

# INTROSPECTION AND PRIVILEGED ACCESS IN FOLK–PSYCHOLOGICAL EXPLANATIONS

---

*Josef Quitterer*

UDK 159.9:398

## *Introspection as a problem for the “theory theory”*

“Folk psychology” is defined here as “a conceptual framework... used by ordinary people to understand, explain and predict their own and other people’s behavior and mental states.”<sup>1</sup> In philosophical naturalism there is widespread consensus concerning the theoretical character of folk psychology. This theoretical or functional interpretation is called “theory theory” (it is the theory that folk psychology is a theory). In the theory theory folk psychology is seen as an explanatory system which can be compared with scientific explanations. Folk psychology — according to this interpretation — is basically an explanatory theory which has a hypothetical character. As in science nomological correlations are relevant for the explanatory power of folk psychology.<sup>2</sup> The general form of folk–psychological explanations is described in the following way:

“When someone is in so–and–so combination of mental states and receives sensory stimuli of so–and–so kind, he tends with so–and–so probability to be caused thereby to go into so–and–so mental states and produce so–and–so motor responses.”<sup>3</sup>

The theory theory has a clear advantage in the naturalization of mental concepts or descriptions: If folk psychology is basically an explanatory theory it can be compared with other theories with regard to its explanatory power. The superiority of the explanations of, for example, neuroscience over folk–psychological explanations could then be proved on the basis of the empirical success of neuroscience.

1 Eckardt (1994), 300.

2 Many theory–theorists have weakened the claim that folk psychology has a deductive–nomological structure. Churchland, for example, now holds a paradigm–based theoretical structure instead of a DN— one; see Churchland (1991) and Caruana (1999), 147ff.

3 Davies & Stone (1995c), 10.

There is, however, a difficulty in this kind of theory comparison: A folk-psychological explanation can be confronted or compared with a scientific one only if there is empirical evidence for testing the different explanatory hypotheses. A lawlike generalization or hypothesis can be tested if and only if it is possible to verify the presence of all relevant *antecedentia* which together form the *explanans*. The intersubjective access to all (relevant) *explanantia* of a theoretical hypothesis is a necessary condition for the objective evaluation of its explanatory power.<sup>4</sup> But the intersubjective access to the antecedent conditions of many folk-psychological explanations is difficult, if not impossible, because they can be verified only in an act of introspection. It is the special label “being conscious of one’s own inner states” which seems to protect first person mental explanations against scientific evaluation and theory comparison. Introspection seems to guarantee an infallible access to the inner mental “material” which can be used as *antecedens* in the *explanans* of folk-psychological explanations.<sup>5</sup> It seems that introspective awareness of the content of mental states establishes explanatory conditions which renders folk-psychological explanations incomparable with scientific explanations. Introspective awareness apparently guarantees immunity against scientific testing.<sup>6</sup>

### *The naturalization of introspection*

The immunity which introspection seems to provide for some *antecedentia* of mental explanations has to be abolished in order to reconstruct folk psychology as a hypothetical theory which can be compared with other hypothetical theories with regard to their empirical adequacy. In philosophical naturalism there are different strategies to overcome the introspective privileged access. These strategies could be used as arguments for the theory theory. I classify these strategies according to two main positions in philosophy of science regarding the relationship between the theoretical and the observational level — the *standard view* and the position of *theory-holism*.

In the standard view it is presupposed that the evidence which provides the empirical basis for the *antecedens* conditions of the *explanans* is equally available for different observers. Empirical evidence, then, provides an intersubjective criterion for the adequacy of a theoretical hypothesis. If the supporters of theory theory share the standard view the strong claim of privileged

4 See Stegmüller (1969), 86ff., and the critical remarks in Davidson (1982), 233.

5 See Evans (1982), 224ff.

6 See Dennett (1991), 70ff.

access of the experiencing and reflecting subject to the mental evidence in an act of introspection could be weakened in the following way: The difference between the access of a single subject to her inner mental states and the intersubjective access to external physical entities should not be seen as an essential one but only as a matter of degree. Armstrong says that introspection cannot claim immunity because it can be erroneous.<sup>7</sup> He refers to cases in psychoanalysis where the presumed certainty of having a certain mental state can be revealed as illusive from outside. For Armstrong these examples justify the assumption that, in the end, evidence brought about by introspection is revealed as much less certain than empirical evidence which is intersubjectively observable:

“The result is that we do not often find cases where ‘the deliverance of introspection’ clashes with what we are quite certain are the facts about our current state of mind. However, we can very easily conceive that, in a future where far more is known than at the present about the workings of the brain, it would be possible to be quite sure that certain introspections were illusory. I might appear to myself to be angry, but *know* myself to be afraid. So the difference between perception and introspection... seems to be a contingent one, and provides no reason to resist the assimilation of introspection to perception.”<sup>8</sup>

Armstrong holds that the intersubjective perception of physical things assures a high level of reliability while the evidence produced by introspection arouses the suspicion of being illusive. One might reply that Armstrong’s argument is only valid for complex psychological states like anger or fear but it cannot be applied on mental states like pain, hunger or colour-vision. The access of the experiencing subject to these qualia-like states is immediate and cannot be “corrected” through external observation.

There is, however, a second position in philosophy of science which avoids this difficulty by relativizing any empirical evidence as theory-dependent. According to this position every perception depends basically on the respective underlying theoretical framework. This position can be called “theory-holism” because it does not distinguish sharply between the theoretical and the empirical domain. In the view of theory-holism sentences like, “there is a subject who is aware of certain mental states”, “my mental states cannot be observed from outside” etc., are expressions of implicit assumptions of an underlying folk-psychological framework. These implicit assumptions can be compared with the paradigmatic principles of scientific theories and

7    Armstrong (1968), 326.

8    Armstrong (1968), 328.

do not only determine the modes of reasoning but also the modes of perception.

One consequence of the influence of the implicitly given theoretical assumptions in folk psychology, according to this position, is the special theoretical character of the perceived evidence: The tacit theory of folk psychology divides perception in two different types — outer perception and inner perception. Awareness or consciousness of one's own inner mental states is only possible within the theoretical framework of folk psychology. According to the theory-holistic position the antecedent conditions of many folk-psychological hypotheses are given primarily through this special sort of theory-laden perception called "introspection".

"Insofar as introspective judgments are just a species of observation judgement then, there is no problem at all about the theoretical nature of the concepts they characteristically involve."<sup>9</sup>

This form of perception presupposes that certain objects (the mental ones) are internal and other objects (the physical ones) are external. It would seem clear that the impression of a "privileged access" to these inner mental objects can arise only in this theoretical context. Because of the intrinsic connection between introspection and the tacit theoretical assumptions of folk psychology it is not necessary to weaken the claim of introspective awareness by revealing the illusory character of some of its inner empirical evidence as was done above in the standard view. Here, instead of a gradual correction of introspection-based knowledge, the entire folk-psychological theoretical system is supposed to be eliminated together with its theory-laden introspective evidence. Within the framework of eliminative materialism the introspected mental states are shown to be theoretically dependent on folk-psychological assumptions, and if we change the underlying folk-psychological framework (for example, in favor of a more sophisticated neurophysiological one) the inner states, that we perceive, will be different (perhaps there are no more 'inner' states in the former sense).

"If our conceptual framework for P-states is an empirical theory, then it is possible, at the limit, that said theory be wholly false, that there are no such things as P-states, that *all* of our introspective judgments have been systematically false by reason of presupposing a false background theory."<sup>10</sup>

9 Churchland (1979), 96.

10 Churchland (1979), 96.

*Kant and the naturalization of introspection*

It is important to note that, despite of many differences, both classical and theory–holistic naturalism have a common view concerning introspection: They both identify privileged access with introspection and they both try to abolish this privileged access arguing that there is no qualitative difference between introspection and perception. Introspection is submitted to the same conditions of possible failure as perception — deception (Armstrong) or theory–dependence (Churchland). It is also remarkable that the main exponents of both positions consider Kant as a pioneer for the discovery that introspection is nothing but a special form of perception.

Churchland: “Kant stands almost wholly alone in insisting that knowledge of oneself is entirely on a par with knowledge of the world external to oneself.”<sup>11</sup>

Armstrong: “... introspection may properly be compared to sense–perception, and Kant’s description of introspection as ‘inner sense’ is perfectly justified.”<sup>12</sup>

It is true that — as Churchland and Armstrong assert — for Kant introspection provides no privileged access towards the ‘inner mental reality’. However, Churchland and Armstrong are not totally right when they specify the reasons for Kant’s being a pioneer for abolishing the privileged access given in introspection. It is not the case that Kant denies the privileged access in introspection because “knowledge of oneself is entirely on a par with knowledge of the world external to oneself” (Churchland) or because “introspection may be compared to sense–perception” (Armstrong). The reason why, according to Kant, introspection does not provide any form of privileged access is the fact that the “proper material” for the “inner sense” is not the inner mental reality but the “representations of the outer senses”.<sup>13</sup> Introspection is for Kant a rather bad instrument to gain knowledge about our inner mental reality. Without the outer sense it provides no objective knowledge at all, and if it is used in the appropriate way (i. e. in combination with the outer sense) it provides no access to the inner mental reality as it is in itself but only insofar as it appears to us.<sup>14</sup> However, Kant’s disdain of introspection as a way to achieve knowledge about the inner mental reality does not imply the abolition of the direct access of the subject towards its own mental states. This

11 Churchland (1979), 99.

12 Armstrong, (1968), 323.

13 Kant (1933), B 67.

14 Kant (1933), B 153.

can be shown by a short analysis of the concept of self-consciousness within Kant's *transcendental apperception*.

### *The “original synthetic unity of apperception”*

“Self-consciousness” is not a basic concept for Kant. According to Kant, the fact that I am conscious of myself can be translated or reduced into the following proposition: I am conscious of the original synthetic unity of apperception under which all representations must stand.

“I am conscious of the self as identical in respect of the manifold of representations that are given to me in an intuition, because I call them one and all *my* presentations, and so apprehend them as constituting *one* intuition. This amounts to saying, that I am conscious to myself *a priori* of a necessary synthesis of representations — to be entitled the original synthetic unity of apperception — under which all representations that are given to me must stand, ...”<sup>15</sup>

Self-consciousness is an expression for the act of synthesis which unites the different representations. Without this unifying synthesis no coherent knowledge could exist because every single representation and sensation would refer to yet another self:

“The thought that the representations given in intuition one and all belong to me, is therefore equivalent to the thought that I unite them in one self-consciousness... For otherwise I should have as many-coloured and diverse a self as I have representations of which I am conscious to myself.”<sup>16</sup>

The act of synthesis not only unites the “manifold of appearances” but also guarantees the unity of the thinking and knowing self. Therefore it is an act of synthesis, of active construction or conjunction which is the first and primary condition for the objectivity of all perceptions and introspections.

“... an *object* is that in the concept of which the manifold of a given intuition is *united*. Now all unification of representations demands unity of consciousness in the synthesis of them. Consequently it is the unity of consciousness that alone constitutes the relation of representations to an object, and therefore their objective validity ...”<sup>17</sup>

15 Kant (1933), B 135.

16 Kant (1933), B 134.

17 Kant (1933), B 137.

The synthesis (or synthetic unity of apperception) is “an act of spontaneity”,<sup>18</sup> but it is not arbitrary. It follows certain rules or *categories* which Kant tries to deduce in the *transcendental analytic*. I don’t want to discuss here the question whether or not the Kantian deduction is true, but it is important to note that for acting according these rules no explicit theoretical knowledge is required; neither can the active synthesis be replaced by a theoretical knowledge of these principles because this knowledge would be empirical and possibly be false.<sup>19</sup> The synthetic unity of apperception connects actively and directly the various perceptions under one principle. Theoretical propositions concerning the perceived or introspected reality are empirical and can be false. The act of synthesis which connects these representations according to certain rules is not empirical and is not submitted to verification or falsification, it is a priori. In the act of synthesis the perceiving or introspecting subject has direct access towards its own mental states: “Being an act of the self-activity of the subject, it cannot be executed save by the subject itself.”<sup>20</sup>

The naturalistic positions of Armstrong and Churchland mentioned above used Kant’s arguments concerning introspection to reveal the erroneous conception of direct or privileged access of the subject towards its inner states. However, Kant’s interpretation of introspection as a form of perception does not entail the elimination or reduction of the immediate or privileged access, but its reevaluation as a necessary condition for every objective knowledge. The abolition of the privileged access in introspection is compensated by the direct access in the synthesis of the perceiving subject. In the *original synthetic unity* there is no distance between the perceiving subject and the perceived object. It is in an act of construction and not of introspection that the single subject has immediate access to its mental states.

In the theory theory of folk psychology there is no place for this kind of active or spontaneous access of the subject towards its own mental states because the theory theory departs from the third person view: “if a person x is in so and so mental state she will act in this way ...”. The first person approach plays no special role in the laws and generalizations assumed in folk-psychological descriptions and explanations because it cannot be integrated in the typological structure of theories. However, for the execution of the act synthesis of representations third person knowledge is not sufficient or even necessary. Special emphasis on the first person perspective is needed because it is the acting subject which is the point of departure in the synthesis of mental

18 Kant (1933), B 130.

19 Kant (1933), B 132–140.

20 Kant (1933), B 130.

representations. This perspective seems to be met by another interpretation of folk psychological explanations — by the simulation theory.

### *Simulation theory — privileged access without introspection*

In the current discussion about folk psychology there is an alternative position, the so-called “simulation–theory”, which rejects the thesis that folk–psychological explanations have the same structure as theoretical or scientific explanations. According to this position, folk–psychological explanations cannot be compared to scientific explanations because they are based mainly on the first person perspective. Folk psychological explanations — at least the most efficient ones — do not presuppose general nomological correlations from certain types of mental events to certain types of behaviors (or other mental events) in order to explain and predict cognitive processes and behavior of human persons. The starting–point of folk–psychological explanations according to the simulation theory is not the question “how do people generally act?” but the question “how would I act, if I were the other person in this situation?”

Simulation theorists claim that we are more successful in the explanation and prediction of our own and other people’s behavior if we simulate how it would be to be oneself or the other person in this or that situation. The main argument for the correctness of simulations is seen in the great certainty of the predictions of behavior in our own case. R. M. Gordon uses very simple examples to demonstrate this:

“I shall now pour some coffee.  
I shall now pick up the cup.  
I shall now drink the coffee... .”<sup>21</sup>

I make these predictions without relying on any nomological correlations between the relevant mental states. On the contrary, if I would rely on the relevant mental states and the typical correlations between them and the resulting behavior I could never achieve this precision in the prediction of my behavior.<sup>22</sup> The predictive precision of first person–declarations of immediate intention is for the exponents of the simulation theory the first evidence

21 Gordon 1995a, 60.

22 Vgl. Gordon 1995a, 61: “... declarations of immediate intention — ‘I shall now X’ — are not products of inference from such premises. Moreover, if they were, one could not account for either their predictive reliability or our *confidence* in their predictive reliability. We are not self–omniscient... But even if we knew all the relevant beliefs and attitudes, our predictions would be at best be qualified and chancy.”



for the non–nomological or non–theoretical character of folk psychology. According to the simulation theory, the explanation of other people’s behavior or attitudes is made by simulating the mental processes of other people from the first person perspective or — as Gordon describes it — as “recentering one’s egocentric map on the other person”.

The priority of the first person perspective in the simulation theory seems to give new weight to introspection. There is a lively debate within the simulation theory whether the great authority of the first person perspective in folk–psychological explanations is based on introspection or not. Goldman and Harris argue for the introspection–thesis;<sup>23</sup> Goldman emphasizes the non–inferential character of introspection. Against Goldman, R. M. Gordon maintains that introspection is not a prerequisite for simulation. For Gordon introspection always presupposes the theory theory because, in order to introspectively identify a mental state, one has to know the type of this state; the knowledge of the type of a mental state is something which presupposes a theoretical framework of mental concepts. According to Gordon, simulation is not based upon theoretical knowledge about one’s own or other people’s mental states or knowledge about the correlations of certain mental states to other mental states or certain kinds of behavior. Simulation requires only an ascent routine of identifying mental states and using them for simulation:

“If an introspectionist version of the simulation theory, such as that of Goldman or Harris, is right, then we should expect children to be able to identify their own present beliefs with a least fair reliability before they can ascribe beliefs to others or to themselves in the past.... But if my view is right, and we ordinarily identify our own present beliefs by using an ascent routine, then there is an important distinction to be made between *comprehending* and *uncomprehending* ascriptions: that is, ascriptions made with and ascriptions made without the understanding that the beliefs ascribed may be false.”<sup>24</sup>

### *Conclusion*

According to Gordon, an immediate identification occurs when the simulation–process is executed. The access to the inner mental states during the simulation does not consist in the awareness of mental states in introspection but in the execution of a certain process in which mental states are involved. Mental “material” is used in a simulation with great success for the prediction

23 Goldman (1995), Harris (1995).

24 Gordon (1995b), 62.

of the behavior of other people. But there is no theoretical knowledge required concerning the logical and causal structures or rules which govern the simulation. On the other hand, theoretical knowledge would not be sufficient to execute the simulation. During the performance of a simulation the question, “does the simulation correspond to reality or not?”, does not occur.<sup>25</sup> Therefore, in FP there is an access of the subject towards its mental states which cannot — like introspection — be relativized as being a wrong or theory-laden identification of a certain mental state.

An analogous situation has to be assumed for Kant’s synthesis of the transcendental apperception by which the manifold of a representations is united.

— a) No explicit knowledge about the principles or rules (categories) which govern the synthesis is needed. According to Kant, objective knowledge has been reached long before he himself established a systematic list of all these principles in the *transcendental analytic*.<sup>26</sup> In physics a successful objective synthesis has been made without the explicit knowledge about these principles.

— b) the rules and principles which govern the synthetic connection of representations cannot be gained by introspection of the representations, they have to be transcendently deduced. Introspection, which is an empirical perception, cannot be the source for the principles of the synthesis, which are transcendental and a priori.

It is clear that Kant’s way to deduce these principles is very problematic and that many aspects of Kant’s critical philosophy are incoherent with the actual knowledge about the human mind. However, what can be shown by this reference to Kant’s philosophy is the fact that introspection is not the basic form through which the subject has direct access to its inner mental states or representations. It is Kant’s position that the transcendental condition for introspection is the spontaneous act of synthesizing these representations according to certain principles or rules. In the synthesis the subject has immediate access to its own inner mental states. This privileged access can be relativized neither by the standard view nor by the position of theory-holism. The reason for the immunity of this form of direct access lies in the fact that the act of synthesis itself cannot be true or false. It is like it is because the acting subject performs it this way. It can violate the principles of the objective synthesis, the synthesis can be “arbitrarily” (Kant), but also in this case the direct influence of the subject on its mental states cannot be questioned.

25 Gordon (1995b), 56.

26 In the preface of the second edition of KrV Kant explicitly states that natural science reached objective knowledge long before philosophy (= Kant) discovered the principles which were the reason for the objective synthesis and the resulting success of natural science in empirical predictions and explanations. (KrV, B VII ff.)

*References*

- Armstrong, D. M. (1968) *A materialist theory of the mind*. Routledge.
- Caruana, L. (1999) Die Elimination des Alltäglichen. In: *Der neue Naturalismus — eine Herausforderung an das christliche Menschenbild*, ed. J. Quitterer & E. Runggaldier. Kohlhammer, 146–161.
- Churchland, P. M. (1979) *Scientific realism and the plasticity of mind*. Cambridge University Press.
- Churchland, P. M. (1991) Folk psychology and the explanation of human behavior. In: *The future of folk psychology. Intentionality and cognitive science*, ed. J. D. Greenwood. Cambridge University Press, 60–69.
- Davidson, D. (1982) *Essays on actions and events*. Clarendon Press.
- Davies, M. & Stone, T., eds. (1995a) *Folk psychology: the theory of mind debate*. Basil Blackwell.
- Davies, M. & Stone, T., eds. (1995b) *Mental simulation. Evaluations and applications*. Basil Blackwell.
- Davies, M. & Stone, T. (1995c) Introduction. In: Davies & Stone (1995b), 1–18.
- Dennett, D. C. (1991) *Consciousness explained*. Little, Brown and Company.
- Eckardt, B. v. (1994) Folk psychology. In: *A companion to the philosophy of mind*, ed. S. Guttenplan. Basil Blackwell, 300–307.
- Evans, G. (1982) *The varieties of reference*. Clarendon Press.
- Goldman, Alvin I. (1995) Empathy, Mind and Morals. In: Davies & Stone (1995b), 185–208.
- Gordon, Robert M. (1995a) Folk Psychology as Simulation. In: Davies & Stone (1995a),
- Gordon, Robert M. (1995b) Simulation without introspection or inference. In: Davies & Stone (1995b), 53–67.
- Harris, Paul L. (1995) Imagining and Pretending. In: Davies & Stone (1995b), 170–184.
- Kant, I. (1933) *Critique of Pure Reason*, transl. by Norman Kemp Smith. Macmillan.
- Stegmüller, W. (1969) *Probleme und Resultate der Wissenschaftstheorie und Analytischen Philosophie. Band I — Wissenschaftliche Erklärung und Begründung*. Springer.