INTEGRATING "DIFFERENT" MODELS IN COGNITIVE PSYCHOLOGY

Alberto Greco

University of Genoa, Italy

Abstract

In seeking to integrate different models, one may encounter two main obstacles, i.e. the models might not share the same sense of "model" and they might not have the same object. In this article it is argued that the first hindrance can be overcome whilst the second cannot. In cognitive psychology, the concept of model oscillates between a psychological sense (a knowledge-representation system) and an epistemological sense (a set of hypotheses, often limited in some aspect). However, these senses can be intermixed, since hypotheses may turn out to be a particular kind of representation or may, themselves, use some kind of representation.

Models which are "different" only in kind may be compared and possibly integrated, but the critical aspect is that they should concern the same psychological function (or functions), either as tools to represent it or as hypotheses on it, independent of implementation. The relevant questions are, then, how to identify which model function is concerned, how psychological variables inside models are defined and in particular what the relationship is between functional variables and their "labels". Examples of these topics in both symbolic and connectionist models are given.

1. Introduction

It is widely recognised that present cognitive psychology proposes a plethora of different "models", sometimes concerning very specific phenomena. But here and there the need is being expressed once again for wider theories, even though such theories seemed to be out of fashion for a while. So the question arises: how to reconcile or integrate different models in cognitive psychology?

Apart from the difficulty of answering this question, what it asks is not so clear. In particular, since the concept of "model" has so many different senses, it seems reasonable at least to try first to understand which sense

Revised version of a paper presented at the Special Workshop of the European Society for the Study of Cognitive Systems, "Models of Cognitive Psychology", Louvain, Belgium, 15-17 July 1992.

is relevant in answering that question. Although different models can be integrated, it might be the case that different *kinds* of models could not be integrated.

Even if one supposes that this matter has been fully cleared up, the concept of "integration" still remains to be clearly defined, because it only proposes some way of "putting together" models, or parts of them, but there may be (in fact there are) very different ideas about how this can actually be done.

The aim of this paper then is to discuss:

(a) what differences between models pose real obstacles to integration;

(b) what patterns of integration are possible in pursuing a unified model of cognition.

A central point to the first issue is what makes us label a model as "psychological"; the second raises the question of what is the real pursuit implied in "unification".

2. Models in cognitive psychology

What is a model? Here we are not interested in a full treatment of the concept of "model" from the epistemological point of view (see also Note 1), but only in considering how this concept is most commonly used in cognitive psychology. In this discipline, models are usually systems for representing scientific knowledge concerning psychological aspects. These systems can be symbol systems (including language), graphic representations or devices that actually work. We shall focus particularly on two senses of the concept:

(i) A representation system is sometimes defined as a "model" when treating complex or entangled processes: we have memory models (in general or concerning particular kinds of memory), we find models of linguistic comprehension, and so on. In this sense, which we shall call *psychological*, the model's peculiarity (as compared with theories, see Note 2) is that it is aimed at representing complex things more clearly, leaving out what is not essential or modifying some aspect of what is represented in order to improve our comprehension of it.

(ii) The other principal case in which the term "model" is commonly used is as an account of psychological aspects, like a "theory", but where this account is not robust, consistent, nor reliable enough to be considered a full theory. In other words, in this sense, which we call *epistemological*, a model *expresses a set of provisional or limited hypotheses*.

The distinction between these two meanings of "model" is important here because we are actually dealing with different kinds of models, which have different scientific goals. Indeed, apart from the general and shared goal of improving knowledge and understanding, in both cases the relevant specific purposes are to *represent* (describe) a phenomenon on the one hand and to *explain* a phenomenon (give reasons why, establish connections with other phenomena) on the other.

We have seen that in sense 1 (psychological) a model is a system for representing a psychological phenomenon more clearly. This can be done by:

(a) using *analogy* (i.e. using non-psychological concepts, which refer to phenomena similar in some aspect to a psychological phenomenon). The representation system is often a physical device and models of this kind are called "simulation" models. In this sense, proposing a model literally means "to model", that is to produce a new phenomenon (much like giving a shape to something shapeless) which must have something in common with the original phenomenon.

(b) *simplifying*, that is leaving out what from a certain point of view is not essential; obviously in this case what is essential changes when the perspective changes and therefore it is crucial to specify the point of view.

These two ways of clarifying representation are fully compatible and, in fact, frequently used at the same time.

We have also stated that in the sense 2 (*epistemological*) a model can be a kind of hypothesis or a set of hypotheses about a psychological phenomenon (and its goal is to explain). As we have outlined, in this sense "model" is similar to "theory" but sometimes contrasted with it because these hypotheses are "limited" in some aspect. In particular, they can be:

(a) a set of tentative, heuristic hypotheses (which typically originate from a still weak empirical support); or

(b) a set of hypotheses about a narrow domain (e.g., the case of "micromodels" in the cognitivist paradigm); or

(c) simply a set of hypotheses or theories from some particular point of view (in this sense "behaviorist model" is simply synonymous with "behaviorist approach").

3. Why models are different

Let us turn now to the other main topic of this paper: in what way models are "different". They could be different because they *do not share the same sense of "model"*, as shown above, or rather because they *have not the same object*. I shall argue that whereas the first kind of difference between models is not crucial, the nature of the object is.

Not sharing the same sense of "model", i.e. being models either as simulations or as hypotheses, is not a crucial difference. In fact, these senses can be intermixed, because sometimes "model" can be used in a psychological context to mean some *subject's* hypothesis (this is, for example, the meaning of the successful expression "mental models": Johnson-Laird, 1983; Gentner, 1983) and sometimes it can be used in an epistemological context to mean a *scientist's* representation system. Moreover, hypotheses are, in the end, a particular kind of representation, or may use some kind of representation, since they are a set of linguistic propositions about some phenomenon expressing what it is or why it is so. Therefore, hypotheses are also representation tools or are based on particular representations of the phenomenon (they also simplify and use analogies). On the other hand, simulative representations act as hypotheses (because they depict a state of affairs *as if it was* in one way or another...). I assume, then, that both senses are relevant when we speak of "models" in cognitive psychology and that this difference can be overcome in seeking integration.

On the contrary, as argued, when the difference between models lies in their not having the *same object*, this is an obstacle that cannot be overcome. To discuss this in greater detail, we need to examine model *objects*.

4. Model objects and psychological variables

What is a model's object? As we have said, a model's object is what the model (as a representation) describes, or what using the model (as a set of hypotheses) one seeks to explain. In short, that is to say *what it is a model of*. It seems obvious to suppose that this should be the first thing required of anyone presenting a set of propositions or a physical system as a model.

In our case, we deal with cognitive models, or - more generally - with psychological models. When it is claimed that a model is *psychological*, according to the distinction previously drawn, it should be either a system of tools to represent a psychological phenomenon, or a system of hypotheses on it. However, either way one must use some psychological label, for example to make reference to a psychological variable which is pre-defined (by making reference to empirical observations, or to psychological theories). This is declared (by the model's builder or by its user) by using labels which make reference to some *psychological variable*.

The point here, of course, is that the decision as to what phenomenon may be called "psychological", or the definition about what a psychological variable is, does not constitute a trivial issue in itself. But saying that different models can be compared only if they share the same object does not necessarily mean that they should share a "psychological" object. More importantly, this means that these models should concern comparable phenomena. In general terms, models should use *comparable variables*; more specifically, they should be about *the same function* (see Note 3). Exactly the same thing can be said of "cognitive" phenomena if we are to adopt the multidisciplinary (but somewhat narrower) perspective of cognitive science: what is cognitive is not easier to define than what is psychological and we should equally identify the relevant variables in order to decide whether they are comparable or not in different models.

Generally speaking, a variable is some aspect of an object or event that can vary, can assume different values; the definition of a variable implies the specification of this aspect. The agreement about what aspects of human behaviour and cognitive processes count as legitimate *psychological* variables is hardly based upon empirical support alone, but it also relies on categorization (abstraction and generalization). Since the behaviorist approach was superseded, the variables giving rise to the most interesting considerations are those referring to internal states. Perhaps one way of classifying different models might be to examine how they are different in assigning a different status to internal (non-observable) variables.

It is obvious that if person A insults person B, then B will get angry and perhaps slap A in the face. We can *represent* this fact by resorting to propositional attitudes or by using categorical labels like "this is an example of revenge" or even by drawing circles standing for A and B and arrows showing the direction of the aggression. We can also *explain* the phenomenon by connecting the insult of A with some psychological event in B (cognition, aggressiveness, and so on) and likewise with the behaviour of B.

The main problem now is: how are the appropriate labels for psychological variables chosen? Most of them come from folk psychology, especially at a macro-level where some processes have standard names: memory, learning, comprehension, etc. But these are often ambiguous names (consider how many definitions of "comprehension" have been given) and in any case they are too broad or generic to be useful for referring to interesting functions. Therefore we have more specialized terminologies (what kind of memory: episodic or semantic? declarative or procedural? etc.). The problem of how these macro-functions are accomplished (e.g. what kind of subfunctions are involved) is crucial in devising models in cognitive psychology. Consequently in many models the representation side is not completely separated from the explication side: it is shown what subfunctions are involved in the process, and at the same time how they work to produce the whole phenomenon.

One of the reasons why different models are hardly comparable is the fact that, after one has chosen labels for their variables, the models talk different languages. This may be due to the fact that, once a modeling system has been constructed, it is too easily considered as a new phenomenon, a new object with its subparts and subprocesses. The risk then is that the reference to the original phenomenon becomes lost or unintelligible because a different language has been used. A different model of the same phenomenon may be constructed as a different system, with its own variables and labels. One additional problem is that, often, already existing labels are still used in different contexts with different meanings or without a complete specification of the differences. For example, compare how the terms "schema" or "frame" are used by Bartlett, Piaget, or Neisser, and also in many current models in cognitive psychology.

5. Relationship between variables and labels

In simulation models the relationship between variables and their "labels" is of particular importance. This relationship can be a problem in comparing different models, both in symbolic and connectionist camps.

Symbolic simulation models

One can take as an example some old-fashioned ways of simulating psychological functions such as a model of paranoia (Colby, 1981). Then, the early idea (certainly naive) was to write a program which would be able to manipulate variables in the computer-programming sense, i.e. by "assigning" or modifying the values of some alphanumeric data (a system based on "label-value" pairs). In Colby's model, the input is processed in a particular fashion so that the program computes values (i.e. numbers) for critical variables like "humiliation" and in this way a paranoid output is finally produced when these values exceed a certain threshold.

One could then construct a model identical to Colby's merely by changing the labels: relabeling the variable that Colby called "humiliation" as "ozonization" one could "show" that paranoia is related to the quantity of ozone in the atmosphere. This paradox shows that giving variables a psychological meaning depends on their *nature*, not their *name*. For example, the role of humiliation in paranoia cannot be shown simply by considering the label attached to a numeric value, but humiliation could be - say - a procedure for interpreting some attributes of the Self in a particular way (seeing the Self as inferior, disparaged, etc.). In short, specifying the nature of a variable means saying exactly what it does, what its function in the whole system is.

Connectionist models

Similar problems concerning the relationship between labels and variables can arise with high-level variables "emerging" from connectionist networks. We know that a variable can be anything which varies, any kind of event. So I can see that, for given inputs, a network changes its states (more or less internally) in specified ways, and reacts again in specified ways; but I can use a psychological label to describe this event only if I am able to identify a psychological function being performed. Otherwise, the best thing I can do with this model is to use it as a toy which I can manipulate and see what happens in the output wires. Put very simply: I can describe my manipulation in psychological terms, and what I observe in the output once again in psychological terms. What is lacking, however, is a psychological description of the process that takes place in the middle, the only thing that makes a psychological model worthwhile.

In functional terms, how sensory information is actually coded is not as important as what its use in the whole system is. Suppose the actual code of a sensory process has been simulated by some activation patterns in a network. In this case, the nature of the particular code is not the crucial thing which makes us call these patterns "visual", "acoustic", "haptic" or use other psychological terms like white or smooth to describe particular states. Sensory inputs are the basis for the "symbol grounding" (see e.g. Harnad, 1990) but are not symbolic in themselves (see Note 4). If a model of them could not be related, sooner or later, to a model of symbolic processes, its interest for psychology would be dubious (unless different levels are involved, as seen below). Hence the problem of giving processes which take place in networks an overall interpretation in psychological terms. When making such an interpretation, one encounters the same problems examined above, in that one has to choose the appropriate labels for processes which take place. In fact even more problems may arise, because often one first has to understand what function is to be "read" into what is going on. What happens here is somewhat similar to when variables emerging from factor analysis are given psychological labels, where statistical results are interpreted using psychological knowledge.

From the consideration of the relationship between variables and labels, both in symbolic and in connectionist models, it becomes clear how important the similarity between model functions is for model comparison. Comparing or integrating different models where variables are labeled in a similar fashion is risky because their similarity might only be apparent. On the contrary, some models could in fact be integrated where their variables, even if labeled differently, have similar psychological functions.

6. Different implementation and different function

It is often the case that different models are hardly comparable, at least at first sight, because they are implemented in different ways (different software, different hardware, different design). Does a different implementation mean a different function? Obviously not always, since we know that the *same* processes or "functions" performed inside a system can have a different implementation and so give rise to different models. Here two important matters are raised: (a) different functions can be considered different *modules* (subparts); b) the same function can be performed at different *levels*.

As to the question of modules, if we are able to identify a specific function, then we can consider the architectural subcomponents which perform it as a subsystem, a module. Functions may simply have subfunctions. This need not necessarily be done in the Fodor (1983) sense, i.e. supposing that each module is logically necessary and sufficient, or that modules are mutually exclusive, or endorsing other similar constraints.

As to the question of levels, some people think that the same function can be accomplished at different levels. The idea that the same function could be described in high-level, intermediate, and low-level terms is very popular; there is no need to refer to Marr (1982) or Newell (1982) to realize this. So in principle it might be possible to compare models concerning processes at different levels on the basis that, after all, they concern the same function.

However, there are also those who think this is wrong: for example, among others, Pylyshyn (1984) has pointed out that a model is best evaluated on "algorithmic" grounds if we want to establish what he calls a "strong equivalence" with a psychological function. As a tool for comparing different programs Pylyshyn proposed the concept of a "virtual machine" which, like a programming language, specifies the primitive operations available in a particular system and provides what he calls the "functional architecture" of a system, replacing its actual (structural) architecture. Indeed, Pylyshyn's lesson is that a comparison between systems (and models) is appropriate only at a specified level.

I basically agree with this view but I would add that this is not the only reason why an "across-levels" comparison is nonsense. A function is not an empty concept. It is always a function *of something* and what this something is depends on a point of view or context in the discourse: the function of a TV set may be to transform radio waves into sounds and images, or to show pictures, or to entertain people or even to influence public opinion. It is not a question of levels but of appropriateness in context. What I maintain is simply that the psychological features must not be lost when identifying the appropriate functions on the basis of which we are to compare different models.

7. Approaches to integrated models of cognition

In this section I shall consider the question of how integrated models of cognition could be possible. I see many ways, some of which are safer than others.

Replacing - The most radical way of proposing a new model (or a new theory) is to set out to use it to replace one or more competitors; usually the new model is claimed to give a better representation or explanation, to fit the empirical data better, to be more parsimonious, etc. In this case integration is clearly not at issue, simply because only one model remains. This case is not very frequent in psychology, where perhaps there has never been a Copernican system that replaces a Ptolemaic system.

Reduction - One model is reduced to another by *translating* each variable of the one into a variable of the other. Some variables cannot completely fit, of course, and so in this process much may be lost. For example some may claim that semantic networks should not be used but rather a frame system. Another classic example: some assert that there is no need for iconic representations, a propositional system is enough. In this case, what a model says about iconic information could be accounted for equally well by a propositional model; therefore it could be "translated" into the language of propositional variables. The reduction fashion of integration is always dangerous, particularly when it happens to be a reduction to a model of a different level (usually lower than the previous one).

Subsumption - In the subsumption case, as in that of reduction, a new model replaces an old one, but *mapping* of variables is used instead of translating: this means that the new variables are adapted to the older ones. The advertising slogan could be: why choose if you can have it all? (for example: semantic network with slots...). Typically, it's not possible to use this approach when models at different levels are concerned.

In reduction and subsumption, one system is pitted against the other and an evaluation is made as to whether the new model works better in fulfilling theoretical requirements or in dealing with constraints related to the overall performance or to the general system architecture. However, if the subsumption method goes on to become systematic, we begin to have integrated systems, in which a "merging" process is used.

Merging - In this case, the new system does not replace the old one, but adds something to it. This is, in my opinion, the case of some models explicitly born as "integrated" systems, like ACT (Anderson, 1983), SOAR (Newell, 1990; Rosenbloom et al., 1991) or DAYDREAMER (Mueller, 1990).

ACT, for example, has several memory systems (a working memory, a declarative memory and a procedural memory) which work differently. Then one can find procedural memories such as production systems that work together with semantic networks. Rules and spreading-activation cooperate in a single model, where also propositional and pictorial representations coexist.

SOAR is Newell's candidate for a unified theory of cognition. It works as a problem-solving system in that it sets its own goals and uses the standard problem-solving strategies (representing initial and final states, representing operators and their effect in a given state, searching in problem spaces, etc.). But it is also a system for reasoning heuristically and learning from experience.

SOAR has memory systems and performs an integration of declarative and procedural aspects similar to that of ACT. Newell also tries to address the important question of different levels and his answer is to keep them separate (e.g. to speak of "spread of activation" at a symbolic level is not a good example of integration).

As another example, take Mueller's system DAYDREAMER, which is less well-known but, in my opinion, a step in the direction of model integration. DAYDREAMER is a system which models "the daydreaming of a human in the domain of interpersonal relations and common everyday occurrences" (Mueller, 1990, p.2). Among other things, it is also a model of problem-solving where not only problem representation is important, as in all standard models, but also associations, goals, even emotions and the stream of thought, which all play a role in modelling "creative" problem-solving. In this model many components are integrated: a conceptual-dependency-based representation system, planning rules, inference rules, goal and emotion management (see Note 5).

Getting back to the general issue: as a matter of fact, in these models the integration is almost mandatory because it comes from their scope, which is rather different and wider than the scope of other systems: they aim to be *general* models of cognition, so they have to account for many features and must have variables for long-term and short-term storage, retrieval processes, etc. But sometimes the integration arises from the consideration that natural processes can use different ways to perform a function. So production systems and semantic networks could be alternatives in a long-term memory model, but trying to use both in a single model could make that model more complete and powerful.

8. A unified model of cognition or correspondence between models?

In this paper the issue of model comparison and integration in cognitive psychology has been considered. Some see the integration as a step towards achieving a unified model of cognition. But, after all, is a unified model of psychological processes really desirable? Newell is perhaps the greatest supporter of a "unified" model of cognition, but he is not alone and not the first: everyone remembers more ambitious attempts like the Neo-empiricist claim for a unified science. We can ask ourselves not only whether a unified psychological theory is more workable than a unified general science but, above all, if a pluralistic system, a competitive system, is not more fruitful and sound. One of the greatest mainsprings of scientific progress is the competition between theories; a monopolistic system is unthinkable.

In fact, a unified model may not necessarily be understood as a *single* model. This may be impossible when seeking to account for complex systems - as is clearly done in psychology. In this case, we cannot but have several models; what we need are tools that enable us to compare them, perhaps by establishing correspondences between these models. In classical epistemology (see e.g. Nagel, 1961) the idea of correspondence rules was proposed in the sense of rules allowing an interpretation of formalized theoretical postulates, by connecting these postulates with the concrete contents of experimental procedures. Now the sense of this idea has become wider; correspondence rules are rules for relating different representations (as suggested in Dalenoort, 1990).

Rather than a sort of a super-model, then, we would need a *strategy* for constructing and considering the relations between different models. This implies making a comparative analysis of variables in models, and recognizing the importance of considering the relations between language (what in this article has been called "labels") and empirical data. The relationship between variables and functions should be also studied. For example, a function performed by all the variables described in one model might correspond to a single variable in some other model. Much work remains to be done in this direction, which could be considered as a new sense of *integration*.

Notes

1. The classical sense of "model" refers to disciplines where, in expressing hypotheses, a certain level of formalization can be reached. In this sense, a model is an interpretation of formal postulates (e.g. see Braithwaite, 1953; Nagel, 1961). A single formal theory, then, may have different models (different interpretations). In psychology such a sense is completely overlooked.

2. The difference between "models" and "theories" can be stated in other terms (see for example point 2 below), or some may even argue that there is no difference at all. However, the terminological aspect is not important here. The important thing is that a large number of "models" proposed in cognitive psychology (perhaps misusing the term "model"), have the features outlined here.

3. The concept of "function" is adopted here in its broadest sense, similar to that of classical functionalists. It is not restricted to single psychological processes like memory or perception, but also includes integrations of different processes which have, for example, an adaptive value or which for some reason can be considered together. Indeed, models concerning the relationships between several processes are not excluded from this analysis.

4. I am not taking a position about the symbolic/subsymbolic/nonsymbolic question. I am only arguing that the important thing is what a model is about, the "functional stance" (Dennett, 1987), not the implementation details. This does not mean being committed to the idea that mental activities should be characterized only in terms of symbols and rules for manipulating symbols.

5. Goals are also found in SOAR, but in this system they are set by the programmer and, as the authors themselves recognize, "a general intelligence must have grounds, that is motivations" (Rosenbloom et al., 1991, p.317). Of course, the actual implementation details of DAYDREAMER do not matter so much and are debatable, such as defining emotion as a mechanism to select goals, but the general architecture seems good.

References

Anderson, J.R. (1983), *The architecture of cognition*, Harvard University Press, Cambridge, Mass.

Braithwaite, R.B. (1953), *Scientific explanation*, Cambridge University Press, Cambridge.

Colby, K.M. (1981), Modeling a paranoid mind, *Behavioral and Brain Sciences* **4** 515-560.

Dalenoort, G.J. (1990), Towards a general theory of representation, *Psychological Research* **52** 229-237.

Dennett, D.C. (1987), *The intentional stance*, MIT Press, Cambridge, Mass.

Fodor, J.A. (1983), *The modularity of mind*, Bradford (MIT Press), Cambridge, Mass.

Gentner, D., & Stevens, A.L. (Eds.) (1983), *Mental models*, Lawrence Erlbaum Associates, Hillsdale, N.J.

Harnad, S. (1990), The symbol grounding problem, *Physica D* **42** 335-346.

Johnson-Laird, P.N. (1983), *Mental models*, Cambridge Univ. Press, Cambridge.

Marr, D. (1982), *Vision: a computational investigation into the human representation and processing of visual information*, Freeman, S.Francisco.

Mueller, E.T. (1990), *Daydreaming in humans and machines*, Ablex Publ., Norwood, N.J.

Nagel, E. (1961), *The structure of science*, Harcourt, Brace & World, New York.

Newell, A. (1982), The knowledge level, *Artificial Intelligence* **18** 87-127.

Newell, A. (1990), *Unified theories of cognition*, Harvard Univ. Press, Cambridge, Mass.

Pylyshyn, Z.W. (1984), *Computation and cognition. Toward a foundation for Cognitive Science*, MIT Press, Cambridge, Mass.

Rosenbloom, P.S., Laird, J.E., Newell, A., & McCarl R. (1991), A preliminary analysis of the Soar architecture as a basis for general intelligence, *Artificial Intelligence* **47** 289-325.

Manuscript received: 15-10-1992, *Revised version*: 3-6-1994.

Address Author:
Dipartimento di Filosofia - Laboratorio Sperimentale di Psicologia
via Balbi, 4 - 16126 Genova, Italy

COGNITIVE SYSTEMS

Journal of the European Society for the Study of Cognitive Systems

The Journal is published two or three times per year. It contains the elaborated versions of papers presented at the workshops, papers submitted for publication, communications, commentaries, and book reviews. The content of papers must concern the cognitive aspects of mind and brain, and must be of multidisciplinary nature.

For the format of papers prospective authors are referred to the papers already published. Typing and correcting manuscripts both cost much time and labour. Therefore, authors who can provide floppy disks of their manuscript, are requested to take up contact with the Editor.

Subscriptions may be obtained from the ESSCS. For members the subscription price is Dfl 60.- per annum (200-350 pages), membership amounts to Dfl 25.- per annum. For non-members the subscription is Dfl 80.-, for institutions Dfl 120.-. These prices include handling and postage, also outside Europe, but airfreight is additional. Details for modes of payment will be provided upon request.

ISSN 0256-663X

European Society for the Study of Cognitive Systems (ESSCS)
Department of Psychology (E&A), University of Groningen
P.O. Box 41096
9701 CB GRONINGEN
The Netherlands