

# How Anti-Introspectionist is Theory Theory?

Josef Quitterer, Innsbruck

Introspection is often seen as criterion to distinguish between theory theory (TT) and simulation theory (ST). Many empirical arguments against ST rely upon the thesis that ST is introspectionist and that it presupposes the Cartesian dictum that the mind is transparent to itself. According to Perner 1999 the capacity to introspect is so important for ST that it can be seen as *the* criterion that distinguishes ST from TT: "These two positions and their sub-varieties differ as to whether they presuppose or explain the ability to introspect. Theory theory is extremely anti-introspectionist. Traditional simulation is based on the ability to introspect one's own mental states." (Perner 1999)

The introspectionism-thesis and the transparency-of-the-mind-thesis together form the basis for different *experimenta crucis* in which the following empirical hypothesis is deduced from ST: If ST is right the child would first develop the capacity to understand its own mental states and only later develop the capacity to understand (or to successfully ascribe) other person's mental states. (Gopnik & Astington 1988, Gopnik & Wellman 1995) The fact that both capacities emerge at the same time is taken as a refutation of ST and a confirmation of TT. (Perner 1999, Gordon 1995)

The thesis that introspectionism is a criterion for the distinction between ST and TT has been criticized by some adherents of ST – like R. M. Gordon. Gordon maintains that introspection is not a prerequisite for simulation. For Gordon, introspection always presupposes a 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 that 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 does not require introspection, it demands an "ascent routine" of identifying mental states and using them for simulation. (Gordon 1995b) Gordon denies the thesis that ST is introspectionist. But it remains undisputed that TT is extremely anti-introspectionist. In my paper I would like to question this anti-introspectionist view of TT. Is there any reason for a theory theorist to be anti-introspectionist? If there is, is it a good reason?

I think the only reason for a theory theorist to be anti-introspectionist is the special character of TT itself. Folk psychology (FP) is defined in TT as "a conceptual framework ... used by ordinary people to understand, explain and predict their own and other people's behavior and mental states." (Eckardt 1994, 300) In TT FP is seen as an explanatory system that can be compared with other theoretical systems (e.g. of natural science).

There is, however, a difficulty with FP when it is compared with other theories: A folk-psychological explanation can be compared with a scientific one only if there is empirical evidence for testing the different explanatory hypotheses. A law-like generalization or hypothesis can only be tested 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 valuation of its explanatory power. (Stegmüller 1969, 86ff.) But the intersubjective access to the antecedent conditions of folk-psychological explanations would be impossible if some of them could be verified only in an act of introspection.

Introspection cannot be accepted as a precondition for a successful mental explanation if FP is basically seen as a theoretical hypothesis for the explanation of human behavior. As a theory it has to be independent of the psychological characteristics of its users, which are available only in an act of introspection. This independence is what A. Goldman calls the attributor-neutral heuristics of FP in TT. (Goldman 2000) Therefore, the reason for the anti-introspectionism of theory theorists is their belief that FP is essentially a theory and as such independent from the properties of the person who is applying the theory. If folk-psychological explanations would presuppose the ability to introspect one's own mental states it would lose its theoretical, intersubjective status.

Is this belief justified? Do folk psychological explanations necessarily lose their attributor-neutrality when they presuppose the ability to introspect one's own mental states? Must everybody who emphasizes the role of introspection in folk psychological explanations share the Cartesian dictum of the transparency of the mind?

As a counterexample I would like to mention Eliminative Materialism (EM) for the following reasons: a) Eliminative materialists are theory theorists. TT is a necessary condition for a successful elimination of FP: Only if FP 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.

b) Eliminative Materialists do not believe at all in the transparency of the mind. In EM introspection is seen as a totally inadequate instrument to gain knowledge about internal cognitive processes.

c) Eliminative Materialists even argue explicitly for the attributor-neutrality of FP:

Knowledge of other minds thus has no essential dependence on knowledge of one's own mind. Applying the principles of our folk psychology to our behavior, a Martian could justly ascribe to us the familiar run of mental states, even though his own psychology were very different from ours. He would not, therefore, be "generalizing from his own case." (Churchland 1990, 208)

An analysis of Churchland's statement makes clear that attributor-neutrality does not imply a non-introspective attitude toward FP. Churchland's Martian acquires a new access to his internal states with the principles of FP. Together with FP he acquires a "special habit of conceptual response to ... [his] internal states", called introspection. This habit of conceptual response "is always contingent on (the integrity of) the acquired

conceptual framework (theory) in which the response is framed." (Churchland 1990, 208)

According to this position every perception basically depends 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 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 this 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 judgment then, there is no problem at all about the theoretical nature of the concepts they characteristically involve." (Churchland 1979, 96) 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, 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 (in favor of a more sophisticated neurophysiological one, for example) 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." (Churchland 1979, 96)

It is clear that the radical position of EM regarding the future of FP is very problematic. In this paper I do not want to comment or criticize Churchland's belief in the elimination of FP through neuroscience. What is important for our question "how introspectionist is TT?", are the following points: In EM we have a conception of TT that presupposes the ability to introspect one's own mental states for folk psychological explanations. But the introspectionism which EM presupposes for FP is neither connected with a transparency-of-the-mind-thesis, nor does it imply attributor-dependent heuristics. Anybody – including Martians – who adopts the conceptual system of FP and uses it for the explanation of cognitive processes

and human behavior will have an introspective access to its mental states.

This (introspectionist) conception of TT could help to interpret the empirical findings of Gopnik & Astington 1988 and Gopnik & Wellman 1995 in a more moderate – that means less anti-introspectionist – way: From the fact that children develop the capacity to understand (or successfully ascribe) other person's mental states together with the capacity to understand their own mental states it need not be concluded that folk psychology does not rely upon the ability to introspect one's own mental states. It could also be interpreted in the following way: Both capacities emerge at the same time because children together with the folk psychological theory also acquire a new access to their own mental states – they learn to introspect. In introspection they obtain an access to some antecedent conditions of folk psychological explanations that they would not have without introspection. Thus, the folk psychological explanations children perform after acquiring this new conceptual framework essentially presuppose the ability to introspect one's own mental states.

## References

- Churchland, P. M. 1979 *Scientific Realism and the Plasticity of Mind*, Cambridge: Cambridge University Press.
- Churchland, P. M. 1990 "Eliminative Materialism and the Propositional Attitudes", in W. Lycan, (ed.), *Mind and Cognition. A Reader*, Oxford: Basil Blackwell, 206-223.
- Churchland, P. M. 1991 "Folk psychology and the Explanation of Human Behavior", in J. D. Greenwood (ed.), *The Future of Folk Psychology. Intentionality and Cognitive Science*, Cambridge: Cambridge University Press, 60-69.
- Davies, M. and Stone, T. (eds.) 1995a *Folk Psychology: the Theory of Mind Debate*, Oxford: Basil Blackwell.
- Davies, M. and Stone, T. (eds.) 1995b *Mental simulation. Evaluations and Applications*, Oxford: Basil Blackwell.
- Eckardt, B. v. 1994 "Folk psychology", in S. Guttenplan (ed.), *A Companion to the Philosophy of Mind*, Oxford: Basil Blackwell, 300-307.
- Goldman, A. I. 2000 "Using Your Own Mind to Read Others", paper presented on the Ninth Annual Conference of the European Society for Philosophy and Psychology (ESPP) - September 2-5, 2000, University of Salzburg, unpublished.
- Gopnik, A. and Astington, J. W. 1988 "Children's Understanding of Representational Change and its Relation to the Understanding of False Belief and the Appearance-Reality Distinction", *Child Development*, 59, 26-37.
- Gopnik, A. and Wellman, H. M. 1995 "Why the Child's Theory of Mind is a Theory", in: Davies and Stone (eds.) 1995b, 232-258.
- Gordon, R. M. 1995 "Simulation without introspection or inference", in Davies and Stone (eds.) 1995b, 53-67.
- Perner, J. 1999 "Metakognition und Introspektion in entwicklungspsychologischer Sicht: Studien zur 'Theory of mind' und 'Simulation'", in W. Janke and W. Schneider (eds.), *100 Jahre Institut für Psychologie und Würzburger Schule der Denkpsychologie*, Göttingen: Hogrefe. 411-431.
- Stegmüller, W. 1969 *Probleme und Resultate der Wissenschaftstheorie und Analytischen Philosophie. Band I - Wissenschaftliche Erklärung und Begründung*, Berlin: Springer.