

Comments on Frith & Happe, "Theory of mind and self consciousness: What is it like to be autistic?" (Mind and Language)

Raffman

For philosophers of mind, what is particularly intriguing about Frith and Happe's discussion is the idea that autistics know their own minds only, or almost only, by observing their own behavior. The authors do not say exactly how the Asperger subjects gain this knowledge of themselves; the clearest statement I can find occurs on p.*: "If the person with autism can judge their own inner states only by their actions, it might be important to teach behaviours which express for oneself what one thinks and feels" [emphasis added—DR]. Nevertheless, I take Frith and Happe's central hypothesis to be that instead of having the direct or immediate or non-inferential self-knowledge we normals take for granted, autistics know what is transpiring in their own minds only by applying an explicit theory of mind *ex post facto* to their own behavior. They do not know their own minds in the normal way, by introspection.

The asymmetry between first-person introspective knowledge and third-person knowledge of the mental states of others has long been a topic of interest to philosophers. The nature and etiology of our knowledge of others' minds seems clear enough: we observe their behavior and infer from that evidence what they think. But the immediacy of our knowledge of our own minds is mysterious: we just know what we are thinking, we don't need to collect evidence. If Frith and Happe are correct, autistics lack this immediate rapport with their own thoughts; they do need to collect evidence. Thus following a long and fruitful tradition of studying abnormal psychological processes in order to understand normal ones, if we can figure out what goes wrong in the autistics, perhaps we can shed some light on the nature of (normal) introspection. That would be exciting indeed.

My role as a commentator, of course, is to raise questions for further reflection. To that end, let me introduce a distinction between two interpretations of the claim that an individual fails to be self-aware. (It is surely not an exhaustive distinction.) On the one hand, such a claim might mean that an individual lacks the concepts of belief and desire and the rest in terms of which we conceive and express our mental states and explain and predict our own behavior and the behavior of others. (Frith and Happe may have something like this interpretation in mind when they write that autistics "lack the cognitive machinery to represent their thoughts and feelings as thoughts and feelings" [p.*].) For all this first interpretation says, if such an individual were to acquire the relevant folk psychological concepts, he would employ them to express self-knowledge gained in the normal way by introspection. On the other hand, an individual might be unaware of himself in a more radical way such that, even if he mastered the folk psychological concepts, still he could apply them to his own case only by observing his own behavior.

The very interesting idea advanced by Frith and Happe, I take it, is that autistics are unaware of themselves in this second, stronger sense: even when they have mastered the relevant concepts, they can apply those concepts in their own case only indirectly. What I will suggest in these brief comments is that the evidence Frith and Happe present may support only the hypothesis that autistics are unaware in the first, weaker sense. Call the latter hypothesis the 'Conceptual Incompetence' hypothesis or 'CI' for short. The authors will want to rule it out as their research progresses.

Two sets of considerations in particular seem to favor CI. First are the autobiographical reminiscences of early childhood recorded by Williams (1994), Jolliffe, Lansdown, & Robinson (1992), Grandin (1984), and Gerland (1997), all cited by Frith and Happe. For example: “I was sick to death of my attention wandering onto the reflection of every element of light and colour...” (Williams); “I sometimes got annoyed once I realized that I was expected to attend to what other people were saying” (Jolliffe, Lansdown, & Robinson); “[t]he feeling was like a constant feeling of stage fright all the time” (Grandin); “[t]ime and again I was very hurt when people said they knew things about me...” (Gerland). These Asperger subjects are currently reporting their past mental states—states they underwent, we are to assume, before they acquired the ability to conceive and report them as wandering attention, annoyance, fright, hurt, and so on. It is hard to see how they could make such reports if they hadn’t been somehow aware of the states in question at the time they underwent them, contrary to what Frith and Happe contend. I suppose it’s possible that these subjects are currently self-ascribing mental states by inference from memories of past behaviors; some of the reminiscences do in fact concern past behaviors (e.g., “As a child, I often talked out loud...”; “I remember minutely observing how the sand flowed...”). But the idea that the autistics’ current self-ascriptions are in general inferred from memories of past behaviors seems untenable. Also, since the ability to self-ascribe mental states was acquired at a time later than the time of the remembered episodes, it can’t be that these subjects are making their current self-ascriptions by remembering past self-ascriptions, i.e., by remembering past verbal behaviors.

Second, Frith and Happe hypothesize that self-knowledge and knowledge of other minds are underwritten by the same cognitive mechanism. In this connection, they cite evidence indicating that the ability to ascribe mental states to oneself covaries with the ability to ascribe mental states to others (p.*, e.g.). As they themselves acknowledge, however, such a hypothesis is at least *prima facie* counterintuitive: intuition suggests that awareness of one’s own mental states, gained by introspection, and awareness of others’ mental states, gained by observation of their behavior, employ radically different mechanisms. Among other things, the seeming immediacy and privilege of introspection find no counterparts in knowledge of others; hence it is difficult to see how the two operations could share a single mechanism.

In contrast, CI leaves room for the possibility that self-knowledge and knowledge of other minds are subserved by different mechanisms in autistics and normals alike. A proponent of CI can say that what self-knowledge and knowledge of other minds have in common is not their etiologies or manners of acquisition, as it were, but rather (roughly) the conceptual structure in terms of which they are formulated and made usable for explaining and predicting behavior. If there is a common mechanism, it is more likely to underlie that conceptual competence and the consequent ability to ascribe mental states to oneself and others. Such a hypothesis is compatible with the evidence for selective impairment of, and covariation between, the ability to ascribe mental states to oneself and the ability to ascribe them to others.

CI also appears compatible with other evidence presented by Frith and Happe. For instance, it sits well with the fact that Asperger subjects acquire the ability to ascribe mental states to themselves and others only with considerable effort and explicit training. Many activities—including “intellectual” activities like reading and mathematical calculation, to say nothing of any kind of scholarly work—can be mastered only laboriously when mastered late, and may never be executed with the ease of one who learns them in early childhood. Folk psychological attribution may well be the same.

I will close by noting that much of the evidence Frith and Happe present to show impaired self-awareness or self-knowledge in autistics seems equally interpretable as showing only impaired (i.e., unusual or even bizarre) experience. The discussion of “abnormal sensory and pain experiences”, “hypo- and hyper-sensitivity to sound, light or touch”, and much if not most of the subject testimony cited by Frith and Happe is interpretable in this way, it seems to me. Of course, a distinction between experience and introspection or self-knowledge (consciousness and self-consciousness, awareness and self-awareness) is difficult to draw even in theory; but even allowing for that, the authors will need to make clear how the evidence they cite pertains to the “second order” states as opposed to the “first order” ones. Be that as it may, the evidence they present, and their discussion of it, will be of significant interest to philosophers of mind.