least one is either AAA or AII, figure one. (Since the A statements have equivalent contrapositives, there are always two figure one versions of AAA, while there is only one of AII.)

So, in the AAA and AII, figure one, version, the syllogism can be described as follows: the minor term makes a "contribution" (of "all" or "some") of S to the "accepting" middle term in the minor premise; then in the major premise the middle term "transmits" that contribution to the final "receiving term," the major term. Finally, for the conclusion, the occurrences of the middle term are eliminated and the "contribution" is made from the minor to the major term directly. So, in all there are the "contributing term," the "accepting middle term," the "transmitting middle term," and the "receiving term" that are involved in this account of syllogistic entailment.

Now, this affirmative, figure-one way of describing a valid syllogism appears preferable to descriptions based on other versions since it "tracks the inheritance" of the quantity from the subject term, through the occurrences of the middle term, to the major term more straightforwardly than do they: if all (or some) S's are M, and all M's are P, then all (or some) S's must be P. And this way of describing them holds for all valid syllogisms even though various versions do not include the mention of the operative terms in the expressions of the propositions (since they mention the complementary terms instead). Still, a term is a contributing term, an accepting middle, a transmitting middle, or a receiving term *relative to the argument*, irrespective of what version the argument is given in. That is, the contributing term is the term which is the minor, the receiving term is the term which is the major, etc., in the AII or AAA, figure one, versions, of the argument. Arguments with universal conclusions, accordingly, have two contributing, receiving, etc., terms since they have two AAA, figure one, versions; that is, S contributes to P through the occurrences of M in

M	A	P
S	A	M
S	A	P

while P′ contributes to S′ though the occurrences of M′ in

$$\begin{array}{ccc} M' & A & S' \\ P' & A & M' \\ \hline P' & A & S'. \end{array}$$

Now an argument may be invalid either because the premises yield no conclusion whatever, or because the conclusion drawn is different from the conclusion entailed. But the premises do yield a conclusion when and only when one of them has an accepting middle and the other a transmitting middle. As the "reduced" versions illustrate, the accepting middle must be the predicate of an affirmative expression, (because if it were a subject it would contribute, rather than receive, and if it were the predicate of a negative expression then no quantity would be contributed to it); and the transmitting middle must be the subject of a universal (because if it were the predicate it would receive, rather than transmit, and if it were the subject of a particular it would not necessarily transmit the quantity contributed to it). But, if either M or M′ is both the predicate of an affirmative expression of one premise and also the subject of the other, when the other is universal, then some conclusion or other is entailed.

Again, it may be the case that the premises do yield a conclusion but the conclusion drawn is not the one that is entailed. In the conclusion entailed, that quantity, which in the premises is transmitted from the contributing term to the receiving term by way of the middle terms, is transmitted from that contributing term to that receiving term directly. That is, the middle terms are merely eliminated and the quantity and extreme terms remain the same. Accordingly, the conclusion may err by being of the wrong quantity, by indicating the wrong term as the contributor, the wrong term as the receiver, or by a combination of these. But these are the only ways that an erroneous conclusion can be drawn.

5 The Use of Distribution Conventional sets of rules specify one distribution requirement for the middle terms (in the premises), and another distribution requirement for the extreme terms (in the conclusion). When the "existential perspective" is allowed, as in the set of rules advanced by Barker ([1], pp. 69–70), the distribution requirements are:

> The middle term must be distributed *at least* once, and A term may be distributed in the conclusion *only if* it is distributed in its premise.

However, from the "hypothetical perspective," as in the set advanced by Sommers (see his [10], p. 34) these may be framed more restrictively as:

> The middle term must be distributed *exactly* once, and A term is to be distributed in the conclusion *if and only if* it is distributed in its premise.

Incidentally Sommers treats inferences by existential presupposition as enthymemic sorites, with "Some T's are T" being the suppressed premise when membership is presupposed for term T. (See also Kelley [8], Chapter 14.)

So the question now concerns how these rules function to distinguish valid from invalid syllogisms. That is, how does conformity to these rules serve to insure valid syllogistic inferences? The answer, in each case, must be advanced from the context of the additional rules of the set under consideration.

5.1 Rule for the Premises Barker's rule of distribution is complemented by a rule of quality, viz., that at least one premise must be affirmative. However, these rules, framed to accommodate the existential perspective, have the undesirable feature of failing to distinguish the sets of conclusion-yielding premises from those that yield no conclusion at all. On the one hand there are premises that do yield conclusions (although not standard-form conclusions because they involve a "fourth term,") that are not admitted because of the rule of quality. For example, M O S and M E P are not admitted, although they entail S′O P (see Johnstone [6]). And on the other hand there are premises, such as S O M and M I P, that yield no conclusion whatever, but which are ruled by Barker's criteria as being admissible premises—that have no admissible conclusions!

Sommers' rules, however, function perfectly in this respect. His rule to complement the distribution requirement is one of quantity, viz., that at least one premise must be universal; and the two requirements together distinguish the conclusion-yielding sets of premises from those that yield none. That is, the two requirements together pick out precisely those sets of premises having statements in which the middle term occurs once as the predicate of an affirmative (the accepting middle) and once as the subject of a universal (the transmitting middle).

5.2 Rule for the Conclusion Barker's rule of distribution for the conclusion is complemented by a rule of quality, viz., that the conclusion is to be negative if and only if a premise is negative. Also, the set contains an optional rule of quantity (for the "hypothetical perspective"), prohibiting the conclusion of a particular from two universal premises.

Now, the application of this distribution rule to the major term prevents an E or O conclusion when it is an A or I that is entailed; however, given the negation rule, this application would seem to be superfluous because such fallacious conclusions should "already" be prohibited by that rule of quality. Indeed, the rule of quality should also prohibit the A or I conclusion when it is an E or O that is entailed. Hence it would seem that the applicability of the distribution rule would be limited to the minor term, to insure the conclusion is of the proper quantity.

Sommers' rule of distribution for the conclusion, on the other hand, is complemented by a rule of quantity, viz., that the conclusion is to be particular if and only if a premise is particular. Hence, here it would seem that the distribution rule would not be needed to prevent illicit process of the minor, because such fallacious conclusions should "already" be prohibited by the rule of quantity.

So, it would seem that if Barker's rule of quality were conjoined with Sommers' rule of quantity there should be no need for the rule of distribution for the conclusion; but this is not the case. For example, A O O, figure three, along with ten other syllogisms, conforms to these rules but yet is invalid. (The other ten are: O E O, figures 1 and 3; I E O, figures 1 – 4, O A O, figure 2; E O O figures 3 and 4; and A A A figure 4.) That is, A O O, figure three,

M	A	P
M	O	S
S	O	P,

yields a conclusion since the middle term is distributed exactly once and (at least) one premise is universal. Furthermore, the conclusion conforms both to Barker's rule of quality and also to Sommers' rule of quantity, but yet both terms of the conclusion stand in violation of Sommers' rule of distribution. (Moreover, the attempt to "correct" the distribution faults by concluding S A P, instead of S O P, is "already" prohibited by the rules of quality and quantity.) So, it is in these cases that the distribution rule for the conclusion is indispensable from the context of the traditional rules.

The situation here is that these premises do yield a conclusion, but each possible standard-form conclusion violates some rule or other. Accordingly, the conclusion that is yielded has to be one that is not in standard form, viz., S′ O P′. And with this conclusion, the rule of distribution, as well as those of quality and quantity are satisfied. That is, since a term is distributed if and only if its complement is undistributed, it follows that S′ is undistributed relative to M O S and that P′ is distributed relative to M A P; hence the conclusion of S′ O P′ conforms to the rule requiring the terms of the conclusion to retain their distribution values from the premises.

Above I proposed that the conclusion may err by being of the wrong quantity, by indicating the wrong term as the contributor, or the wrong term as the receiver. Now, in response to the question as to what it is that the distribution rule for the conclusion does, the answer is that it prevents the wrong term, i.e., the wrong complement, from being selected both for the contributing and for the receiving term in the conclusion. Or, if nonstandard-form conclusions are allowed, it serves to identify the contributing and receiving terms (complements) for the conclusion. That is, in the example above, S′ I M is the operative statement of the minor premise (M O S); accordingly, S′ is the contributing term, while M (of S′ I M) is the accepting occurrence of the middle; then M (of M A P) is the transmitting occurrence of the middle, and P is the receiving term. Then, with the middle terms eliminated, the conclusion is S′I P (or S′ O P′). Incidentally, the rule of quality becomes superfluous when in the conclusion the proper quantity is assigned from the proper contributing term to the proper receiving term.

6 Heuristic, Specious, and Essential Criteria Again, I propose that (i) a syllogism is valid if and only if it has a version in which the middle term occurs both as the predicate of an affirmative statement and as the subject of a universal, while the contributing and receiving terms of the premises retain their respective roles as such in the conclusion, but Barker and Sommers contend that (ii) a syllogism is valid if and only if it conforms to certain conditions of distribution (and quality or quantity). Of course, both approaches pick out the same syllogisms and, furthermore, completely different strategies could be devised that would accomplish the same discrimination. For example, all the syllogisms might be arranged in alphanumerical order—from A A A-1 through O O O-4—and numbered from 1 through 256, and perhaps then a mathematical formula, F, could be found that would identify just the valid ones. If so it could then be said that (iii) a syllogism is valid if and only if formula F picks it out of the alphanumerical lineup. Now, I propose that these three criteria are "essential," "specious," and "merely heuristic," respectively.

Firstly, a criterion is merely heuristic if the identifying feature to which it appeals is contingent, so that its applicability is accidental. Hence, for example, in a situation in which all and only those triangles that are equilateral are colored blue, the criterion,

> A triangle is equilateral if and only if it is blue,

is merely heuristic, because the unique color is accidental to the unique shape. Likewise, the criterion,

> A syllogism is valid if and only if formula F picks it out of the alphanumerical lineup,

would be merely heuristic since the names that establish the alphanumerical sequence that allow F to apply are accidental to the syllogism.

Secondly, a criterion is specious if the identifying feature to which it appeals is necessary, but nevertheless is a feature that is not directly related to the discrimination that its application makes. So, for example, I suggest that the criterion,

> A triangle is equilateral if and only if it has equal angles,

is specious. That is, the criterion appeals to a necessary feature of equilateral triangles since such triangles have equal angles necessarily. Indeed, this feature even seems to constitute the *ground* of a triangle's equilateral figure, in that it is *because* its angles are equal that their opposing sides are equal (and vice versa). But, it is not by reason of its equal angles that the triangle is equilateral since if, *per impossible*, the angles were unequal while the sides were equal, then the triangle would nevertheless be equilateral. Likewise, I propose that the criteria,

> A syllogism is valid if and only if it meets requirements of distribution (and quality or quantity),

is specious. That is, the pattern of the distribution values of terms in proposition are what they are necessarily; accordingly, the patterns of distribution values in each of the 256 standard-form syllogisms based on mood and figure, including the 15 valid ones, are thus also what they are necessarily. Accordingly, distribution rules appeal to a unique feature that valid syllogisms necessarily have. But, I propose that it is not by reason of its possession of this feature that a syllogism is valid since if, *per impossible*, the feature were absent in AAA-1, for example, this syllogism would nevertheless be valid.

Finally, the criterion,

> A triangle is equilateral if and only if all its sides are equal,

illustrates the third, essential kind of criterion, because the identifying feature to which it appeals is both necessary, and also is the feature by reason of which an equilateral triangle is equilateral. Accordingly if, *per impossible*, the angles of the triangle were equal while its sides were unequal, it would then not be an equilateral triangle after all since, again, it is by reason of the equality of its sides, rather than its angles, that it is equilateral. Likewise, I propose that the criterion,

> A syllogism is valid if and only if the contributing and receiving terms of the premises retain their roles as such in the conclusion after the receiving middle (which is predicate of an affirmative) and the transmitting middle (which is the subject of a universal) are eliminated,

is essential. That is, I propose that the identifying feature to which this criterion appeals is that necessary feature of valid syllogisms by reason of which they are valid. Accordingly I propose that if, *per impossible*, the distribution rules (along with those of quality or quantity) held for AAA-2, for example, it would nevertheless be invalid since it does not possess that feature to which the above criterion appeals.

7 The Irrelevance of Distribution In conclusion I defend the claim that the criterion I propose is essential, and that the criteria based on distribution are specious. In this defense I appeal to the following assumption:

> If it is possible to see that something holds necessarily and yet not know whether it has a certain property, then that necessity does not hold by reason of that property.

So, for example, it is possible to see that the sides of an equilateral triangle have to be equal without knowing whether its angles are equal (because, for example, one may never have thought about it); accordingly, it follows from the above assumption that this necessity does not hold by reason of the equality of a triangle's angles.

But, if this assumption is unacceptably strong, the following, weaker version is still relevant to the issue:

> If it is possible for someone to see that something holds necessarily and yet not know whether it has a certain property, then that person has good reason to doubt that the necessity holds by reason of that property.

Again, a term is distributed (undistributed) in a proposition if and only if that proposition logically implies every other that is formed by replacing the term with one of its subterms (superterms) by virtue of that replacement. Furthermore, a term is made to be distributed (undistributed) by being the predicate of a negative (affirmative) statement of a proposition. That is, being the predicate of a negative (affirmative) statement is the ground of a term's replaceability capacity. Accordingly, I shall say that the distribution value of a term is "directly related" to the syllogism if that relationship involves its replaceability capacity, that it is "indirectly related" if the relationship involves its role as the predicate of a statement, and that it is "totally unrelated" if it involves neither of these. I conclude that the distribution requirements are never directly related to the syllogistic validity, but are always either indirectly related, or totally unrelated, to it instead.

Since Sommers' distribution requirements include those stipulated by Barker, I proceed by considering Sommers' requirements and, in the process, cover Barker's. Sommers' requirement for the premises is that one middle term (the accepting middle) be undistributed and that the other (the transmitting middle) be distributed.

Now, concerning the accepting middle term, it is clear that it must be the predicate of an affirmative statement in order to receive the quantity from the contributing term; and since this is to possess the ground of undistribution, then it must also be undistributed as well. However, it seems clear that what is essential for the syllogism is that it possess this *ground* of undistribution, rather than actual undistribution, because if, *per impossible*, the term could be the predicate of an affirmative statement without being undistributed, it clearly would work just as well as an accepting middle term. For example, we may see that the argument,

All M's are P
At least nine S's are M
———
At least nine S's are P,

is valid before we come to realize that the M of the minor is undistributed. Indeed, even if we mistakenly concluded that it is distributed it would not shake our confidence in the validity of the argument at all. Instead, we see that it works because the M accepts the contribution from S (i.e., that it is the predicate of an affirmative), whether it is distributed or not. And since this is so, we must conclude that the distribution value is only indirectly related to the syllogism, since it is the ground, rather than the undistribution itself, which is relevant.

On the other hand, the requirement that the transmitting middle term be distributed is totally unrelated to the working of the syllogism. That is, it is not indirectly related since what is necessary is that it be the subject of a universal, and this—contrary to the traditional view—is not even the ground of distribution. And it cannot be directly related since if, *per impossible*, the transmitting middle term should be the subject of a universal premise while being undistributed the premise would still yield its conclusion just as readily. In fact, in the wider scheme of entailment of which the syllogism appears to be only one instance (see Murphree [9]) it is not necessary that the transmitting middle be fully universal; instead, it is necessary only that it be sufficiently nearly universal—i.e., that its exceptions be sufficiently limited—so as to insure the transmission of the quantity from contributing to the receiving term that the conclusion alleges. So, for example the following argument,

At least all but six M's are P
At least nine S's are M
———————————
At least three S's are P,

is valid. And, in fact, we may see that it is valid without realizing that the M of the major is distributed. Indeed, again, even if we mistakenly concluded that it is undistributed it would not shake our confidence in the validity of the argument at all. Instead, we see that it works because the M is sufficiently nearly universal—i.e., that it admits of sufficiently few exceptions—that it necessarily transmits at least three elements of the contribution to the receiver, whether it is distributed or not. And since the case is the same with the standard syllogism (where the only difference is that the exception to the universality of the transmitting middle is zero, rather than nine), we must conclude that the distribution value is totally unrelated.

Sommers' distribution rule for the conclusion, again, requires that the terms keep the same value they had in the premises. Concerning the receiving term, it is clear that it must be the predicate of an affirmative in the premise in order to receive the contribution from the transmitting middle term, and also that it must be the predicate of an affirmative in the conclusion in order to receive the contribution from the contributing term directly. So since in both cases it must possess the ground of undistribution, its distribution value necessarily remains the same. So what the rule of distribution insures is that the major term will play the same role—either as the receiving term, or as the complement of the receiving term—in the conclusion that it does in its premise. However, what is essential for the syllogism is that the major term retain this same ground of distribution or undistribution in the conclusion that it has in its premise. That is, if being the predicate of an affirmative statement, or being its complement, should remain constant while the distribution value changed, then we would assess its validity on the basis of the former condition, rather than the latter. For example, the argument,

Exactly all but six M's are P
At least nine S's are M
At least three S's are P,

is obviously valid, although P is neither distributed nor undistributed in the premise while it is undistributed in the conclusion. Hence, the relevance of the distribution requirement for the receiving term is only indirect; in fact, the requirement is useful only as a way of determining that the ground has remained constant in standard-form syllogisms (which insures that the receiving term is not confused with its complement).

(Admittedly, it seems that "Exactly all but six M's are P" should be analyzed as the conjunction of "At least all but six M's are P" and "At most all but six M's are P" and that, therefore, when they are separated the M of the former is distributed while the M of the latter is undistributed. Then the objection might be advanced that the M of relevant conjunct, "At least all but six M's are P," does turn out to conform to the distribution requirement after all. However, I think this does not weaken the example since it is designed to show that we would not doubt the validity of the argument even if we thought the distribution value failed to hold constant.)

On the other hand, although in the case above the relevance of the distribution requirement for the receiving term is only indirect, the requirement that the distribution value of the contributing term remain the same is totally unrelated to the syllogism. First, it is not related in its use for determining the quantity of the conclusion, since quantity is not the ground of distribution. Moreover, a rule of quantity, like the one advanced above, would prohibit errors of quantity. But it is also totally unrelated as it is used to distinguish the contributing term from its complement since, as a contributing term, it is the subject of the contributing statement while its ground of distribution is the fact that it is the predicate of some other statement. Of course, it does have a distribution value as the contributor in the premise and, granted that the quantity of the conclusion is correct, it will retain that value as the subject of the conclusion. Hence, requiring its distribution value to remain constant turns out to be a convenient way to guard against confusing it with its complement, which has the opposite value. But any other way of insuring this would work just as well because, again, if the value could change while the term remained identical we would surely assess the validity on the basis of the latter condition, rather than the former. For example, the argument,

At least all but six M's are P
Exactly nine S's are M
At least three S's are P,

is obviously valid, although S is neither distributed nor undistributed in the premise while it is undistributed in the conclusion.

(As with the example above, it seems that "Exactly nine S's are M" is properly rendered as "At least nine S's are M" and "At most nine S's are M," and that S is undistributed in the former conjunct, which is the operative one for the argument. But, as with the example above, I believe this does not weaken the point.)

So, I propose that we can see that the conclusion of a valid syllogism follows necessarily from its premises by seeing that the quantity assigned from the contributing term to the receiving term in the conclusion is that quantity assigned from that contributing term to that receiving term, by way of the accepting and transmitting middle terms, in the premises. And I propose that this can be seen without knowing whether the syllogism conforms to all, or any, of Sommers' distribution requirements. So granted the strong assumption above, viz.,

> if it is possible to see that something holds necessarily and yet not know whether it has a certain property, then that necessity does not hold by reason of that property,

it follows that the entailment does not hold by reason of the syllogism's conformity to the distribution requirements. Or, on the weaker, alternative assumption stated, it follows that at least we have good reason to doubt that the entailment holds by reason of that conformity.

Accordingly, I conclude that Sommers' distribution rules are never directly related to the syllogism. Rather, as noted above, I conclude that they are indirectly related to it in two of their applications, and that they are totally unrelated to it in the other two.

Furthermore, from the same considerations I conclude that Barker's distribution rules are only indirectly related to the syllogism in the one application that prohibits illicit process of the major, while the other applications, like the analogous applications of Sommers' rules, are totally unrelated to it.

In summary, I propose the distribution and undistribution of terms to be, respectively, their replaceability by subterms and superterms; and I propose that such replaceability is grounded in the principle that what is denied of a term is denied of all its subterms, and that what is affirmed of a term is affirmed of all its superterms. As such, I contend that distribution and undistribution are coherent logical notions. Moreover, I suggest that they are important notions in the analysis of such nonsyllogistic arguments as

All colored things are visible
Some squares are blue things
———————————————
Some rectangles are perceivable,

where the terms, in fact, are replaced.

However, the fact is that no such replacement ever occurs in the syllogism. Accordingly, I propose that any attempt to explain the syllogism by appealing to this systematically unused capacity of its terms inevitably tends to conceal, rather than reveal, the nature of the entailment actually involved.

Why it is that the assigning term of the premises must also be the assigning term of the conclusion can be clearly seen; why it is that the transmitting middle term must be the subject of a universal can be clearly seen; why it is that both the accepting middle term and the receiving extreme term must be the predicates of affirmative statements can be clearly seen; but why it is that these terms must have certain replaceability capacities as well is far from clear. The fact is that although the standard application of the syllogism is limited to propositions whose terms necessarily have these capacities, syllogistic entailment, itself, does not require them at all.

Acknowledgments I am very grateful to an anonymous referee for insightful comments on, and suggestions for, an earlier draft of this paper.

REFERENCES

[1] Barker, S., *The Elements of Logic*, McGraw-Hill, New York, 1965. Zbl 0168.00303 1, 5

[2] Englebretsen, G., "Defending Distribution," *Dialogos*, vol. 45 (1985), pp. 157–159. 1, 3

[3] Geach, P., *Reference and Generality: An Examination of Some Medieval and Modern Theories*, Cornell University Press, New York, 1962. 1

[4] Geach, P., "The Doctrine of Distribution," *Mind*, vol. LXV (1956), pp. 67–74. 1

[5] Geach, P., "Distribution: A Last Word," *Philosophical Review*, vol. LXIX (1960), pp. 396–398. 1

[6] Johnstone, H., "Pairs of Negative Syllogistic Premises Yielding Conclusions," *Logique et analyse*, vol. 85–86 (1979), pp. 109–110. 5.1

[7] Katz, B., and A. Martinich, "The Distribution of Terms," *Notre Dame Journal of Formal Logic*, Vol. XVII (1976), pp. 279–283. Zbl 0283.02016 2

[8] Kelley, D., *The Art of Reasoning*, Norton, New York, 1990. 5

[9] Murphree, W., *Numerically Exceptive Logic: A Reduction of the Classical Syllogism*, Peter Lang, New York, 1991. 3, 7

[10] Sommers, F., "Distribution Matters," *Mind*, vol. LXXXIV (1975), pp. 27–46. 1, 3, 5

[11] Toms, E., "Mr. Geach on Distribution," *Mind*, vol. LXXIV (1965), pp. 428–431. 1

[12] Williamson, C., "Traditional Logic as the Logic of Distribution Values," *Logique et analyse*, vol. 14 (1971), pp. 729–746. 3

Department of Philosophy and Religion
Mississippi State University
P.O. Box JS
Mississippi State, Mississippi 39762–5833
email: wam1@ra.msstate.edu