BRAIN AND COGNITION 5, 41–66 (1986)

On Drawing Inferences about the Structure of Normal Cognitive Systems from the Analysis of Patterns of Impaired Performance: The Case for Single-Patient Studies

ALFONSO CARAMAZZA

The Johns Hopkins University

An analysis of the logic of valid inferences about the structure of normal cognitive processes from the study of impaired cognitive performance in brain-damaged patients is presented. The logic of inferences from group studies and single-case studies is compared. It is shown that given certain assumptions, only the single-case method allows valid inferences about the structure of cognitive systems from the analysis of impaired performance. It is also argued that although the single-case approach is not entirely problem-free, the difficulties encountered are relatively minor.

In this paper, I consider some fundamental issues of method in Cognitive Neuropsychology. By method, I mean, broadly, the criteria of adequacy for relating data to theory, that is, the rational basis for supposing that an explanatory account can validly be applied to the phenomena of interest. It is well known that the literature on the philosophy of science is full of unresolved and acrimonious debates on what constitutes an explanation and the conditions to be satisfied for the formulation of a valid explanatory account.[1] And yet, scientists go on with their work

The research reported here was supported in part by NIH Grant NS 14099 to The Johns Hopkins University. A slightly modified version of this paper was read at the III European Workshop on Cognitive Neuropsychology, Bressanone, Italy, January 20–25, 1985. In the development of the ideas contained in this paper, I have benefited from the animated but friendly criticisms of my colleagues Bill Badecker, Neal Cohen, Howard Egeth, Gary Hatfield, Mike McCloskey, Gabriele Miceli, and Domenico Parisi; their insistence that I get the question right and that I answer the right question has been invaluable. I am sure they have saved me from many errors. I thank them for their help. I also thank Kathy Sporney for her help in the preparation of this manuscript. Requests for reprints should be addressed to Alfonso Caramazza, Department of Psychology, The Johns Hopkins University, Baltimore, MD 21218.

[1] A *caveat*. I have not always been careful to explicitly distinguish the technical (in the philosophy of science) from the ordinary language use of such terms as "*a priori*," "valid,"

0278-2626/86 $3.00

unperturbed and perhaps even unaware of the complex problems that they are asked to solve daily—In what way and in virtue of what do the observations collected inform the theories designed to explain the phenomena of interest? The fact that scientists, for the most part, ignore these thorny philosophical issues does not appear to have hampered scientific progress. Perhaps it is best not to ponder too deeply issues of method—let's get on with our work and all will turn out fine in the long run. However, I believe that ignoring issues of method exacts its price, sometimes one we can ill afford. Consider J. E. Gordon's account, in his *Structures, or Why Things Don't Fall Down,* of the medieval mason's achievements in building the impressive cathedrals we so much admire today. He writes:

> On the face of it, it would seem obvious that the medieval masons knew a great deal about how to build churches and cathedrals, and of course they were often highly successful and superbly good at it. . . .
>
> Naturally, the buildings we see and admire are those which have survived: in spite of . . . their skill and experience, the medieval masons were by no means always successful. A fair proportion of their more ambitious efforts fell down soon after they were built, or sometimes during construction. However, these catastrophes were just as likely to be regarded as sent from Heaven, to punish the unrighteous or to bring sinners to repentance, as to be the consequence of mere technical ignorance. . . . So long as there was no scientific way of predicting the safety of technological structures, attempts to make devices which were new or radically different were only too likely to end in disaster." (pp. 26–27)

I am afraid that, *mutatis mutandis,* we can say the same thing for scientific practice. So, although like most of you I am much more at home doing science than reflecting on complex problems of method, I think that Cognitive Neuropsychology might not progress into a mature discipline without adequate consideration of method.

In this paper, I argue that serious consideration of this issue commits us to the conclusion that valid inferences about the structure of cognitive systems from the analysis of the performance of brain-damaged patients are only possible for observations of individual patients' performances. My comments are organized as follows. First, I describe briefly what I take to be the proper form of an explanatory account in cognitive science. I then present in as neutral a form as possible the structure of valid arguments for relating theory and evidence in the case of normal subjects and brain-damaged patients. In this section, I show that given certain assumptions, the single-case approach in the analysis of cognitive impairments allows valid inferences about the structure of normal cognitive systems. In the next section, I consider the form of a possibly valid

---

and "inference." I do believe, however, that the contexts in which these terms are used are sufficient to indicate the intended sense.

argument for relating patient–group data to cognitive models and conclude that this is not possible for the type of explanatory accounts we are seeking in Cognitive Neuropsychology. In this section, I provide a critique of classificatory schemes in cognitive neuropsychology. I conclude with a discussion of various objections that have been raised against the single-case approach and show that where these objections are coherently formulated, they do not pose insurmountable problems. In a *coda,* I discuss some practical issues concerning the role of single and double dissociations and co-occurrence of symptoms in developing and testing specific hypotheses about the structure of cognitive systems.

1. Explanation in Cognitive Neuropsychology

The critical issue to be addressed concerns the nature of the evidence that bears on questions of interest. We cannot begin to deal with this issue without having at least a relatively clear idea of the kind of questions we are trying to answer and, conversely, without having specified the general form of answers we are to consider appropriate for these questions.

There certainly is agreement on what we consider to be the kinds of questions we wish to answer. These are questions concerning how we perform such tasks as understanding sentences, adding numbers, writing words, drawing pictures, and so forth. I shall assume, without supporting argument, that the kind of answers we are to consider appropriate consist of computationally explicit, information-processing accounts of cognitive performance. In other words, our objective is to specify functional architectures which permit computationally explicit accounts of the input/output relationship for specific cognitive systems. The degree of detail of a proposed functional architecture is determined by the range of input/output pairs we take into consideration, or, more generally, by the range of evidence we deem relevant to the formulation of our explanatory account. However, what will count as relevant evidence is not independent of the kind of theoretical account we are formulating. In other words, there is a complex, mutual interdependence between theory and evidence.

Consider, as an illustrative example of the kinds of explanatory accounts we are seeking and of the complex interdependence of evidence and theory, how we might proceed in developing a model of single-word oral reading. What kind of evidence should we bring to bear on this issue? Obviously, the centerpiece of the evidential basis should be the pattern of reading performance—the structure of the graphemic input/phonological output relationship in normal reading. The objective is to propose a functional architecture which specifies in computationally explicit terms the relationship between *any* particular word input and the produced, appropriate phonological response. Although we can go some way in articulating a model of reading with just this data base (i.e., normal

performance on various words), it is relatively unconstraining. These data allow at least two distinct functional architectures (F.A.) as depicted in Fig. 1. F.A.1 postulates at least three distinct processing components related sequentially as indicated; F.A.2 also postulates three distinct components, but the relationship among components takes a different form. More importantly for our purposes, however, the data base considered thus far does not allow significant insights into the structural organization of the hypothesized components in the two functional architectures. Indeed, it is not clear that we would want to take the schematic representations in Fig. 1 as constituting an explanatory account of even the relatively impoverished data base we have considered thus far. A minimal requirement to be satisfied by any explanatory account of the sort considered here is that it make substantive, nontrivial claims about the nature of the input/output relationship and associated computational structure for each of the postulated components (i.e., G, S, and P) in a proposed functional architecture. Thus, we would only take F.A.1 and F.A.2 as constituting explanatory accounts of oral reading if an explicit formulation were to be provided for the processing burden carried by the components labeled "Graphemic Input Lexicon," "Lexical Semantic Component," and "Phonological Output Lexicon" within each of the proposed functional architectures.

Suppose that we were to expand the data base to include oral reading of orthographically regular nonwords. Would these data be relevant to the construction of a model of word reading? An affirmative or negative answer to this question depends on the type of functional architecture proposed. If we were to entertain a dual-route model of reading—i.e., postulate distinct reading processes for words and nonwords—the answer

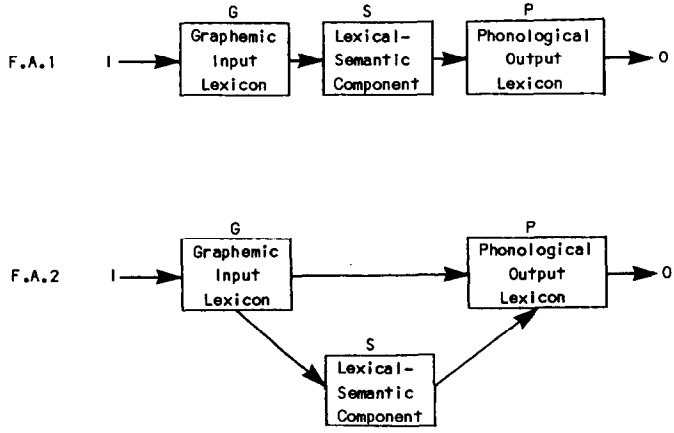

FIGURE 1

would most likely be no, although it would allow us to articulate the functional architecture of oral reading as in Fig. 2 where F.A.1.1 and F.A.2.1 now include a separate processing component for converting nonlexical graphemic representations into phonological representations. By contrast, if we were to adopt a single-route model of reading, the mere fact that we can read nonwords would be relevant to the organization of the word reading system. First, it would exclude as possible models of the reading system functional architectures of class F.A.1. In addition, the nature of the nonword input/phonological output relationship would be informative with respect to the internal functioning of at least two of the hypothesized components of processing in functional architectures of class F.A.2—the computational structure of components G and P must be of such a nature as to allow input/output mappings not only for words but also for nonwords. This latter requirement severely restricts the type of computational structures for components G and P.

Suppose that we were to expand still further the evidential base to include reaction time (RT) data for oral word reading. Are these data

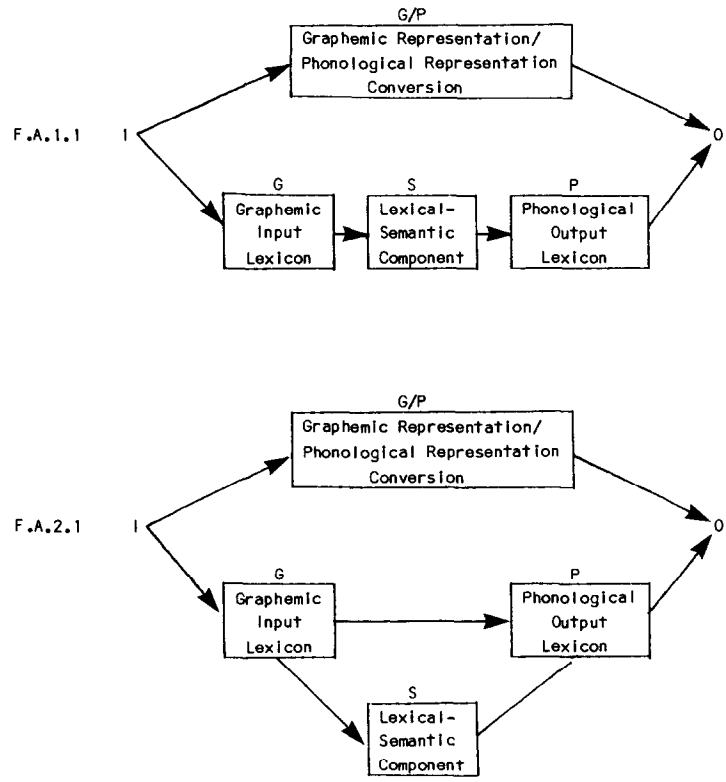

FIGURE 2

relevant to the development of a model of reading and, if so, in what way? First of all, we should emphasize that it is not at all obvious that we ought to answer the first part of our question affirmatively even though that is precisely what we normally do. The reason we assume that oral reading RT data are relevant to the formulation of a model of reading is that we expect that the structure of these data will reflect the influence of putatively relevant theoretical factors. But, it might not be so—oral reading RT could be randomly related to all theoretically relevant factors in the reading process. In this specific instance, oral reading RT data *is* affected by such lexical factors as word frequency and possibly morphological structure, providing the pretheoretical justification for considering these data as potentially relevant to an explanatory account of the reading process. Ultimately, however, the evidential role of these data for a model of reading depends crucially on our ability to articulate in detail some argument that establishes an explicit relationship between the processing structure of a component (or components) of the proposed functional architecture and oral reading RT performance—that is, we must make explicit the hypotheses that justify the use of an RT measure to inform and constrain a model of oral reading. To the extent that we succeed in making explicit these auxiliary hypotheses of cognitive performance, we are justified in making use of patterns of oral reading RTs as part of the evidential basis for confirming hypotheses about the structure of components in a proposed functional architecture. Thus, for example, the fact that word frequency affects oral reading RT performance could be taken to support some hypothesis about the mechanisms of access to the graphemic input lexicon.

There are obviously many other kinds of evidence that could be relevant to the development of a model of oral reading. Thus, for example, performance in reading visually degraded stimuli or RT performance in deciding whether or not a string of letters forms a word could be brought to bear on specific hypotheses about the structure of components of a proposed functional architecture of the reading system, provided that we make explicit the auxiliary hypotheses that link performance on these tasks to relevant components of the reading process. But, there are also observations that we would not credit with providing an evidential basis for a model of reading even when the performance implicates directly the reading system. Thus, for example, if a subject, because of some personal aesthetic proclivity, were to read some words in a whisper and others in a loud voice, this would not be taken as being relevant to any significant hypothesis about the structure of components of the reading system.

In this section, I have dealt with two issues. The first concerns the structure of explanatory accounts in cognitive science. I have assumed that cognitive processes are to be considered as the functioning of in-

formation processing systems. Adoption of this position implicates two interdependent levels of analysis for cognitive systems—the formulation of a functional architecture which specifies the componential structure of a cognitive system and the specification of the computational structure of the individual components of processing that comprise the functional architecture of the system.

The second issue I have dealt with, rather obliquely to be sure, has been the complex relationship between evidence and theory. A major point to be emphasized is that what is considered to be relevant evidence is not theory-independent: The type of explanatory framework we adopt partially determines what will count as the relevant phenomena to be explicated by our theories; observations do not carry on their sleeves signs indicating whether or not they constitute relevant evidence in some domain of investigation. An especially important point is that a specific set of observations (e.g., oral reading RT data) will assume evidential status with respect to some model only if we are able to provide adequate arguments (what I have called auxiliary hypotheses) to explicitly link the type of observations in question to the component or components of processing being investigated. In other words, a set of observations is relevant to a model only if the model is elaborated in sufficient detail to make clear the relationship between the data and the components of the model for which it is putatively relevant.

I have dwelt so long on these issues, at the risk of pedantry, because I want to emphasize that what ultimately gets considered as relevant evidence in a field of study depends crucially on the type of explanatory account we are trying to formulate. A serious discussion of method in Cognitive Neuropsychology is only possible in the context of an explicit recognition of the type of questions we are trying to answer and the type of answers we will deem appropriate to these questions.

I turn now to a discussion of the structure of valid arguments for relating evidence to theory. The quasi-formal characterization of these arguments is intended to be relatively neutral (if that is possible) with respect to current debates on the nature of explanation.

II. On Relevant Evidence for a Model

I have argued that evidence is in a complex relationship to theory, involving, as it does, intermediate hypotheses used to motivate the relevance of a particular type of observation for the model being evaluated. This complex relationship can be summarized in quasi-formal terms as follows:

> *(1) A set of observations E_i obtained under initial condition c_1, c_2, c_3 . . . c_n constitutes relevant evidence in support of hypothesis H and related auxiliary hypotheses a_1, a_2. . . a_n in model* M *provided that E_i can be derived computationally from H in* M *and the specified initial conditions* (see Fig. 3).

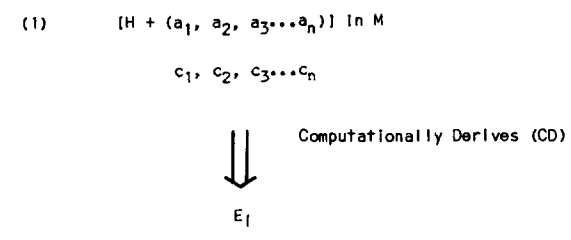

FIGURE 3

This rather opaque statement may be rendered more transparent by instantiating the variables *E*, *c*, and so forth. Thus, consider again the example on the use of oral reading RTs to address particular hypotheses about the structure of the reading process. The set of observations E_i would correspond to a pattern of RT results; the initial conditions c_1 through c_n would correspond to some lexical factor manipulated in the experiment (e.g., morphologically simple vs. morphologically complex words) and other experimental conditions (e.g., tachistoscopic exposure, speeded responses, etc.); the hypothesis H would correspond to some hypothesis about the graphemic input lexicon such as, for example, that access requires morphological decomposition; model M would be an explicitly articulated functional architecture of which the Graphemic Input Lexicon is a component; and the auxiliary hypotheses a_1 through a_n would correspond to various assumptions we would have to make to relate RT measures to the hypothesis of interest. So, for example, the obtained RT results might constitute support for the hypothesis tested if response latencies were to be longer (or shorter) for morphologically complex words than morphologically simple words when all other relevant factors (e.g., frequency) are taken into consideration.

In this account of the logic for relating data to models, I have used the rather cumbersome locution ". .support hypothesis H and related auxiliary hypothesis a_1, a_2, a_3. . .a_n in model M. . ." because that is, in fact, the way we relate data to models. However, for simplicity of exposition, I will reduce this complex phrase to just M with the proviso that it be understood in its expanded form. Thus, the modified schematic representation of the structure of valid arguments for relating data to models will be represented as in Fig. 4.

There are a number of obvious but important points that should be noted about this schema for relating evidence to models. Here I focus on three points that play an important role in subsequent discussion. One point to note is that if the computationally derived E_i does not occur, or if an E_i occurs that cannot at present be computationally derived, it does not mean that we must reject M in its totality—H could be false or some of the auxiliary hypotheses could be false. We retain confidence in a particular model if, on the whole, this model does the *best* job among

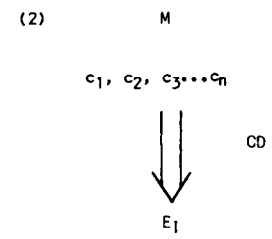

FIGURE 4

available models in explaining the phenomena of interest. Another point is that in this schema, the initial conditions are known by the investigator. That is, we know independently of our observations E_i what initial conditions we have. This situation allows us to computationally derive E_i given our model M and the initial conditions c_1 through c_n. It is important to note that in this schema there are no "unknowns" to undermine the computational derivation of E_i from M.[2] The final point to take note of concerns the nature of E_i and M in this schema. E_i is *any* relevant set of observations within the scope of M—in this case, various types of cognitive performance. M is, in our case, the functional architecture of a cognitive system, along with a computational characterization of its components, which is assumed to characterize the actual functioning of the mind/brain. A crucial feature of M is the *assumption of universality;* that is, the assumption that M is true of "normal" human mind/brains in general and, therefore, of *any* individual normal mind/brain. Of course, it is clear that we are going to have to place some restrictions on what will count as "normal human mind/brain," but it is equally clear that if we were not to accept the assumption of universality, we would negate the possibility of scientifically investigating the mind/brain.[3]

[2] Further clarification may be needed to explicate what I mean by the claim that there are "no unknowns" in the inferential schema under consideration. By this claim I mean no more than that all the factors needed to derive an E_i can be explicitly stated *in advance* of the data collection process. Of course, it could turn out *in practice* that we are mistaken about some assumption or other and that we must make the appropriate changes in light of new theoretical and empirical developments. But once we have made these changes, we can then use these modified assumptions in deriving, in appropriate contexts, *any* other E_i. Later in this paper, I propose that there are inferential schemata in which we cannot claim to "know" all the relevant factors for deriving an E_i in advance of actually obtaining E_i. This new inferential schema is one in which the derivation of E_i depends on *retrodicting* some hypothesis, L_i, on the basis of E_i and other known factors.

[3] This claim may be too strong. Thus, it might be possible to scientifically investigate some domain of natural phenomena where we could not make the assumption of universality (e.g., if there were several kinds of human minds). However, for such a program to be successful, we would at least need to be able to unambiguously distinguish among the various kinds of phenomena to be explicated. To the extent that we successfully distinguish among the various kinds of phenomena, we can proceed with a scientifically sound program of investigation. But note that it is now possible to make the assumption of universality for each of the various kinds identified in our domain of investigation.

The assumption of universality plays a critical role not only in justifying the possibility of scientifically investigating the mind/brain but also in motivating a particular practice in experimental investigations of the structure of cognitive systems. By the latter, I am referring to the practice of using group data to evaluate and develop hypotheses about the structure of cognitive systems. The justification for using the performance of groups of subjects in our experimental investigations is based on the assumption that the averaged performance of the group essentially reflects the performance of *any* individual in the reference population from which the group was drawn. Thus, any conclusions arrived at for the group of subjects tested will be assumed to be true of all individuals in the reference population. This argument is only valid if the assumption of universality is true. (Of course, there is another assumption that is made in motivating the use of group data to test hypotheses; namely, whatever the individual processing differences, if any, among members of a group, these are theoretically irrelevant and randomly related to the "modal" processing structure of the cognitive system being investigated. The force of this assumption is to motivate the expectation that individual differences will tend to cancel each other out in group performance. I will return to the problem of individual differences in Section IV.)

Let us make sure that it is clear what it is we get through the assumption of universality. First and most importantly, we are allowed to consider all the *E*'s obtained in any "well-designed" experiment to constitute potentially relevant evidence for M for a given reference population. Second, we are allowed to consider the averaged performance of a group of subjects as representative of any individual in the reference population. Of course, in practice the problem of identifying the reference population and related M is not unproblematic. Thus, consider once again as an example the case in which we are trying to formulate a model of reading, and let us suppose that we had good reasons for believing that the functional architecture of the reading system contained a Graphemic Input Lexicon which was addressed through a morphological decomposition procedure. Let us also assume that the empirical evidence on which we originally based this belief consisted of experimental investigations carried out with English-speaking subjects. The problem we now have is, "what is this model a model of?" Well, obviously it is a model of the cognitive structure of the reading process of English speakers. However, we would also be justified, barring plausible counterevidence, in concluding that the proposed functional architecture characterizes the reading process of speakers of any language containing inflected or morphologically derived words. Thus, we could generalize our conclusions to speakers of Italian, French, Hebrew, and so on, but not to speakers of Chinese or Vietnamese, which are languages that contain only monomorphemic words. This example illustrates an important point; namely, not only is the question of

what observations constitute relevant evidence theory-dependent, but also, the definition of a reference population is theory dependent. Thus, ignoring various practical and theoretical considerations, while in principle it would be justified to average the performance of English and Italian speakers in an experiment on reading, it would not cross the mind of any reasonable investigator to average the performance of English and Chinese speakers in this experiment. The basis for this conclusion is that we have reasonable grounds for believing that the observations obtained with English and Italian speakers are homogeneous with respect to the model being investigated while the observations obtained with English and Chinese speakers violate this assumption of homogeneity with respect to the cognitive system under consideration.

I trust that my comments thus far are a relatively accurate rendition of the general metatheoretical and methodological principles that guide our work as cognitive scientists. Assuming it to be so, I propose that these principles can be naturally extended to cover the case in which the cognitive performance of brain-damaged patients is used to infer the structure of normal cognitive systems.

Again, as in the case of normal performance, the relationship between evidence from the performance of brain-damaged patients and normal cognitive systems is a complex one. In the present case, the relationship is further complicated by the presence of a nonindependent hypothesis that must be included in the inferential schema for us to draw valid inferences about the structure of normal cognitive systems from patterns of impaired performance. We can summarize the evidential role of the performance of brain-damaged patients for models of normal cognitive systems as follows:

(2) *A set of observations* E'_i, *that is, the cognitive performance of a brain-damaged patient, is evidence for* M *just in case it is possible to computationally derive* E'_i *from* M, L_i, *and initial conditions* c_1 *through* c_n; *where* L_i *is a complex hypothesis about the locus of damage to a functional architecture and associated assumptions concerning the effects of the hypothesized damage on the cognitive system* (see Fig. 5).

Thus, for example, L_i could be the hypothesis that a patient has damage to the graphemic input lexicon, and the effects of this damage are that the cognitive system computes phonological representations for words not by the normal process of addressing preassembled, lexical–phonological representations, but by a nonlexical procedure that converts submorphemic graphemic representations into phonological representations. In this example, the patient's reading performance should be directly explicable in terms of the computational structure of component G/P in F.A.1.1 or F.A.2.1., and, if so, we can take this pattern of performance as evidence in support of the proposed functional architectures.

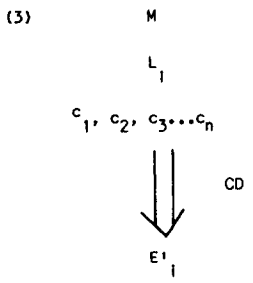

FIGURE 5

There are two important features of this schema for relating data to theory that I would like to focus on. One of these two I will only discuss briefly here because I assume it to be relatively uncontroversial among Cognitive Neuropsychologists. You have noticed that I have chosen to represent the inferential schema for cognitive performance of brain-damaged patients as in (3) as opposed to (4) (Fig. 6).

This difference in schematism is supposed to capture what I have called elsewhere the *transparency assumption*. This assumption essentially says that the cognitive system of a brain-damaged patient is fundamentally the same as that of a normal subject except for a "local" modification of the system represented by L_i in our inferential schema (3). This assumption rejects the possibility that brain damage results in the *de novo* creation of cognitive operations resulting in a cognitive system M′ (schema (4)) that has a nontransparent relationship to M. Instead, I assume, as shown in schema (3), that the cognitive performance of brain-damaged patients can be directly related through L_i to M—the cognitive system of the normal human mind/brain. The transparency assumption must be accepted if we are to use the performance of brain-damaged patients to inform and constrain theories of normal cognitive processing.[4]

The second and more controversial issue I want to discuss concerns a specific consequence of the proposed inferential schema for drawing conclusions about normal cognitive systems from the analysis of impaired cognitive performance. Let us consider once again the inferential schemas for normal and impaired cognitive performance—schemata (2) and (3),

[4] My formulation of the transparency assumption implies that E_i can only be related to M when the damage to the system is "local." This assumption may be too strong as an *in principle* claim—nonlocal, very general modifications of the system may still allow the possibility of relating E_i to M. However, *in practice*, given the tremendous complexity of the systems we are dealing with, it may *only* be possible to draw meaningful conclusions from impaired performance to normal cognitive systems under a restricted sort of condition. It is certainly the case that the identification of L_i becomes increasingly problematic the more general and extensive the damage to the cognitive systems.

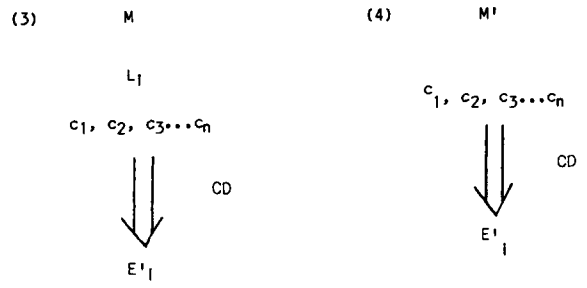

FIGURE 6

respectively (see Fig. 7). The critical difference between these two schemata is the presence of L_i in schema (3). Thus, whereas in schema (2) E_i is related to M through known initial conditions c_1 through c_n, E'_i in schema (3) is related to M through initial conditions c_1 through c_n *and* L_i. Now, an important characteristic of L_i is that it is not known *a priori* but must itself be posited to computationally derive E'_i. In other words, L_i has to be inferred from E'_i, M, and initial conditions c_1 through c_n. To the extent to which we can successfully specify some L_i given E'_i, we have support for model M. A consequence of adopting this inferential schema is that we can validly consider the set of observations E_i as evidential support for M *only* for single-case studies. The reason for this conclusion is straightforward and compelling.

It will be recalled that in experimental investigations with normal subjects, the assumption of universality is used to motivate the use of group data to draw inferences about M. Thus, in schema (2) we assume that M is true of *any* individual drawn from a reference population, and given that initial conditions c_1 through c_n are constant for any group of individuals, our expectations are that the observed individual performances E_1, E_2, E_3. . . E_n that are averaged to produce E_i, are homogeneous—i.e., we assume that the expected observations for individual subjects will be equivalent in theoretically relevant respects. We cannot construct a parallel

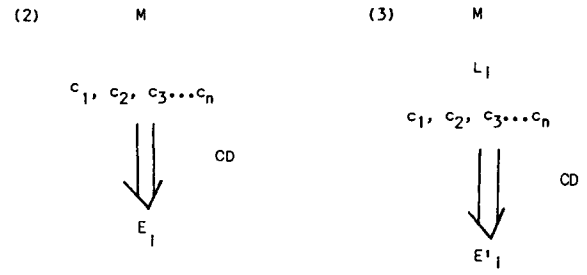

FIGURE 7

argument for brain-damaged patients. Although we can make the same assumptions about M and known initial conditions in investigations with brain-damaged patients as in the case of normal subjects, there remains the fact that we cannot assume the form of L_i, which is needed to justify the assumption of homogeneity for patient groups, independently of E'_i. Thus, consider the inferential schemata for individual patients P_1 through P_n shown in Fig. 8.

The expected pattern of performance, E_i, for individual patients, under constant initial conditions, is a function of the nature of damage, L_i, that characterizes each patient. We would be justified in averaging the performance of a group of subjects *only* if we could assume that the nature of damage, L_i, to a particular cognitive system in each patient is identical in *all theoretically relevant respects*—that is, *only* if $L_1 = L_2 = \ldots L_i = \ldots L_n$, in which case we would expect the pattern of performance for each patient, E_i, to be homogeneous with respect to M. Notice, however, that a determination of whether $L_1 = L_2 = \ldots L_i = \ldots L_n$ can only be made *a posteriori*, that is, once we have E'_i. Now, my contention is that the inferential schema for relating brain-damaged patients' performance to the normal cognitive system is only valid if we are able to specify some L_i which allows us to derive the observed pattern of performance E'_i from M and specified initial condition. But, since we cannot specify L_i independently of E'_i, this is equivalent to saying that there are no sufficient, *a priori* conditions that can be specified which would allow us to assume the theoretically relevant identity of L_i's for a group of brain-damaged patients. Without the assumption of identical

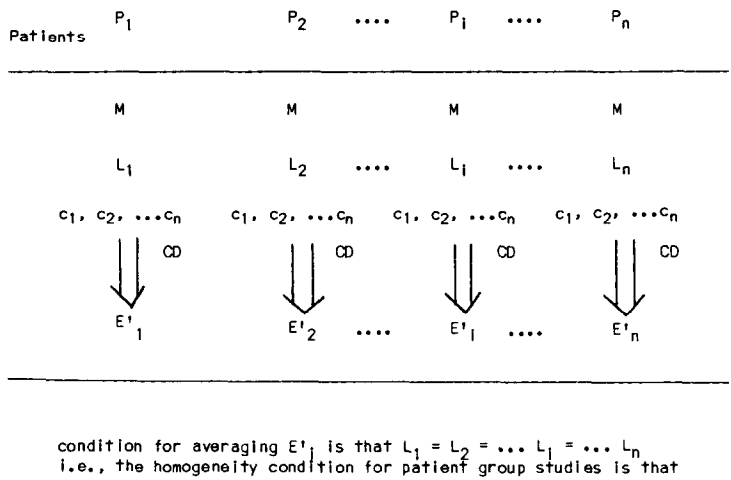

FIGURE 8

L_i's in theoretically relevant respects, the grouping of patients' performance results in meaningless entities, and we must conclude, therefore, that we cannot construct a valid inferential schema for relating grouped observations from brain-damaged patients to models of normal cognitive systems. (Of course, it could be argued that there are means for devising situations in which the assumption of homogeneity for groups of brain-damaged patients is justified. Thus, it might be proposed that grouping patients by syndrome type would satisfy the assumption of homogeneity. I will show in the next section that the logic of patient classification is such that there is no *a priori* classification schema that can be used to justify the assumption of homogeneity and allow valid inferences from patient group data.)

In this section I have shown that we can draw valid inferences from the analysis of the performance of brain-damaged patients to the structure of normal cognitive systems, but that such inferences are only valid for single case studies and not for group studies. However, my comments are open to the charge that given the latitude I have permitted myself concerning the kinds of questions that are worth asking, the kind of answers we deem appropriate, and the type of inferential schema we consider valid, it is relatively trivial to debunk the group study approach in Cognitive Neuropsychology. To this charge, I plead guilty. My only defense is that I don't know of any other way of dealing with issues of method without specifying the type of explanatory account we are to consider appropriate. And surely my characterization of the accepted explanatory goals of Cognitive Neuropsychology is not that inaccurate. But it is true that I have presented a characterization of the kinds of experimental questions that can be reasonably addressed in Cognitive Neuropsychology which may not be shared by all cognitive neuropsychologists. Thus, I have not even considered questions of the type, "Is it the case that patients of type *R* also manifest property *y*?" where *R* is a clinical category and *y* is some pattern of performance. Perhaps if our objective were to answer questions of this type, we would find it necessary to do group study research.

III. Patient Classification Schemes and Group Studies

Is it possible to relax the strict criterion of homogeneity adopted in the preceding section and still draw meaningful conclusions from group studies? What if we adopted a weaker criterion of homogeneity which only required of these studies that the patients in a group satisfy the criterion of being members of some antecedently specified reference class—e.g., agrammatics, deep dyslexics, phonological dysgraphics, and so on. What kinds of questions could we address with this modified criterion of homogeneity, and what kind of inferential schema is appropriate for this approach in Cognitive Neuropsychology?

It seems to me that only "statistical" questions can be addressed in this paradigm, that is, questions of the type, "What is the probability that patients of type R manifest property y?" and not questions of the type, "Is it the case that patients of type R always manifest property y?"

Let us consider first questions of the type, "What is the probability that patients of type R manifest property y," where $p < 1.0$. Questions of this type would only be meaningful if we were to believe that the phenomena of interest to be explicated were of an *intrinsically probabilistic* nature, that is, if we believed that answers of the type, "the probability that an agrammatic patient also presents with asyntactic comprehension is 0.9," or, "the probability that a patient who makes semantic paralexic errors also presents with a form class effect is 0.8," and so on, were the basic evidential statements to be explicated by our theories of cognitive functioning. The explanatory strategy for this type of question is fairly straightforward. We start with an initial reference class, A (e.g., brain-damaged patients), and partition it into subclasses by invoking statistically relevant factors, C_i (e.g., omission of grammatical markers, production of semantic paralexias, and so forth), such that the resulting subclasses, $A \cdot C_j (= R_i)$, will be homogeneous with respect to some set of relevant phenomena to be explained. However, for this strategy to be applied meaningfully, we must assume that no statistically relevant partitions are possible for the identified reference subclasses. This requirement of homogeneity is needed because we do not want to account for just random statistical occurrences but for *intrinsically probabilistic* statements about natural phenomena—e.g., that for an atom of carbon 14 the probability of decay in 5730 years is 0.5, and, there is no statistically relevant partition of carbon 14 that would alter the posterior probabilities of whether an atom will decay in the given time frame. That is, given

P(decay in 5730 years | carbon 14 atom) = $\frac{1}{2}$,
there is no C_i such that
P(decay in 5730 years | carbon 14 atom $\cdot$ C_i) = P_i,
for all i will result in $p \neq 0.5$.

In effect, then, this approach too requires that grouping of data satisfy a criterion of homogeneity. However, in this case, even when the reference class is perfectly homogeneous, we will only obtain probabilistic statements.

There are two related points I would like to make about statistical questions in Cognitive Neuropsychology. First, even if we were to grant that such questions are meaningful, we would have to provide independent theoretical justification for the selection of the reference categories for which we wanted to determine probabilistic relationships to particular events. That is, we would have to theoretically motivate the decision to consider specific patient groupings (agrammatics, deep dyslexics, etc.) as theoretically basic. Second, if we were to consider statistical statements

to be the basic evidential material to be explicated, our theories of cognitive functioning would have to be developed in such a way as to explicate the obtained probabilistic statements. However, I not only do not know of any work which is explicitly concerned with determining the probabilistic (less than 1.0) relationship between a category type and some event, but I am equally unfamiliar with any theoretically coherent account of the structure of cognitive systems that purports to explicate probabilistic statements of the sort we have considered here.

In short, then, if we were to be interested in answers to statistical questions of the type "What is the probability that patients of type *R* manifest property *y*?", the patient–group study method would be appropriate for these questions. However, my contention is that questions of this type are not relevant to the theoretical concern of determining the functional architecture of a cognitive system. (We are not interested in determining the specific probability values for an event given some antecedently specified category, although I do not rule out the possibility that these kinds of statements could be of some value in addressing theoretical concerns having to do with the distribution of cognitive systems in the brain.) Rather, if there is to be a meaningful question concerning the relationship between an antecedently specified patient type and some pattern of performance, it will have to be of the form "Is it the case that patients of type *R* also *necessarily* manifest property *y*?" If my formulation of the kind of question we want to answer is correct, then group studies cannot be used to answer questions of this type. Answers to this type of question require that *y* be true of *any R*, which entails a case-by-case analysis! (See Caramazza (1984) and Badecker & Caramazza (1985) for detailed criticisms of cognitive neuropsychological research based on patient classification.) In other words, we must determine that for any individual patient of type *R*, it is true that the patient will manifest property *y*. I elaborate on this claim below.

Suppose we start with two patients both classified as being of type *R* (which we take to reflect a disruption at some specified level of a cognitive system), and we are interested in determining whether they will perform in particular form *y* on some task. There are several possible outcomes; I will consider only two here. Both patients could present with performance *y*. If *y* were *necessarily* required of *R* given some hypothesis about L_i in model M, then this outcome will increase our confidence in L_i and M. But note that now the criteria for inclusion in *R* as a reflection of L_i in M will have to include *y*, and any patient who does not manifest *y* cannot be classified as *R*. Obviously, there is nothing special about the set of criteria initially chosen to define *R* since *y* will not have the same evidential status as the other factors included in the set of criteria for picking out the members of the category. Thus, consider the following example. Suppose we started with two patients classified as agrammatics—

as indexed by the omission of grammatical markers (i.e., function words and inflections) in spontaneous speech. Suppose also, for the sake of argument, that we were to take this pattern of performance to reflect a lesion to a syntactic processing device which is used both in comprehension and production of sentences. On this hypothesis about L_i, not only is it our expectation that both patients should present with asyntactic comprehension, but it must necessarily be the case that both agrammatic production and asyntactic comprehension will be present in a patient for us to be able to infer that the locus of damage in the system is at the level of the hypothesized syntactic processing device. (I ignore in this example the fact that no theoretically plausible account has been presented which explicates how a single syntactic processing system may be involved both in sentence comprehension and in sentence production.) Thus, if there is to be a category R which reflects damage to a syntactic processing device of the hypothesized sort, it must be indexed by both agrammatic production and asyntactic comprehension.

Another possible outcome is that only one of the two patients presents with performance y; the other presents with performance z. What can we conclude from this outcome? First of all, these results can be taken as counterevidence to M (or, more precisely, to some hypothesis H in M). More importantly for our present purposes, however, we are now confronted with the problem of determining what to do with our original category R—where we had a single category, we now have two categories, $R \cdot y$ and $R \cdot z$. Our model of the cognitive system being investigated must account for both patterns of performance even if the distribution of $R \cdot y$ and $R \cdot z$ were to predominantly favor the occurrence of the first over the second. Let us briefly return to our example concerning agrammatism and asyntactic comprehension, and let us suppose that one of the two agrammatic patients did present with asyntactic comprehension, but not the other. On the basis of this result, we could simply conclude that our model of language processing was wrong—that contrary to our hypothesis, there are independent syntactic processes for comprehension and production. In this case, the existence of R and non-y is taken to reflect the dissociability of the two independent syntactic processing components. But note that this conclusion does not depend on the probability of the event (0.5 in this case); we would have been forced to reach the same conclusion even if the probability of $R \cdot \sim y$ were quite small. Note also that if we were to average the comprehension performance of agrammatic patients, we would both violate the assumption of homogeneity and preclude the possibility of rejecting a false hypothesis, especially when the probability of an event such as $R \cdot \sim y$ is relatively small compared to $R \cdot y$.

What are the implications of these arguments for the possibility of drawing valid inferences from patient–group data where the criterion for

grouping performance is based on patient type? What I have shown, I believe, is that this type of grouping allows us to establish homogeneity only over the domain of performance that comprised the classification criteria (e.g., agrammatism); it does not guarantee homogeneity over the new performance being investigated (e.g., asyntactic comprehension). But, valid inferences from the grouped performance to a proposed model of cognitive functioning are only valid if we can also establish homogeneity over this new domain of performance. This criterion of homogeneity can only be satisfied *a posteriori*—that is, on the basis of the experimentally relevant performance of individual patients. Thus, the homogeneity assumption for patient–group studies is satisfied only by carrying out a series of single-case studies to establish that the nature of cognitive damage is the same for each patient in a group. In other words, either we satisfy the homogeneity assumption by carrying out a series of single-case studies, which makes the concept of patient-group study vacuous, or we fail to carry out a patient-by-patient analysis, in which case we cannot draw valid inferences from patient group data because we cannot assume homogeneity over the relevant experimental observations. So, we are back to our original formulation of the valid inferential schema for drawing inferences from impaired performance to normal cognitive systems, an inferential schema that allows valid inferences only for single-case observations (or a series of single cases).

In this section, I have argued that the only kinds of questions we can ask for patient–group studies are statistical ones, but that these questions are not appropriate if our explanatory goal is that of developing theories of the structure of cognitive systems. I have also argued that even when we attempt to formulate nonstatistical questions by reference to patient categories, the relevant information we are seeking is provided by the pattern of performance of individual patients. I turn now to possible objections to the single-case study method in Cognitive Neuropsychology.

IV. On Objections to the Single-Case Study Approach in Cognitive Neuropsychology

Various objections have been raised against the single-case study methodology. Some of these are poorly thought out, gut reactions to something that does not conform with the received canons of experimental methodology in Cognitive Psychology and Neuropsychology. I will dispense with two such objections fairly quickly. However, a more serious objection has been raised concerning the problem of replication in single-case studies. This objection is not without some force, but, as I hope to show, it does not raise insurmountable difficulties.

An often-voiced objection to the single-case study methodology takes the following form: "You cannot construct a theory on the basis of a single-case study." That is indeed correct, just as one cannot construct

a theory on the basis of a single experiment with a group of subjects or patients. This objection clearly reflects a monstrous misapprehension of what single-case study methodology is all about. No one has proposed that we construct a *different* theory for each case studied, just as no one would propose that we construct a different theory for each experiment with a group of subjects or patients. What *is* being claimed is that the performance of each patient potentially provides relevant evidence for a model. Therefore, the performance of all individual patients (as well as the performance of normal subjects) must be considered in the evaluation of a proposed model of a cognitive system. The pattern of performance of a single patient can be used to propose a specific hypothesis, but ultimately, the evaluation of a model is based on the full range of available relevant evidence, including the performance of other patients and normal subjects.

Another less-than-clear objection that has been voiced against the single-case study methodology takes the following form: "You cannot generalize from single-case studies." Generalize to what? If, on the one hand, by this objection is meant that one cannot generalize to a particular patient population, then so be it. Of what use would generalizing to a patient population be in any case? None that I can think of. If, on the other hand, by this objection is meant that we cannot generalize to the normal cognitive system, then it is not clear what the basis for this objection might be. Is the objection that one can generalize to normal cognitive models from group studies but not from single-case studies? If this is the intended sense of the complaint, then surely it must be rendered clear what it is specifically about single-case studies that leaves them open to this objection.

Actually, I think that these poorly formulated objections may be on to something that is not all that uninteresting. Perhaps what is intended by the two objections raised thus far against the single-case study methodology is that the performance of individual patients is *too* specific to allow meaningful generalization to the reference population. That is, the objection could be that focusing on the performance of individual patients accentuates the evidential role that might be assigned to theoretically uninteresting, idiosyncratic performance. This objection could become a serious problem when considered in light of another objection to the single-case methodology—that of the impossibility of controlled replication in single-case studies. Let me turn, then, to this other objection, and we will discuss the problem of overspecificity of observations together with other potential difficulties engendered by the impossibility of replication.

The potentially most troubling objection to the single-case study method concerns the claim that because single-case studies do not permit replication, the approach is methodologically unsound. To better understand the force of this objection, we should consider the role of replication in

scientific investigations. Replication is deemed important because of the possibility that uncontrolled factors may have contaminated the results obtained in any single experiment. This is equivalent to saying that our initial conditions may be ill-specified. Replication allows us to determine whether a specific result is reliably related to the initial conditions we thought were present in some experiment. The important role of replication becomes apparent when we consider those cases when, in fact, we have uncontrolled factors in an experiment. In these cases the derivation of *E* from M is invalid. If we don't have a way of distinguishing these contaminated results from the body of relevant evidence, we either will not be able to develop a model that can account for the putatively relevant evidence, which will include false evidence, or we will construct a model that does what it is not supposed to do—explain false evidence.

The single-case methodology does not permit pure replications in a controlled fashion. Of course, over the course of time, we can find patients who perform in qualitatively identical respects over a theoretically relevant domain. But these "replications" only allow us to increase our confidence in a particular pattern of performance. There is no possibility for finding "replications" which allow us to *reject* a previously reported result. This situation, together with the previously cited concern over the potential accentuation of idiosyncratic performance in single-case studies (perhaps exacerbated by the sociological factor that favors the tendency to report "new" observations over already "known" results), could spell serious difficulties for this methodology. But how serious is the difficulty posed by the fact that we cannot do controlled replications in single-case studies? The problem is certainly not negligible, but neither is it devastating. In fact, with appropriate precautions, we can easily do without controlled replication in our efforts to use patterns of impaired cognitive performance to develop models of cognitive systems. This optimistic evaluation is based on the following considerations.

Consider, first of all, the evidential weight of any single experiment or case study. The evidential weight of any single result is determined by the total body of evidence available to us at any point in the course of the scientific enterprise—in general, the larger the available body of relevant evidence, the smaller the weight assigned to any single result. Of course, it is possible for a single result to have a relatively determining effect, but this would only be possible if the single result were congruent, under some explanatory account, with the vast majority of established evidence in the field.

Consider next the expected pattern of results with single-case studies. If our assumptions of universality and transparency (see Section II) are correct, then our expectation is that the pattern of results obtained should converge on a single, best theory of the structure of cognitive systems. That is, the pattern of results should generally be congruent since, by

assumption, the patterns of impaired performance reflect the functioning of a fixed, common set of cognitive systems. This does not mean that all of the performance of each patient or any of the performance of all cases should be congruent with the overall pattern of evidence in the field. Indeed, if our concern about the overspecificity of results and accentuation of idiosyncratic performance in single-case studies is valid, then we *do* expect that there will be patterns of results that do not fit the overall structure of accumulated evidence. However, our expectations about these discrepant sets of results is that they should *diverge* from each other, weakening the value of any single discrepant result. Thus, if we take together the fact that the weight of any single result is a function of the total body of available evidence and our assumption that anomalous or idiosyncratic results will diverge from the pattern of "true performance," then as the total body of evidence relevant to a cognitive system increases, the potentially deleterious effects of overspecificity and idiosyncracy will be proportionately reduced.

Can we actively contribute to the reduction of the effects of overspecificity and idiosyncracy? Yes, we can. We can pay closer attention to the premorbid abilities of the patients we study. If there is any suspicion that a patient may have presented with anomalous performance premorbidly (e.g., developmental dyslexia, exceedingly low level of education for some domains of investigation, and so on), then this fact should be given appropriate weight in reporting the case. Another important measure we can take is to obtain accurate and detailed control data on the performance of normal subjects, paying special attention to the range and type of variation, if any, among normal controls. This type of precaution is especially important for those tasks for which we expect high variance in performance in normal controls (e.g., nonword reading performance, which may be especially sensitive to the educational background of a subject).

In this section, I have considered possible objections to the methodological soundness of the single-case study approach in Cognitive Neuropsychology. Clearly, this approach is not problem-free. However, I have suggested that the potential difficulties that we may encounter because we cannot carry out controlled replications do not constitute a serious threat to the possibility of drawing valid inferences from single-case studies. In short, then, it seems to me that given that the potentially serious objections raised against the single-case study methodology turn out to be relatively minor, and given that the patient group study methodology does not allow valid inferences about the structure of cognitive systems, we have little choice but to conclude that Cognitive Neuropsychological research is only meaningful when carried out in the framework of the single-case study methodology.

CODA: ON SINGLE AND DOUBLE DISSOCIATIONS AND ASSOCIATION OF SYMPTOMS

The basic form of observation in Cognitive Neuropsychology consists of patterns of responses which deviate from the expected pattern for normal subjects. However, these observations would be of little theoretical value if the pattern of deviant responses were to be unconstrained; that is, if patients' performances were to be deviant on all relevant tasks. Patterns of deviant responses only become meaningful when interpreted in the context of patterns of spared performance in the same patient. In other words, the critical information on which we base our inferences about cognitive systems consists of differential patterns of spared and impaired performance or dissociations. For example, a *selective* difficulty in processing polymorphemic words in reading *could* be taken as evidence for a model of reading that implicated morphological decomposition at some stage of processing—that is, if the proposed functional architecture for a cognitive system contained a component of processing which operated over morpheme as opposed to word representations.

There are two methodological issues I would like to consider in the context of the role of dissociations to inform and constrain models of cognitive systems. The first concerns the distinction between single and double dissociations. "Single dissociation" refers to a differential pattern of spared and impaired performance; "double dissociation" refers to the occurrence of both a particular pattern of spared and impaired performance and the reverse pattern in which the previously impaired performance is now intact and the previously spared performance is now impaired. It is commonly assumed that the double dissociation is of more value in drawing inferences about the structure of cognitive systems than the single dissociation. I would like to suggest that this assumption is, under one reading, trivially true, and, under a second reading, false or, at least, unmotivated.

If the difference between single and double dissociation is construed as merely one of amount of information—one pattern of results vs. two patterns of results—then double dissociations *are* more useful than single dissociations. Under this reading, the presumed superiority of double over single dissociation is trivially true: $E_1 + E_2$ is more informative than just E_1. However, there is a second reading possible for the presumed superiority of the double over the single dissociation; namely, a pattern of *reverse* dissociations (read double dissociation) is more useful than just two single dissociations. This claim is, at best, unmotivated. As I have already argued, the fundamental logic for relating evidence to theory in Cognitive Neuropsychology is to articulate the structure of a cognitive system such that when "lesioned" appropriately it leads to a specific pattern of impaired performance. On this view, a dissociation is taken

to reflect the functioning of the normal cognitive system under specified conditions—viz., conditions that specifically indicate the nature of the disruption to the proposed functional architecture. Conversely, a dissociation constitutes evidential support for a model if together with the model it retrodicts the hypothesized functional lesion. Now, it should be clear that on this account, if the proposed model can explicate not only the occurrence of one single dissociation but of a second single dissociation, our confidence in the model is appropriately increased. Is there any reason for supposing that a reverse dissociation is of greater evidential value than just another single dissociation? No logically compelling reason has been provided to support this supposition. Instead the argument is often made by means of examples.

Consider in this context a quotation from a recent paper by Max Coltheart (1985): "It is generally considered that the double dissociation is of more value than the single dissociation. If one finds patients, for example, with a deficit of the non-lexical procedure and sparing of the lexical procedure, but not the reverse, the claim that this indicates the existence of two separate procedures could be disputed by the counter-claim that reading non-words is simply more difficult or less familiar a task than reading words even if the two tasks use a common procedure, and neurological damage may compromise difficult tasks more than easier ones. This counter-claim is not applicable to inferences based on a double dissociation—that is, the observation in some patients of selective sparing of the lexical procedure and in others of selective sparing of the non-lexical procedure." Notice that the force of Coltheart's argument for supposing the greater value of the double over the single dissociation is ambiguous. It could mean either that the double dissociation provides two patterns of performance which *together* rule out a particular account which was possible when the only information available was the single dissociation (i.e., the difficulty or familiarity hypothesis), or it could mean that the *reverse* dissociation plays a *special* role in rejecting the alternative account which is possible on just the single dissociation. I have already commented that the first of the two ways in which we could interpret the presumed superiority of double dissociation has nothing to do with *reverse* dissociation and that it is trivially true that the more relevant information we have, the better off we are. Here I want to consider the second interpretation—that it is the reverse nature of the dissociation that is critical. There are two implicit assumptions in Coltheart's example that should be made explicit. The first is that the available evidence, the single dissociation (plus other relevant evidence), is not sufficient to distinguish between (equally plausible) alternative accounts of the functional organization of the reading system. The second assumption is that the reverse dissociation does a better job of ruling out one of the two alternative functional architectures than some other dissociation. This second as-

sumption is unmotivated. Thus, consider the following. Suppose we start with a single dissociation consisting of a patient who can read words but cannot read nonwords. Let us suppose also that there is a model which is able to account for how normal subjects read nonwords (e.g., Marcel, 1980). Let us further suppose that we entertain as reasonable Coltheart's hypothesis that ". . .reading nonwords is simply more difficult or less familiar a task than reading words even if the two tasks use a common procedure, and neurological damage may compromise difficult tasks more than easier ones." We now get two additional dissociations, a reverse (double) and a single dissociation—the reverse dissociation would be the existence of a patient who can read nonwords but who cannot read words. The new, single dissociation consists of a patient who again can read words but who, on reading nonwords, makes systematic errors which are explicable on the basis of the computational structure of a hypothesized component that converts graphemic representations into phonological representations (e.g., the patient systematically produces within-phonological-class errors in reading nonwords, "peert" $->$ /tirg/).

Which of these two new sets of results provides stronger evidence in favor of the single-route model of reading? We cannot say! An answer to this question depends crucially on the nature of the models being evaluated. It is certainly not impossible to construct an argument that would allow the single-route model to account for the relatively impoverished evidence we have considered in the double dissociation discussed. Furthermore, it may turn out that the new single dissociation may be especially difficult to accommodate within a single route model of processing.

In short, then, the evidential role of a pattern of performance is not independent of the total body of theoretically relevant evidence, nor is it independent of the nature of the models proposed to explain cognitive processing. For these reasons we cannot assume that double dissociations are intrinsically more useful than patterns of single dissociations.

The second issue I wish to discuss, but only very briefly here, is the evidential role of the much maligned association of symptoms. It is common practice to point out the danger of a too easy interpretation of the co-occurrence of symptoms as reflecting a functionally necessary association of symptoms. This warning should not go unheeded, but it should be considered in its proper perspective. There are situations where our models do require that we predict the co-occurrence of symptoms given some hypothesis of the nature of the functional lesion in a proposed model (see Caramazza, Miceli, & Villa, in press). For example, if the proposed functional architectures of the reading and writing systems were to contain a phonological buffer which was used in both processing activities—in reading words and nonwords and in writing nonwords, for example—we would expect impaired performance in word reading and

nonword reading and writing on the assumption that the phonological buffer were damaged. If we were then to find a patient in whom reading and writing performance were completely dissociated—for instance, poor reading but normal writing—this would not necessarily constitute counterevidence to our proposed models of reading and writing. As in our discussion of single and double dissociations, the evidential role of various patterns of dissociations has to be assessed in terms of whether the obtained results allow retrodiction to a hypothesized functional lesion to the model being considered. In other words, it may very well be the case that the new dissociation results from damage to some processing component other than the one of interest in the original case of association of symptoms. Ultimately, what we should be cautioned about is not the evils of using single dissociations or overinterpreting association of symptoms, but the evils of not developing a sufficiently detailed model of the cognitive systems of interest to guide the search for richly articulated patterns of performance in brain-damaged patients. To the extent that our models of cognitive functioning are well developed, we will be able to make efficient use of single and double dissociations and association of symptoms.

REFERENCES

Badecker, W., & Caramazza, A. 1985. On considerations of method and theory governing the use of clinical categories in Neurolinguistics and Cognitive Neuropsychology: The case against agrammatism. *Cognition,* **20,** 97–125.

Caramazza. A. 1984. The logic of neuropsychological research and the problem of patient classification of aphasia. *Brain and Language*, **21,** 9–20.

Caramazza, A., Miceli, G., & Villa, G. The role of the (output) phonological buffer in reading, writing, and repetition. *Cognitive Neuropsychology*, in press.

Coltheart M. 1985. Cognitive neuropsychology and the study of reading. *Attention and performance XI.* Hillsdale, NJ: Lawrence Erlbaum.

Gordon, J. E. 1978. *Structures, or why things don't fall down.* Middlesex: Penguin.

Marcel, T. 1980. Surface dyslexia and beginning reading: A revised hypothesis of the pronunciation of print and its impairments. In M. Coltheart, K. E. Patterson, & J. C. Marshall (Eds.), *Deep Dyslexia.* London: Routledge & Kegan Paul.