THE CONCEPT OF REPRESENTATION IN PSYCHOLOGY

Alberto Greco

University of Genoa, Italy

Abstract

An analysis is given of the status of the concept of representation in psychology, and the various ways it is used, including its explanatory status and its use as a causal agent.

1. The concept of Representation

1.1. *Representation: Representing and Representations*

This paper concerns the status of the concept of representation and the psychological use of this concept as an explanatory construct. As we have seen in the *Introduction* of this issue, representation – in its most general sense – seems to mean "something" that *substitutes* something else. However, it is enough to look up any dictionary definition of representation to realize that the term is not always used to refer to "something" but also to some activity or *operation*. A rather basic distinction to make, indeed, is that representation may be either the *act* of representing or the *product* of representing. This distinction is not only relevant from a lexical point of view, but also turns out to be a *psychologically* important distinction, that is the distinction between *process* and *content*.

Process and content were first distinguished as far back as Brentano. Typically, when talking about content one can use the plural form ("representa*tions*"), but not when talking about process, which we may call "represent*ing*". We can find corresponding distinctions also in other psychological processes: thinking and thoughts; perceiving and percepts; storing and memories, etc. Here, in order to stress this difference, I propose to adopt a simple terminology: in the *functional* sense of the process of representation I shall use the term "representing" and in the *structural* sense of content of representation I shall use the term "representations"; I shall continue using the word "representa*tion*"

Revised version of a paper presented at the 12th Annual Workshop of the European Society for the Study of Cognitive Systems, Camogli (Genoa, Italy), 12–14 April 1994.

in the most general (or ambiguous) sense, which includes both. This terminology perhaps is not particularly new, but I find it very useful.

I am stressing this difference because the psychological significance is different in the two cases. If attention is directed towards *representations* as products, then the main focus concerns their *form,* their *structure*. On the other hand, in the "representing" sense, attention is shifted to *dynamic* and *time–related* aspects, which seem more natural (in my opinion, the time dimension continues to be neglected in psychology). Furthermore, represent*ing* seems more appealing from the neural point of view. But I shall return later to the subject of why representing may be more appealing than having representations.

1.2. *Representation as "Internal Entities or Events"*

Now I want to continue examining other aspects of the basic meaning of the term representation. We started with the basic definition as "standing for something else". The psychological use of the concept of representation, however, seems more general than this, and indeed it might not be limited to substituting.

More generally, in the psychological sense representations are entities we postulate as *internal* to the organism. Thus representations can be conceived as *internal entities or events*. If we speak of substitution or of substitutes, we are actually speaking of *one* possible function of these entities or processes; but it is questionable whether this function of substitution is the only one or the most relevant. But this point concerning the *function* of representation deserves discussion in itself, something that we shall do later.

2. Reasons for postulating representations

2.1. *Everyday Psychology: Phenomenological Evidence – Intentional Explanation*

Before, we should first examine a more basic question: why internal entities or events are postulated? I shall mention some reasons, first in everyday or commonsense psychology and then in scientific psychology. In everyday psychology, an early boost for postulating representation, or rather internal representations, comes from subjective experiences. There are naive or folk reasons: I feel the evidence of psychological states inside me (thinking, emotions, images ...) and therefore I suppose that something in the head of other persons must explain their behaviour and also their psychological states, that they report to have. So there is a *phenomenological evidence* of meaning, or the evidence that mental activities persist over time (the phenomenon of memory). It has to be remarked that, in everyday psychology, representat*ions* are almost always preferred to represent*ing*: it seems easier to consider *ready–made* thoughts rather than processes of thinking. Here, then, representation is certainly a useful concept in that it constitutes the basis for the intentional system of

explanation, typical of commonsense psychology (desires, beliefs, etc. represent internal states; images do represent objects, etc.).

2.2. *Scientific Psychology: Mediating Between Stimulus and Response*

A second reason for postulating representation is more technical and belongs to psychologists. The early, naive "something in the head" in scientific psychology becomes less vague, but unfortunately also more confusing, because in psychological literature the term may refer to different phenomena, and to different levels. A list of examples of representations (or of sets of representations) could be endless: I shall mention only some of them: linguistic symbols, mathematical symbols, visual patterns, even visual fields or images. Also high–level concepts (sometimes similar to those of everyday psychology) are considered as working as representations, both in natural and artificial systems: so we have categories, beliefs, propositional attitudes, schemata, semantic networks, and so on. Recently representation is also spoken of in connectionist models, and some even speak of "neural" representation. We have to admit that the confusion is overwhelming.

The best thing to do would be to evaluate the explanatory role of mental representation not "in the vacuum" but in some particular theoretical framework. (The philosopher Cummins has suggested a similar point of view). But this is practically impossible here, because virtually all theories in the history of psychology have had to come to grips with representation and could be mentioned.

Anyway, a good starting point for discussion seems to be the observation that representations are introduced in scientific psychology essentially as mediating entities. In order to explain behaviour and mental states, psychology needs entities or processes which *mediate between stimuli and responses*, or between inputs and outputs or, more widely, between situations and behaviour (at least, if humans are to be understood as systems which are not completely determined by their environment). This idea of an internal mediation between the environment and the organism's action is typical of modern psychological conceptions, but is not necessarily the only one possible. For example, the most classical philosophical positions have considered representation rather like a sort of internal reality which we find naturally inside us (no matter if there was the problem of comparing it to the so–called "true" reality... Hence the well–known classical dualism, and hence many philosophical discussions concerning the adequacy of internal reality with respect to the so–called "true" reality, the problem of misrepresentation, and similar arguments that amuse philosophers so much).

As is well known, the idea of "mediating entities" historically came in psychology from the need to overcome the typical *impasse* of the behaviorist position concerning the "gap" between stimuli and responses. In substance, cognitivists said: the gap between stimuli and responses can be filled if we consider that

stimuli do not act as such, but are manipulated, processed. Hence the need for representations.

We have up to now considered representation as an internal event (a process, or a product of a process) which can explain psychological phenomena by virtue of its working as a causal connection between stimuli and responses. Now I want to ask two questions, which are related to each other:

1) where does the causal power of representations come from?

2) what function (or functions) is (are) attributed to representations by psychological theories?

We shall examine these two issues in the next two sections

3. Causal Power of Representations

We have seen that, in general, representations are given an explanatory power inasmuch as they act as causes. The paradigm is the same as that in physics, with the difference that here one speaks of internal rather than external causes. It is interesting, therefore, to know where the causal power of representations comes from.

3.1. *Symbolic representation*

3.1.1. *Personality and Social Psychology: Representation as "Subjective" Reality*

In some areas of psychology (especially in personality theories or in social psychology) the concept of "representation" is often used (to tell the truth, not so differently from everyday psychology) to express the assumption that individuals do not act on the basis of "objective" patterns of the world, but on the basis of their so–called "internal representations" of it, which do not necessarily correspond to what actually happens in the world, but can be abstractions, simplifications, perhaps misrepresentations or even illusions. Even if this does not necessarily involve the earlier–mentioned risk of a dualism between reality as it is and as it appears, representation as a "subjective" reality here is strongly opposed to the "objective" reality, (which – by the way – usually happens to be the one of psychologists). Moreover, representations in this sense are, in general, complex ones (e.g. categories, expectations, interpretations, scripts), and they can explain fears, problems, behaviour – only if they are properly interpreted and connected with the right meaning. Here the risk is of being too general, vaguely confusing representation with having any idea, knowledge, or thinking. It was to avoid these very risks that scientific psychology abandoned the naive way of explaining.

3.1.2. *Cognitive Psychology: Representations as Symbols to be Interpreted*

Cognitive psychology perhaps is better placed because it restricts itself to the elements that make thinking possible. We know that in cognitive science it has become commonplace to see cognitive activities as information processing. Cognitivism considers knowledge acquisition and management in terms of symbol manipulation which follows formal rules. Representation in this perspective is the postulation of a set of internal entities standing for other things, separate from these other things which they represent.

But, if this is true, in this aspect representation from the standpoint of cognitive science is not different from other psychological theories, because the causal power of representations still comes from their interpretation. Their simple existence is not enough to give representations a causal power. The causal power of representations does not come from their mere existence, but rather from their interpretation. In other words: because they are symbolic. The medium that conveys meaning may have an arbitrary form, and there are formal rules that make going from the sign to its interpretation possible. The model is language, and indeed this perspective is often called the one that resorts to a "language of thought".

3.2. *Non–Symbolic Representation*

Psychoanalysis: Simple Existence

In psychological history, of course, this has not been the only proposal, even if we have to admit that it has become the most fashionable. We can take – as an example of a different approach – psychoanalysis, which is interesting because it has been influential in general psychological culture. The peculiarity of Freud's system is that *representations which are non–accessible but still have a causal power* are allowed. Here representations are meaningful ideas, but different from those of commonsense. They are symbols, in a different sense than in cognitive theory, symbols as contents which replace other contents, and this replacement occurs because the original contents are not acceptable, hence not representable. In this sense, differently from the cognitive perspective, *the causal power of representations comes from their mere existence*. We shall see later other examples of so–called representations which act by mere existence, one example is the connectionist perspective.

The important thing now is that the question we asked earlier (where does the causal power of representations come from?) can be answered in two ways: according to some perspectives, it comes from the *interpretation* of internal entities; according to other perspectives, the causal power comes rather from their *mere existence*. We can call these two cases *symbolic* and *non–symbolic* representation.

In the next section we shall examine the second question previously asked, about the functions that psychological theories attribute to internal entities.

4. Functions of Representations

4.1. *The Substitution Function of Internal Entities: Making Actual (Storing and Anticipating)*

We have said that, in the most common definition, representations are not just any kind of internal entities but are considered *internal substitutes*. But what does it mean to have internal substitutes? What are they meant to substitute? Clearly, substitution is necessary when something is not present but still required. Then, in principle, if representation is substitution, we do not need representation to manage present, *actual* events. Rather, we need representation to cope with *past* and *future* events. On the one hand we have to resort to representation when we need to store and retrieve information about something that has already happened. On the other hand we need representation when we need information to be used for something that has not yet happened, but which we ourselves can control, that we can make happen, that we can construct. In other words, our behaviour (of course in the widest sense, including language for example). We can find this sense of representation in famous psychologists like Bruner and Piaget.

Following an old but influential distinction put by Bruner et al. (1966), we can describe a not particularly common kind of representation, the one that he calls "active representation", which has a different function from the usual one, it has the function of anticipating *doing* or action. It is an internal organization of behaviour occurring *before* behaviour. A similar proposal, perhaps in clearer terms, can be found in Piaget's theory, where representation at the beginning reflects action and afterwards becomes more abstract (according to Piaget, as we know, the organization of thought is based on the earlier organization of action) In these senses, representation does not imply symbol manipulation, but there is a connection with action.

Then, if we are to postulate internal representation as "internal substitution", this can be understood as *a function of "making actual" what is not actual.* And, more precisely, this function includes two subfunctions:

1) *storing* past perceptions or past behaviour;

2) *anticipating* perception or behaviour (antecedents of our behaviour, which in some way is relevant to "plan" performance: e.g. motor schemata, pre–linguistic representations, attitudes etc.).

Why should we need something internal that substitutes behaviour? It is the same reason why *planning* is necessary for any intelligent system. Actually

trying any possible action would simply not be economic. Hence we need motor schemata for planning action at the lowest level, and also mental operations or "moves" when we have to solve a problem.

In what sense do these representations substitute action? They do in the sense that they have the function of making action possible only in the mind.

It is interesting to note that the same functions also hold for artificial systems: in order to work properly, they have to store past events and also instructions on how to produce output states. But this is not true of all systems, because some of them (neural networks), as we know, do not need instructions.

4.2. *The Correspondence Function of Mental Entities*

Now I shall mention a different perspective in considering the function of internal entities, which gives up the idea of substitution. One can say: perhaps internal events, in fact, do not substitute anything, do not anticipate anything. According to this perspective, the main function of representation is the one I shall call *correspondence*.

According to this idea, the effect of stimulation starts with transduction and hence gives rise to a modification of neural states. Something happens, in accordance with physical or physiological constraints. There is a variation of internal states *corresponding* to a variation of external (or bodily) states: in other words, a covariation.

From the psychological point of view, a similar idea was put by Gestalt psychologists as the isomorphism postulate: this means that when having the *same* perceptions we always have the same internal processes, *whatever they are*. However, the usefulness of such a concept is dubious if we say *whatever they are,* and we are not able to identify what these processes are. Even if we could describe them accurately as neural processes, the same old problem of connecting these processes to precise *psychological* processes would still remain. This idea is relevant in neural networks. It may be claimed that networks *do represent* because they change their state in a non–random way, closely related to changes in the input.

In some connectionist systems it is claimed that symbolic representations are not required: in these systems, at least in multi–layered and subsymbolic ones, there is still a sort of internal storage, which is in the weights of connections, but there is not an understandable relationship either with the input (the past) or with the output (the future). By "not understandable" I mean that the function of these internal states, as they are not symbolic, cannot be understood by someone who inspects the network while it is working. Hence the well–known problem of deciding whether these internal states are still to be called representations or not. Why resort to non–symbolic representation? If it is synonimous with a generical internal event, it could be called "non–symbolic *processing*"

It is easy to see that the *correspondence* function of representation is perfectly compatible with the earlier–mentioned non–symbolic representation. This is what we "find" in our senses, not what we construct. We know that there is experimental support to the isomorphism hypothesis (for example, the findings of Hubel and Wiesel). Is this enough? Perhaps not, since we know from experimental psychology that there is wide support also for the symbolic representation hypothesis. We need both but we do not know how to consider them in a single theoretical framework. The real problem is that we do not know how symbolic representation arises from the non–symbolic one, or anyway how they are related.

5. Representing as using representations or constructing representations

From what has been said up to now, it's easy to see that in psychology greater importance has been given to represent*ations* rather than to represent*ing* (linguistic symbols, beliefs, schemata, etc.: what else are they, if not ready–made represent*ations*?) This approach of giving more importance to represent*ations* has a respectable tradition; we know that especially the psycholinguistic approach is involved in this tradition. Sometimes internal representations have been called "tokens" or something similar, and this seems to reflect a topographic conception of cognition, which postulates ready–made pieces of meaningful material, of meaningful building blocks. This is the conception of representation as a language of thought, and this conception is also bound to the notorious "compositionality" requirement that some philosophers have pointed out.

Perhaps "representing" is avoided because it is difficult to consider a process in abstract terms, without specifying its content. And perhaps there is another reason. When one postulates a *process* it seems natural to imply an "agent" that carries out this process. Of course, this agent need not be a homunculus inside our skulls, but may take more sophisticated forms, like origins or causes, or forces (and so on) which drive the process. But the problem, in any case, is that in making our theories, as students of psychological processes, we can only use our logic tools: we cannot avoid considering processes or events as if they were predicates. Being predicates, they require arguments (what is the subject, what is the object) and these arguments are hard to specify when talking about mental processes. In my opinion, this is one of the reasons why understanding internal processes in *non–subjective* terms is so difficult.

However, we cannot deny that, unlike the commonsense perspective, in scientific psychology also a *process* of represent*ing* is sometimes put into the field, once again in order to account for the gap between situation and behaviour. But this option of conceiving representation as a process of represent*ing* is rarer. The main perspectives that come to mind are: personal styles in social psychology (not *what* one believes but *the way* one believes) or, of course, the typical position of cognitivist psychology (for example, not parti-

cular schemata but the general way schemata are constructed or managed). Here we are dealing with the sort of cognitive machinery that has been called (Pylyshyn, 1984) the *functional architecture*, which makes representing possible.

But here we must be careful, because in most cases the difference between represent*ations* and represent*ing* is only apparent, because represent*ing* often is merely a process of managing ready–made represent*ations*. In this case, it only means using two different ways to say, in fact, the same thing.

In speaking of represent*ing*, instead, I mean a radically different conception: I am referring to a process not of *using* representations but of *constructing representations*. The gist of the previous discussion is that when those internal entities called representations are used to *substitute* external events (to store information or to anticipate action), a symbolic interpretation is required. These representations seem to be constructed, sometimes with effort. On the contrary, when internal entities are used to *reflect* external events, there is no such need for a symbolic interpretation but their effect seems to be found, already ready. Hence the feeling of an automatic internal reality. In my opinion, most of the problems of the concept of representation come from the fact that, according to what we know up to now, these seem to be *two* concepts of representation, and there is no way to consider them in an overall framework. This is the difficulty of reconciling finding and constructing (passive/active, meaningless/meaningful, nonsymbolic/symbolic).

6. Conclusion: Proposing a New Metaphor

In conclusion I would like just to outline (really in a flash) a working hypothesis about how the symbolic function of represent*ing* could arise not in managing represent*ations* but in "constructing" them. All evidence coming from neuropsychology tells us that this constructing should be conceived as a process in which there are first internal events which simply reflect stimuli (at a low–level, a neural level). What happens of these corresponding or isomorphic events? Subsequently they should detach themselves from the simple correspondence function, they must stop working as simple mirrors and start to substitute.

Perhaps a new metaphor to be explored may be proposed, coming from biology. The process we are discussing, might be similar to the process whereby cells copy themselves while reproducing, but with errors, or with sorts of "mutations", and from these errors information and organization arise. This idea has to be explored, and neural networks might be a good field to try it. I am working in this direction but perhaps I need help from biologists.

References

Bruner, J.S., Olver R.R., Greenfield P.M. et al. (Eds.) (1966) *Studies in cognitive growth*, New York: Wiley.

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

Manuscript received in final form: 11–04–1995

Address author:

Dept. of Philosophy – Lab. of Psychology, University of Genoa,
Via Balbi, 4 ; 16126 Genova, Italy.

E–mail: greco@igecuniv.cisi.unige.it