

DOCUMENT RESUME

ED 469 176

TM 034 481

AUTHOR Plumer, Gilbert E.
TITLE Testing for Structure Recognition. LSAC Research Report Series:
INSTITUTION Law School Admission Council, Newtown, PA.
REPORT NO LSAC-RR-98-02
PUB DATE 1999-08-00
NOTE 15p.
PUB TYPE Opinion Papers (120)
EDRS PRICE EDRS Price MF01/PC01 Plus Postage.
DESCRIPTORS *College Entrance Examinations; *Persuasive Discourse; Test Construction; Test Items
IDENTIFIERS Argument Research; *Task Structure

ABSTRACT

The nontechnical ability to identify or match argumentative structure is considered by many to be an important reasoning skill. Instruments that have questions designed to measure this skill include major standardized tests for graduate school admission, for example, the Law School Admission Test (LSAT), the Graduate Record Examination (GRE), and the Graduate Management Admission Test (GMAT). Writers and reviewers of such tests need an appropriate foundation for developing such questions. They need a proper representation of phenomenological argumentative structure, for legitimacy, and because these tests affect people's lives. This paper attempts to construct an adequate and appropriate representation of such structure, the logical structure that an argument is perceived to have by mature reasoners, albeit reasoners who are not trained in logic. (Contains 1 table and 22 references.) (Author/SLD)

PERMISSION TO REPRODUCE AND
DISSEMINATE THIS MATERIAL HAS
BEEN GRANTED BY

J. VASELECK

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- This document has been reproduced as received from the person or organization originating it.
- Minor changes have been made to improve reproduction quality.
- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

■ **Testing for Structure Recognition**

Gilbert E. Plumer
Law School Admission Council

■ **Law School Admission Council**
Research Report 98-02
August 1999



A Publication of the Law School Admission Council

The Law School Admission Council is a nonprofit corporation that provides services to the legal education community. Its members are 196 law schools in the United States and Canada.

Copyright© 1999 by Law School Admission Council, Inc.

All rights reserved. This book may not be reproduced or transmitted, in whole or in part, by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without permission of the publisher. For information, write: Communications, Law School Admission Council, Box 40, 661 Penn Street, Newtown, PA 18940-0040.

LSAT® and the Law Services logo are registered marks of the Law School Admission Council, Inc.

This study is published and distributed by the Law School Admission Council (LSAC). The opinions and conclusions contained in this report are those of the author and do not necessarily reflect the position or policy of the Law School Admission Council.

Table of Contents

Executive Summary.....	1
Abstract.....	1
Introduction.....	1
Question Format.....	2
Formal Structure.....	4
Informal Structure.....	8
References.....	11

Executive Summary

The nontechnical ability to identify or match argumentative structure is considered by many to be an important reasoning skill. Instruments that have questions designed to measure this skill include major standardized tests for graduate school admission, for example, the Law School Admission Test (LSAT), the Graduate Record Examinations (GRE), and the Graduate Management Admission Test (GMAT). Writers and reviewers of such tests need an appropriate foundation for developing such questions—they need a proper representation of phenomenological argumentative structure—for legitimacy, and because these tests affect people's lives. This paper attempts to construct an adequate and appropriate representation of such structure, that is, the logical structure that an argument is perceived to have by mature reasoners, albeit ones who are untrained in logic.

Typical questions on the standardized tests mentioned that ask the test taker to identify or match structure consist of a short argumentative passage, a question stem on the order of either (*I*) "The argument's method of reasoning is" or (*M*) "The pattern of [flawed] reasoning in the argument above is most similar to that in which one of the following arguments?" and five answer choices. Since all answer choices must be cast in ordinary nontechnical prose, questions of type (*I*) generally concern only the grosser features of an argument's structure. Questions of type (*M*), however, can pertain to much more subtle features (since the test taker is not asked to explicitly identify them), and it is this type that constitutes our focus.

Regarding formal or deductive structure, the principle that is argued for is this: In the construction and defense of questions of type (*M*), when a question stem emphasizes reasoning structure by the use of a phrase such as "pattern of reasoning" or "parallel reasoning," more weight can legitimately be assigned to reasoning structure than to surface logical structure and the structure of the argument's terms. Yet these latter must still be taken into account in determining overall (phenomenological) argumentative structure. In this way we adopt the principle that the proper or "best formal representation will be the one that exhibits the most structure."

It is undeniable that in ordinary life we routinely evaluate arguments as invalid or fallacious. But if a certain theoretical view that alleges a strong asymmetry between showing validity and showing invalidity were right, many, if not all, of these judgments would be illegitimate. A case is developed against this view, and this helps to legitimate questions on an exam like the LSAT that ask test takers to match *flawed* patterns of reasoning.

Regarding the informal or nondeductive structure of an argument, my thesis is that it can include any of the argument's general elements that figure in the purported cogency of (that function in purporting to advance the conclusion in) any pattern of reasoning. The proper representation of a given argument's phenomenological argumentative structure will include these elements whether or not the given argument exhibits the pattern of reasoning in question. This point regarding informal structure corresponds to the point before regarding formal structure that such features as surface logical structure and the structure of the argument's terms need to be taken into account. But also as before, more weight can legitimately be assigned to the general elements that actually figure in the purported cogency of the given argument. It is shown how this approach has more substance to it than might be evident. In the first place, the approach rules out purely syntactical features, such as the location of the argument's conclusion, as immaterial. Secondly, it coheres well with the established tradition in informal logic that the cogency of a nondeductive argument is largely a matter of its form. Finally, it is shown how the approach can be used to help explain both the statistical success and failure of nondeductive test items of type (*M*).

Abstract

The nontechnical ability to identify or match argumentative structure is considered by many to be an important reasoning skill. Instruments that have questions designed to measure this skill include major standardized tests for graduate school admission, for example, the Law School Admission Test (LSAT), the Graduate Record Examinations (GRE), and the Graduate Management Admission Test (GMAT). Writers and reviewers of such tests need an appropriate foundation for developing such questions—they need a proper representation of phenomenological argumentative structure—for legitimacy, and because these tests affect people's lives. This paper attempts to construct an adequate and appropriate representation of such structure, that is, the logical structure that an argument is perceived to have by mature reasoners, albeit ones who are untrained in logic.

Introduction

What is the proper representation of phenomenological argumentative structure? By "phenomenological argumentative structure" I mean the logical structure that an argument is perceived to have by mature

reasoners—yet ones who are untrained in logic. Except for a few remarks, this paper will not be concerned with whether this informal ability to identify or match argumentative structure is an important reasoning skill; rather, it will be primarily concerned with *judging* or attempting to measure this skill. Instruments that have questions designed to do this include major standardized tests for graduate school admission, for example, the Law School Admission Test (LSAT), the Graduate Record Examinations (GRE), and the Graduate Management Admission Test (GMAT). Writers and reviewers of such tests need an appropriate foundation for developing such questions—they need a proper representation of phenomenological argumentative structure—for legitimacy, and because these tests affect people’s lives.

A further motivation is cost. A single question on these tests is very expensive to develop, so it is not a trivial matter when a test item is miscast and fails psychometric statistical review. Even given this, however, it may be that an attempt to represent phenomenological argumentative structure through (probably costly) empirical studies would not be advisable. The results could be bewildering and not generalizable (one study found that the diagrammatic aids examinees drew when taking like tests tended to be quite idiosyncratic—Cox & Brna, 1995). Instead, the approach that this paper will take will be mainly philosophical rather than empirical.

It would certainly appear that the informal or nontechnical ability to identify or match argumentative structure is fundamental to reasoning well. With only one putatively clear kind of exception, the validity (for deduction), or more broadly, cogency (for both deduction and nondeduction), of an argument is entirely (for deduction) or largely (for nondeduction) a function of its logical structure or form (cf., e.g., Sainsbury, 1991, ch. 1; also Walton, 1995, ch. 5 for a distinction of 25 nondeductive argument structures or “schemes”). The same applies to the *invalidity* or lack of cogency of an argument. The only arguments that supposedly constitute an exception are those that proceed through conceptual analysis, that is, those that are termed “materially” valid or invalid; a classic example is “this is red all over, so it is not blue all over” (e.g., Read, 1994). So apart from such arguments, and apart from conversational and rhetorical matters and matters related to the actual truth values of premises and conclusions, to perceive the logical structure of an argument is to perceive that in virtue of which the argument is good or bad (deduction) or is to perceive much of what makes the argument good or bad (nondeduction). Naturally, then, a principal way of assessing the cogency of a given argument is to match its structure with that of an argument whose cogency is known or obvious. In the case of showing lack of cogency, this tactic is called “refutation by logical analogy.” (Some of the presuppositions of these remarks will be defended below.)

Question Format

Typical questions on the standardized tests mentioned that ask the test taker to identify or match structure consist of a short argumentative passage, a question stem on the order of either

- (I) The argument’s method of reasoning is
or
- (M) The pattern of [flawed] reasoning in the argument above is most similar to that in which one of the following [arguments]?

and five answer choices. Since all answer choices must be cast in ordinary nontechnical prose, questions of type (I) generally concern only the grosser features of an argument’s structure. Questions of type (M), however, can pertain to much more subtle features (since the test taker is not asked to explicitly identify them), and it is this type that will constitute our focus.

Notice that (M) questions create a somewhat artificial setting that usefully restricts the task in a number of ways. That the text in the passage (and normally in each of the answer choices) is supposed to constitute an argument is settled, although clearly in ordinary discourse “it is not always easy” to determine whether this is the case (Baum, 1981, p. 91). Moreover, whether or not the argumentative structure is supposed to lack cogency is normally given in the question stem by whether or not a term such as “flawed” appears in the stem. This can make a great difference in the argumentative structure that people perceive. Now before any test item is used in a scored section of an LSAT, it appears in an unscored section of a previous LSAT; this is known as

Earlier versions of this paper were presented at Law School Admission Council, Newtown, PA, and at the Fourth International Conference on Argumentation in Amsterdam. The paper has benefited from discussion on these occasions and with Kenneth Olson, and from written comments by Deborah Kerman and Stephen Luebke.

“pretesting.” The purpose is to determine the item’s psychometric statistical characteristics so that if these are acceptable, the item can later be incorporated according to specification into a section that will be scored. Example (1), with “flawed” appearing in the question stem and (B) as the credited response, performed on the LSAT at pretest in a statistically acceptable, albeit marginal, fashion:

*Example (1)**

John is an excellent member of the team. All the members of the team are fathers. Therefore, John is an excellent father.

Which of the following exhibit the same [flawed] logical structure as that exhibited in the argument above?

- I. This is a fake diamond. All diamonds are hard. Therefore, this is hard.
- II. This is a red apple. All apples are fruits. Therefore, this is a red fruit.
- III. This is a big flea. All fleas are pests. Therefore, this is a big pest.

- (A) II only
- (B)* III only [*Denotes the credited response.]
- (C) I and II only
- (D) II and III only
- (E) I, II, and III

Havoc ensued, however, when, with the same credited response, “flawed” was taken out. The reason seems plain: In the first case pretest examinees naturally took the rather informal fallacy of distributing an attributive adjective (“excellent,” “big”) across two different noun phrases as part of the argumentative structure. (Recently, this fallacy has apparently been adequately formalized in first-order predicate logic for some types of attributive adjectives; see Ben-Yami, 1996.) In the second case, with “flawed” out, many pretest examinees interpreted the structure more formally and saw the passage, II, and III as exhibiting the same underlying “logical” (as opposed to “illogical”?) structure; so they picked option (D). Hence in general, insofar as test takers can depend on the fallaciousness of the passage’s argument being noted in the question, the matter of whether to interpret the argument charitably basically becomes irrelevant.

These factors direct and limit the interpretative task for test takers. Variations on such factors include leaving out the phrase “pattern of” (or an equivalent) for arguments in which formal structure is not prominent or those in which conceptual connections are prominent; using a term such as “questionable” instead of “flawed” for suspicious, but not clearly fallacious, arguments; and specifying the number of flaws (e.g., “Which one of the following exhibits both of the logical flaws exhibited in the argument above?”). But the wording of the question stem is not the only kind of constraint that defines the interpretative task; the other major constraint lies in how the passage argument and (especially) the correct answer choice are constructed. Other than that, obviously, they must be constructed to accurately reflect the stem’s wording (and vice versa), I think that this constraint principally amounts to the injunction that the arguments normally not be substantially enthymematic. Arguments that were substantially enthymematic could be too subject to variance in the perception or analysis of their structure to be fair and defensible material. Moreover, measurement of the ability to match structure could be confounded by the additional task of dealing with unstated premises or conclusions.

It might be wondered whether such constraints create a setting that is so artificial that the ordinary nontechnical ability of mature reasoners to identify or match argumentative structure is not really being measured. It seems, however, that these constraints, common to standardized tests, that function to direct and limit the interpretative task for the test taker, are probably just harmless *context surrogates*. For it is an argument’s context and background information specific to its presentation that generally decides such matters as whether the discourse is supposed to constitute an argument or whether to apply a principle of charity and take an ostensibly fallacious argument as a cogent enthymeme (assuming that it has not had an “undeserved

* All test items reproduced in this paper are © Law School Admission Council.

persuasive power on an audience”—Adler, 1994, p. 276). Standardized tests that are not unduly long generally could not provide realistic surrounding context for arguments and still be reliable, since a test’s reliability is an asymptotic function of the number of questions it has (assuming they are of approximately equal quality) (e.g., Gulliksen, 1987, ch. 8). In addition, if a large amount of text were provided as surrounding context, the skills measured would be less definite insofar as the test taker would have more opportunity to apply unintended skills. An indication that context surrogates are harmless is a high correlation between performance on the test and the performance that the test is used to predict. On the LSAT, questions of type (*I*) and (*M*) appear in “Logical Reasoning” sections, which have a (very high) average correlation of .483 with first-year law school grades (Roussos & Norton, 1999).

Formal Structure

My thesis here is that if the passage and answer choices in a question can be formally analyzed at all, the formal analysis that is the proper representation of phenomenological argumentative structure is normally that which *departs least* from what actually appears in these arguments, but with a special consideration given to elements that figure in the arguments’ purported validity or (more broadly) cogency. This seems correct for at least two reasons. First, almost any departure from actual text is *prima facie* questionable (cf., e.g., Sainsbury, 1991, ch. 6). A common departure is taking ordinary language universal or existential quantifications that are not in conditional or conjunctive form as if they were in these forms since that is how they are translated in first-order predicate logic. Of course the alternative that is closer to the actual text insofar as it expresses the surface logical structure is that of Aristotelian or syllogistic logic; and this alternative is preferable so long as it adequately expresses purported validity or cogency. So for example, the proper representation of the phenomenological structure of “Some people are fools” is *Some P’s are F’s*, not $\exists x (Px \ \& \ Fx)$. In a question of type (*M*) that was recently pretested on the LSAT and that failed statistically, the major premise in the passage was “children would be proud of themselves if their teachers were proud of them.” This was supposed to be matched in the credited response with “any biography that flattered its subject would be liked by that person.” Possibly, the difference between the two forms that these sentences exhibit, among other things, contributed to the lack of success of the question.

A second reason for understanding the proper representation of phenomenological argumentative structure generally to be that which departs least from the actual text is that this approach is logically *inclusive*. It respects and attempts to take into account all of the text that could reasonably be taken into account in light of the various established logics—syllogistic, propositional, first and higher order predicate logics; tense, modal, deontic, epistemic, relevance, and probabilistic logics; logic with generalized quantifiers; logic of indexicals; and so on—within the discipline of logic. A pragmatic side benefit is that a test taker who happened to be trained in logic could legitimately appeal to any of these logics in answering or later challenging a test question. However, since a test taker need not have any training in logic, it would be inappropriate to *de facto* require the test taker to have mastery of and endorse some particular logic or formal analysis by, for instance, insisting on a formal analysis that incorporates certain putative logical constants to the exclusion of others. This makes it critical that the formal analysis employed by test writers and reviewers be inclusive and close to the actual text.

For any argument, there is a strong temptation to proceed as if standard predicate logic can adequately or exhaustively represent its structure or form, probably because this logic is firmly established and very familiar. But standard predicate logic does not incorporate quantification over properties (as does a second- or higher-order logic). Nor does it incorporate generalized quantifiers (e.g., “the,” “few,” “most”) or modal (e.g., “necessarily,” “can”), tense (e.g., “in the future,” “now”), deontic (e.g., “should,” “permissible”), epistemic (e.g., “knows,” “guesses”), or probability operators (e.g., “likely,” “there is a chance that”)—all of which are quite reasonably regarded as logical constants. And so on. Hence, standard predicate logic, with its limited supply of logical constants (“all,” “not,” “if . . . then,” etc., interpreted in the classical narrow way), can yield representations of structure that depart dramatically from actual text.

The appropriate recognition of the power and appeal of first-order predicate logic, as well as syllogistic and propositional logic, seems to be to give these logics priority over less well-established logics in the formal representation of phenomenological argumentative structure. Certainly, “deviant” logics on the order of many-valued and intuitionistic logics fall under the latter category. They have a substantial history now of attracting few advocates; so if even logicians are generally repulsed, it is hard to see how such logics could shed light on how ordinary mature reasoners perceive argumentative structure. Furthermore, first-order predicate logic has shown itself to be remarkably adaptable and extendable—from Russell’s Theory of Descriptions and

Davidson's proposal about adverbial modification (involving quantification over events) to an extension such as quantified modal logic. Indeed, there has even been work on "a unified account of a fairly wide range of logical systems," including "classical logic, relevant logics such as Anderson and Belnap's R, close relatives of fuzzy logic, some modal logics and many weaker, but still interesting, nonstandard systems" (Slaney, 1990, p. 74).

In formally representing phenomenological argumentative structure it would be too simplistic to follow any such principle as Haack's (1978, pp. 24-25):

... the optimal formal representation [is] the one which reveals the least structure consistently with supplying a formal argument which is valid in the system if the informal argument is judged extra-systematically valid. This is Quine's *maxim of shallow analysis* ... "where it doesn't itch, don't scratch."

Compare Luebke (1995, p. 40):

When can we say that two such arguments in ordinary language have the same argument structure? Must they be identical in respect of every one of their logical elements? The answer to this last question seems clearly no, for some of the logical elements of an argument function to advance the conclusion of the argument and others do not.

One problem with this kind of view is that it does not cover *fallacious* arguments; so there at least would have to be amendment in terms of "invalidity in the system," "judged extra-systematically invalid," and "purporting to advance the conclusion." A more serious problem is indicated in how Haack is a little misleading with respect to Quine's view. Quine's "*maxim of shallow analysis*" actually says "*expose no more logical structure than seems useful* for the deduction or other inquiry at hand" (1960, p. 160). The inquiry at hand *here* is the proper representation of the logical structure that an argument is perceived to have by mature yet untrained (in logic) reasoners. Of course *we* (logicians) can distinguish between, on the one hand, the reasoning structure in an argument—how the "logical elements" function in purporting to establish the conclusion—and on the other hand, such features as surface logical structure and the structure of the argument's terms. But there does not appear to be any acceptable way of requiring untrained reasoners to take only the former into account.

Surely, question stems of type (*M*) do not suffice; and if these do not, nothing will (for instance, the question stem in Example (1) is worse). After all, although these questions explicitly ask the test taker to focus on the "pattern of reasoning," they do not say anything to the effect that to determine this pattern one should ignore term structure and go logically deep when and to the extent necessary. They cannot do this since these concepts are technical or relative to a specific system of logic. Consider:

Example (2) (2/94 LSAT)

Government official: Clearly, censorship exists if we, as citizens, are not allowed to communicate what we are ready to communicate at our own expense or if other citizens are not permitted access to our communications at their own expense. Public unwillingness to provide funds for certain kinds of scientific, scholarly, or artistic activities cannot, therefore, be described as censorship.

The flawed reasoning in the government official's argument is most parallel to that in which one of the following?

- (A) All actions that cause unnecessary harm to others are unjust; so if a just action causes harm to others, that action must be necessary.
- (B) Since there is more to good manners than simply using polite forms of address, it is not possible to say on first meeting a person whether or not that person has good manners.
- (C) Acrophobia, usually defined as a morbid fear of heights, can also mean a morbid fear of sharp objects. Since both fears have the same name, they undoubtedly have the same origin.
- (D)* There is no doubt that a deed is heroic if the doer risks his or her own life to benefit another person. Thus an action is not heroic if the only thing it endangers is the reputation of the doer.
- (E) Perception of beauty in an object is determined by past and present influences on the mind of the beholder. Thus no object can be called beautiful, since not everyone will see beauty in it.

The credited response, (D), is a fairly straightforward instance of one variety, namely, *if r then h, therefore if not r then not h*, of the formal fallacy of confusing necessary and sufficient conditions. The argument in the passage may also be said to exhibit this particular fallacy, but notice that the term corresponding to *r* is propositionally disjunctive only in the passage and that the passage's conclusion is a categorical statement, not a conditional as in (D). (The negation of *r* is also clearer in (D)'s conclusion than it is in the passage—but this has more to do with the reasoning structure.) My point is, such differences must be taken into account in the writing and review of matching structure test questions. Phenomenologically, the flawed reasoning in the passage of Example (2) is not exactly parallel to that in (D) because of such differences. It might seem to be exactly parallel if one puts undue emphasis on the word “flawed” in the question stem; in fact, the words “reasoning” and “parallel” are equally (un)emphasized there. Also, one might be fooled by the preceding propositional representation of the fallacy. But the kind of structural differences in question constitute a matter of degree, and they can *accumulate* to the point where the test item becomes dubious or indefensible. For instance, suppose that the passage in Example (2) consisted entirely of categorical statements. It is at least questionable whether the categorical error (*all R's are H's, therefore all non-R's are non-H's*) is the same as the propositional error.

To take another kind of example, suppose the passage and a noncredited response were a Modus Ponens and a Modus Tollens, respectively, the terms of which were all atomic statements (plus some negation). Suppose also that the credited response was a Modus Ponens, but its terms were all *really* complex compound statements. Would such a test question be defensible on the grounds that only in the passage and credited response are the patterns of reasoning the “same,” even though any Modus Ponens (Modus Tollens) can be turned into a Modus Tollens (Modus Ponens) simply by the application of contraposition to the major premise? The test question would at least be problematic.

In Example (2) I think that (A) is, among the noncredited options, the one that is closest to being correct. But it is not correct. And the same reason yields both of these judgments, namely, that (A), where at least the categorical statement is translated as a conditional (or vice versa), is a contrapositive inference, which of course is valid. The actual structural differences between the passage and (D) are relatively insignificant when one considers that the question stem asks one to pick the option with the most parallel flawed reasoning—so the focus is on reasoning structure—and that (A) is formally valid (and is not informally fallacious either), whereas the passage and (D) are formally invalid.

The principle that is emerging is this: In the construction and defense of questions of type (*M*), when a question stem emphasizes reasoning structure by the use of a phrase such as “pattern of reasoning” or “parallel reasoning,” more weight can legitimately be assigned to reasoning structure than to surface logical structure and the structure of the argument's terms. Yet these latter must still be taken into account in determining overall (phenomenological) argumentative structure. In this way we adopt the principle that Haack rejects, namely, the proper or “best formal representation will be the one that exhibits the most structure” (1978, p. 24); it involves at least the argument's logical constants (broadly construed) and the logically significant pattern of occurrence of these logical constants, individual constants, variables, and predicate terms. Such a fine-grained notion of structure means that passage/credited response pairs in good matching structure test questions generally will not consist of arguments with *identical* structures. Accordingly, question stems should be cast in terms of reasoning or reasoning patterns that are *most similar* or *most parallel* to one another, like (*M*) and as in Example (2), rather than in terms of identity, as in Example (1). The weaker terminology also has the advantage of hedging one's bets against unnoticed structural differences.

Differences in term structure can themselves signal differences in reasoning structure, so we ignore the former to our peril. Luebke (1995, p. 40) says:

- (a) if p then q, p, therefore q
- (b) if (r and s) then (t or u), r and s and y and z, therefore t or u or v or w

These two arguments do not have exactly the same logical elements, but the pattern of reasoning that establishes the conclusion is the same in each case—*modus ponens*. Both arguments argue for their conclusion in the same way. So the argument structures, as opposed to the term structures, are the same.

In fact, (b) does not exhibit Modus Ponens since in (b)'s conclusion the consequent (“t or u”) of the conditional that constitutes the major premise is not affirmed; rather, the much weaker “t or u or v or w” is affirmed. For (b) to instantiate Modus Ponens, its conclusion would have to read *therefore t or u, therefore* [by twice applying the rule of inference of Addition] *((t or u) or v) or w*—but then, the overall pattern of reasoning is not simply

Modus Ponens. Even aside from this, it is questionable whether (a) and (b) exhibit the same pattern of reasoning because the rule of inference of Simplification must be applied (twice) to (b)'s minor premise in order for it to be clearly the case that the antecedent ("r and s") is affirmed. (Technically, this discussion is rendered somewhat indeterminate by the fact that in (b) the minor premise and conclusion are not even well-formed formulas.)

It will prove useful to examine the following case discussed by Massey (1995, p. 161):

Example (3)

If something has been created by God, then everything has been created by God.
Everything has been created by God.

Something has been created by God.

Massey says of this argument that it "instantiates ... *affirmation of the consequent*" yet it "is valid." The reference to God in the argument is not essential; alternatively, the argument could be cast "if something is physical, then everything is physical" (which, indeed, is one way of expressing a part of Bishop Berkeley's philosophy), and so on. Massey uses this case to try to help establish what he calls "the asymmetry thesis" (1975, p. 66):

To show that an argument is *valid* it *suffices* to paraphrase it into a demonstrably valid argument form of some (extant) logical system; to show that an argument is *invalid* it is *necessary* to show that it cannot be paraphrased into a valid argument form of *any* logical system, actual or possible.

I think Massey is wrong on all counts.

Example (3) would be regarded as a valid argument in standard predicate logic. As expressed in that system, the conclusion follows from the minor premise since $\forall x \exists y (x = y)$ ("everything exists") is a theorem; and although the major premise is not used in drawing the conclusion, this does not matter formally since the system is monotonic (i.e., "if you start with a deductively valid argument, then, no matter what you *add* to the premises, you will end up with a deductively valid argument"—Sainsbury, 1991, p. 11). This last point itself indicates a problem with Massey's account. If it is not the case that the conclusion is being drawn *through* affirming the consequent of the conditional (major) premise, in what sense could Example (3) "instantiate" the "so-called formal fallacy" (1995, p. 160) of affirming the consequent? (The fallacy is "so-called" for Massey since it is clear that the necessary condition he proposes for showing that an argument is invalid could never be satisfied.) If the machinery of standard predicate logic were all that we had at our disposal, we could still say that Example (3), understood as valid, commits a gross *informal* fallacy of irrelevance (of its major premise) (or we could say instead that the argument actually consists just of the minor premise and the conclusion). And in relevance logic, this fallacy is treated as a *formal* fallacy (e.g., Haack, 1978, p. 199).

In a particularly plausible version of free logic, $\forall x \exists y (x = y)$ is not a theorem; the logic does not require that every domain of interpretation be nonempty. This is plausible because it is hard to see the fact that there is something rather than nothing as a truth of *logic* (cf. Sainsbury, 1991, pp. 205-10). Standard predicate logic's requirement that every domain be nonempty seems to be merely a simplifying assumption that is innocuous for most purposes. But then this falsifies the sufficient condition, proposed by Massey, for showing that an argument is valid; "paraphrased" in a respectable system of (free) logic (if not also in relevance logic) Example (3) is invalid, although it is valid as paraphrased in standard predicate logic. Also falsified is the necessary condition for showing invalidity, since this is more or less just the contrapositive of the validity sufficient condition.

The strongest principle that Massey is entitled to, one that is true as well, is relativized to a system of logic:

An argument is valid (invalid) *in a system of logic S* if and only if there is some (no) valid argument form in S that the argument instantiates.

This is perfectly adequate to handle all the stock cases; for example, in propositional logic we would not want to say that a case of Modus Ponens is invalid merely on the grounds that it also instantiates the invalid form $r, p,$

therefore q . Moreover, the asymmetry this indicates between showing validity and invalidity seems offset by the opposite asymmetry that it is possible to show that an argument is invalid, but not that it is valid, simply by considering the actual truth values of its premises and conclusion—if it has true premises and a false conclusion, the argument is invalid. So contrary to Massey, it is not true that “our ability to prove invalidity is markedly more circumscribed than our ability to prove validity” (1995, p. 164). What is true is, as Govier (1995, p. 175), puts it, “*formal analysis presupposes nonformal judgment as to the appropriacy of a paraphrase and the correctness of the logical system to which the argument is referred.*”

As expressed in propositional logic, Example (3) is a clear case of the invalid form of affirming the consequent. So what is the proper representation of Example (3)’s phenomenological argumentative structure? I think that for cases like Example (3) “*nonformal judgment*” must say that the matter is seriously indeterminate. We cannot merely analyze the argument propositionally because there is logical structure (repeated from the major premise) in the minor premise and in the conclusion, and it functions in purporting to establish the conclusion. But as expressed in one respectable system of logic that takes account of this structure, the argument is valid (although informally fallacious); in at least one other respectable system it is invalid. A variation on Example (3) that is in some ways more interesting is “if Lyra is a female sibling then she is a sister, Lyra is a sister, therefore she is a female sibling.” This argument is *materially* valid by virtue of the analytic truth that a sister (in the relevant sense) just is a female sibling, yet as expressed in propositional logic the argument is invalid. (One might want to say that the conditional here is somehow “really” a biconditional; but notice that the same might be said of Example (3) and the “physicalist” variation that I initially gave of it. However, in testing using short fixed texts, as in much communication such as legal contracts, the focus must be on what is actually said and not on anything like divining author meaning. Cf. Adler 1994, pp. 275-76.) So to avoid confusion or *de facto* requiring test takers to endorse a particular system of logic, it seems that no such seriously indeterminate argument should appear in a question of type (I) or (M) on an exam like the LSAT. Simply not indicating that the reasoning is flawed could very well engender a statistically dreadful performance, as with Example (1).

The other moral to draw from this consideration of Example (3) has to do with the undeniable fact that in ordinary life we routinely evaluate arguments as invalid or fallacious. If Massey were right, many, if not all, of these judgments would be illegitimate. But he is not right, and this is especially telling since he presents perhaps the strongest theoretical case for the kind of view in question. The positive alternative that is particularly appropriate for the study of phenomenological argumentative structure is a kind of “transcendental argument for arguments having a certain kind of structure: this is the structure arguments need to have in order for us to assess them in the ways in which we do” (Parsons, 1996, p. 174). Needless to say, this helps to legitimate questions on an exam like the LSAT that ask test takers to match *flawed* patterns of reasoning.

Informal Structure

I think that, phenomenologically, the informal logical structure of an argument can include any of the argument’s general elements that figure in the purported cogency of (that function in purporting to advance the conclusion in) any pattern of reasoning. The proper representation of a given argument’s phenomenological argumentative structure will include these elements whether or not the given argument exhibits the pattern of reasoning in question. This point regarding informal structure corresponds to the point before regarding formal structure that such features as surface logical structure and the structure of the argument’s terms need to be taken into account. But also as before, more weight can legitimately be assigned to the general elements that actually figure in the purported cogency of the given argument.

This approach has more substance to it than might be evident. In the first place, it rules out purely syntactical features, such as the location of the argument’s conclusion, as immaterial: these do not figure in the purported cogency of any pattern of reasoning.

Secondly, it coheres well with the established tradition in informal logic that the cogency of a nondeductive argument is largely a matter of its form. Salmon, for instance, indicates that a nondeductive argument is cogent if “the argument has a correct form, and . . . the premises of the argument embody all available relevant evidence”; so for example, the “correct” form of the “argument from authority” is *x is a reliable authority concerning p, x asserts p, therefore p* (1973, p. 91; cf. Walton, 1995, ch. 5). Here, as is typical of informal structure, general elements that are not topic neutral (the concepts of a reliable authority and of asserting) are treated as logical constants. But this is hardly radical; it is a move that is routinely made even in *formal* (e.g., tense and deontic) logic. This is a fundamental point that appears to be insufficiently appreciated by those who,

like Lambert and Ulrich (1980, ch. 1, sec. 3; cf. Massey, 1995, pp. 159-60), hold that informal fallacies cannot be structurally defined. Their ostensibly “formalist” view involves the claim that validity precludes fallaciousness, which is about as (im)plausible as its corollary, namely, that nondeductiveness precludes cogency (for more argument against the view in question, see, e.g., Johnson, 1989; Govier, 1995). In any case, notice that a consequence of the present approach seems to be that the fact that an argument purports to proceed through conceptual analysis (as with materially valid arguments) should be counted as a (informal) structural feature.

Regardless of the theoretical debate about the extent to which informal fallacies can, or should (Berg, 1987; Brinton, 1995), be structurally defined, there generally seems to be little difficulty in attributing and relying on such structure in practice—at least on major standardized tests for graduate school admission. Consider:

Example (4) (6/93 LSAT)

Genevieve: Increasing costs have led commercial airlines to cut back on airplane maintenance. Also, reductions in public spending have led to air traffic control centers being underfunded and understaffed. For these and other reasons it is becoming quite unsafe to fly, and so one should avoid doing it.

Harold: Your reasoning may be sound, but I can hardly accept your conclusion when you yourself have recently been flying on commercial airlines even more than before.

Which one of the following relies on a questionable technique most similar to that used in Harold’s reply to Genevieve?

- (A) David says that the new film is not very good, but he has not seen it himself, so I don’t accept his opinion.
- (B) A long time ago Maria showed me a great way to cook lamb, but for medical reasons she no longer eats red meat, so I’ll cook something else for dinner tonight.
- (C) Susan has been trying to persuade me to go rock climbing with her, claiming that it’s quite safe, but last week she fell and broke her collarbone, so I don’t believe her.
- (D)* Pat has shown me research that proves that eating raw green vegetables is very beneficial and that one should eat them daily, but I don’t believe it, since she hardly ever eats raw green vegetables.
- (E) Gabriel has all the qualifications we have specified for the job and has much relevant work experience, but I don’t believe we should hire him, because when he worked in a similar position before his performance was mediocre.

Here I would say that the appropriate representation of the informal fallacy is *s does not heed s’s own credible advice a, therefore a is unacceptable*. This is appropriate in that it is cast at the right level of specificity and generality so that it applies to both the passage and the credited response—here, (D)—yet does not apply to any noncredited response. If it were more specific, it might not do the former; if it were more general, it might not do the latter. There is fairly good indirect evidence that test takers perceive such fallacies in the manner indicated, and so, that such patterns belong to the proper representation of phenomenological argumentative structure. For example, on the LSAT for the period June 1991 to June 1997, pretest questions of type (M) with a term such as “flawed” included in the stem were statistically rejected at a rate of 10.7%, which is not particularly high considering that pretest statistical rejection rates for the other question subtypes in Logical Reasoning sections ranged from 2.1% to 12.3% (source: Law School Admission Council statistical databases; a total of 3,312 Logical Reasoning questions were pretested).

A third indication of the substance of the present approach is that it *helps* to explain the lack of success of some intended measures of the ability of mature yet untrained (in logic) reasoners to match argumentative structure, such as:

Example (5)

Professor X: The predictions made by professional economists concerning future economic conditions have not proved to be accurate and reliable, so despite the many contributions they make in keeping track of the economy, professional economists have only a limited understanding of the complicated causal structures that determine economic outcomes. For if one is unable to make accurate and reliable predictions about some subject area, one's understanding of the forces involved is probably quite limited.

Which one of the following arguments uses a pattern of reasoning that is most similar to that used in Professor X's argument?

- (A) Economists have a limited understanding of the causes of economic events, so their long-term predictions are not reliable. As a result, their main contributions probably consist in keeping track of how the economy is doing.
- (B)* Some students do not find advanced mathematics easy to master, so they will not pursue the study of mathematics beyond its more elementary phases. For if a person does not find a subject easy to learn, he or she will probably not pursue the study of it.
- (C) Predictions made by astrologers only seem to be reliable, so astrologers do not really know what is going to happen in the future, despite the fact that many people take their predictions quite seriously. For the predictions astrologers make probably seem to be reliable only because they are very general and vague.
- (D) Astrologers make predictions about future events in which people have a keen interest, so they are likely to be believed by many people, despite the fact that their predictions are not very reliable. For it is easy to fool people when their emotions become involved.
- (E) Astronomers make accurate predictions about phenomena such as eclipses and the appearance of comets, so they must understand the causes of such phenomena. For if one understands the causes of a range of phenomena, one will probably be able to make accurate predictions about those phenomena.

The psychometric statistical characteristics of this question, pretested on the LSAT, were very bad. A relatively straightforward indication of this is what is called a "fifths" table (source: Law School Admission Council statistical databases):

Fifths table for example (5)

Option	Bottom 5th	4th 5th	3rd 5th	2nd 5th	Top 5th	Total
(A)	80	77	76	61	52	346
(B)*	135	113	104	109	116	577
(C)	179	190	174	164	85	792
(D)	77	69	50	39	47	282
(E)	107	149	199	231	309	995
Omit	44	24	19	18	13	118
Total	622	622	622	622	622	3,110

*Denotes the credited option.

The 3,110 examinees who took this question are divided into five groups ("fifths") based on their performance on the two scored Logical Reasoning sections (which comprise a total of about 50 questions). The columns in the fifths table show how many of each fifth chose the various answer options (e.g., in the bottom fifth, 80 test takers chose (A)). As judged by this fifths table, the question would be a fairly good one if (E) were the credited response—but (B) is. For instance, of the test takers in the top fifth, a full 50% chose (E), whereas only 19% chose (B). Both the passage and (B) exhibit the simplified nondeductive *reasoning structure*: *if p then probably q, p, therefore q*. Option (E)'s major (conditional) premise has the same structure at this level of analysis as that in the passage and (B), yet with respect to this premise (E) exhibits an informal variant of the fallacy of affirming the consequent. However, in *conditionalized* form (in its major premise) option (E)

embodies the reasoning pattern *s understands the causes of x, therefore probably s can make true predictions about x*. Surely, this is a common reasoning pattern. The only other argument in the test item that has the general elements of this pattern is the passage, where the conditional is a probabilistic contrapositive of the conditional in (E). (The corresponding conditional of the first sentence in option (A) differs in that it is not general, not probabilistic, and is the fallacious reversal of the conditional in the passage.) Again the point to make is that embedded structure that has nothing to do with a given argument's cogency (here, the passage and (E)) nevertheless must be taken into account in determining that argument's phenomenological argumentative structure.

References

- Adler, J.E. (1994). Fallacies and alternative interpretations. *Australasian Journal of Philosophy*, 72, 271-83.
- Baum, R. (1981). *Logic* (2nd ed.). New York: Holt, Rinehart and Winston.
- Ben-Yami, H. (1996). Attributive adjectives and the predicate calculus. *Philosophical Studies*, 83, 277-89.
- Berg, J. (1987). Interpreting arguments. *Informal Logic*, 9, 13-21.
- Brinton, A. (1995). Analysis of argument strategies of attack and cooption: Stock cases, formalization, and argument reconstruction. *Informal Logic*, 17, 249-58.
- Cox, R. & Brna, P. (1995). Supporting the use of external representations in problem solving: The need for flexible learning environments. *Journal of Artificial Intelligence in Education*, 6, 239-302.
- Govier, T. (1995). Reply to Massey. In H.V. Hansen & R.C. Pinto (Eds.), *Fallacies: Classical and contemporary readings* (pp. 172-80). University Park: Pennsylvania State University Press.
- Gulliksen, H. (1987). *Theory of mental tests*. Hillsdale, NJ: Lawrence Erlbaum.
- Johnson, R.H. (1989). Massey on fallacy and informal logic: A reply. *Synthese*, 80, 407-26.
- Haack, S. (1978). *Philosophy of logics*. Cambridge, UK: Cambridge University Press.
- Lambert, K. & Ulrich, W. (1980). *The nature of argument*. New York: Macmillan.
- Luebke, S. (1995). Informal logic issues in a practical testing context. In F.H. van Eemeren, R. Grootendorst, J.A. Blair, & C. A. Willard (Eds.), *Proceedings of the Third ISSA Conference on Argumentation Vol. IV* (pp. 31-41). Amsterdam: International Centre for the Study of Argumentation.
- Massey, G. (1975). Are there any good arguments that bad arguments are bad? *Philosophy in Context*, 4, 61-77.
- Massey, G. (1995). The fallacy behind fallacies. In H.V. Hansen & R.C. Pinto (Eds.), *Fallacies: Classical and contemporary readings* (pp. 159-71). University Park: Pennsylvania State University Press.
- Parsons, T. (1996). What is an argument? *Journal of Philosophy*, 93, 164-85.
- Quine, W.V.O. (1960). *Word and object*. Cambridge, MA: The M.I.T. Press.
- Read, S. (1994). Formal and material consequence. *Journal of Philosophical Logic*, 23, 248-65.
- Roussos, L.A. & Norton, L.L. (1999). *LSAT item-type validity study* (Technical Report 98-01). Newtown, PA: Law School Admission Council.
- Sainsbury, M. (1991). *Logical forms*. Oxford: Basil Blackwell.
- Salmon, W.C. (1973). *Logic* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Slaney, J. (1990). A general logic. *Australasian Journal of Philosophy*, 68, 74-88.
- Walton, D. (1995). *A pragmatic theory of fallacy*. Tuscaloosa: University of Alabama Press.



U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)



NOTICE

Reproduction Basis

X

This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").