A Note on the Relation between Action and Perception*

M. T. Turvey†
Haskins Laboratories, New Haven, Conn.

I would like to explore the thesis that, in principle, the fundamental problems to be solved and the concepts that will provide their solution are the same for both the Theory of Action and the Theory of Perception. Consider the following transcription task. A person sees a written capital A and is required to respond by writing the letter that she has seen. There can be, of course, different tokens of the letter seen, and there can be a variety of manners in which our person is requested to make her written response. We may partition a perceptual-motor occurrence of this kind into three phrases. The first consists of a set of functions that map states of the optic space (o) into states of the perceptual space (p); the second consists of a set of functions that map states of the perceptual space into states of the act space (a); and the third is a set of functions that map states of the act space into states of the motor space (m). We can represent the three phases as: \( F_1 = \{f|f:o \rightarrow p\}, F_2 = \{f|f:p \rightarrow a\}, \) and \( F_3 = \{f|f:a \rightarrow m\} \).

The thesis to be explored suggests that we should look for similarities between the functions on the perceptual end and those on the action end of our transcription task, i.e., between \( F_1 \) and \( F_3 \). In addition, it suggests that we should ask in what way(s) the perceptual and act spaces may be similar. The present paper is a response to these suggestions.

The concept of action

I am going to assume that our collective intuitions about the concept of perception will probably suffice for the present purposes. However, I will not make the same assumption for the concept of action, primarily because there has been far less hue and cry about this concept in theoretical psychology. Consequently, its character is less well articulated—witness the tendency to equate action with response in comparison to the tendency to equate perception with stimulus. We do the latter rarely, and the former frequently. Action, like perception, is an abstract relation between the organism and the environment. Just

---


†Also University of Connecticut, Storrs.

[HASKINS LABORATORIES: Status Report on Speech Research SR-34 (1973)]
as we are unable to point to perception, we are unable to point to action, although, of course, we might be able to detail the parameters of stimulation, and to describe fully the muscles and joints involved in an exhibited movement.

A number of philosophers (see Care and Landesman, 1968) have sought to lay bare the concept of action. Here I will report only briefly on their endeavors merely to portray two rather important characteristics of the concept. The first is that nerve impulses and muscle contractions—though necessary conditions for action—are more accurately considered accompaniments of action than characteristics of action. This point can be defended on at least two grounds: first on the notion of intentionality, and second (to be explored at some length below) on the issue of movement equivalence, or constancy. Clearly, it is perfectly reasonable to say that one intends to kick a football, but it is not reasonable to say that one intends to contract and to relax one's biceps femoris and one's rectus femoris, respectively, to this or that degree. Generally, one cannot choose or intend to transmit a nerve impulse or to contract a certain set of extrafusal and intrafusal muscle fibers. Intentionality is a defining characteristic of acts, but not of muscle contractions.

Another defense of the notion that the concept of action cannot be reduced to bodily movement is that any particular constellation of muscular contractions and joint motions brought into play when one performs an act (say, reaching for and lifting up a cup) cannot be said to be identical to the act. A radically different configuration of muscles and joints could just as easily have been used to achieve the same result. An act is Gestalten, that is to say, in a variant of the hackneyed phrase of Gestalt Psychology: an act is more than the sum of its constituent movements.

The second important aspect of the concept of action is that consequences, i.e., changes wrought in the environment by a configuration of movements, are integral to the concept of action in that no reliable distinction can be drawn between the concepts of action and consequence. Consider the following: George kicks the football (of the round kind), and scores the goal that wins the championship. Now we could say that George kicked the football and that a consequence of this act was that a goal was scored. Or, we could say, just as appropriately, that George scored a goal with championship-winning consequences. "Scores the goal," therefore, can be viewed either as consequence or as action. We might suppose that there are criteria available to determine what occurrences should receive an action label, and what occurrences should receive a consequence label. Unfortunately, the criteria that have been advanced have not been greeted with universal approval.

The problem of constancy in perceiving and acting

Let us now return to the phases in transcription referred to above, in particular $F_1$ and $F_3$. Consider that a visually presented capital A can occur in various sizes and orientations and in a staggering variety of individual scripts. Yet in the face of all this change, the identification of the letter remains constant; we see through the variations to the canonical form.

This phenomenon of constancy is not limited to the domain of perception, but is equally characteristic of action. Thus, the letter A may be written without moving any muscles or joints other than those having to do with the fingers. Or,
it may be written through large movements of the whole arm with the muscles of the fingers serving only to grasp the writing instrument. Or, more radically, one can write the character without involving the muscles and joints of either arms or fingers, by clenching the writing instrument between one's teeth or toes. It is evident that a required result can be attained by an indefinitely large class of movement patterns.

On examination of the phenomenon of constancy we might raise the query: How can these indefinitely large classes of possible shapes, and of possible movement patterns, be stored in memory? The answer is that they are not. Clearly, I do not have on record in memory all possible visual versions of A, since I have never experienced most of them. And similarly, I do not have memorized all possible temporal sequences of all possible configurations of muscle motions that write A; indeed, I have yet to perform them and by all accounts I never will. The essential question about our transcription task, therefore, can be stated more fundamentally: How can I recognize and produce the indefinitely various instantiations of A without previous experience of them?

In response to this question let us turn our attention to linguistic theory. The point of departure for transformational grammar is that our competency in language is such that we can produce and understand a virtually infinite number of sentences. As Weimer (1973) has pointed out, there are echoes of Plato's paradoxes in Chomsky's (1965) claim that our competence in language vastly outstrips our experience with it. Chomsky's claim is motivated by the observation that experience with a limited sample of the set of linguistic utterances yields an understanding of any sentence that meets the grammatical form of the language. To explain this competency is, for Chomsky (1966), a central problem in the Theory of Language. But given the points advanced above, the constancy function in action and perception is likewise indicative of a competency that exceeds prior learning. The child, we may note, learns to write A under conditions which restrict her to a small subset of the very large set of A-writing movements. But she is able subsequently to write A with practically any movement pattern she chooses, i.e., she can write A in novel ways. Similarly, limited visual experience with some A's is sufficient to allow the child to identify virtually any A. Thus, acting and perceiving are creative in the sense that language is creative, and I would submit, therefore, that the explanation of this creativity is central to the theories of action and of perception, and at the very heart of our understanding of perceptual-motor skill.

The search for a workable account of the creativity manifest in language has led transformational grammarians to what has been aptly described as "the explanatory primacy of abstract entities" (Weimer, 1973). The idea is that the speaker-listener has at his disposal an abstract system of rules or principles referred to as the deep structure that allows him to generate and to understand an indefinitely large set of sentences referred to as the surface structure. This distinction, drawn in linguistic theory, between deep and surface structure applies to our present concerns in two important respects. The first is the transformational grammarian's view that deep structure is far removed from surface structure; it is argued that although the deep structure determines the surface structure it is not manifested in the surface structure. The importance of this view is that it concurs with Bernstein's (1967) general conclusion in his classic analysis of the coordination and regulation of movement. Referring to the engram or motor-image of an act Bernstein comments: "The higher engram,
which may be called the engram of a given topological class, is already structurally far removed from any resemblance whatever to the joint-muscle schemata...." (p. 49). The essence of Bernstein's (1967) view is that the central substrate for a pattern of movements is a representation of the environment. Following Evarts' (1967) work, Pribram (1971) has argued that the cortical representation can be thought of as "a 'mirror image' of the field of external forces" (p. 246). Thus, the underlying structure is best described as an Image-of-Achievement since it encodes environmental contingencies. An interesting upshot of this view is that action and consequence, which prove to be inseparable conceptually, are also inseparable neurophysiologically.

The second characteristic of the surface-deep structure distinction I wish to touch upon is that the child must come to determine the nature of the underlying deep structure from a limited experience with surface structures. It is assumed by Chomsky and his colleagues that the child essentially "looks through" the utterances she hears to the abstract form behind those utterances. The child is said, therefore, to construct a theory of the regularities of her linguistic experience. Similarly, the child seeing capital A's must determine an abstract representation that will extend over an indefinitely large set of instantiations of that character. And, by the same token, our hypothetical child learning to write the letter A must determine from her limited experience with the set of A-writing movements a theory of how to write A. The abilities to recognize indefinitely various A's, and to write A in indefinitely various ways are based on representations that are abstract and generative, like the grammar Chomsky has in mind for language. We should not be surprised by this conclusion: there is no reason why the nervous system should not solve similar problems in similar ways.

The mathematical group as an example of an abstract structure

Clearly, the form of the representation that allows for the writing of A in novel ways is not motor. That is, it cannot be said to consist of programs of muscle innervation. In the same way, the abstract representation that affords the identification of novel A's cannot be sensory, i.e., it cannot be described as any circumscribed set of sensory properties. We should note that the constancy function in the identification and in the writing of A reveals an indifference of both modes to metrical variation, and suggests rather strongly a dependency of both modes on topological properties (Bernstein, 1967). Thus, common to all capital A's viewed and written is that they are members of a single topological class, while the differences between capital A's, both viewed and written, would be determined by topological differences of a higher order (Bernstein, 1967).

On the foregoing considerations we should argue that the action concept of A and the perception concept of A cannot be represented, respectively, as a particular aggregate of motor elements and as a particular aggregate of sensory elements. Instead, they are more accurately viewed as injunctions or rules specifying how a set of elements should relate, whatever those elements might be.

In attempting to account for constancy in visual perception several students of the problem have appealed to the mathematical concept of group (e.g., Cassirer, 1944; Pitts and McCulloch, 1947). Essentially, a group is any set or collection of elements (and they need not be specified) which can be combined according to a
law such that any combination of them produces an element belonging to the set
itself. The set, therefore, is said to be self-contained or closed.

More formally, a group may be defined as a set $G$ together with a composition
rule which generates for each pair $a$ and $b$ of elements of $G$ a third element $ab$ of
$G$ for which the following conditions hold:

1. The composition rule is associative: For any three elements $a$, $b$, $c$ of $G$: $(ab)c = a(bc)$.

2. There exists an element $i$ in $G$ such that $a \cdot i = i \cdot a = a$. The
element $i$ is known as the identity element.

3. To each element in $G$ there corresponds an inverse in $G$ such
that: $a \cdot a^{-1} = a^{-1} \cdot a = i$.

If we now define a generic concept as a group of transformations (Cassirer,
1944), then it can be said that there is a group $G$, which defines the action
concept of $A$ and another group $g$, which defines the perception concept of $A$.
However, an interesting property of groups is that two groups can be isomorphic,
that is, they can represent the same abstract group, if the manner of internal
interlocking of elements is the same in both cases, even though the elements of
the two groups may differ radically from one another in other respects. Thus,
although the perception concept of $A$ and the action concept of $A$ would appear to
differ because of the different elements with which they work (sensory proper-
ties on the one hand and muscle contractions on the other) they may, in fact, be
identical. The idea is that the two groups, $G$ and $g$, which define the two con-
cepts, have the same internal structure. This speculative conclusion can be
stated more usefully as follows: the abstract structure that affords the identi-
fication of optical instantiations of $A$ also affords the production of motor
instantiations of $A$.

Conclusion

I have considered certain characteristics of a "simple" transcription task
in order to argue that the problems that beset the perception theorist and those
that beset the action theorist are very similar in nature, and thus similar in
the principles needed for their solution. On a less general level I have spec-
ulated that for this transcription task (and I suspect for others) perception and
action may be related through a common abstract structure indigenous to neither.
And finally, I have expressed, implicitly, the view that the Theory of Action is
as much a part of Cognitive Psychology as are the Theory of Perception and the
Theory of Language. If you remain unconvinced of the abstract nature of the
knowledge underlying so-called perceptual-motor activity, consider the following
description of balancing on a bicycle presented by Michael Polanyi (1964). As
the cyclist starts to fall to the right he turns the handlebars to the right,
deflecting the bicycle along a curve to the right. The result of this maneuver
is a centrifugal force pushing the cyclist to the left and offsetting the grav-
tational force pulling to the ground to the right. Consequently the cyclist is
thrown out of balance to the left and responds by turning the handlebars to the
left deflecting the bicycle along a curve to the left, which results in a centri-
ugal force pushing him to the right, etc., etc. In the course of these maneuvers
the cyclist is obeying the following injunction: "adjust the curvature of the
bicycle’s path in proportion to the ratio of unbalance over the square of the
speed." Keeping one's balance on a bicycle is a very cognitive act.
REFERENCES


