Phenomenology and the Cognitive Sciences **3:** 195–214, 2004.

Intersubjectivity of cognition and language: Principled reasons why the subject may be trusted

NINI PRAETORIUS
Department of Psychology, University of Copenhagen, Njalsgade 88, DK-2300 Copenhagen S, Denmark (E-mail: nini.praetorius@psy.ku.dk)

Abstract. The paper aims to show that scepticism concerning the status of first-person reports of mental states and their use as evidence in scientific cognitive research is unfounded. Rather, *principled* arguments suggest that the conditions for the *intersubjectivity* of cognition and description of publicly observable things apply equally for our cognition and description of our mental or internal states. It is argued that on *these* conditions relies the possibility of developing well-defined scientific criteria for distinguishing between first-person and third-person cognition and description. The paper concludes by outlining the consequences for cognitive research and for functional theories of mind.

Key words: consciousness, cognition, intersubjectivity, language, subjectivity

Introduction

In the debate concerning the use of introspective reports in cognitive neuroscience, two distinctly different problems, each requiring distinctly different solutions, are often confounded. The first problem, which I shall call the *methodological* problem of introspective reports, may be expressed as follows.

1. How can correlational studies of brain processes and concurrent conscious mental states be considered *scientific*, given that the descriptions of brain processes rely on rigorously developed methods of observation and precise, well-defined terms to describe such processes and structures, whilst observations and reports by subjects about their conscious mental states lack this rigour of observation and precision of description?

Concerns over the lack of rigour and precision in observing and describing conscious mental phenomena have been systematically addressed and solutions attempted by a number of cognitive scientists.[1] The solutions suggested typically include the development of adequate conditions and training of subjects to observe and be aware of their "mental goings-on," as well as of appropriate vocabularies and terms by which to describe them. The assumption is that, granted this pre-experimental training as well as

adequate conditions to observe and describe their conscious mental states, then experimental subjects will be able to learn how to observe their mental states during the experiments and to use these terms both reliably and correctly to describe them – as well as "what it is like" to have them.

The second problem, which I shall call the *epistemological* problem of introspective reports, is as follows.

2. How can knowledge and descriptions of something to which only the subjects have access and may observe, i.e. their conscious mental states, be considered as data for *scientific* research, given the requirement in such research of public criteria to determine *what* is being observed, and third-person validation of the correctness of descriptions of such states?

This problem is a much deeper than the first, the methodological problem, in that the very *status* of introspective reports as scientific evidence is called into question. Unlike the first problem, this is not a problem that may be solved by the employment of appropriate procedures and methods for first-person observation and description of conscious mental states – be they phenomenological or otherwise – nor can it be solved by appeal to phenomenological evidence. The significance of a solution of this problem for cognitive research is clear; for if the absence of third-person validation disallow introspective reports as evidence, then it is hard to see how the scientific study of consciousness could get off the ground at all.

Concerns over this absence of third-person validation have given rise to "resistance to introspective evidence" (Jack and Roepstorff 2002), in particular among cognitive researchers versed in the methods of the natural sciences – and underlie the contention of Dennett that, "First-person science of consciousness is a discipline with no methods, no data, no results, no future, no promise. It will remain a fantasy" (Dennett 2001; quoted in Gallagher 2003). A solution of this problem would require that – despite not being accessible to third-person validation – it can somehow be shown that conscious mental states – nevertheless – may be just as amenable to *reliable* observation and *correct* descriptions as are observations and descriptions of publicly observable things or states – and hence that observation and description of conscious mental states may indeed fulfil the requirement of scientific evidence at par with evidence from traditional areas of scientific research.

However, for such a solution to convince the sceptics among cognitive scientists and philosophers holding the views of Dennett, it needs in my view to be based on principled arguments – as opposed to methodological considerations, or to the outcome of present or future empirical research within cognitive neuroscience. The latter has been proposed by, among others, Baars. In "How Brain Reveals Mind" (Baars 2003), he argues that although "it is still true that brain evidence has greater credibility than subjective reports, no

matter how reliable," new brain studies of a wide variety of conscious states show "increasing convergence between subjective experience and brain observations;" indeed, such studies "support the central role of consciousness." However, the argument that evidence from observations of the brain vindicates the existence and role of conscious states, as well as the reliability of the subjective reports about them, suffer in my view from circularity in that the *correlational* studies cited by Baars of brain states and concurrent conscious states all *depend* on and hence *presuppose* both the existence of the conscious states being investigated, and that the subjects' introspective reports about them may be relied upon to be in general correct and reliable. Thus, to the extent that the experimental determination of neuronal mechanisms giving rise to or underlying conscious states necessarily relies on the reliability of introspective reports by subjects about their conscious states during the experiments conducted, it can never be an empirical *outcome* of such experiments that these states actually exist and that the subject's reports about them are in fact reliable. Indeed, as should become clear in what follows, the fact that introspective reports of mental phenomena may be both reliable and correct is not something which may be vindicated empirically, but which, on the contrary, has to be taken for granted in all empirical studies within cognitive neuroscience as well as in scientific research in general.

Consequently, rather than providing evidence from empirical research, a solution based on *principled* arguments will be attempted in this paper – a solution, which in turn may serve as a firm foundation for the solution of the *methodological* problems of introspective reports outlined above. I shall proceed by first presenting some of the main problems concerning the nature and status of reports about our internal and mental states as set out within traditional philosophy. Next, an extended analysis of the position on these problems defended by Wittgenstein in his famous "Beetle in the box" example will be carried out – a position which is echoed in the scepticism expressed by many present-day cognitive scientists. It is the aim of this analysis to show that sceptical arguments against the possibility of observing and communicating reliably and correctly about our internal or mental states necessarily relies on this very possibility, i.e. relies on the fact that, despite not being publicly observable, our knowledge and experience of such states are indeed amenable to intersubjective discourse.

In the sections that follow I shall attempt to clarify what it means that our cognition and use of language is *intersubjective* and to show that the very same assumption of intersubjectivity is a condition which applies equally for knowledge, description and communication of things in publicly observable material reality and for our non-publicly observable mental or internal states. This condition may be expressed by the general presupposition that *if other people could be in our situation*, *and had the same possibilities of observing the things and states that we do*, *then they would observe what we do*, *and use*

language to describe them as we do – and vice versa. That is, what is true or correct about the things or states we observe and experience would also be true or correct to others, could they observe and experience the things or states that we do – and vice versa. I shall argue, that this assumption, and hence that we may indeed use language to talk correctly and reliably about *both* knowledge of that which is publicly observable and shared *and* knowledge of what is personal, is *logically* necessary for any rigorous determinations of and thus of any well-defined distinctions between what is *publicly observable and what is not*. In other words, I shall argue that it is *because* of this presupposed intersubjectivity, fundamental for the cognition and use of language of persons, that a distinction between first-person and third-person cognition and description is possible at all. And that it is possible *therefore* for people, e.g. scientists, together to develop well-defined notions about scientific *validity* and *objectivity*, as well as methodological procedures and criteria to ensure such validity and objectivity in scientific research.

The paper concludes by outlining the consequences of the intersubjectivity of cognition and use of language for research within cognitive science as well as for functional computational theories of mind.[2]

The epistemological problems concerning the status of introspective reports

Among the main issues debated within philosophy concerning the status of knowledge and description of our conscious mental states and other "internal goings-on," are the following: Given that mental or internal states such as thoughts, emotions and feelings of pain, are not publicly observable, but may only be known or experienced by the persons who have them, how then can we be certain that these states exist and may be observed in the same way by different people? Furthermore, in view of the lack of public criteria or standards, how then can we be certain that assertions put forward about such states have the same implications and use for different persons and language users? Or be certain that we are using language in the same way – or indeed using the *same* language – as we do when talking about and describing things in public material reality? Indeed, given that internal states are not the sort of things which are publicly *shareable*, how do we ever come to learn and talk about them in a language we do share?

That solutions of these issues have serious consequences not only for the possibility of scientific cognitive research but for a science of psychology in general, is obvious when we consider that a crucial part of the knowledge a person has of himself, of his mind, body and acts – and, not the least, *of his mind and body being his*, and *of his acts being acts intended and controlled by himself* – rests on observations and experiences to which only the person

himself has access. No one except the person himself will ever be able to experience what it is like to be the person he is *in the way he experiences* it, or to know what it is like to have his perceptions, thoughts, or feelings of pain in his body *in the way he does*, because no one except the person himself can see with his eyes, think his thoughts or feel his pain.

To this uniqueness of our personal experience must be added the differences in our knowledge and conception of ourselves, of others and the world in which we find ourselves, due, for example, to differences in our upbringing, training, education or cultural background. And yet, neither the knowledge of persons which is uniquely personal, nor the differences in knowledge and conceptions of themselves, of others and the world, due to their different backgrounds, can be said to be *private*. For, despite these differences in our personal experiences and conceptions, to be a person is something fundamentally social. Indeed, no one can be a person, and thereby someone who may realise that he or she is uniquely different from other persons, *without* other persons from whom he or she may differ – nor without having possibilities of determining *how* one differs from others – whether such differences concern one's notions and experiences of things in material reality, or one's inner feelings, thoughts, beliefs, emotions or pains. Indeed, it would seem that no one could be a person, that is, someone who *differs* and *knows* that one differs from others, without being able to communicate and talk with others about *how* one differs from them.

In the next section I shall first argue that despite not being publicly observable, it seems that we shall have to assume that we may both have knowledge about and use language to talk just as correctly about our mental or internal states as about things in publicly observable reality.

Conditions for description of mental or internal states

One of the significant differences in the conditions for cognition and description of things in publicly observable reality and of our non-publicly observable internal states concerns the *procedures* for determining the truth and correct application of descriptions of such things and states. In the case of descriptions of publicly observable things, for example, we will be able to take part in a procedure for determining whether the implications of the descriptions hold true for the things in question, and thus to observe the result of a test as to the correct application of the description. However, in the case when somebody says, e.g. that he has a pain in his finger, no such public procedure exists for determining whether a "pain-description" is a true description of what the person feels in his finger. For only he knows and may observe what he feels in his finger.

But given that no public procedure exists for determining the correct use of our description of internal states such as feelings of pain, how then can we

be sure that when different people talk about pains, they are talking about the same kind of "thing"? How, more precisely, *can I be sure that when I use the term 'pain', I use this term to refer to the same kind of "thing" as others do when they use the same term?"* Or, conversely,*"when other people use the term 'pain' do they then use this term to refer to the same kind of "thing" as I do when I use the same term?*

This would seem a perfectly sensible question. However, in his classical "Beetle in the box" example, in which Wittgenstein attempts to make clear what this question entails, it seems to be a question to which, for obvious reasons, no sensible answers may be given. Thus, Wittgenstein writes:

> If I say of myself that it is only from my own case that I know what the word 'pain' means – must I not say the same of other people too? And how can I generalise the one case so irresponsibly?
>
> Now someone tells me that he knows what pain is only from his own case! Suppose everyone had a box with something in it; we call it a 'beetle'. No one can look into anyone else's box, and everyone says he knows what a beetle is only by looking at his beetle. Here it would be quite possible for everyone to have something different in his box. One might even imagine such a thing constantly changing. But suppose the word 'beetle' had a use in these people's language? If so it would not be used as the name of a thing. The thing in the box has no place in the language game at all; not even as something: for the box might even be empty. No one can 'divide through' by the thing in the box; it cancels out, whatever it is (Wittgenstein 1945/1953, para 293, p. 100).

In what follows I shall try to show that the question about 'pain', posed in this *radical* sense, is not only a question which cannot be answered, but more importantly, it is also a question which cannot be *asked*. For it is a question in which the very condition for putting it forward is itself questioned – thereby debarring it of any sensible meaning.

Let me begin by noting that Wittgenstein does not only *suppose* that the word pain "has a use in people's language" – he knows it for a fact. Indeed, when he or anyone else asks the question, "do I use the term 'pain' to refer to the same 'thing' as others do when they use the term," he and they are obviously asking this question in a language of which the term 'pain' is part – and thus is a term which is supposed to have a meaning and use which he and other people with whom he shares the language know *in virtue* of being speakers of that language, just as they know to what the term may be correctly applied, i.e. some particular unpleasant sensations somewhere in one's body. If Wittgenstein did not know this – and did not *presuppose* that all other speakers of the language knew this – neither he, nor they would have any idea what he is asking about, and no further discussion of the question would seem possible.

Now, it has to be admitted that we may have all kinds of difficulties in giving adequate accounts in words of the pain we may feel, and that we often feel uncertain about the choice of appropriate terms. Is this pain, for example,

a sharp, piercing, dull, shooting, tender, searing pain, or is it a nagging or stabbing pain? Indeed, we may have great difficulties in conveying precisely to others the suffering we endure when we are in pain. However, these difficulties of *adequately* describing pains are not relevant to, nor do they invalidate the point just made. Indeed, these difficulties of adequate descriptions of pains, and discussions about them, could not take place *unless* people discussing them had a concept of and a term for pain which they shared, and knew of what it may be used correctly to refer to.[3]

We may contend then that Wittgenstein's discussion of the "language game" of giving expressions to pain relies on this knowledge and these presuppositions, and hence on the assumption that sensations and feelings of pain are the sort of "things," which do indeed exist and which we may use language to refer to. Furthermore, it relies on the assumption that other people would use the term 'pain' to refer to the feeling of pain that he may have in his body – could they feel what he feels – and hence that he may indeed generalise from himself to others, when they use the term 'pain' – just as they may generalise from their use of the term to his and the use of the term by others.

However, contrary to the presuppositions on which his whole discussion of the question of 'pain' rely, Wittgenstein suggests, for the sake of argument, that because his feeling of pain, and those of others, are not publicly observable, his feelings of pain may be completely different from others and, thus, he as well as others may be using the term 'pain' to refer to completely different things – or to none at all. Indeed, he invites us to assume that in the language he shares with others it would be perfectly sensible so to say and suggest. However, it does not make sense to suggest that we may talk about, let alone *determine* any individual differences in our feelings of pain or in our use of the term 'pain', *unless* the implication of the term 'pain' is shared by everyone involved, and unless, furthermore, this term is used to refer to the same sort of "thing." Without these presuppositions, any discussion of the question of 'pain' disintegrates into nonsense.

In summary, it would seem that the question of 'pain' as set out by Wittgenstein is *obtuse* in the sense that putting it forward presupposes that we know the meaning and use of the terms of the question and also to what it may be correctly applied. But then we are asked to forget or ignore this presupposed knowledge, and to pretend that it is immaterial for a discussion of what the terms 'pain' may be correctly used to refer to – or whether it may be used correctly to refer to anything at all. But it is a question which can only be asked and discussed granted we have already learned a language which we may use to talk about pain, and thereby granted pain to be the kind of "thing" which exists as something we may talk *correctly* about and *refer* to. Hence, if we do ask this question, the answer is logically implied: an affirmation would be redundant, while a denial would be contradictory.

Now, if the same sceptical question had concerned the use of the terms 'cups' or 'neurones' instead of 'pains', it would have been obvious why such question would not make sense. Indeed, it would probably have been so obvious that we would hesitate to ask it. For if we did ask this question about the use of the terms 'cups' and 'neurons', we might as well ask the same question about all other terms in our language – and we would be well on the road to asking whether we can be sure that we may use terms in our language to refer to and talk correctly about *any* objects in material reality. However, although in *particular* cases we may be in doubt as to whether a particular term may be correctly used to refer to some particular thing, i.e. a thing which has been identified in a shared public world, we cannot doubt that as language users taking part in this discussion, we do know (other) correct terms for the thing (i.e. those forming part of the identification of the thing), nor doubt that we know how to use *these* terms correctly. One cannot doubt the necessity of these conditions for settling the question under discussion – unless, of course, one has been seriously contaminated with scepticism, and mistakenly assumes that one may get away with using language to question the very possibility of using language to talk correctly about anything.

However, similar conditions seem to apply to the question of 'pain' – and by extension to questions concerning other internal or mental states such as emotions, thoughts and belief – in the sense that scepticism as to whether we may use language to talk correctly – or at all – about and refer to such states, presuppose that we know the meaning of terms for these states, and that together with other language users we may determine what they may be correctly used to refer to. So rather than questioning the existence of such internal or mental states and the possibility of having knowledge about and of correctly describing and being able to communicate about our experiences of such states, this questioning logically rests on the assumption of both their existence and this possibility. Indeed, without these assumptions concerning the *intersubjectivity* of both our experiences of, and of our use of terms to talk about and refer to our non-publicly observable internal or mental states, neither ordinary everyday communication nor philosophical discussions about such states would be possible.[4]

In the section which follows I shall further clarify what it means that our cognition and use of language is *intersubjective*, just as I shall show that the very same assumption of intersubjectivity is a necessary condition for knowledge, descriptions and communication of both that which is and that which is not publicly observable and shared.

The intersubjectivity of public and personal knowledge and experiences

It has to be admitted that it is somehow puzzling that pains and other internal states, which are only directly observable to the persons who have them, and

thus are not *observationally* shareable by others, are nevertheless things which we may communicate about in a language we *do* share with other persons. So, apparently, shareability in the sense of being *publicly* observable and known cannot be a condition for the possibility of communication among persons about things which only they may directly observe, experience and know about.

In the case of observing, experiencing and communicating about objects in material reality, our situation is arguably significantly different. Take, for example, two people sitting on either side of a table with cups and plates, a teapot, a bowl of sugar and a bottle of milk. All these things exist in a shared public world and are perfectly observable to both persons; they may together determine the things on the table and also whether the descriptions they put forward about them are correct. And yet, it could be argued that how these things are observed and appear to them from their different vantage points on either side of the table is different, i.e. due to the fact that the parts and features of the things which are directly observable to the one, are not the same as those which are so observable to the other. However, this does not present any serious difficulties since, first, it is part of our knowledge of things having been identified as particular things, that they will appear differently when, e.g., we move around and look at them from different vantage points – and that, generally, things do appear differently when observed with regard to different possibilities of observations and action. And these differences of perception and experience do not represent any serious problems since, secondly, none of the particular ways of perceiving and experiencing the things on the table, and none of the descriptions by either person of their experiences of the cups, plates, etc., are unique to him or her. Indeed, it is *assumed* that they are not, just as in general any person and language user will assume that if other people could look at the things from his or her vantage point(s), they would observe what he or she does, and report that they perceive the same features and properties of the things, and describe what they perceive as he or she does. *If we could not count on this assumption*, *communication and action between persons about things in the world would be impossible*.

This assumed *intersubjectivity* of our perception and cognition, however, is not only basic to everyday action and communication between people, but is *presupposed* in any instruction given to subject taking part in cognitive experiments concerning those features of the experimental set up which they are supposed to attend and respond to – verbally or non-verbally. The purpose of instruction, as pointed out by Jack and Roepstorff (Jack and Roepstorff 2002; Roepstorff and Frith 2003), is to provide the subject and experimenter with a shared representation of the experiment and of the nature of the task to be carried out and reported by the subject. Instruction in the form of such shared "scripts" is a prerequisite for setting up cognitive experiments. However, without the general assumption, shared by experimenter and subject,

that what each of them perceives, experiences and describes from their different positions, would also be perceived, experienced and described as they do by others, if they were in their positions, neither instruction to subjects nor any shared "scripts" between experimenter and subject would be possible, let alone make sense (see also Praetorius 2000, chapter 4).

But there are numerous other ways in which the knowledge and description of persons concerning things in publicly observable reality may differ. Just think of the differences due to our different background, education, previous history of experiences, and the opportunities to observe and describe such things which are or have been available to us. Examples are *legion* – I only have to think of the knowledge I have of aeroplanes, their construction and how to fly them compared to that of a pilot. Or, conversely, think of the knowledge I have as a clinical psychologist about the transference phenomena occurring between client and therapist during psychotherapy compared to that of an aeroplane pilot, who has not encountered such phenomena, and who does not have the language and terms to describe them that I have. And yet, despite the fact that our knowledge of these and numerous other matters are not exactly the same, and probably never will be, we are in fact able to make available and to communicate the knowledge that each of us has about aeroplanes and transference problems and those other matters, and thus to share our knowledge of what in this respect is personal to each of us. But if the condition for the intersubjectivity and shareability of knowledge and description in the actual case as well as between persons in general, is not and cannot be that persons have exactly the same knowledge and experience of things, nor the same possibilities of describing things, since this condition is only rarely if ever met due to their different background, education, history of experiences, and so forth, on *what* then relies this intersubjectivity of knowledge and description of persons?

It relies no doubt on the fact that, apart from differences in our knowledge, experience and background, we do share a substantial amount of knowledge and description of the world in which we live and act, of the things with which we may act, of ourselves and of the persons with whom we may co-act. However, to say so does not of course add anything to our notions of 'shared knowledge and description', that is, knowledge and descriptions of things which have been available to shared inspection. Nor does it suffice to account for how knowledge of things which may only have been available to ourselves,[5] and which in this sense is personal, is amenable to description and intersubjective communication. It does not do so unless it is assumed that vantage points, background, and situations we may be in, are *in principle* shareable, and thus that other persons *could be* or *could have been* in the same situations. But nor does it suffice independently of assuming that, granted other people had had the same background, or had been in the same situations that we ourselves have been or are in, then they would have the knowledge of

the situation and the things that we have, and describe them the way we do. However, this suffices, indeed it will *have* to suffice to say that it is fundamental to the cognition and experience of persons that, although other people may not be in our situations, and may not have, or may not have had, exactly the same experiences and knowledge that we have or have had, they would – *could* they be, or *had* they been, in our situations. Likewise, it suffices, and will *have* to suffice to say that to be language users and to share a language with other persons *logically* implies and presupposes that other language users, granted that they could be in our situation and have the experience, knowledge, background, points of view, etc. that we have, would use language to describe what we experience, know of, etc., in those situations as we do. Or, they would consent that the descriptions we put forward about our experience and knowledge are correct and correctly applied.[6]

Now, if we can agree that these presuppositions must be fundamental to the cognition, use of language and communication of persons, and indispensable for any meaningful discussion among persons about what they know and how they describe what they know, I think we shall also have to agree that this intersubjectivity of cognition and language relies on a notion of 'truth' which implies that what is true or false, correct or incorrect, is also true or false, correct or incorrect for other persons, or just for 'others'. That is, this intersubjectivity relies on a notion of 'truth' which *logically* implies and presupposes a notion of 'others'. This assumed intersubjectivity must be the rock bottom, the very point of departure from which any discussion about our knowledge and description must be based and proceed – whether such discussions concern our knowledge and description of thing in publicly observable reality, or our internal states, such as our emotions, thoughts or feelings of pain.[7]

In what follows I shall argue that the intersubjectivity of human cognition, language and of the notion of 'truth' is a condition for the possibility rigorously to determine and thus to distinguishing between knowledge and descriptions of that which is publicly observable and that which is not – and hence for people together to develop conventions, criteria or standards for the validity and objectivity of their knowledge and description of things and events in publicly observable reality. And I shall argue, moreover, that on these conditions relies the possibility of characterizing and determining individual differences and variations in the cognition and description of different persons – concerning both what is publicly observable and shared and *personal*.

Conditions for distinguishing what is publicly observable from what is not

It is well known that considerable variations may exist in how things are described by different persons, and by the same person in different situations,

not only when the things concerned are "internal" states and events, but even things in material reality. Not only our intentions, purposes and possibilities of observation and action, but even our moods and temperaments may at times determine what we experience, and how we experience and describe ourselves and the rest of reality. So, would it not then be reasonable to reserve the notion 'true descriptions and assertions', and 'knowledge of what exists objectively' for cases in which no such individual differences and variations exist, and to reserve the terms 'exist' and 'determinable' for things and events about which no uncertainty prevails – because they belong to what is materially and publicly observable?

This solution has been attempted, notably by the logical positivist and by the radical and logical behaviourists in both psychology and philosophy, who aimed to establish a firm epistemological basis for scientific research. In this pursuit they argued that only *that* exists objectively, and hence can be the object of scientific research, which may be determined by rigorous public criteria and standards, and that only knowledge and description of what had been thus determined and observed, may be said to be meaningful and true. In effect, any determinate notions of the truth and meaning of statements and knowledge would have to derive from observations fulfilling such criteria and determinations. Consequently, what cannot be subject to rigorous public scrutiny and observation fulfilling such standards and criteria does not exist, nor can description of things which cannot be so observed be true; hence, descriptions and the existence of knowledge of such things may be discarded as non-sensical.

However, against such arguments we only have to consider that the very process by which we identify and determine a situation, and what in this situation is materially and publicly observable, presupposes that something is the case or true about the situation and things being determined and observed, which is also the case or true to other people – *in casu* the people involved in the observation. Thus, it is not because situations exist or may be arranged, in which things are publicly observable, and which we may come to agree to describe in particular ways, that the notion "emerges" of what – for everyone involved – is true or correct about things and situations. It is the other way round – for no such determinations of correctness of knowledge and description of things and situations could be agreed upon, let alone be arranged and function as criteria or standards for correctness, unless it was presupposed that *when* arranging and determining these criteria or standards, we *already* have a concept of 'truth' which we know how to use correctly; a concept of 'truth', furthermore, which is such that what is true or false, correct or incorrect, is also true or false, correct or incorrect for others.

Thus, the point, so easily overlooked, is that even in a situation in which the things and events being described are publicly observable and identifiable, it is logically implied and presupposed of descriptions put forward and

being understood by others, and of these others, being in the same situation in which we are, that they will have the knowledge we have about these things and events, and will describe them as we do. In general, it is presupposed that our notion of 'correct' or 'true' is such that what is true or correct, is also true or correct for others. *This* presupposition granted, it is possible for persons – in spite of the differences and variations which exist in their knowledge, background, opportunities for observation and action – to arrange conditions and develop criteria and standards for what may count as correct, *objective* descriptions and *publicly* existing things, that is, possible *to determine and distinguish between conditions under which such criteria or standards apply*, *and in which they do not*. Furthermore, it is *because* of this presupposed intersubjectivity that it is possible to determine individual differences in the cognition and descriptions of different persons – both concerning that which is publicly observable and that which is only observable to the persons themselves, such as their mental or internal states. In other word, it is because *both* the knowledge shared by different persons *and* the knowledge of persons which is personal, are perfectly sensible issues of intersubjective discourse *that it is possible to distinguish between what is publicly observable and shared and what is personal*.[8]

Let me illustrate this point by giving the following example. One of my patients tells me: "I am terribly depressed; everything looks so grey and colourless – even the trees and flowers look grey and colourless." Now, could it not reasonably be argued that at least in this case we are not talking about a "public" issue, but rather of something "private," and also that it is a situation in which it would make no sense to maintain that my patient is still using language correctly? Is it not a situation in which any well-defined notions of correct or true assertions have been suspended? Not at all. For one thing, I do understand what my patient is saying. I am perfectly able to communicate with him about his – in this case – curious experience of the colours of trees and flowers. But a condition for maintaining that I understand what he is saying, and for communicating with him about his curious experiences of the colours of trees and flowers is, naturally, that he still uses language correctly when talking about his experiences of these things. That is, it is a condition that he knows the correct implications of terms for various colour categories, and that he knows how to apply them correctly. And it is a condition that what he is talking about is something he may refer to, and about which true and false assertions may be made, i.e. his curious experience of the colours of trees and flowers. Thus, it is a condition that both of us are still using the same language – indeed the very same language that he and I use under normal circumstances to talk about quite ordinary everyday matters; and it is a condition that we are together able to determine *what* he is talking about. If we are able so to do, and thus able to talk about his experiences, however curious, it has to be maintained that he is using language correctly when describing his experiences.

However, it is quite clear that his description of the colours of the trees and flowers is not of *general validity*, and I do not take his description as an attempt on his part to produce descriptions of general validity – i.e. descriptions which would be correct under normal everyday conditions of observation. But an important part of the message he is trying to convey to me – and which I understand – is precisely that his situation is *not* normal, i.e. that his depression affects his perception and description of things in ways which differ from how he normally perceives and describes them. A fact of which he himself is perfectly – and probably painfully – aware.

So, although my patient may feel eminently on his own with his unusual and personal experiences during his depression, neither his experiences nor his descriptions of them are *private*. They are perfectly understandable to others because it is presupposed – by us and by him – that he is using language when describing what he experiences as others would – could they be in his situation and experience what he does. Indeed, our communication about what he experiences relies on the presupposition that what to him is the case or true about his experiences would also be the case or true for others, had they his experience.[9]

Let me conclude my discussion of this example by saying that it shows that individual differences in the cognition and description of reality of different persons may indeed exist and be determinable. However, it also shows that a condition for these differences between persons and their cognition and description to exist, is that persons and language users, despite such differences, share a vast amount of knowledge and correct descriptions of reality. And it shows, furthermore, that the possibility of determining and of talking correctly about such differences relies on the presupposition that, when we describe what we know of or experience, we use language to describe it as others would, if they had our knowledge and experience; however, this in its turn both presupposes and implies that the notions of 'correct' and 'true' of persons are such that what is correct or true, is also correct or true to others. If these presuppositions and implications concerning the intersubjectivity of our cognition and language did not apply, no personal differences in knowledge and experiences could exist, nor any possibility to determine or talk sensibly about such differences – whether these differences concern what is or what is not publicly observable. Conversely, granted these presuppositions, and hence that situations, points of views and observations are in *principle* shareable, it is possible for persons to determine and distinguish between individual differences and variations in their cognition and description concerning both what is and what is not publicly observable, and to determine under what circumstances such individual differences occur. And it is possible *therefore* for persons together with others to develop criteria and standards with which rigorously to determine and distinguish between first-person and third-person observation and description, so crucial for the possibility of developing sciences.

Consequences for cognitive research

Whereas scepticism is no longer in fashion when it comes to the possibility of so-called *third-person* description and cognition of publicly observable reality being true or objective, uncertainty still prevails as to the status of our *first-person* description and cognition of our non-publicly observable internal or mental states, such as our emotions, thoughts and feelings of pain. The aim of my discussion has been to provide *principled* arguments which refute this scepticism.

I have done so by arguing that everyday communication about our internal or mental states in general, as well as any philosophical discussion as to the nature and status of our cognition and descriptions of them, necessarily rests on the presupposition that as users of the language we share with others, we know correct implications of terms for such states, just as we know their correct application. Indeed, so I argue, for the same reasons that we have to assume that we may observe and use language to talk correctly about things in publicly observable material reality, we also have to assume that we may observe and talk correctly about our non-publicly observable mental or internal states. Likewise, in both cases it is logically presupposed that what to a person is true or correct about the things or states he or she observes and experiences, would also be true or correct to other persons, could they observe and experience the things or states that he or she does. These assumptions and conditions of intersubjectivity granted, there is no reason why first-person verbal reports by subjects about their internal or mental states may not be just as correct and reliable data for research in cognitive science as are third-person reports about publicly observable objects and events.

This said, it does not, of course, mean that no *practical* problems may exist in observing and giving adequate or satisfactory descriptions of our internal or mental states. How these problems may be overcome, i.e. the problems referred to as the *methodological* problems of introspective reports, are thoroughly discussed by, among others, the researchers referred to in endnote 2 of this paper.

It is important to stress, moreover, that the epistemological points about the intersubjectivity of language and cognition developed in this paper, do not in themselves entail guidelines, prescriptions or rules for how to set up experiments, nor do they offer suggestions as to techniques and methods to use in order to get at the reliable introspective data needed in cognitive research. Indeed, in general, epistemological points and theories are silent about matters of methodological concern, i.e. how in actual situations we go about acquiring knowledge – or fail so to do. Epistemological theories and points are about the *general* conditions and possibilities of having and acquiring knowledge at all, and about the conditions and possibilities of language to communicate and describe (truthfully, correctly, etc.) what we know of and describe.

It could be objected that although both first-person and third-person data are determined on the common presupposition of intersubjective communication, third-person data are still different in regard to their public accessibility, and that it is *this* difference which is relevant to the debate and scepticism about how they may count as scientific data. And it could quite rightly be objected, furthermore, that the epistemological points made in this paper about the intersubjectivity and communicability of introspective reports does nothing to eradicate scepticism on the grounds that such reports cannot be subjected to third-person validation. This scepticism, so it would seem, will therefore in the final analysis have to be overcome *methodologically*, i.e. by showing that cognitive research using introspective data measures up to the methodological standards and criteria which apply for scientific research in general – such as the possibility of precise description of initial conditions and reproducibility of data (that is, the same introspective reports may be reproduced under the same experimental conditions either by the same subject or by different subjects).

However, firstly, someone being sceptical about the use of introspective reports as scientific data on the grounds that what they are about is not publicly accessible, would not be convinced by the fact that they are reproducible – *any more* than a sceptic concerning the stories he reads in a newspaper would be convinced about their truth by repeated reading. Secondly, although the principled arguments about the intersubjectivity and communicability of conscious mental states do not render reports of such states open to third-person validation, the points argued still stands – namely that these states may be just as amenable to *reliable* observation and *correct* descriptions as are observations and descriptions of publicly observable things or states – and hence may indeed fulfil the requirement of scientific evidence. Since this assumption, as argued above, is a necessary condition for any rigorous determinations of and hence any well-defined distinctions between what is publicly accessible and what is not, *it must be among the very basic conditions on which science itself is founded*.[10]

Consequences for cognitive theory building: Conclusion

However fundamental – and almost embarrassingly banal – the assumption of the intersubjectivity of human cognition and use of language may seem, it has been widely overlooked within philosophy of mind and consciousness. According to the traditional assumption, shared by many philosophers even today, we all start out as "Cartesian subjects," having knowledge and experience of the content of our own mind, i.e. our sense data, perceptions, thoughts, emotions and feelings. From this supposed private, though certain knowledge and experiences "from our own cases" it is believed to be possible to work towards true knowledge of the nature of what causes this content

and the rest of the objective order of reality, including other persons, and to develop a language in which we may talk of this knowledge and experience. This same assumption also seems to inspire and lie behind functional models and accounts of the cognition and use of language of people currently being developed within Cognitive Science – be those models computational *cum* representational or connectionist. However, it would seem to be a insoluble problem for such models to account for how the cognition of an individual – formed in "splendid solipsistic isolation," and with no notion of the intersubjectivity of its own cognition and that of others – would ever come to accord with the cognition of others; let alone account for how any differences in the cognition of different individuals may occur, and how they may come to realise the existence of such differences.

In the current debate "on the nature of consciousness"[11] it has been proposed – in defence of such functional models – that it is both intuitively plausible and logically possible that a *zombie*, defined as a "being" who is physically and functionally identical to a human being,[12] but lacking any conscious experiences whatsoever, will nevertheless be able to behave in ways which are indistinguishable from that of a human being. Indeed, it is claimed to follow *logically* from the fact that this "being" is physically and functionally identical with a human being that "it" will be *conscious,* albeit only functionally, i.e. "it" will be awake, and be able "functionally" to perceive and observe both "its" internal "goings on," such as pains in "its" body, and things in the external world; "it" will even be able to give detailed reports of the content of "its" mental states in ways indistinguishable from those of a human being, and to communicate with others about "its" and their observations and perceptions of both their internal states and things in the world, just as "it" may co-act with others in the world of physical things in ways which are deceptively similar to that of a human being (Chalmers 1996).

In view of the arguments presented in this paper, however, it could well be asked, how such a "being," lacking a notion of the intersubjectivity of "its" own mental states and those of others, be they of internal "goings on" or of things in the external world, could possibly develop or acquire a language being shared with others to report the content of "its" mental states, and hence of "its" perceptions and observations of these "goings on" and things. That is, it could be asked how such a "being" – lacking the notion that what "it" knows, observes and perceives, would also be known, observed and perceived by others entertaining the same mental states – could nonetheless have or come to acquire knowledge shared with others of the correct implications and applications of terms to report the content of "its" mental states, and to communicate with others about the content of theirs. And one may well ask, moreover, how a zombie, lacking the notion of the intersubjectivity of observations and perceptions of the world, and hence lacking the notion of a *publicly shared world*, could possibly develop or acquire the ability to *co-act*

with others, let alone together with others determine the things in the world that their co-action concern. Indeed, ask how a zombie would ever be able to distinguish between what is observationally shared by others and what is personal, i.e. without being able to determine with others *what* in a situation is publicly known and available, and what is only known and available to oneself, and hence without the notion that *knowledge*, *points of view*, *observations and perceptions*, *public as well as personal*, *are inherently intersubjective*.

The lack of these notions and conditions in causal, functional theories of cognition, use of language and behaviour should suffice as reasons why it is neither intuitively plausible, nor logically possible for a zombie to behave, form judgements, and acquire and use language to communicate in ways which are indistinguishable from conscious human beings. What is lacking in such theories which starts from the position of the individual alone set against the rest of the world, is not just a social *context* of others, which enables the individual to confront and compare his knowledge with the knowledge of others with the purpose of determining, for example, whether his or her knowledge is in accordance with theirs, and hence may be considered of general validity and objective, or whether it relies on one's subjective dreams, illusion or imagination. What is lacking is precisely the presupposed intersubjectivity inherent in cognition and use of language *which makes such determination and distinction possible at all*.

This is just another way of saying that a cognitive science which claims that cognitive and linguistic theory building may make do with causal, functional accounts and explanations, whilst ignoring the requirements and conditions for the intersubjectivity of knowledge and experiences, has nothing to do with the cognition and use of language peculiar to human beings.

Notes

1. Ericsson and Simon (1993), Varela (1996), Vermersch (1999), Roy et al. (1999), Lutz and Thompson (2003), Jack and Roepstorff (2002), Overgaard (2001), Gallagher (2003) and Ericsson (2003).
2. It is worth noting in advance, I think, that the aim of the arguments in this paper, most of which are *reductio ad absurdum* arguments, is akin to that of Ryle's in "The Concept of Mind" (Ryle 1949), namely to explore the "logical geography" of the concepts and propositions of mind, and "ways in which we may operate consistently with them, what propositions follow from them and from what propositions they follow."
3. To spell it out, behind questions such as "do we mean the same thing when we talk about 'pain' – or, for that matter about other mental or internal states such as 'memory' or 'recognition' – lies the assumption that such terms are no more randomly applied to mental or internal states of human beings than 'oak' and 'ash' are randomly applied to trees. Indeed, to take such questions seriously is already, albeit implicitly, to endorse certain assumptions of *how language functions in communicative contexts*. Thus, it is implied and assumed that it is possible for language users together to identify – if only rudimentarily – what they

are talking about (e.g. some particular states or properties of our mind, or some sensations felt somewhere in our body), and that, on the basis of such common consent, it is possible to investigate whether the implications of those terms are in fact the same for everyone – and whether we in fact use these terms to refer to the same sort of "things." Conversely, such questions cannot be asked in any sensible way, nor may these terms be "mentioned" without or independently of how language and its terms are used to refer to actual things or events.

4. I am not saying that the fact that we have a language with terms for both objects in reality and pains in bodies *proves* the existence of objects in reality and pains in bodies, nor that all and every concrete statement we put forward about either objects or pains are always or infallibly correct. What I am saying is that we cannot begin to discuss or investigate language and the use of language to talk about such things as objects or pains, without assuming, generally, that both objects in reality and pains in bodies exist as things that we may have knowledge of and use language to talk correctly about. One cannot take part in this kind of philosophers' discussion without committing oneself epistemologically.
5. That is, due to differences in our possibilities of observing these things, our educational background, cultural differences, etc.
6. The importance of this intersubjectivity of human cognition, communication and co-action, becomes clear when we consider that a substantial part of our knowledge of both material reality and of the societies in which we live and co-act with others, does not rely on first-hand personal experiences, but rather is knowledge we have adopted or acquired from others. In this sense, knowledge of reality thus acquired resembles knowledge communicated to us by others about their mental states, and in the sense, furthermore, that both kinds of knowledge relies on and presupposes the conditions of intersubjectivity of cognition and use of language outlined above.
7. As argued extensively elsewhere (Praetorius 2000), this assumed intersubjectivity of cognition, language and of the notion 'truth' can neither be proved nor doubted without being conceded, and hence will have to be taken granted as a *principle*.
8. Empirical studies of the development of cognition and language in children by Tomasello and Rakoczy (2003) seems to corroborate the points made above. Thus, according to Tomasello and Rakoczy, "to understand beliefs young children must learn to differentiate [. . .] between the mental perspective of an individual and "reality." And reality is not just the child's individual perspective of the moment, which may conflict with another person's, nor an intersubjectively shared perspective with other persons, but rather it is objective in the sense that no one perspective is privileged [. . .]. *The notion of reality*, *subjective beliefs*, *and intersubjective perspectives thus form a logical net that can only fully be grasped as a whole*. Comprehending this net as a whole takes children, apparently, several years [i.e. the first 4–5 years of their lives] to accomplish" (Tomasello and Rakoczy 2003, pp. 133–134, italics added).
9. There are of course cases in which this intersubjective understanding breaks down. Indeed, differential diagnosis between different psychiatrical disorders, such as depression, schizophrenic psychosis and autism, as well as research into their neuronal and biochemical basis, relies in large part on the various ways in which and to what extent intersubjective communication and cognition appears to be compromised, and consequently the ability of the patients to experience themselves as coherent, autonomous "selves" *vis-a-vis* others and the world around them (see the discussion in Zahavi and Parnas 2003). And so does the choice of possible neuro-pharmacological and other therapeutic interventions for these groups of patients.
10. I am indebted to an anonymous referee who prompted me to clarify this point, and for pointing out that Erwin Schrödinger (1935) presented similar augments in his discussion

of the conditions on which scientific knowledge within physics are based (Schrödinger, 1935).

11. Contributions in Block et al. (1997) and Chalmers (1996).
12. That is, "he will be processing the same sort of information, reacting in a similar way to inputs, with his internal configurations being modified appropriately" (Chalmers 1996. p. 95).

References

Block, N., Flanagan, O. and Güzeldere, G. 1997. *The Nature of Consciousness. Philosophical Debates*. Cambridge, MA: MIT Press.

Chalmers, D. 1996. *The Conscious Mind. In Search of a Fundamental Theory.* New York: Oxford University Press.

Baars, B. 2003. How brain reveals mind: Neural studies support the fundamental role of conscious experience. *Journal of Consciousness Studies* 10: 100–114.

Dennett, D. 2001. The fantasy or first-person science. Nicod Lectures. Private circulation.

Ericsson, K. A. 2003. Valid and non-reactive verbalization of thoughts during performance of tasks. *Journal of Consciousness Studies* 10: 1–18.

Ericsson, K. A. and Simon, H. A. 1993. *Protocol Analysis: Verbal Reports as Data*, revised 1st edn., 1984. Cambridge, MA: MIT Press.

Gallagher, S. 2003. Phenomenology and experimental design: Towards a phenomenologically enlightened experimental science. *Journal of Consciousness Studies*. 10: 85–99.

Jack, T. and Roepstorff, A. 2002. Introspection and cognitive brain mapping: From stimulus – response to script-report. *Trends in Cognitive Science*.

Lutz, A. and Thompson, E. 2003. Neurophenomenology. *Journal of Consciousness Studies* 10: 31–52.

Overgaard, M. 2001. The Role of Phenomenological Reports in Experiments on Consciousness. *Psycoloquy* 12 (29).

Praetorius, N. 2000. *Principles of Cognition, Language and Action. Essays on the Foundations of a Science of Psychology*. Dordrecht/Boston/London: Kluwer Academic Publishers.

Roepstorff, A. and Frith, C. 2004. What's at the top in the top – down control of action? Script-sharing and top – top control of action in cognitive experiments. *Psychological Research* 68: 189–198.

Roy, J. M., Petitot, J., Pachoud, B. and Varela, F. J. 1999. *Naturalizing Phenomenology*. Stanford, CA: Stanford University Press.

Ryle, G. 1949. *The Concept of Mind.* London: Hutchinson.

Schrödinger, E. 1935. Quelques remarques au sujet des bases de la connaissance scientifique, *Scientia* LVII: 181–191.

Tomasello, M. and Rakoczy, H. 2003. What makes human cognition unique? From individual to shared to collective intentionality. *Mind & Language* 18: 121–147.

Varela, F. J. 1996. Neurophenomenology: A methodological remedy to the hard problem. *Journal of Consciousness Studies* 3: 330–50.

Vermersch, P. 1999. Introspection as practice. *Journal of Consciousness Studies* 6: 17–42.

Zahavi, D. and Parnas, J. 2003. Conceptual problems in infantile autism research: Why cognitive science needs phenomenology. *Journal of Consciousness Studies* 10: 53–71.