

## MEDIATION THEORY AND THE PROBLEM OF PSYCHOLOGICAL DISCOURSE ON 'INNER' EVENTS: PART II\*

*Gottfried Seebass*

The present article attempts to investigate the 'philosophical foundations' of psychology and thereby of the social sciences in general with regard to a central problem, viz. the question of the 'inner'. It does this with special critical reference to an authoritative psychological theory, viz. the so-called 'mediation theory', and tries to show the necessity of interdisciplinary clarification. In the first part mediation theory was introduced as a variant of psychological behaviorism which attempts to substitute for the untenable total neglect of 'inner' events a way of talking about them which is secured methodologically and retains a behavioral foundation. Of the three crucial questions raised by the project the second, viz. question (B), which concerns the theoretical status of the assumed inner mediators could be answered to the effect that they must be "hypothetical constructs" (in the sense of MacCorquodale and Meehl 1948) having the nature of real, phenomenally verifiable events, *if*, indeed, they are to figure as links in (causal) S-R-chains. Conceiving of them as "intervening variables" (and this is theoretically still an open possibility, as a consequence of a purely dispositional description of the phenomena), such a conception would exclude this function. The discussion of 'internalization' of behavior sequences as assumed by mediation theory has led to the conclusion that the relevant phenomena *cannot*, in this case, justify the necessity of a non-dispositional explanation which goes beyond 'external' performances (question (A)). Likewise the subject matter of question (C) remained undecided; that is, the possibilities of phenomenally verifying the assumed 'inner' mediation processes which are only asserted to exist by the theory but are unspecified in content. Therefore, the observable performances in the case of sequential behaviour do not support the mediation theoretic solution of the problem of the 'inner'.

### 5. THE PROBLEM OF 'MEDIATED GENERALIZATION'

This can be expected more easily from the second relevant domain of phenomena: 'secondary' ('mediated') generalization, which has been the main focus of experimental research. The principle according to which the experiments are set up is, notwithstanding numerous modifications in the

\* Part I appeared in the previous issue of *Ratio*, XXIII, 2, 1981.

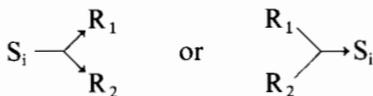
experimental design and evaluation, the same throughout: two or more S-R-connections acquired earlier with at least one common link lead to a new one which was not learned as such. This happens either *directly*, manifested by its more or less probable spontaneous occurrence in the individuals under investigation, or *indirectly*, manifested only by significantly easier learning in subsequent special training. Mediation theory attributes both to an unobservable 'mediation' by the common link or links and is regarded as a demonstration of the occurrence of inner stimuli and responses. Some experiments are implausible a limine because of obvious simplifications. Out of the plausible ones five prominent examples are to be examined:

- (1) To which words a certain stimulus gives rise is established by free association. Two relatively strong associations of this kind are selected, e.g. 'stem' → 'flower, and 'flower' → 'smell'. An artificial expression, 'YOF', is introduced by appropriate training (paired-associate learning) as strongly associated stimulus-word for 'stem'. And now we find that learning the so far unknown connection 'YOF' → 'smell' has been significantly facilitated compared with any other connection 'YOF' → 'X' (Russell/Storms 1955; Mednick/Freedman 1960, Richardson 1962, Berlyne 1965, 53f.).
- (2) An unconditioned stimulus, e.g. blue light, is connected in an associative way with a certain verbal reaction, e.g. the word 'blue'. By way of conditioning the same stimulus becomes a trigger for a second reaction, e.g. salivary secretion. Afterwards a significantly high percentage of the test persons showed a tendency to associate the so far unconnected events, viz. 'blue' → salivary secretion (Osgood 1953, 701ff.; Staats/Staats 1964, 147ff.; Berlyne 1965, 58f.; Hörmann 1967, 186f.).
- (3) A child learns to call another child in its neighborhood 'friend' and soon afterwards also makes use of this title in the case of adult friends and acquaintances of its parents. If it also learns to shake hands with the neighbour's child one may observe, when the occasion arises, that it quite independently transfers this behaviour to adult persons (Osgood 1953, 359; Duncan 1965; cf. Hull, 1943a, 192f. 198; Osgood 1953, 405; Staats/Staats 1964, 99ff.; Hörmann 1967, 185f.).
- (4) Stimulus-words, e.g. 'tomato' and 'cherry', which are (as is shown by appropriate prior investigations) jointly but to different degrees associated with certain other words, e.g. to a strong degree with 'red' and to a lesser degree with 'edible', have to be learned in the main test as joint subordinated terms associated with them. The success in

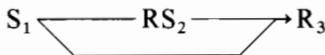
learning increases in proportion to the association strength (Underwood/Richardson 1956, Griffith/Spitz/Lipman 1959).

- (5) A sequence of (appropriately selected) words is presented, and the test person has to decide in each case (of a new occurrence of a word) whether that word has already occurred or not. The number of wrong recollections increases in a significant way in the case of those words which have (according to prior investigations) a strong associative connection with a word that had in fact been presented (Underwood 1965; Hörmann 1967, 153).

Why are these and similar observations observations about generalization performances? And why about generalization performances that rest on the intervention of inner mediators? Mediation theory interprets them as follows. (1) is an example of a chain in which the observable connecting links between initial stimulus and final response cease to exist. The three existing associations jointly constitute a four-part chain 'YOF' → 'stem' → 'flower' → 'smell'. If the new association 'YOF' → 'smell' occurs it is assumed that parts of the responses connected with 'stem' and 'flower' — parts of their "meaning" — function as unobserved mediators, i.e. that covertly the whole chain reaction takes places (Staats/Staats 1964, 168; Berlyne 1965, 53.85f.; Hörmann 1967, 189). It is a generalization performance because the result consists in a joint connection of one stimulus with several responses or one response with several stimuli, respectively. In the first case people speak of "*response generalization*", in the second of "*stimulus generalization*":



The connections in the experiment admit of both conceptions: if one concentrates on the shared reaction ('smell') one gets a stimulus generalization, if one concentrates on the shared stimulus ('YOF') one gets a response generalization. Following Berlyne we can speak in both cases of generalization in accordance with the principle of "*chaining*"<sup>9</sup>:



<sup>9</sup> Cf. Berlyne 1965, 53f. and 85f. Strangely misjudging the fact that "stimulus-" and "response-chaining" share their theoretical fundamentals (a fact which is recognizable in his own diagrams) Berlyne treats them as two kinds of generalization existing separately beside one another and applies different experiments to them. This conceals that the apparent distinction is only a matter of different 'focus'.

In (1) 'RS<sub>2</sub>' is merely split up into two further links: 'stem', 'flower'. According to the simple explanation which Underwood himself gives for his experiment (1965, 382f. 390f.), (5), too, has to be analyzed in accordance with the chaining principle. In the case of wrong recollection the test person reacts with 'Yes, it has occurred already' to a stimulus-word that had in fact not been presented. The immediate trigger for this reaction is not the word itself but the part of its meaning which it shares with a stimulus word that has actually occurred. The latter functions as inner mediator: the new word leads to the relevant partial 'meaning reaction' which in turn (thanks to prior presentation of the other word associated with it) leads to the objectively wrong recollection response.

What about example (2) in this regard? Here, too, an analysis according to the chaining principle would be imaginable provided one begins with the (not implausible) presupposition that the unconditioned associative connection also holds in the converse direction to the one mentioned above: in this case perception of the word 'blue' could trigger the inner perception of blue light which, on account of the conditioning process, leads to salivary secretion. In fact, mediation theory does give this explanation but only as part of a more complex analysis which becomes necessary because the inner reaction must not be tied without exception to the occurrence of blue light (cf. Osgood 1953, 701ff.; Staats/Staats 1964, 147ff.; Berlyne 1965, 58f.). Analyzed completely, there are not two but three associative connections which give rise to the new association 'blue' → salivary secretion: blue light → "blue-meaning", 'blue' → "blue-meaning" and blue light → salivary secretion. But these connections cannot be combined in one continuous chain. We have the first link of the required chain, viz. 'blue' → "blue-meaning", but the second is missing: "blue-meaning" → salivary secretion. In order to render the mediating function of the inner reaction "blue-meaning" comprehensible we have to assume that the connection which forms the second link in our chain arises as a consequence of the two existing connections, i.e. in general: the connection of a stimulus with two different responses leads to an associative connection between both response-events (cf. Hörmann 1967, 186). We can speak in this case of generalization due to the principle of "transfer". Only after this principle has come into force can the chaining principle become operative and lead to the desired result.

The combination of both principles yields the schema, called "secondary" or "mediated generalization" in the precise sense in mediation-theoretic literature. Just as in the case of simple generalization according to the principle of chaining, we are free to choose viewing it as stimulus- or response-generalization, depending on which event we are concentrating

on.<sup>10</sup> Besides (2) examples (3) and (4) are also recognized as cases of mediated generalization. In the case of (3) we can see this at once since here again we have a stimulus (neighbour's child) which is connected with different responses ('friend', shaking hands) and a second stimulus which is initially connected with one of them only and which is independently associated with the other one because of joint association with the first stimulus. More difficulties arise in the case of (4). Typically the cited authors do not provide an exact analysis of the assumed mediation occurrences. To be sure, it is *possible*, to analyze this example in the manner of (2) and (3). We may say that two stimuli ('tomato', 'cherry') are associated with a certain inner response ("red-meaning") and that during the learning phase of the main test one of them is connected with a second response ('red') in the sense of a 'learning hypothesis' and that subsequently an independent transfer to another stimulus takes place manifesting itself in faster progress in learning—provided the hypothesis was correct.

But this analysis which is circumstantial and goes far beyond the observations mentioned in (4) already shows that there is a considerable disproportion between empirical evidence and mediationtheoretic explanation. This is even clearer in the case of (5) where only two things are observed (in the relevant respect) in the experiment itself: the stimulation of the test person by a word and his later response 'Yes, it has occurred already' to a different word which the experimenter knows is relatively strongly connected with the first word in the average speaker. Underwood's explanation, which invokes the chaining principle, is unnecessarily complicated and at the same time not sufficiently sophisticated. It is not sophisticated enough because it does not take into account the methodical recourse to memory and the complications arising from that (time lag, memory capacity and sincerity of the test persons etc.). It is more complicated than necessary due to its recourse to inner mediators where dispositional explanations would suffice.

A stricter analysis of the logic of the experiment is as follows. It is presupposed that each stimulation by a word — let us say, 'A' or 'B' — leaves a memory trace in the test person which is present during the whole time of the experiment and which is marked in its time position, viz. as 'having come into being earlier' (notation:  $[X_{t,n}]$ ). Further it is presupposed that the test person, sincerely obeying certain verbal instructions, reacts by

<sup>10</sup> Berlyne (1965, 54ff. and 87ff.) treats them, too, as different kinds of response. But the fact that the theoretical fundamentals are the same is impressively demonstrated by the fact that the same experiment cited by Hull (1943a, 192f. 198) as an example for secondary stimulus generalization could be cited by Hörmann (1967, 185ff.) as an example for secondary response generalization.

uttering 'Yes, it has occurred already' ('C') if and only if the memory-trace which is now produced by a word-stimulation is of the same type as the earlier one. That is, we presuppose the validity of the general principle:  $[X_{vi-m}] \wedge [X_{vi}] \rightarrow C$ . (' $\wedge$ ' denotes the logical conjunction 'and'.) It is then possible to say more on the basis of the test observations than is externally visible. At time  $t_1$   $A \rightarrow [A_{v1}]$  occurs and at time  $t_2$   $B \rightarrow [B_{v2}]$  and  $[B_{v1}] \wedge [B_{v2}] \rightarrow C$  occur; the latter is inferred from the actual stimulation by B and the subsequent occurrence of C. The theoretically interesting question is: how do we obtain  $[B_{v1}]$  which has to be presupposed in this case? Because the earlier B-stimulations are now missing, and also since there do not seem to be any explanatory alternatives, one takes recourse to the word 'A' that had in fact occurred and which has the strongest possible connection with 'B' among the other test words, and infers backwards (for time  $t_1$ ) something like the following:  $A \rightarrow [A_{v1}] \rightarrow [B_{v1}]$ , or:  $A \rightarrow ([A_{v1}] \wedge [B_{v1}] \wedge \dots)$ . As one can see many presuppositions are required in order to make the substitution of an austerer description of the occurrences at time  $t_1$  by a richer one appear to be conclusive. At any rate, the scheme of the experiment would have to look much more complicated than the explanation referring to the chaining principle given above. On the other hand even if it were conceded that the description had to be expanded with regard to time  $t_1$  this would not commit us to the description which is essential for mediation theory, viz. the one which refers to inner mediation processes. For the assumed influence of earlier stimulation on later response can be represented dispositionally, too:  $A \rightarrow (B \rightarrow C)$ .

Thus here, just as in the case of behaviour sequences the question arises whether people do not overestimate the weight of the evidence available or even whether the explanations apply to the phenomena at all. Is there, e.g., really in all cases an independent *transfer* to a new case and consequently a genuine generalization performance or is it perhaps only another *application* of quite familiar capacities? The latter is probable in the case of (4) since what is established by prior investigation is not the joint association of both stimulus-words with a particular inner response ("red-meaning") but only the association with just that word ('red') which is sought in the main test, and it is not a great surprise if those associations that earlier proved to be the strongest continue to do so later on. But even if we consider only those cases in which a new connection in fact turns up, two critical questions remain: Is it really *necessary* to explain this new connection in terms of 'mediation' via a jointly associated further event? If yes, how is the recourse to *inner* mediators to be justified?

Let us first address ourselves to the question of mediation as such. A more precise analysis of generalization due to the chaining principle indicates that it is dispensable. As long as one presupposes that the com-

plete chain consisting of three links occurs (overtly or covertly) it is clear that a connection between stimulus  $S_1$  and response  $R_3$  comes into being, namely as a consequence of the existing partial connection  $S_1 \rightarrow RS_2$  and  $RS_2 \rightarrow R_3$ . But the mere fact that there is a *manifold* of causes and effects which is simply a consequence of there being a causal chain consisting of several links must, of course, not be taken as a criterion of the occurrence of generalization performances. This is justified only if both events which are causally connected with a third offer *alternative* possibilities of behaving, i.e. if the linear chain consisting of three links is replaced by a ramified chain consisting of two links of the kind explicated on page 31. Mediation theory gets this result by means of its principle of "short-circuiting" according to which (under certain conditions) a response which occurs in the sequence is completely or partly, overtly or covertly, anticipated at an earlier place of the chain. If we apply this to our special case we get the required direct connection (cf. Staats/Staats 1964, 168 Fig. 4.22 and above all Berlyne 1965, 53 Fig. 3-1 and 85 Fig. 4-1):



But this means that even according to the explanation which is proposed by mediation theory itself a generalization obtains only if the relevant event  $RS_2$  does *not* 'mediate' any longer! If the short-circuiting principle operates mediation becomes redundant. One could, at most, say that without prior mediation by  $RS_2$  the short-circuited connection  $S_1 \rightarrow R_3$  would not have become *possible*. In this way one could save mediation with respect to the *first* occurrence of a new performance but forfeit all further claims.

But even then it remains a merely theoretical attachment which is unnecessary for the explanation of the experiments. A dispositional explanation is sufficient. Generalization due to the *chaining-principle* is based on the (inductive) statement that if somebody is equipped with two associations of the form ' $S_i \rightarrow RS_j$ ' and ' $RS_j \rightarrow R_k$ ' which have a certain strength then he is also equipped with the association schema ' $S_i \rightarrow R_k$ ' of equal or lesser strength. But from a 'molar' point of view each association is a disposition. If we make use of a schematic representation, this principle can be formulated as 'material' implication (notation: ' $\supset$ ') between three dispositions:

$$(P1) \quad D_{[S_i \rightarrow RS_j]} \wedge D_{[RS_j \rightarrow R_k]} \supset D_{[S_i \rightarrow R_k]}^{11}$$

<sup>11</sup> To be read as quantified sentence (cf. note 2): *whenever* an organism is in a dispositional state in which provided it were stimulated by an event of type  $S_i$  it would react with behaviour of type  $RS_j$  and at the same time in a dispositional state in which stimulation of type  $RS_j$  would give rise to behaviour of type  $R_k$  in it, *then* it is also in a dispositional state in which provided it were stimulated by an  $S_i$ -event it would react with  $R_k$ -behaviour. — Any further complication may be neglected here.

But then the 'mediating' function of  $RS_j$  is no longer mentioned. Corresponding things may be said about generalization due to the *transfer principle*. In this case it is the matter of a connection of the form ' $RS_j \rightarrow RS_k$ ' or ' $RS_k \rightarrow RS_j$ ', which is dependent on two connections of the form ' $S_i \rightarrow RS_j$ ' and ' $S_i \rightarrow RS_k$ '. (The mutual character of the dependent connection arises from the fact that each concrete event may be formally viewed as both ' $RS_j$ ' and as ' $RS_k$ '. In order to get the desired mediated generalization a causal relation in one direction would be sufficient.) We write:

$$(P2) \quad D_{[S_i \rightarrow RS_j]} \wedge D_{[S_i \rightarrow RS_k]} \supset D_{[RS_j \rightarrow RS_k]} \wedge D_{[RS_k \rightarrow RS_j]}.$$

In doing so we make no more reference to mediation processes than in (P1). But by applying both these principles we can explain all the examples mentioned above insofar as they describe genuine generalization performances at all.

Consequently the assumption of mediation is not *necessary*. The most that can be said in favour of mediation theory is that it attempts to develop a further-reaching explanation in order to give a '*deeper*' account of the facts which are only superficially described by (P1) and (P2). But why do we need such an explanation? One motive could be the general interest in supplementing a 'molar' behaviour theory with a more sophisticated 'molecular' one. Let us accept this for a while even if, at the same time, we have to maintain that the *observations* we have referred to so far do not commit us to any recourse to the 'inner'. The question remains how we are to provide a positive justification for the special mediationtheoretic explanations. *Why* should we believe that the dispositional connections, which are described by (P1) and (P2), can be made comprehensible only if we assume at least initial mediation by event  $RS_j$  or  $S_i$ , respectively which occurs in the antecedent of (P1) and in that of (P2)? *Why* is it superficial to assume that the observed generalization performances occur 'immediately'? Insofar as mediation theory attempts to find an answer to this question (cf. e.g. Hull 1943a, 194; Osgood 1953, 359ff.; Staats/Staats 1964, 147; Berlyne 1965, 53f.) it refers to different degrees of similarity. The generalization performance of an animal conditioned to react to a crimson stimulus and later on to transfer the relevant response to a brick-red or orange stimulus, may be described as 'primary' (unmediated) generalization because the stimuli stand in obvious though continuously decreasing sensory similarity relation to each other. That mediation does not actually occur is indicated by the fact that the strength of the association decreases in proportion to the decrease in similarity. But on the other hand, it is assumed that a human being cannot but use a mediator if he shows the same response to perceptually quite dissimilar stimuli and if

the strength of association is not affected by continuous decrease of similarity up to a certain abrupt limit. Is this a plausible argument?

Hardly. The criterion of graded perceptual similarity is imprecise and always dependent on the underlying system of perceptual qualities. (The quality-space of a dog whose capacity of olfactory discrimination exceeds its visual one ought to be fundamentally different from the quality-space of a human being who is predominantly visually oriented.) There is no 'ultimate', 'objective' system. It is an open question whether the system applied by psychologists is sophisticated enough to capture all qualitative differences of the organism under investigation which belong to a certain class of stimuli and to which the organism reacts in the same way. Another open question is the converse, namely whether the organism in fact realizes all those differences which are recognized by the psychologist (possibly by means of expensive experimental apparatus). Corresponding things may be said about responses. *Every* S-R-connection rests on qualitative differentiations and non-differentiations. It goes without saying that on an elementary level internal differences are *not* differentiated from each other, just as the marked *distinction* of differentiated qualities is self-evident. Assume that an organism starts with a purely yellow stimulus (as seen by the psychologist) and abruptly finishes generalizing at the point where greenish-yellow becomes yellowish-green. Then this does not demonstrate that the organism connects these stimuli with a joint mediator before it reacts but only that it draws a sharp qualitative boundary *here*. Were this not the case, one would also have to regard sameness of reaction to 'purely' yellow stimuli as mediated since it is possible to introduce further differentiation within this relatively narrow range. At any rate, we would have to extend the operation of mediation up to a limit where the psychologist speaks of 'qualitatively indistinguishable' stimuli. S-R-theory would imply a kind of qualitative 'atomization' that would render it totally impractical. Even if one were prepared to concede this, the question remains whether the introduction of a mediator could be of any use in explaining higher generalization performances. *If* it is a fact that the relevant stimuli and responses are dissimilar, then, of course, their joint connection with the mediator is no less problematic than their connection with the stimulus or the response itself. If there is a problem at all it is only shifted.

We have to assume, therefore, that there are no *general* reasons which would justify the allegedly 'deeper' mediationtheoretic explanation of the connections as described in (P1) and (P2). Such an explanation is possible only if we show, in *each* relevant case, that a mediation process actually takes place. There are cases in which it can be observed *externally* (cf. Hull 1943a, 192f.; Berlyne 1965, 85). Of course, no objection can be raised against a mediationtheoretic explanation in such cases. But then it

loses its special theoretical advantage over usual S-R-explanations. The crucial cases are those in which external observation is lacking. Consequently, the main problems arise regarding the justifiability of the assumption of *inner* mediators. As general considerations are not sufficient for this purpose we can rely only on inference from analogy from observable to unobservable cases as long as inner mediators cannot be directly verified (by 'breaking up' the 'black box'). Inference from analogy has been explicitly acknowledged by some authors as the basis of their assertions about inner mediation processes in the case of generalization performances (e.g. Underwood 1965, 382f.). But now the same difficulties arise as in the case of inference from observed behaviour sequences to 'internalized' chains. Transfer from observably mediated to other performances is not justified. We cannot exclude the possibility of changes or abridgements if 'internalization' occurs. Besides this the principle of short-circuiting shows that mediation theory itself recognizes this possibility. Furthermore, the nature of the assumed events remains undetermined in this case, too. The requisite specifications could be provided only *after* the 'molar' behaviour theory had been supplemented, phenomenally as well as theoretically, by an appropriate 'molecular' theory. But this move is not open to the mediation theorist because of his restriction to a *behaviorist* basis.

## 6. THEORETICAL CONSEQUENCES

Thus all our questions concerning mediation theory have been answered. Neither learned behaviour sequences nor the phenomena of 'secondary' generalization necessitate recourse to the 'inner' (question (A)). Both are model cases of mediationtheoretic explanation with exemplary importance. But one cannot exclude the possibility that 'breaking up' the 'black box' and introducing 'inner mediators' might be necessary in other cases. If we want to do that in accordance with a 'molecular' supplement of a 'molar' S-R-theory the mediators must be real events which can be causally connected with external events. In this case they would have the theoretical status of hypothetical constructs (question (B)), and it would have to be possible in principle to exhibit them in the physical (physiological) or mental 'inner' of the beings concerned. If such a proof of their existence had been produced or were at least acknowledged as a verifiable empirical consequence of the theory, mediationtheoretic explanations could retain their importance even if recourse to the 'inner' could not be shown to be necessary in any single case. But under such conditions the behaviorist basis would definitely have been abandoned. Mediation theory faces a dilemma: as a behaviorist theory it cannot give an account of why

we have to go beyond the domain of observable behaviour; as a theory wishing to include reference to the unobservable it is handicapped by its behaviorist preconceptions. The strange ambiguity of its declarations about the theoretical status of mediators is a direct consequence of this intermediate position. Phenomenal verification of the inner performances assumed by mediation theory (question (C)) remains remote precisely because it tries to escape this obligation. But if it had tried to bring about the desired verification it would have found itself with the very methodological difficulties which it wanted to avoid. Thus it is obvious that mediation theory has failed to solve the problem of talk about the 'inner'.

More important than this statement itself are the conclusions we have to draw from it. One tends to be misled here. In European psychology behaviorism has never played as dominant a role as in America. A representative of American 'cognitive psychology' will also be tempted to regard the negative result as confirmation of his own position and to say that he had long known that an adequate treatment of 'inner' phenomena is impossible on a behaviorist basis and that for this reason he had abandoned its methodical restrictions a long time ago. This would certainly be the wrong reaction. In fact, it is just the other way round: if a theory can be induced to make unjustified statements about the 'inner' although its very prejudice is in favour of the 'outer' which should prevent it from doing so, the temptation is even larger for theories that are devoid of such scruples. An example may show that this suspicion is fully justified. Every adult human being can (under normal circumstances) distinguish a voluntary movement of his own body (e.g. violently bending the right arm) from an involuntary one, even if there is no difference in his external behaviour. 'Wanting' and 'wishing' are apparently events in the mental 'inner' which are accessible only through 'introspection'. Probably one may take it for granted that they do not yet occur on more elementary phylogenetic or ontogenetic levels. The exact boundaries are unclear. However, we reach a point where it is possible to describe behaviour performances as 'desired', 'wanted' etc. In ethology and developmental psychology this happens without reflection (cf. e.g. Piaget 1950, ch. III; Rensch 1973, chs. 4-5). But we lack nearly all conception of *what* it is that is attributed to animals or children in such cases and under what observable *conditions* these attributions are permissible. While mediation theory has at least attempted to base its statements about the inner on experiments ethologists and development psychologists have simply lost sight of the problem. Reference to the 'inner' has become something trivial which is no longer the subject of theoretical reflection.

If one does face the problems there remain three theoretical possibilities. First of all, one can adopt the procedure of logical behaviorism and try to

*reduce* the phenomena which seem to necessitate recourse to the 'inner' to *behaviour*. This has been suggested by Ryle (1949, ch. 3) and Wittgenstein (1953, secs. 611ff.) in very general terms for 'volition'. The simplest method of carrying out this programme in greater detail would probably be by means of an explication of 'goal-directed' action with the help of non-intentional teleological explanatory models, and one can find several proposals of this sort in recent literature (Rosenbluth/Wiener/Bigelow 1943; Sommerhoff 1950; Braithwaite 1953, ch. X; Nagel 1961, ch. 12; Taylor 1964, ch. 1). But they are not satisfactory either, in several respects (Woodfield 1976), and even (to my mind) the most sophisticated proposal of this kind (Bennett 1976, chs. 2-3) proves to be gratuitous on closer inspection. Whether the behaviorist reduction<sup>12</sup> can ever be successful (either on the whole or at least for the psychologically most interesting cases) is—cautiously expressed—as yet completely unknown. But as long as it has not been brought about anybody who does not want to transcend the confines of physicalism has only the second theoretical possibility at his disposal, viz. total *abandonment* of scientific investigation of the 'inner' or *postponing* it until sufficiently complex neurophysiological descriptions are available. This self-imposed restriction is honest and respectable. But it implies the deliberate methodical neglect of the most interesting questions for psychologists which do not, incidentally, reappear by themselves in the course of neurophysiological progress. For how is it possible to look for the physiological basis of something which cannot any longer be identified even in a prescientific form? The basis of doing psychology is still methodical behaviorism the 'counterintuitive' phenomenal restrictions of which provided the main motivation for mediation theory's attempt to integrate the 'inner' into a 'molar' behaviour theory.

The way out of this dilemma can only lie in positive reference to just those 'intuitions' which make behaviorism appear inadequate independently of progress which is to be expected in neurophysiological research or of the possibility of logical reductions. This is the third theoretical alternative. Strict physicalism has in this case been abandoned from the outset. To some people this may seem to be high treason. But so far I cannot see how it can be avoided if one does not simply want to close one's eyes to the problems. The phenomenal irreducibility at least of the mental seems to be irrefutable. It is more appropriate to see a fundamental

<sup>12</sup> The behaviorist *reduction* should not be mistaken for the strengthening of the behavioural conditions on which the external *attribution* of (irreducible) inner performances is dependent. All writings cited have made essential contributions towards the latter task and offer at least a considerably more solid basis than the uncritical procedure of ethology and developmental psychology mentioned above.

difficulty for this position in the fact that one has obviously returned to 'introspection', i.e. with just that method which behaviorism rejected for good reasons. That would in fact be fatal, but a simple relapse can be avoided. It is unnecessary to gather 'introspective' evidence exclusively under artificial experimental conditions as was the norm at the beginning of this century. On the contrary, as 'cognitive psychology' has recently successfully demonstrated it can be collected in just those situations where it naturally occurs in human life. A much greater danger than a nostalgic return to Würzburg is presented by the following difficulty. We have deliberately put the expressions 'introspection' and 'intuition' in quotes. In both cases we are concerned with metaphors which are far more a sign of embarrassment than of an adequate description of the subject matter. If talk about the 'inner' is not to be taken in the literal spatial (physical) sense it becomes ambiguous (cf. Ryle 1949, 41ff.). One cannot speak of 'intuition' in its usual sense even where mental 'images' are concerned — as in the case of the classical mnemotechnic procedures which have only recently been more thoroughly examined. This applies, of course, to an even greater extent to representations of other sensory fields, sensations and emotions. Above all, the metaphors are totally inadequate in an area which has played the most important role in reconsidering 'intuitive' and 'introspective' evidence: language.

If Chomsky (1965, ch. I, sec. 4; 1969, 81) takes 'intuitions' and 'introspections' of a competent speaker besides observations about linguistic behaviour as a basis for a theory of grammar, there can, of course, be no talk about the informants or the linguists "*looking into*" their mental 'inner'. On the contrary, the concern is with (meta-) linguistic *propositions* about the well-formed character of sentences or structural relations which are called 'intuitive' only because they cannot be verified merely by observing linguistic behaviour. But in this negative statement the decisive factors remain obscure. *What* is the basis of grammatical judgments? Chomsky and his followers have spoken of a special form of 'knowledge'. But apart from the fact that this explanation is also more or less negative its theoretical significance is doubtful simply for the reason that the 'intuitive' judgments of (apparently) competent speakers are actually very diverse, and cannot be made compatible with each other simply by 'recourse' to their verification conditions. The situation becomes much more complicated if one advances from the area of phonology and syntax, in which at least the *objects* of 'intuitive' propositions are 'external', to the area of semantics where the largest part of that which belongs to the speaker's 'knowledge' does not manifest itself in behaviour (cf. Seebass 1980, 457ff.). One is inevitably led astray if one takes the metaphor 'inner vision' as a model for the theoretical explication of 'understanding'

expressions (cf. Tugendhat 1976, lectures 9–10). But it is certainly not obvious how one can conceive of it positively if the 'innerliness'-aspects have been acknowledged. The temptation to seek comfort in the alleged certainties of behaviorism is just as large as the temptation to declare 'intuitive' propositions to be unquestionable. But that the latter cannot be a way out is one of the insights nobody can escape after Wittgenstein. Even if the question of *verifying* such propositions is taken to be senseless the question of their *intelligibility* remains (in form of the question what the conditions for *learning* the relevant expressions are) leading back to the old problems. The problem of psychological talk about the 'inner' has not been solved in any of the classical positions. On the contrary, it is only our seeing that and why they are insufficient which promises to give rise to a formulation of the problem which makes all the difficulties involved visible. The solution is a task which must be shared by all the disciplines concerned.

(translated by Graham Spinks and Eckart Hoffmann)

#### REFERENCES

- B. Aune, 'On Thought and Feeling', *Phil. Quart.* **13**, 1–2. (1963).  
 B. Aune, *Knowledge, Mind, and Nature*, New York (1967).  
 Jonathan Bennett, *Linguistic Behaviour*, Cambridge (1976).  
 D. E. Berlyne, *Structure and Direction in Thinking*, New York (1965).  
 C. V. Borst (ed.), *The Mind-Brain Identity Theory*, London (1970).  
 R. B. Braithwaite, *Scientific Explanation*, Cambridge (1953).  
 C. S. Chihara and J. A. Fodor, 'Operationalism and Ordinary Language: a Critique of Wittgenstein', *Amer. Phil. Quart.* **2**, 281–295. (1965).  
 N. Chomsky, 'A Review of B. F. Skinner's *Verbal Behavior*', orig. in: *Language* **35** (1959), repr. in J. A. Fodor and J. J. Katz (ed.), *The Structure of Language*, pp. 547–578. Englewood Cliffs (1964).  
 N. Chomsky, *Aspekte der Syntax-Theorie*, orig. Cambridge/Mass. (1965), dt. Frankfurt (1969).  
 N. Chomsky, 'Linguistics and Philosophy', in S. Hook (ed.) *Language and Philosophy*, New York (1969).  
 C. P. Duncan, 'Mediation in Verbal Concept Learning', *J. Verb. Learn. and Verb. Behav.* **4**, 1–6 (1965).  
 H. Feigl, *The 'Mental' and the 'Physical'*, Minneapolis (1967).  
 J. A. Fodor, *Psychological Explanation*, New York (1968).  
 C. F. Graumann (ed.), *Denken*, Köln (1965).  
 B. C. Griffith, H. H. Spitz and R. S. Lipman, 'Verbal Mediation and Concept Formation in Retarded and Normal Subjects', *J. Exp. Psych.* **58**, 247–251 (1959).  
 E. v. Holst and H. Mittelstaedt, 'Das Reafferenzprinzip', *Die Naturwiss.* **37**, 464–476 (1950).  
 H. Hörmann, *Psychologie der Sprache*, Heidelberg (1967).  
 C. L. Hull, 'Knowledge and Purpose as Habit Mechanisms', *Psych. Rev.* **37**, 511–525 (1930).  
 C. L. Hull, *Principles of Behavior*, New York (1943a).  
 C. L. Hull, 'The Problem of Intervening Variables in Molar Behavior Theory', *Psych. Rev.* **50**, 273–291 (1943b).  
 G. Humphrey, 'Some Reflections upon Gilbert Ryle's Considerations', *Acta Psych.* **9**, 197–200 (1953).  
 K. S. Lashley, 'The Problem of Serial Order in Behavior', in L. A. Jeffress (ed.), *Cerebral Mechanisms in Behavior*, pp. 112–136, New York (1951).

- R. A. Littman, 'Mr. Ryle on "Thinking"', *Acta Psych.* **10**, 381-384 (1954).
- K. MacCorquodale, 'On Chomsky's Review of Skinner's *Verbal Behavior*', *J. Exp. Anal. Behav.* **13**, 83-99 (1970).
- K. MacCorquodale and P. E. Meehl, 'On the Distinction Between Hypothetical Constructs and Intervening Variables', *Psych. Rev.* **55**, 95-107 (1948).
- N. Malcolm, 'The Myth of Cognitive Processes and Structures', in Th. Mischel (ed.), *Cognitive Development and Epistemology*, New York/London (1971).
- S. A. Mednick and J. L. Freedman, 'Facilitation of Concept Formation Through Mediated Generalization', *J. Exp. Psych.* **60**, 278-283 (1960).
- R. Meili and H. Rohracher, *Lehrbuch der experimentellen Psychologie*, Bern (<sup>2</sup>1966).
- G. A. Miller, E. Galanter and K. H. Pribram, *Plans and the Structure of Behavior*, New York (1960).
- E. Nagel, *The Structure of Science*, New York (1961).
- U. Neisser, *Kognitive Psychologie*, orig. New York (1967), dt. Stuttgart (1974).
- U. Neisser, *Kognition und Wirklichkeit*, orig. San Francisco (1976), dt. Stuttgart (1979).
- J. O'Connor (ed.), *Modern Materialism: Readings on Mind-Body Identity*, New York (1969).
- Ch. Osgood, *Method and Theory in Experimental Psychology*, New York (1953).
- J. Piaget, *Der Aufbau der Wirklichkeit beim Kinde*, orig. Neuchâtel (1950), dt. Stuttgart (1975).
- B. Rensch, *Gedächtnis Begriffsbildung und Planhandlungen bei Tieren*, Berlin/Hamburg (1973).
- J. Richardson, 'The Learning of Concept Names Mediated by Concept Examples', *J. Verb. Learn. und Verb. Behav.* **1**, 281-288 (1962).
- A. Rosenblueth, N. Wiener and J. Bigelow, 'Behavior, Purpose, and Teleology', *Phil. Science* **10**, 18-24 (1943).
- W. A. Russell and L. H. Storms, 'Implicit Verbal Chaining in Paired-Associate Learning', *J. Exp. Psych.* **49**, 287-293 (1955).
- G. Ryle, *Der Begriff des Geistes*, orig. London (1949), dt. Stuttgart (1969).
- G. Ryle, *Denken*, orig. in: *Acta Psych.* **9**, (1953), dt. repr. in Graumann (ed.) *loc. cit.*, 461-466 (1965).
- G. Seebaß, *Das Problem von Sprache und Denken*, Frankfurt (1981).
- W. Sellars, 'Empiricism and the Philosophy of Mind', *Minn. Stud.* **1**, 253-329 (1956).
- W. Sellars, 'Language as Thought and as Communication', *Philos. Phen. Res.* **29**, 506-527 (1968-69).
- B. F. Skinner, *Science and Human Behavior*, New York (1953).
- B. F. Skinner, *Verbal Behavior*, New York (1957).
- G. Sommerhoff, *Analytical Biology*, London (1950).
- A. W. Staats and C. K. Staats, *Complex Human Behavior*, New York (1964).
- Ch. Taylor, *The Explanation of Behaviour*, London (1964).
- E. Tugendhat, *Vorlesungen zur Einführung in die sprachanalytische Philosophie*, Frankfurt (1976).
- B. J. Underwood, 'False Recognition Produced by Implicit Verbal Responses', *J. Exp. Psych.* **70**, 122-129 (1965).
- B. J. Underwood and J. Richardson, *Das verbale Lernen von Begriffen als Funktion der Instruktion und des Dominanzniveaus*, orig. in: *J. Exp. Psych.* **51**, (1966), dt. repr. in Graumann (ed.) *loc. cit.*, 444-457 (1965).
- W. E. Vinacke, *The Psychology of Thinking*, New York (<sup>1</sup>1952).
- J. B. Watson, Psychology as the Behaviorist Views it, orig. in *Psych. Rev.* **20**, (1913), dt. repr. in: Watson a.a.o., 13-28 (1968).
- J. B. Watson, 'Is Thinking Merely the Action of Language Mechanisms?', *Brit. J. Psych.* **11**, 87-104 (1920-21).
- J. B. Watson, *Behaviorism*, orig. Chicago (1924-25), dt. Köln (1968).
- W. M. Wiest, 'Some Recent Criticisms of Behaviorism and Learning Theory', *Psych. Bull.* **67**, 214-225 (1967).
- L. Wittgenstein, *Philosophische Untersuchungen*, Oxford (1953).
- A. Woodfield, *Teleology*, Cambridge (1976).