Borrower: RLG:NJPG
Lender String: RLG:GAEG

ILL Number: 129376

Patron: McGeer, Victoria


Volume: Issue: 2003

Page(s): 1-15

Call #: BF723 .P48 153 2003
Location: 9

Ariel: 128.112.205.74
Fax: 609-258-0441

Shipping Address:
Princeton University
Princeton University Library
Firestone Interlibrary Services - Borrow
One Washington Road
Princeton, NJ 08544-2098 USA

Your Maxcost: 0.00

Charges: $ Exempt

THIS IS NOT AN INVOICE !!
INVOICE WILL BE SENT SEPARATELY
Theory of Mind in Autism and Schizophrenia
A Case of Over-optimistic Reverse Engineering

PHILIP GERRANS
University of Adelaide

VICTORIA McGEER
Australian National University

CONSTRAINTS ON MODULAR THEORIES:
HORIZONTAL INTEGRATION, VERTICAL INTEGRATION,
AND THE THEORY-OF-MIND MODULE

Although autism and schizophrenia present as widely differing disorders, there are intriguing connections between them, which have led to a number of speculative attempts at theoretical unification. First, the term autism, now confined to autistic subjects, was first coined by Eugene Bleuler in the 19th century to capture the social isolation and lack of impetus to engage with the world characteristic of schizophrenics experiencing what today would be called “negative symptoms.” Second, both autistics and schizophrenics have sensory-motor disorders including perceptual abnormalities, stereotypy and disorganization, unusual patterns of affect, attentional and executive deficits, and social problems. Third, these are both disorders where the range of symptoms extends in a characteristic pattern across many cognitive domains, although general intelligence is sometimes spared.

Explaning these disorders, either separately or together, poses a two-staged challenge of horizontal and vertical integration (Robbins, 1997). Horizontal
integration means accounting for and so unifying a pathological set of behavioral and phenomenological symptoms by reference to a cognitive model of normal function. Vertical integration refers to the way in which the symptoms thus unified are linked to neurobiology by a theory that shows how the relevant cognitive function is normally implemented, hence pathologically, in human neurobiology. Thus cognitive models initially constructed by abstracting from neurobiology at the stage of horizontal integration are ultimately mapped to neurobiological function at the stage of vertical integration via an implementation theory. This two-staged project of horizontal and vertical integration is complicated by the developmental dimension of neurocognitive disorders (Karmiloff-Smith, 1998; Thomas & Karmiloff-Smith, in press). In disorders acquired during adulthood, theorists can be more confident of the ways in which distinctive abnormalities result from damage to, or disruption of, normally articulated cognitive systems, whereas the distinctive abnormalities of developmental neurocognitive disorders inevitably reflect a long history of organismic adaptation and compensation.

In this chapter, we examine a theory of cognitive function that aims to resolve these difficulties for both autism and schizophrenia. This theory proposes, first, a cognitive account of the core symptoms of each disorder, thereby meeting the goal of horizontal integration. Second, it suggests a way of mapping a range of diverse symptoms onto a highly specific neural substrate, thereby meeting the goal of vertical integration. The theory, originally proposed by Simon Baron-Cohen and collaborators for autism and extended by Chris Frith for schizophrenia, is that both disorders result from the malfunction of a single cognitive system, incorporating, most importantly, the “theory-of-mind” module (or ToMM) (Baron-Cohen, Leslie, & Frith, 1985; Frith, 1992). According to these theorists, the ToMM is required for an understanding of intentional agency because it underwrites autistic peoples’ capacity to conceptualize mental states, specifically beliefs and desires, as part of a theory of behavioral explanation. In brief, ToMM theorists suggest that if this specific cognitive module fails to develop properly (perhaps due to failure of subsidiary cognitive modules that play an important role in triggering ToM), the result is autism; if it fails in maturity, the result is schizophrenia. The idea is that the autistic child’s specifically social failures and the schizophrenic person’s disrupted sense of agency may have a common causal-cognitive structure basis in malfunction of a module specialized for the representation of mental-state concepts. There are, of course, many differences between these disorders despite the similarities just noted. However, as Christopher Frith reminds us, this need not tell against the project of finding a unified theory encompassing both. Given that autism is a developmental disorder, whereas schizophrenia usually occurs first in adulthood, we might reasonably expect significant (although cognitively related) differences in their typical symptoms. As Frith put it:

Despite its initial appeal, we argue that schizophrenia fails. There is no single ToM social reasoning and behavior in normal ToMM module, realized in neural substrate, broken down in schizophrenia. In other words, we are in the first stage of horizontal integration.

CONSTRANTS ON MODULARITY

Cognitive Domains: Actual or Virtual?

Our objection to the ToM modularity hypothesis is for or against neural localization. We accept realized in distributed neural architecture constraits modularity hypotheses must meet be an actual cognitive domain subserved by information specific to that domain, it must have a neural assembly—that is, an assembly that If this constraint cannot be met, then even if it comprises a pathological behavioral domain, to be a cognitive disorder is merely virtual modularity but the appearance is misleading, the apparent ToM deficits are a consequence and (possibly) noncognitive malfunction.
It is likely that the cognitive deficit in autism is present from birth, although not reliably detectable until about the third year (Schopler & Mesibov, 1988). As a consequence the whole course of development must be abnormal. There is evidence that a proportion of schizophrenic patients show signs of social abnormalities during childhood (Castle, Wessely, & Murray, submitted). However in most cases of schizophrenia development appears to be entirely normal until the first breakdown, typically in the early 20s. My proposal is that people with schizophrenia resemble people with autism in that they too have impairments in the mechanism that enable them to empathize. However in most cases this mechanism was functioning adequately until their first breakdown. Given these very different developmental histories this deficit will be manifest in different ways. The autistic person has never known that other people have minds. The schizophrenic knows well that other people have minds but has lost the ability to infer the contents of these minds: their beliefs and intentions. They may even lose the ability to reflect on the contents of their own mind. However they still have available ritual and behavioural routines for interacting with people which do not require inferences about mental states. (Firth, 1992, p. 121)

Despite its initial appeal, we argue that the ToMM theory of autism and schizophrenia fails. There is no single ToM module responsible for successful social reasoning and behavior in normal subjects; hence there is no dedicated module, realized in neural substrate, that fails to develop in autism or breaks down in schizophrenia. In other words, we argue that the ToMM theory fails at the first stage of horizontal integration.

CONSTRAINTS ON MODULAR THEORIES

Cognitive Domains: Actual or Virtual

Our objection to the ToM modularity hypothesis does not hinge on evidence for or against neural localization. We accept that cognitive functions can be realized in distributed neural architectures. In our view, the only structural constraint modularity hypotheses must meet is the following: In order for there to be an actual cognitive domain subserved by a dedicated module for processing information specific to that domain, it must be realized in a functionally specific neural assembly—that is, an assembly that serves no other cognitive function(s). If this constraint cannot be met, then even supposing a particular disorder comprises a pathological behavioral domain, the domain specificity of the underlying cognitive disorder is merely virtual: There may be an appearance of modularity but the appearance is misleading. In both autism and schizophrenia, the apparent ToM deficits are a consequence of a disunified array of cognitive and (possibly) noncognitive malfunctions. This implies that whatever unity
obtains in the domain of social cognition is merely virtual, the result of a number of interacting subsystems whose interaction is not governed by the operations of a single "theory-of-mind" mechanism.

**Cognitive Capacities: Developmentally Set or Developmentally Constructed**

To say that a cognitive domain is actual—that is, subsumed by a dedicated module—is a synchronic hypothesis about cognitive architecture. It does not depend on claims about the diachronic history of cognitive development. Thus objections to modular nativism should be disentangled from objections to modularity per se (Karmiloff-Smith, 1998; cf. Elman et al., 1996). A cognitive capacity may be modular in the sense of functionally discrete without such modularity being genetically predetermined. If so, there can be failure of typical modularization without failure of a genetically specified module.

To illustrate this difference, consider two alternative explanations of specific language impairment (SLI), a deficit in language ability that spares other cognitive capacities. The classical approach to this disorder, on which nativist theories of ToM deficits are modeled, is one of straightforward reverse engineering: In order to account for the unified and specific behavioral deficit, classical theorists posit a genetically presupposed dedicated language module for deriving grammatical rules that fails to mature normally. Evidence of heritability is often taken as further support for the innate modular theory (Gopnik & Goad, 1997; Van der Lely, 1997).

Another possibility is that SLI is not the result of failure in a genetically presupposed modular language capacity at all. It is, rather, the absence of a modularized capacity—hence, a specific language impairment—due to the developmental impact of a subtle hearing deficit (Donnai & Karmiloff Smith, 2000; Tallal, 1985, 1988; Thomas & Karmiloff-Smith, in press). In order to read linguistic structure into the acoustic stream, children need to be able to detect significant acoustic variation in that stream, but the SLI subject is unable to do this. For example, the SLI child might just hear "ough" instead of "D" "O" "G" where "dog" comes in the middle of an acoustic stream. Wright, Lombardino, Puranik, Leonard, and Merzenich (1997) found that in order for SLI subjects to distinguish linguistically significant variations against the masking effects of surrounding sounds, they needed to be amplified 45 dB above the surrounding stream (see also Tallal, Miller, & Fitch, 1995, on the ameliorating effects of lengthening phonemic transitions). The hearing of these subjects would not show up as abnormal in standard tests because it is their ability to overcome masking effects that is the problem, not auditory function per se. (Subjects can hear the sounds perfectly well as long as they are not masked by preceding and succeeding sounds.) If this explanation is right, then the link between SLI as a behavioral domain and the underlying cognitive domain is not, as it initially appears, the result of an impairment in the product of a developmental cascade, which has nothing to do with language proper like that of Stephen Pinker that SLI is every specified module devoted to syntactic development (1998, 2002; Karmiloff-Smith, 1998; Karmiloff-Smith, in press). This explanation will not undermine integration for SLI as a developmental neural substrate of a specific grammatical development, theorists ought to be looking at a cognitive system dedicated to processing phonology.

We call a cognitive capacity developed if a module whose cognitive architecture is such a cognitive capacity is developmentally constructed. In other cases, that achievement is progressively strong and environmental interactive development is the result of experience in consequence of lower order competencies fail to develop as a result of environment in consequence of lower order competencies fail to develop as a result of environment in consequence of lower order competencies fail to develop as a result of environment in consequence of lower order competencies fail to develop as a result of environment in consequence of lower order competencies fail to develop as a result of environment in consequence of lower order competencies fail to develop as a result of environment in consequence of lower order competencies fail to develop as a result of environment in consequence of lower order competencies fail to develop as a result of environment in consequence of lower order competencies fail to develop as a result of environment in consequence of lower order competencies fail to develop as a result of environment in consequence of lower order competencies fail to develop as a result of environment in consequence of lower order competencies fail to develop as a result of environment in consequence of lower order competencies fail to develop as a result of environment in consequence of lower order competencies fail to develop as a result of environment in consequence of lower order competencies fail to develop as a result of environment in consequence of lower order competencies fail to develop as a result of environment in consequence of lower order competencies fail to develop as a result of environment in consequence of lower order competencies fail to develop as a result of environment in consequence of lower order competencies fail to develop as a result of environment in consequence of lower order competencies fail to develop as a result of environment in consequence of lower order competencies fail to develop as a result of environment in consequence of lower order competencies fail to develop as a result of environment in consequence of lower order competencies fail to develop as a result of environment in consequence of lower order competencies fail to develop as a result of environment in consequence of lower order competencies fail to develop as a result of environment in consequence of lower order competencies fail to develop as a result of environment in consequence of lower order competencies fail to develop as a result of environment in consequence of lower order competencies fail to develop as a result of environment in consequence of lower order competencies fail to develop as a result of environment in consequence of lower order competencies fail to develop as a result of environment in consequence of lower order competencies fail to develop (Gopnik & Goad, 1997). Yet another possibility is that SLI is not the result of a genetic deficiency per se, but rather, the result of a genetic predisposition, which is later modified by an experience or environment that is genetically predisposed to the capacity in question.

**Cognitive Versus Mechanical Explanation Guarding Against Methodological Bias**

Explaining any cognitive disorder begins with a descriptive pattern of disabilities and spared abilities. There, theorists hypothesize underlying theoretical profile, testing these hypotheses against sources of evidence (e.g., cognitive-behavioral, computational models), and then modify these hypotheses as necessary. This methodological practice of reverse engineering produces a bias in explanation toward theories of cognitive structures. This bias is not universal—some theories may be on the right track (e.g., computational models), and therefore more concretely induced bias can easily slide into alternative explanations are not appropriately considered when it comes to experimental design, and are more elegant to postulate an X-module to explain...
is merely virtual, the result of a num-
ber of interaction is not governed by the opera-
tion of the system.

**Set**

That is, subserved by a dedicated mod-
lified cognitive architecture. It does not de-
scribe the story of cognitive development. Thus
a set of straightforward reverse engi-
eered and specific behavioral deficit, clas-
sified as a language module for
ature normally. Evidence of heritabil-
ity of the innate modular theory (Gopnik &
Meltzoff) at all. It is, rather, the absence of a
language impairment—due to the de-
cision (Donnai & Karmiloff-Smith, 2000; Smith, in press). In order to read lin-
early, children need to be able to detect
words, but the SLI subject is unable to do
ear “ough” instead of “D” “O” “G”
auditory stream. Wright, Lombardo,
(2015) found that in order for SLI subjects to
perform significantly better than the surrounding
1995, on the ameliorating effects of
hearing of these subjects would not
because it is their ability to overcome
auditory function per se. (Subjects can
they are not masked by preceding and
right, then the link between SLI as a
cognitive domain is not, as it initially
appears, the result of an impairment in native linguistic capacity. Instead SLI is
the product of a developmental cascade based on a subtle perceptual deficit
which has nothing to do with language per se (Wright et al., 1997). Thus, theses
like that of Stephen Pinker that SLI is evidence for the presence of a genetically
specific module devoted to syntactic processing are undermined (Gerrans,
2002; Karmiloff-Smith, 1998; Karmiloff-Smith et al., 2003; Thomas & Karmiloff-
Smith, in press). This explanation will naturally complicate the project of verti-
cal integration for SLI as a developmental disorder: Instead of searching for the
neural substrate of a specific grammatical capacity in early stages of linguistic
development, theorists ought to be looking for the neural substrate of a cogni-
tive system dedicated to processing phonological distinctions.

We call a cognitive capacity _developmentally set_ insofar as it is achieved by
a module whose cognitive architecture is genetically prespecified. By contrast,
a cognitive capacity is _developmentally constructed_ if the mechanism or mech-
isms that achieve it are progressively structured in consequence of an organism’s
environmentally interactive developmental history. Thus it counts against a ded-
cated modularity hypothesis as a developmental thesis about high-level cogni-
tive capacities such as syntax or ToM if it is shown that such specialized
competencies fail to develop as a result of the way organisms interact with their
environment in consequence of lower order abnormalities. However, as in the
case just described, it may be that some cognitive capacity is still an actual cog-
nitive domain, capable of being selectively impaired in adulthood, even if it
depends developmentally on mediating systems that are cognitively unrelated to
the capacity in question.

**Cognitive Versus Mechanical Explanations: Guarding Against Methodological Bias**

Explaining any cognitive disorder begins with its behavioral profile: the distinc-
tive pattern of disabilities and spared (or sometimes superior) abilities. From
there, theorists hypothesize underlying structures that account for this behav-
ioral profile, testing these hypotheses against further external and/or internal
sources of evidence (e.g., cognitive-behavioral and/or neurological studies, as
well as computational models), and then modifying or replacing their theories.
This methodological practice of reverse engineering is indispensable yet intro-
duces a bias in explanation toward theories that postulate dedicated higher or-
der cognitive structures. This bias is not necessarily a bad thing, because the
theories it favors may be on the right track. However, we note that this method-
ologically induced bias can easily slide into a reverse-engineering fallacy if al-
ternative explanations are not appropriately considered. This is particularly true
if alternatives seem theoretically more complicated and empirically less tract-
able when it comes to experimental design. Still, although it may be more el-
egant to postulate an X-module to explain a set of symptoms, the truth for any
particular disorder may often lie in a complex multiplicity of overlapping cognitive and noncognitive causes. To keep the full range of conceptual possibilities in mind, we think a distinction between cognitive and purely mechanical explanations of cognitive function can be usefully clarified.

**Cognitive Explanations.** An explanation for a disorder counts as genuinely cognitive if it makes essential or theoretically ineliminable reference to a system's design by way of invoking a malfunctioning cognitive mechanism or mechanisms (Dennett, 1978). For example, we lose explanatory power if we explain visual neglect purely in neural terms without reference to the visual task the neural system implicated normally performs. Of course, because cognitive mechanisms are realized in neural substrate, cognitive malfunctions are inevitably problems at the level of neuronal functioning. However, a cognitive theory initially abstracts from neural realization in identifying a cognitive function and then maps that function to its neural substrate via an implementation theory. Consider an analogous case: A computer malfunction is correctly explained at the design or programming level (analogous to the cognitive level) if its occurrence is contingent upon some flaw in the program no matter how that program is physically realized—differently, as it may be, in two different computer systems (cf. below note 1).

Thus, in terms of the taxonomy introduced earlier, a cognitive module, putatively identified as the cause of a disorder, may be actual or virtual. If it is actual, then there is selective damage to a dedicated cognitive mechanism, making the reverse engineering inference from behavior to mechanism relatively straightforward. In this case, the module implicated in the disorder is architecturally real (see Figure 12.1). By contrast, the postulated unifying cognitive function may be virtual: It gives the appearance of being achieved by a dedicated mechanism, but the underlying architectural reality is of a number of interacting, possibly lower order, quasi-independent cognitive subsystems. The proposed cognitive function is descriptively too abstract to serve as an implementation theory and, in this sense, the modularity hypothesis misdescribes the organization of the agent's cognitive system (see Figure 12.2).

What kind of evidence could distinguish between these two alternatives? One important piece of evidence exploits the putative link between autism and schizophrenia. For, should it turn out as we suspect in schizophrenia, that postdevelopmental abnormalities in lower order systems suffice to produce the pathological symptoms, this is good reason to question the existence of a module specialized for cognizing that domain, hence one that could be differentially affected in the course of autistic development.

**Mechanical Explanations.** So far, we have claimed that reasoning backward from a behavioral profile may misleadingly invite high-level modularity hypotheses for specific cognitive disorders, either diachronically, as a developmental
complex multiplicity of overlapping cognitive functions. The full range of conceptual possibilities extends from cognitive and purely mechanical explanations to be usefully clarified.

A disorder counts as genuinely eliminable reference to a system’s functioning cognitive mechanism or mechanism. We lose explanatory power if we explain without reference to the visual task the system performs. Of course, because cognitive substrates, cognitive malfunctions are inevitable, but the extinction of malfunctions is correct explained at a level analogous to the cognitive level if its occurrence in the program no matter how that program may be, in two different computer systems.

As introduced earlier, a cognitive module, virtual or actual, may be actual or virtual. If it is a dedicated cognitive mechanism, made from behavior to mechanism relatively unimplicated in the disorder is architectural reality is of a number of independent cognitive subsystems. The system is too abstract to serve as an implementation module hypothesis misdescribes the system (see Figure 12.2).

How distinguish between these two alternatives? The putative link between autism and schizophrenia (see Figure 12.3) is the more than the existence of a modality, hence one that could be developmental.

We may have claimed that reasoning backward from symptoms to substrate often invite high-level modularity hypothesis, or synchronically, as a hypothesis about contemporary architectural organization. However, the postulation of dedicated cognitive entities can also mislead if it turns out that the disorder is not cognitive at all. This is the case with structureless neuropathologies such as Alzheimer’s disease. The interest of this type of case for our purposes is that structureless neuropathology can sometimes produce behavioral outcomes that misleadingly invite cognitive unification at higher levels and thereby virtualize certain cognitive feats.

FIGURE 12.1. Reverse engineering: Mapping symptoms to neural substrate via a cognitive model.

FIGURE 12.2. Reverse engineering fallacy 1: Lower level cognitive explanation.
Consider, for instance, the disorder phenylketonuria (PKU). Although PKU has many symptoms that are similar to autism, it is produced not by the failure of a specific cognitive module, but by the absence of an enzyme that synthesizes an amino acid. In this case, postulating the failure of a higher order mechanism is ultimately mistaken, not because the postulated module fractures into a subset of cognitively or developmentally more basic ones, but because the cause of the disorder is essentially noncognitive: It does not arise through the breakdown of any cognitive system qua cognitive system. Because there is no cognitive story to tell in this case, the correct approach for achieving vertical integration is bottom up: We explain the neural malfunction and treat the resultant behavior as the outcome of haphazard interference with the development of a number of arbitrarily involved cognitive functions. The correct explanation, as illustrated in Figure 12.3, is entirely at the level of neural mechanism.1

Autism and schizophrenia are interesting cases precisely because their diversity of symptoms and lack of uniformity from case to case continue to encourage the idea, popular among pioneers of neurology, that the deficits in question may be essentially mechanical rather than cognitive. Nonetheless, clinicians are reluctant to endorse this conclusion because the disorders do seem to form a pathological domain, rather than a random collection of symptoms. Hence, unifying cognitive theories, such as ToMM, continue to be attractive. However, ToM theorists go too far in proposing that high-level processing of social information constitutes a developmentally set or, indeed, architecturally real cognitive system whose malfunction is implicated in both autism and schizophrenia.

![Diagram](https://example.com/diagram.png)

**FIGURE 12.3.** Reverse engineering fallacy 2: Mechanical explanation.

In both autism and schizophrenia we presume the ToMM hypothesis can be explained in terms of malfunction at sensorimotor or perception of a lower order cognitive module, not as a higher order social cognition. In the case of autism, we remain nonplussed because determining the nature of the social developmental precursor to autism is a complex matter. However, we do think that ToM, on early sensorimotor malfunction involves a number of higher order social cognition. In the case of schizophrenia we develop alternatives to a ToMM account thinking with option (a), the motor control produces a significant set of symptoms, whose neural substrate is the prefrontal cortex. In the diffuse cortical and subcortical effects.

These arguments are directed at a specific form of ToMM which claims that a malfunctioning ToMM is the symptoms of these disorders. The ToMM hypothesis claim that autism and schizophrenia involve which only an essential core, typically four ToMM deficits. However, this weaker claim is supported by evidence of dissociation between symptoms occurring in both these disorders. ToMM were there convincing evidence of dissociation, or even, for example, there be some evidence of dissociation between symptoms occurring in both these disorders. The so-called core (i.e., ToM) aspects of the core deficits are often problems to constitute an autistic (or schizophrenia) core deficits can be produced in either case. There is no reason to save the ToMM hypothesis, apart from the argument against strong ToMM for the more diffuse case.

In light of these difficulties, we do claim that all symptoms, core and peripheral, are not dependent on the same module, but specific the cognitive abnormalities seen in the autism and schizophrenia disorder involves a number of subsystems like ToMM.
phenylketonuria (PKU). Although PKU autism, it is produced not by the failure in the absence of an enzyme that synthesizes the failure of a higher order mechanism postulated module fractures into a suite of basic ones, but because the cause of the core loss: It does not arise through the breakative system. Because there is no cognitively approach for achieving vertical integration dysfunction and treat the resultant behavior with the development of a number of functions. The correct explanation, as the level of neural mechanism. 1

In resting cases precisely because their diversity from case to case continue to endeavor of neurology, that the deficits in rather than cognitive. Nonetheless, conclusion because the disorders do seem than a random collection of symptoms. Such as ToMM, continue to be attractive. Proposing that high-level processing of mentally set or, indeed, architecturally is implicated in both autism and schizo-

STRICT AND WEAK TOMM HYPOTHESIS

In both autism and schizophrenia we propose that the symptoms unified under the ToMM hypothesis can be explained in either one of two ways: (a) as a result of malfunction at sensorimotor or perceptual levels (i.e., in terms of malfunction of a lower order cognitive module), or (b) as a result of purely mechanical malfunction.

In the case of autism, we remain neutral between the two alternatives because determining whether the type of sensorimotor disorder we identify as a developmental precursor to autism is low-level cognitive or mechanical is not a simple matter. However, we do think that a developmental cascade consequent on early sensorimotor malfunction is a sufficient explanation of mentalizing abnormalities, especially once the role of social interaction, modulated by sensory experience, is given its due in characterizing the development of higher order social cognition. In the case of schizophrenia, there are two well-developed alternatives to a ToMM account that fit the pattern we outlined. In keeping with option (a), the motor control account locates the malfunction that produces a significant set of symptoms in a lower order cognitive subsystem whose neural substrate is the premotor cortex. In keeping with option (b), the misconnection account treats schizophrenia as a mechanical malfunction with diffuse cortical and subcortical effects.

These arguments are directed at a strong version of the ToMM hypothesis, which claims that a malfunctioning ToM module accounts for all significant symptoms of these disorders. The ToMM theorist may also defend a weaker claim that autism and schizophrenia involve a conglom of cognitive capacities of which only an essential core, typically having to do with social cognition, are ToMM deficits. However, this weaker claim is open to the following objections: First, if the affected capacities are genuinely independent, then there should be some evidence of dissociation between so-called core and peripheral symptoms occurring in both these disorders. This has yet to be established. But even more, there being evidence of dissociation, the ToMM hypothesis could no longer play an explanatory role in autism (or schizophrenia) per se, but only for so-called core (i.e., ToM) aspects of the disorder. This only postpones the question of why these “core” problems are often enough conjoined with other problems to constitute an autistic (or schizophrenic) behavioral type. Moreover, if the core deficits can be produced in either of the ways we suggest, then there is no reason to save the ToMM hypothesis at all. This in effect replays the argument against strong ToMM for the more restricted set of “core” symptoms.

In light of these difficulties, the weak ToMM theorist may revert to the claim that all symptoms, core and peripheral, are connected even though they are not dependent on the same module. But this is to concede that however specific the cognitive abnormalities seem in autism (or schizophrenia), the disorder involves a number of subsystems linked by neural architecture or neuro-
chemical modulation. In other words, as in the case of PKU or "misconnection" explanations of schizophrenia, the correct level of explanation is mechanical.

In sum, the weak ToMM hypothesis is not a stable fallback position. But the strong ToMM hypothesis—that all significant symptoms of these disorders are in fact cognitively unified by the ToM module—is far too strong to remain defensible, even in the eyes of some of its main proponents. We conclude that although the ToMM hypothesis has generated much valuable research, it is in the end an instance of overly optimistic reverse engineering.

THE CASE OF AUTISM

Autistic individuals share a distinctive triad of impairments in social, communicative, and imaginative capacities (the latter demonstrated by the absence of pretend play in childhood and restricted interests and activities that persist throughout life) (Rutter & Schopler, 1987; Wing & Gould, 1978, 1979). Although 75% of individuals diagnosed with autism are intellectually handicapped in a general way (as reflected in low IQ scores), the remaining 25% have normal to high IQs and often perform well, and sometimes better than average, on reasoning tasks that do not require any understanding of the mental life of agents. In contrast, on so-called "theory-of-mind" tests, these "high-functioning" autistic children are significantly impaired when compared with normal children and even those with Down's syndrome who are matched with them by mental age (for a review of research, see Baron-Cohen, 2001). For instance, on first-order false-belief tasks, which require subjects to predict another's behavior on the basis of attributing them to a false belief, children will normally pass by a mental age of 4 to 5 years (Wimmer & Perner, 1983). Autistic individuals, if they pass at all, only do so when they are considerably older: on average, at a verbal mental age of 9 (Happé, 1995).

A prime example of this dissociation between social and nonsocial reasoning skills involves the Zaitchik "false-photograph" task, which is modeled on the standard false-belief task except insofar as it tests children's ability to reason about physical (photographic) instead of mental representation (Zaitchik, 1990). Using a simplified version of this task, Leekam and Perner (1991) tested a group of high-functioning autistic teenagers in two conditions, one testing false-belief understanding and the other, photographic "misrepresentation." In both conditions, participants were shown a doll (Judy) wearing a red dress. In the false-belief condition, a second doll (Susan) sees Judy in the red dress and then leaves the room. Judy's dress is changed from red to green, and subjects are asked: "What color does Susan think that Judy is?" In the false photograph condition, a Polaroid photo is taken of Judy in the red dress. While the photo is developing, her dress is again changed from red to green, and subjects are asked: "In the picture, what color is Judy?" Only 25% of autistic participants were correct on the false-belief question, but almost all those tested passed the false-

photograph question (Leekam & Perner, 1992). This pattern suggests that passing false-photograph tasks does not rely on understanding the theory of the nature and the operations of the ToM module, an "innate, isolable construct" (Baron-Cohen, 1995; Leslie & Thaiss, 1992).

Results like these strongly suggest that autistic individuals are especially impaired in their reasoning abilities, and the idea that autism results from an "innate, isolable construct" has profound implications for research on clinical profiles of autistic children. For example, the notion that autistic individuals have primitive, or "atavistic," forms of social interaction and interaction (Baron-Cohen, 1995) is not supported by current research. In contrast, the evidence suggests that autistic individuals have a specific pattern of impairment in the cognitive processes that underlie social interaction and social understanding.

Consider, for instance, the characteristic social behavior of autistic children with autism. These might easily be confused with each other, especially if there are other children who show no interest in the environment and even in other people. They show no tendency to engage in social interaction or to direct another's attention toward an object or person, or even to gather information about it. They may not even recognize the actions of others or how others' actions affect them or how their actions affect others or how others' actions affect them. They may be confused by what other people do, by intentionally misleading behavior, or whether or not the behavior is experienced by other people's physical "antisocial" actions. They might be confused by other people's physical "antisocial" actions, for example, or even by other people's physical "antisocial" actions.

Communicative abnormalities may be more pronounced. Language skills vary widely across the autistic spectrum, but those who develop good language skills tend to be more fluent and expressive than those who develop poor language skills. Communicative abnormalities may be more pronounced. Language skills vary widely across the autistic spectrum, but those who develop good language skills tend to be more fluent and expressive than those who develop poor language skills. Communicative abnormalities may be more pronounced. Language skills vary widely across the autistic spectrum, but those who develop good language skills tend to be more fluent and expressive than those who develop poor language skills.
is in the case of PKU or "misconnection" not a stable fallback position. But significant symptoms of these disorders - and autism's main proponents. We conclude that it is far too strong to remain and sometimes better than average, on understanding of the mental age of agents. *High-functioning* autism, when compared with normal children (who are matched with them by mental age, Pfenninger, 1995). For instance, on first-prime tests, these "high-functioning" autistic children show no tendency to judge behavior, children will normally pass a false belief test by false belief (Happ, 1983). Autistic individuals, if considerably older: on average, at a photographic question (Zettl, Perner, 1991). Similar results were obtained by Leslie and Thais (1992). This pattern of dramatically failing false-belief while passing false-photograph tasks does not occur in normal 4-year-olds.

Results like these strongly suggest that autistic individuals are not generally impaired in their reasoning abilities, but rather have a specific inability to reason about, and perhaps even conceptualize, mental states and processes—hence the idea that autism results from the developmental failure of a so-called ToM module, an "innate, isolable component of the mind which embodies the theory of the nature and the operations of mind" (Baron-Cohen, 1996, p. 258; see also Baron-Cohen, 1985). Moreover, as an important extension of their theory, ToMM theorists argue that a dysfunctional ToM module can account not just for the deficits identified experimentally, but for the clinical profile collected under the headings of social, communicative, and imaginative abnormalities.

Consider, for instance, the characteristic social abnormalities associated with autism. These might easily be connected with an inability to attribute mental states to others, especially if these abnormalities reflect an apparent indifference or insensitivity to what others are thinking and feeling. Thus, autistic children show no interest in, and even a positive aversion to, meeting another's eyes. They show no tendency to engage in social referencing behaviors, that is, directing another's attention toward an object in order to share their interest in it or gather information about it. They show little understanding of how their actions affect others or how others' actions are meant to affect them. They may often be confused by what other people do, but show little capacity to be hurt by intentionally malicious behavior, or touched by intentionally kind behavior whether or not the behavior is experienced as beneficial. They may be amused by other people's physical "antics," even when those antics betray extreme distress or pain. They understand sabotage, but are blind to deceit and other forms of slyness.

Communicative abnormalities may also be rooted in this mentalizing deficit. Language skills vary widely across the autistic population. But even among those who develop fair linguistic capacity, typical problems remain. These are connected in particular with communicative and pragmatic aspects of language use that depend on the speaker's awareness of the conversational situation, including especially the listener's point of view: abnormal prosody (rhythm, stress, tone), abnormal shifts in topic, inability to give and receive conversational cues, abnormal accompanying gestures and facial expressions, pronoun reversals ("I" for "you"), idiosyncratic use of words, abrupt interruptions and terminations of conversation, insensitivity to taboos on personal topics, and so forth. Autistic individuals also tend toward extreme literal-mindedness—showing an insensitivity to metaphor, irony, sarcasm, even idioms as idioms: To autistic individuals, "he went the whole nine yards" means, literally, "he went nine whole yards." There is little or no understanding that others may intend to convey by their words something more or other than just what their words mean.
Some of these communicative abnormalities are closely related to the final element in this triad of deficits: autistic lack of imagination. From early childhood, autistic individuals show a notable absence of spontaneous pretend play, as if it never occurs to them to think about things (represent them) other than as they are. Instead, they will engage in repetitive, stereotyped activities such as sorting objects or lining them up in rows. They also tend to show limited or absent interest in the larger meaning of things (function, associations, symbolic properties) but focus instead on superficial details, with obsessive interests that are circumscribed accordingly. It may be memorizing bus routes, timetables, birth dates, or even door colors. Many autistic individuals are notable for their rote memory skills, even though they show little concern with focusing on what's worth remembering for other cognitive purposes. Perhaps this is because they have a limited capacity for imagining what those purposes might be, hence a limited capacity for opportunistic planning (for a discussion of planning deficits as connected with theory-of-mind capacities, see Currie, 1996; for an alternative perspective, see Russell, 1997).

So prima facie ToMM is a very attractive unifying hypothesis, but we should note that there are other autistic abnormalities that seem to have little to do with "theory-of-mind" capacities. These include sensory-motor problems: for example, extreme and unusual physical sensitivities and insensitivities; slowed orienting of attention; oddities of posture and gait; ties, twitches, and unusual mannerisms; and stereotypies such as rocking, hand-flapping, spinning, thumb-twiddling, and echolalia. They also include abnormalities in perceptual processing, leading to a characteristic autistic profile of assets and deficits on various perceptual tasks: for example, insusceptibility to certain perceptual illusions, superior performance on finding embedded figures within a larger design, superior visual memory and capacity for rendering scenes in precise detail, perfect pitch, difficulties with “gestalt” perception—seeing whole figures or scenes as opposed to their parts, absence of perceptual “switching” with ambiguous figures such as the duck-rabbit, and so on.

How can ToMM theorists account for these additional symptoms? As we saw in the introduction, a weak ToMM view would involve conceding that these abnormalities are fundamentally unrelated to the core ToMM deficit, so that autism is, in effect, an association of relatively independent disorders resulting from multiple failures across a variety of distinct neurological systems. However, if these neurological systems are genuinely independent, then either we should expect some evidence of dissociable breakdowns or, in lieu of that, some explanation for why breakdowns in multiple systems co-occur. The weak ToMM theorist encounters trouble either way: The first possibility seems empirically unindicated and the second undercuts the explanatory power of the weak ToMM hypothesis.

Consider the first possibility: If ToM deficits are essentially unrelated to other characteristic symptoms in autism, then we should expect to see a relatively pure ToM-impaired autistic type. Such autistics would most nearly resemble individuals who have developed an unusual cognitive brain and body systems, evident because of acquired neurological or psychological processes. Yet there seems to be no evidence of this in the autistic population. Whatever range of profile is quite unlike that of brain-damaged sensory-motor and perceptual difficulties. For us, these differences in profile should not be neurodevelopmental disorder.

Interestingly, neurodevelopmental disorder may be more diffuse cortical damage, which suggests widespread, affecting complex neural areas. Furthermore, across a range of animal species, not only are cortical systems atypical subcortical structures are involved, both the cerebellum and limbic system. Some findings suggest deviations in brain development and cannot be easily classified as lesions. Our theories of structural brain abnormalities and disorders will have to incorporate these facts rather than relying on more established facts.

More likely, then, is the second possibility: These symptoms across various distinct neurological systems are functionally unrelated; their failure is a mechanical one. As with PKU and its effect on haphazard interference with the development of cognitive functions, including the hypothesis that ToMM is a particular deficit in the productive work in autism, whatever it may do to social cognition. (This is not to say we entirely dismiss the role in the explanation of autism: social cognition is normally accomplished by a network of systems and individuals, not just a single module.)

Faced with these difficulties, the ToMM theorist makes this stronger claim: Many if not all “peripheries” in terms of a malfunctioning ToM module are general to defects in social cognition at all, some more metarepresent mental states. Uta Frith and others made this suggestion, building on an idea of Gopnik: that a ToM capacity implies no assumptions of mental states. Thus, the mind would be just as impossible to understand as the mind of a child is to a child.
Theories of mind in autism and schizophrenia

Normalities are closely related to the final lack of imagination. From early childhood absence of spontaneous pretend play, about things (represent them) other than repetitive, stereotyped activities such as rows. They also tend to show limited or of things (function, associations, symbolic details, with obsessive interests that be memorizing bus routes, timetables, autistic individuals are notable for their low little concern with focusing on what's the purposes. Perhaps this is because they what those purposes might be, hence a thing (for a discussion of planning deficits facilities, see Currie, 1996; for an alternative unifying hypothesis, but we should normalsities that seem to have little to Doyle include sensory-motor problems: for actual sensitivities and insensitivities: slowedature and gait; ticss, twitches, and unusual locking, hand-flapping, spinning, thumbs include abnormalities in perceptual profile profile of assets and deficits on various possibility to certain perceptual illusions, situated figures within a larger design, surrendering scenes in precise detail, percepception—seeing whole figures or scenes perceptual "switching" with ambiguous an.

What for these additional symptoms? As we view would involve conceding that these related to the core ToMM deficit, so that relatively independent disorders resulting of distinct neurological systems. How-genuinely independent, then either we available breakdown or, in lieu of that, some triple systems co-occur. The weak ToMM The first possibility seems empirically the explanatory power of the weak ToMM ToMM deficits are essentially unrelated to them, then we should expect to see a relation. Such autistics would most nearly re-ssemble individuals who have developed normally but show impairments in social cognition because of acquired neurological damage in their frontal or temporal lobes. Yet there seems to be no evidence of such specific abnormalities in the autistic population. Whatever range of symptoms autistics manifest, their profile is quite unlike that of brain-damaged adults, especially with regard to sensory-motor and perceptual difficulties. However, as Tager-Flusberg reminded us, these differences in profile should not be surprising given that autism is a neurodevelopmental disorder:

Interestingly, neurodevelopmental disorders are more often associated with diffuse cortical damage, which suggests the impact of such disorders is more widespread, affecting complex neural systems rather than simple localized areas. Furthermore, across a range of developmental syndromes, we find that not only are particular cortical systems affected but often associated atypical subcortical structures are involved as well. For example, in autism both the cerebellum and limbic system show significant abnormalities. These findings suggest deviations in brain development that begin early in embryology and cannot be easily classified and interpreted as later acquired focal lesions. Our theories of structural brain abnormalities in neurodevelopmental disorders will have to incorporate these kinds of developmental complexities rather than relying on more established studies from work with adults. (Tager-Flusberg, 1989, p. 3)

More likely, then, is the second possibility: Autism involves multiple failures across various distinct neurological systems that co-occur for a reason. If these systems are functionally unrelated, as the weak ToMM theorist avers, then it seems the only kind of account that would make sense of this multiple failure is a mechanical one. As with PKU, autistic behavior might be the result of haphazard interference with the development of a number of arbitrarily involved cognitive functions, including those supporting social cognition. If so, hypothesizing a particular deficit in the putative ToM module does no explanatory work in autism, whatever it may do for explaining acquired disorders in social cognition. (This is not to say we endorse such a hypothesis, only that it plays no role in the explanation of autism even under the assumption that social cognition is normally accomplished by a modularized neural system.)

Faced with these difficulties, the ToMM theorist might well consider the stronger claim: Many if not all "peripheral" autistic symptoms can be explained in terms of a malfunctioning ToM module, in which case they are not peripheral to deficits in social cognition at all, since they too are produced by failure to metarepresent mental states. Uta Frith and Francesca Happé (1999) recently made this suggestion, building on an idea repeatedly emphasized by Alison Gopnik: that a ToM capacity implies no asymmetry between first- and third-person ascriptions of mental states. Thus, the autistic subject's understanding of her own mind would be just as impoverished as her understanding of other minds.
Frith and Happé speculated that because autistic subjects cannot metarepresent their sensory and perceptual processes as states of their own mind, they are at the mercy of them in a way that could produce the characteristic profile of autistic abnormalities. For instance, on motor tests involving monitoring and correction of action, autistics perform poorly when compared with normal controls (Russell & Jarrold, 1998, 1999). Frith and Happé offered the following ToM explanation: “without self-awareness, an individual might not know how she is going to act until she acted, nor why she acted as she did. . . . A person who lacks self-consciousness may be unable to distinguish between her own willed and involuntary actions” (Frith & Happé, 1999, p. 8). Alternatively, there might be improved performance on tasks where action without in-depth conscious reflection is superior to consciously performed action. This might explain autistic individuals’ relatively good capacity to perform routinized action coupled with a poor capacity to act flexibly and imaginatively (ibid., p. 10). It might also account for autistic insusceptibility to certain visual illusions, such as the Titchener circles (or Ehbbinghaus illusion), where subjects’ “superior performance in verbal response is not contaminated by conscious reflection” (ibid., p. 10). [In support of this contention, recent evidence shows that when even normal subjects respond motorically (and apparently unconsciously) in their reaching behavior to the correct size of the circles, nevertheless they continue to (consciously) judge the size of the circles incorrectly (Aglioni, DeSouza, & Goodale, 1995)].

One notable and particularly salient characteristic of autism is its extreme and unusual sensory experience. Frith and Happé relate this to a dysfunctional ToM as follows:

If low-functioning autistics are unable to reflect on their inner experiences, then they would be unable to develop over time the richly connected semantic and experiential associations which normally pervade our reflective consciousness. Observation by parents suggests that the awareness of sensations and experiences may be peculiar in children with autism. Anecdotal reports of abnormal sensory and pain experiences are on occasion quite extreme. . . . One anecdotal example is the case of a young girl with autism who was found to have suffered acute appendicitis, but had not complained of pain and, when asked how she felt, did not report anything wrong. Abnormal response to heat and cold, as well as hypo- and hyper-sensitivity to sound, light or touch are frequently reported. . . . Such responses might be expected if there was an inability to reflect on inner experiential states. Of course, normal pain perception is greatly affected by attribution and expectation. These individuals might feel immediate pain in the same way as everyone else, but would not be able to attribute to themselves the emotional significance that normally accompanies pain. This might explain why they do not complain about it. We may speculate that the self-conscious person reflects not only on the pain but also on the experience of pain. This person is feeling “misery” in addition to feeling pain. (Frith & Happé, 1999, pp. 10–11)

Frith and Happé’s view implies that if autistic subjects are hypersensitive, then they may be at the mercy of their own sensory experiences, because their sensory experiences are extremely vivid. Autistics also appear to be abnormally aware of their sensory experiences. In the extreme, they are aware of them as mental events (Frith & Happé, 1999, p. 8). For instance, they are able to account for it in terms of a general deficit of self-consciousness; autistic abnormalities in this terms become understandable for them in terms of local motor and sensory systems.

Our conclusion is that Frith and Happé’s view of autism reflects these difficulties because it requires a wrong direction of explanation between capacities. Reverting to the SLJ case, it is not the case that autistic cases are explanations of autistic experiences. Rather, autistic cases are explanations of autistic experiences. The higher order abnormalities of other developmental disorders are far more complex, and are not as well understood. It is not clear whether autistic individuals are the result of abnormalities in the SLJ. (Smith, 1998, cf. Tager-Flusberg, 1999).}

To make theoretical sense of this autistic sensory and perceptual disturbances, it is necessary to develop a capacity for understanding others as people. The first step involves restoring reverse engineering reasoning, of the role of developing a child's capacity for normal and interactions with others. The development of a child's capacity for understanding others is a necessary condition for the development of ToM. Additionally, those relating to mind reading, a capacity for understanding others as people, developing a capacity for mind reading, the capacity to be minded like others, and the capacity to read others' psychological self-regulation and behavior towards them (predictable, rational) agency. If the information is abnormal, his or her motivations and the types of interaction that lead to development are disrupted from the very beginning. Hence, autistic individuals may not be so much abnormal in capacity as the failure to develop their capacity.
that because autistic subjects cannot conceptual processes as states of their own in a way that could produce the character. For instance, on motor tests involving autism, Frith and Happé offered that self-awareness, an individual might be she acted, nor why she acted as she did. Goodness may be unable to distinguish by actions” (Frith & Happé, 1999, p. 8). Performance on tasks where action with other to consciously performed action. A relatively good capacity to perform rou- sibility to act flexibly and imaginatively as autistic insusceptibility to certain visual (or Ebbinghaus illusion) where subjects’ size is not contaminated by conscious re- this contention, recent evidence shows and motorically (and apparently uncon- the correct size of the circles, nevertheless, the size of the circles incorrectly (Aglioti, 1999).

A new characteristic of autism is extreme Frith and Happé relate this to a dysfunctional to reflect on their inner experiences, inner conceptual frameworks that are disconnected social deficits. The first step involves restoring the emphasis, sometimes lost in reverse engineering reasoning, of the role intersubjective encounters play in developing a child’s capacity for normal agency. Reciprocal, affectively patterned interactions with others are important determinants of infant experience and a necessary condition for the development of higher order cognitive capacities, especially those relating to “mind reading” (here, minimalistically conceived as a capacity for understanding others as psychological agents like oneself). In our view, developing a capacity for “mind reading” goes hand in hand with developing the capacity to be minded like others—that is, with acquiring habits of psychological self-regulation and behavior that conform to shared norms of sensible (predictable, rational) agency. If the infant’s sensory-motor and perceptual systems are abnormal, his or her motivation and ability to engage with others in the types of interaction that lead to developing such capacities of agency will be disrupted from the very beginning. Hence, the difference between autistic and normal individuals may not be so much the lack of some specifically social capacity as the failure to develop their capacities within an intensely social context.

Frith and Happé’s view implies that autistic subjects would be rather unreflective about their sensory experiences. But, as they themselves noted, autistic hypersensitivity is as dramatic as autistic hyposensitivity. In fact, both third- person observation and first-person report indicate that autistic subjects are abnormally aware of their sensory experiences, and aware of them mainly because their sensory experiences are extreme and persistently captivating. Furthermore, they are aware of them as mental experiences. It’s hard to see how this can be accounted for in terms of a general deficit of self-consciousness stemming from a dysfunctional ToM mechanism. And, of course, if there is no general deficit of self-consciousness; accounting for autistic motor and perceptual abnormalities in these terms becomes equally suspect (as opposed to accounting for them in terms of local motor and perceptual system dysfunctions).

Our conclusion is that Frith and Happé’s strong ToMM hypothesis encounters these difficulties because it reverses, temporally and conceptually, the correct direction of explanation between lower and higher order cognitive capacities. Reverting to the SLI case, it is as if someone tried to explain auditory failures and related behavior as the consequence of defective grammatical processing rather than defective grammatical processing as the consequence of early auditory difficulties. The higher order cognitive difficulties of autism and other developmental disorders are far more likely to be the cascading effects of relatively early, lower order sensory and perceptual abnormalities than lower order abnormalities are to be the result of a failure in higher cognition (Karmiloff-Smith, 1998; cf. Tager-Flusberg, 1999). If this developmental connection makes theoretical (and empirical) sense, then the explanatory difficulty of integrating the various symptoms of the disorder disappears.
specialized machinery for attending to or using such machinery, with the consequence of creating an integratively powerful source of stimuli that quickly becomes overstimulating. In any case, the devastating effects of the sensory dysregulation rather than a normal one are as follows: (1) Autistic individuals would be engaging their sensory experiences perhaps without incoming sensory stimuli in ways that lead to a number of characteristic autistic behavioral patterns, including self-stimulatory behaviors such as rocking, echolalia, head-banging, and biting. Being excluded from the regulative influence of others, they develop habits of agency that conform to a primitive sense of the world, think, and act in recognizable normative ways. For example, the mutual giving, receiving, and sharing of the very kinds of interactions that give rise to one's own disability reflected in the perplexing nature of their own perplexity at the behavior of others.

In sum, we think that strong ToMM connection between a capacity for self-awareness and self-governance and weak ToMM theorists, we also think stronger ToMM theorists may have difficulty understanding the foundational role of the initial problems, neglect to consider the environmental supports he or she may require, and may produce capacities that are sufficient to task. Such capacities have been found for some children and adults with autism, though they vary widely and do not lead to genuine mental states in the usual sense. However, in our view, this is not evidence that such individuals then develop such capacities in the usual sense.

There are a number of advantages to this view.
Elementary developments of this view, precisely repeat, a theme emphasized by parents to a child's normal social cognition of his or her intersubjective experience (1979; Trevarthen & Hubley, 1978). Between self and other is sustained by expressed feelings of others: smile for a fearful look. This makes others poten- tially at once: Not only do they provide experience; they also serve as a critical resource, for instance, a mother may com- pliment one in face and voice expressions that are her own distress, then modulating these in group. The child, carried along by its innate direction of the mother's expression of its own distress in consequence, may be reflected in them, rather than any direct epistemic as with responsive others.

The infant is learning how to be itself in this way; these structured interactions, first mediate via the mediation of others, systemic front as well, for they allow the use in ways that are conducive to develop- predictive place. The normal child of other people thus derives a double self-world, including the progressively more her people, is made open to manageable people become known to the child his or her "mind-reading" capacities.

What happens in normal development, it is to sensory disturbances and their fail- to developing normal capacities of a, basic social capacity, or even drive to (Meltzoff & Gopnik, 1993). Indeed, many selective parroting tendencies. Nor selective disorder, as Hobson and others engage in intersubjective encounters may if autistics find their contact with be tolerated. Indeed, in an effort to individuals might need to shut other fear from indicating that they lack any specialized machinery for attending to others, this may well show that they have such machinery, with the consequence that others constitute a disproporti- onately powerful source of stimuli that quickly become overwhelming for them. In any case, the devastating effects of finding in others an abnormal source of sensory dysregulation rather than a normal source of helpful regulation are two- (1) Autistic individuals would be cast back on their own resources for managing their sensory experiences perhaps by reducing, repeating or drowning out incoming sensory stimuli in ways they can control. This could explain a number of characteristic autistic behaviors that range from being seemingly dull and repetitive to bizarrely self-stimulatory and even self-abusive: lining up blocks, counting and calculating, repetitively flushing toilets, examining grains of sand, chewing things regardless of taste or danger, spinning, hand-flapping, rocking, echolalia, head-banging, biting and slapping oneself, and so forth. (2) Being excluded from the regulative influences of other people, autistics will not develop habits of agency that conform to shared norms of what it is to experience, think, and act in recognizably normal ways. Hence, they will be deprived of the very kinds of interactions that give rise to ordinary capacities of agency, a disability reflected in the perplexing nature of their own behavior as well as in their own perplexity at the behavior of others.

In sum, we think that strong ToMM theorists are right to emphasize the connection between a capacity to kno other minds and the sorts of capacities for self-awareness and self-governance that make for normal agency. As against weak ToMM theorists, we also think strong ToMM theorists are right to empha size the connection between higher order cognitive abnormalities in autism and so-called peripheral symptoms. However, in postulating a single higher order cognitive deficit, they neglect to consider how basic sensory and perceptual problems can have cascading developmental effects, particularly with respect to sociocognitive development, by disrupting an infant's normal environment of regulative interactions with others. From this perspective, autistic mentalizing deficits are part of a more general pattern of deficits that constitute from birth to maturity an abnormal developmental trajectory that leads to a distinctive cognitive style or set of cognitive styles. In some cases, depending on the severity of the initial problems, the child's compensatory abilities, and the kind of environmental supports he or she may find, this developmental trajectory may even produce capacities that are sufficient to pass some, if not all, "theory-of-mind" tasks. Such capacities have been found amongst high-functioning autistics, although they vary widely and do not lead to normal social behavior. Are these autistic individuals then not genuine mentalizers? Certainly they are not mentalizers in the usual sense. However, in our view, this is not because their ToM module is inoperative or only partially operative; rather, it is because the capacities they develop bear only, and to varying degrees, a family resemblance to capacities that characterize the cognitive styles of normally developing individuals.

There are a number of advantages to the bottom-up developmental ap-
proach we suggest. We have discussed two of them and alluded to a third. These include: (1) respecting the distinction between “peripheral” (or lower order) and “core” (higher order cognitive) abnormalities, and yet accounting in an integrated way for their comorbidity; (2) making sense of the neurofunctional differences between a disorder like autism and acquired pathologies consequent upon localized brain damage; and (3) providing a natural way to account for the range of mentalizing capacities found amongst high-functioning autistics without having to posit a late-developing ToM mechanism that is only partially operative.⁸

A fourth and final advantage of our approach bears special mention because it makes sense of an otherwise puzzling phenomenon. If autistic sensory disturbances do indeed lie at the developmental core of later “mindblindness” and other higher order cognitive abnormalities, then other clinical populations with early sensory problems ought to show similar kinds of deficits. And indeed this is the case. Deaf children of hearing parents as well as congenitally blind children show autistic-like abnormalities in social, communicative, and imaginative skills, as well as selective incapacity to pass reasoning tasks with a metalinguistic component (Brown et al., 1997; Hobson, 1993; Peterson & Siegal, 1998, 1999; Peterson, Peterson, & Webb, 2000).⁹ We think the parallels among these populations are so stunning as to call for a unifying explanation. It follows from the account developed here that any child will be unable to develop mind reading, as long as it is impossible for him or her to make good regulative use of other people. This may stem from having a missing sensory avenue to others, as much as it may stem from having one’s sensory avenues to others overwhelmed by the overstimulation involved in sustained exposure to them.

In terms of the taxonomy we laid out in our introductory remarks, we have argued that the mentalizing deficits of autism constitute at the very least a developmentally constructed cognitive domain. Whatever neural specialization for social reasoning occurs in normally developing children, we believe this is the outcome of an ongoing process of functional development in which these higher order cognitive capacities depend on a child’s normal engagement with others in a structured social environment. Such normal engagement is naturally mediated by a child’s sensory-motor and perceptual systems, among other things, so if profound disruptions occur in these systems, it is not surprising that subsequent neurofunctional development will be dramatically affected. Do we therefore think that normal development leads to the kind of modularization of “theory-of-mind” capacities that nonnativist ToMM theorists suggest? Do we think, in other words, that normal ToM capacities constitute an actual cognitive domain, subserved by an architecturally real neural system? Nothing we have said so far argues strongly against this possibility. However, do we think that the ToMM account of an adult-onset disorder like schizophrenia, with symptoms that are interestingly similar to and interestingly distinct from autism, is sufficiently problematic that it points to the general conclusion that social reasoning constitutes an architecturally virtual as well as a developmentally constructed domain.

In its strong form, the ToMM theory of symptoms of the disorder derive from an accompanying theory that schizophrenia performs (Langdon & Coltheart, 1999) and that in cases of severe negative symptoms, an impetus (and challenge) for the ToMM account of these ToM deficits that the range of disparate symptoms not observed in cognition. For example, it is not clear how thought disorder, and inappropriate affect (avolition, apathy, flat affect) could all be the problem.

The first stage in achieving horizontal arguments is to focus on the nature of the deficit in the disorder. Typically these concern the delusions that the subject is the author of some confabulated narratives. In some cases the schizophrenic feels as if these are, so to speak, “in her mind” are the owner of these experiences but not (2000b; Gerrans, 2001). Similarly, schizophrenic thoughts are being influenced b others or in some cases of objects in the schizophrenic herself. Bovet and Parnas thought that his urinations launched beliefs the universe is expanding or contracting peristalsis is well known to clinicians.

These phenomena are all “first rank” symptoms described by Carl Schneider as a “loss of reality” symptoms concern the subject’s awareness.
two of them and alluded to a third. These between “peripheral” (or lower order) normalities, and yet accounting in an inte-

<table>
<thead>
<tr>
<th>MIND</th>
</tr>
</thead>
<tbody>
<tr>
<td>THEORY OF MIND IN AUTISM AND SCHIZOPHRENIA 289</td>
</tr>
</tbody>
</table>

THE CASE OF SCHIZOPHRENIA

In its strong form, the ToMM theory of schizophrenia hypothesizes that all symptoms of the disorder derive from an acquired deficit in ToM function. It is certain that schizophrenics perform very poorly on standard ToM tests (Langdon & Coltheart, 1999) and that their social interactions, especially in cases of severe negative symptoms, are severely impaired. However, the real impetus (and challenge) for the ToMM theory derives less from giving an adequate account of these ToM deficits than to providing an integrated account of the range of disparate symptoms not obviously connected with problems in social cognition. For example, it is not clear how hallucinations and delusions, formal thought disorder, and inappropriate affect and the psychomotor poverty (alexia, avolition, apathy, flat affect) could all be attributable to one underlying cognitive problem.

The first stage in achieving horizontal integration of schizophrenia symptoms is to focus on the nature of the delusions and hallucinations characterizing the disorder. Typically these concern the subject’s sense of agency—that is, the feeling that the subject is the author of his or her own thoughts or actions. In some cases the schizophrenic feels as if auditory experiences (“hearing voices”) that are, so to speak, “in her mind” are nevertheless not produced by her. She is the owner of these experiences but not the author of them (Gallagher, 2000a, 2000b; Gerrans, 2001). Similarly, schizophrenics may feel as if their own thoughts are being influenced by someone else (someone else is making them think certain thoughts) or actually being inserted into their minds by a kind of psychokinesis. In the reverse case, schizophrenics may feel as if they are inserting their thoughts into the minds of others (thought broadcast).

A similar dissociation between ownership and authorship is characteristic of the schizophrenic experience of action. Schizophrenics may feel as if their body is not under their own volitional control, as if they were a kind of inert marionette moving at the will of someone else. Once again, they own these actions in the sense of knowing that the movements are movements of their own bodies. But because they are not aware of intending to perform the movements or of their voluntary control, they do not feel as if they author them. Voluntary control of action can also fail in a way that is analogous to thought broadcast. In these cases the schizophrenic may feel as if the movements of others, or in some cases of objects in the external world, are caused by the schizophrenic herself. Bovet and Parnas (1993) reported a case of a subject who thought that his urinations launched bombing raids, and the delusion that the universe is expanding or contracting according to the subject’s own bodily peristalsis is well known to clinicians.

These phenomena are all “first rank” (core) symptoms of schizophrenia described by Carl Schneider as a “loss of ego boundaries.” These Schneiderian symptoms concern the subject’s awareness of her own cognitive or practical
agency, which seems to affect her ability to correctly discern who is the author of particular experiences/thoughts or actions.

Once the cases are described in this way, ToMM theorists claim that their hypothesis gains plausibility due to the conceptual structure of folk psychological ascriptions of agency (Frith, 1992; Leslie, 1994). According to a philosophical analysis of this structure, agency is the intentional control of bodily movement: What distinguishes an action from mere bodily movement is that the agent's intentions play a causal role in producing the action. So attributing agency involves a cognitive act of attributing intentions to the author of the relevant bodily movement, whether the author is oneself or someone else. Further, because intentions decompose into constituent beliefs and desires, attributing intentions requires a cognitive capacity to attribute beliefs about the world in which an agent's desires are to be realized, and desires to change the state of the world. The desires give the ends, and the beliefs give the means.

Support for this view comes from picture sequencing ToM tests. Schizophrenics, like autistic subjects, are good at completing sequences where the relevant movements can be captured in nonintentional terms (as in someone's being pushed off a seat and starting to cry, a strictly causal chain of events); but they are poor at completing sequences in which the story requires interpreting the characters' intentions, as in stories of deception. Further clinical and experimental data show that the schizophrenic ability to attribute agency, internal (self) or external (other), is very fragile. Perhaps then the schizophrenic makes an observation or has an experience that stands in need of belief/desire interpretation in order to be correctly attributed to an agent but lacks the cognitive capacity to do so. Hence the schizophrenic will be very poor at linking action to its governing intention in both internal and external cases, because the schizophrenic has an acquired deficit in the module that represents the constituents (beliefs and desires) of an intention.

Such a deficit might also explain so-called negative symptoms of the disorder in terms of an inability to generate the requisite intentions or to make use of them to terminate or adjust an action (failures of willed action, perseveration, Parkinsonianism). Finally, and most speculatively, if we conceive of thought itself as intentionally guided, a deficit in the device that monitors intentions might explain formal thought disorder and inappropriate affect. Perhaps the schizophrenic who initiates a train of thought "loses" access to its guiding intention, with the consequence that that sequence of thoughts becomes rambling, incoherent, or fixated and disconnected from the appropriate emotions (Campbell, 1999; Frith, 1992; Gallagher, 2000a).

The foundational idea behind all these speculations is that an agent who cannot (meta)represent beliefs and desires, the task supposedly performed by the ToM module, will be unable either to determine agency in conditions where attribution is an issue or to exert appropriate agential control in cases where she needs access to her original governing intentions. As Christopher Frith put it in 1992, "resolution may be achieved at the price of defects of will and defects in inferring a cognitive deficit" (Frith, 1992, p. 122). Frith's notion is work as "doubt overinclusive." But have the cognitive domain subserved by the ToM really provided an explanation of a pathological behavioral pattern?

One prominent refinement of this version, to Frith himself, and from our point of view, was the weakening of the ToMM hypothesis (by weakening the cognitive deficit) for all schizophrenic symptoms and a shift to lower order cognitive systems. The set of Schneiderian symptoms is likely predating that initially postulated by ToMM system. Various "misconnection" or "disconnection" or "dysconnection" further: They explain the variety of symptoms and the need for cognitive explanations at a lower order.

It is worth pointing out that, depending on the theory, these mechanical and cognitive accounts of the symptom, such as the mechanism of misconnection or disconnection of cortical areas, then one consequence might be that the premotor cortex, an architecturally, is not involved. In our introduction, we would still consider less the possibility that the negative symptoms to be a basic noncognitive deficit of Parkinson's disease, which, en route, produces the unusual grammatical disorder (Mann, 2000).

We now describe the motor components in slightly more detail, showing how they underlie these symptoms without implicating a dysfunction.

Forward Models and the Sense of Agency

Frith originally argued that delusions of control would be essentially failures to correctly attribute agency, failures to correctly cognize intentions.

However, a more economical explanation of the sense of control implicates a more basic control of action. That subsystem, neurally, is a forward model for motor control.

The idea of a forward model goes back to the 1970s and is captured in Figure 12.4. This model builds...
MIND

ToM theorists claim that their conceptual structure of folk psychology (Leslie, 1994). According to a philosophically intentional control of bodily movement: the bodily movement is that the agent's doing the action. So attributing agency intentions to the author of the relevant bodily self or someone else. Further, because beliefs and desires, attributing intentional beliefs about the world in which desires to change the state of the world is give the means.

One prominent refinement of this "overinclusive" formulation owes much to Frith himself, and from our point of view, it is interesting that it involves both a weakening of the ToM hypothesis (a dysfunctional ToM does not account for all schizophrenic symptoms) and a shift of explanatory emphasis from higher order to lower order cognitive systems. Frith now claims that an important subset of Schneiderian symptoms is likely produced at a much lower cognitive level than that initially postulated by ToM theory: namely, the motor control system. Various "misconnection" or "dysmetric" theories of schizophrenia go even further: They explain the variety of symptoms at a mechanical level, bypassing the need for cognitive explanations at all.

It is worth pointing out that, depending on the actual mechanisms involved, these mechanical and cognitive accounts could be made to fit together. If, for example, the mechanism of misconnection is neurochemical transport between cortical areas, then one consequence might be local over- or underactivation of the premotor cortex, an architecturally discrete subsystem. As we explained in our introduction, we would still consider the ultimate cause of schizophrenic positive symptoms to be a basic noncognitive mechanism that happens to selectively affect one subsystem first en route to its more global disabling effect. Another example might be a dopamine failure in the basal ganglia, characteristic of Parkinson's disease, which, en route to its more global manifestation, produces the unusual grammatical disorders characteristic of Parkinsonians (Pinker, 2000).

We now describe the motor control and misconnection approaches in slightly more detail, showing how they explain important schizophrenic symptoms without implicating a dysfunctional ToM module.

Forward Models and the Sense of Agency

Frith originally argued that delusions of control or influence in schizophrenia are essentially failures to correctly attribute authorship of actions resulting from failures to correctly cognize intentions.

However, a more economical explanation of problems with the attribution of authorship implicates a more basic cognitive subsystem dedicated to the control of action. That subsystem, neurally realized in the premotor cortex, is the forward model for motor control.

The idea of a forward model goes back at least to the 19th century and it is captured in Figure 12.4. This model builds on the idea that any bodily move-
The forward model is a computational control. Although not all of its elements are substrates, there is a large body of cor-

idea captures the essential cognitive elements (Blakemore, Goodbody, & Wolpert, 1998; Jahanshahi & Frith, 1998; Decety et al.) evidence to suggest that current and predictive parietal regions, and that actions are indexed by activity in fronto-parietal areas by activity in frontal areas (Jeannerod, 1998).

Apart from its empirical credentials, the forward model is well placed to explain the way in which our actions are perceived and attributed to someone else.

Option 1: The schizophrenic might believe that the act of another. Note that in macaque monkeys as part of the normal control system that signals that a movement is self-initiated are unavailable to the system. The observed action to herself (Jeannerod, 1998).

These hypotheses are supported by delusions of influence. They experience connectivity between the motor system and the parietal motor control task. For example, schizophrenia control when asked to move a joystick.
pated (bodily) consequences of the movement (for instance, that a movement of the hand 15 degrees to the left will produce a corresponding change in visual and proprioceptive information representing its orientation). This is the forward model. Now if the hand moves too far, say 20 degrees to the left, the reafferent information can be compared to the prediction, thereby yielding an error of 5 degrees. The system then generates a new inverse model, motor instruction, and efference copy to initiate and control a corrective movement of 5 degrees to the right.

The forward model is a computational solution to problems of motor control. Although not all of its elements have been decisively mapped to neural substrates, there is a large body of convergent research to show that the basic idea captures the essential cognitive properties of the motor control system (Blakemore, Goodbody, & Wolpert, 1998; Frith, Blakemore, & Wolpert, 2000; Jahanshahi & Frith, 1998; Decety et al., 1994). In particular, there is good evidence to suggest that current and predicted states of limbs are represented in parietal regions, and that actions are initiated in frontal brain regions. The most likely mechanism for the comparator system, then, is the inhibition of parietal areas by activity in frontal areas (Jeannerod, 1994).

Apart from its empirical credentials, the forward model is theoretically well placed to explain the way in which we establish ownership and authorship of actions. For any action, it will be ours if it corresponds to an efference copy that our motor system has generated (Frith, 1987; Frith & Done, 1989). Schizophrenic misattribution of agency might then arise at this level of processing in one of two ways:

Option 1: The schizophrenic performs an action but the forward model system contains no matching prediction—in such cases the schizophrenic attributes the action to someone else.

Option 2: The schizophrenic might activate the prediction system by observing the action of another. Note that this phenomenon has been established in macaque monkeys as part of the normal observation of action. It seems that merely observing an action can provoke neural activity in that part of the monkey’s control system that signals that a movement is an action: namely, the forward model (Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Gallese, Fadiga, Fogassi, Luppino, & Murata, 1997; Rizzolati, Fadiga, Gallese, & Fogassi, 1996). Due to hypoactivity in the inhibitory system, or lack of connectivity between controller and predictors, the normal signals that tell the system that the action is not self-initiated are unavailable to the system. The schizophrenic then attributes the observed action to herself (Figure 12.5) (Dapriati et al., 1997; Georgieff & Jeannerod, 1998).

These hypotheses are supported by experiments with schizophrenics who experience delusions of influence. They show unusual patterns of activity and connectivity between frontal and parietal areas and failures of attribution in motor control tasks. For example, schizophrenics report vivid sensations of alien control when asked to move a joystick. In other cases, when provided with am-
If the positive symptoms of schizophrenia are classified as a disorder of a ToMM hypothesis (explored by Frith), then it seems that a weak ToMM hypothesis (explored by Frith), now interpreted as failures of working memory, as a result of ToM malfunction—for example, and act on intentions. However, as we have seen, there are candidates for a mechanical explanation, and is left without any real work to do.

In any case, to repeat our argument, positive and negative symptoms are not cleanly separable, and negative symptoms are implicated in schizophrenia. But this is because there is only one system involved in the negative symptom hypothesis suggested), or it is because the interaction of the two negative systems is one of mechanism. The obvious hypothesis is one of mechanisms that cause disruption in a number of characteristic array of positive and negative symptoms (see the misconnection theory of Schizophrenia is ultimately a cognitive abnormality. Most importantly, the idea that seemingly discrete domains at the level of the architecture of the system is an Herulean task. In particular, it is not to be a real, but only architecturally virtual concept.

**Misconnection and Dysmetria**

Motor control theories of misattributions hypothesis for schizophrenia by accounting for a subset of symptoms in terms of a maladaptive system, rather than one that is devoted to a second way of opposing the strong ToM hypothesis. In this case, even if the sensorimotor system is specific to the sensorimotor system per
If the positive symptoms of schizophrenia can be explained in this way, then it seems that a weak ToMM hypothesis has not much role to play in accounting for the disorder. After all, Schneiderian symptoms are generally held to constitute the core of the disorder. Of course, another way to defend a weak ToMM hypothesis (explored by Frith, 1992) is to argue that the core of the disorder is actually the negative symptoms (psychomotor poverty, avolition, apathy), now interpreted as failures of willed action, and attempt to explain these as a result of ToM malfunction—for example, as failures to form, metarepresent, and act on intentions. However, as we show, the negative symptoms are prime candidates for a mechanical explanation. If so, then the weak ToMM hypothesis is left without any real work to do.

In any case, to repeat our arguments made in the introduction, if positive and negative symptoms are not cleanly dissociable, this suggests that the cognitive system or systems implicated in schizophrenia are affected together. Either this is because there is only one system involved (as the discredited strong ToMM hypothesis suggested), or it is because more than one system is involved (as the weak ToMM hypothesis suggests). But then the problem becomes one of accounting for the connection between them: Why do they fail together? The most obvious hypothesis is one of mechanical failure: Some neurological process causes disruption in a number of cognitive systems, which produces the characteristic array of positive and negative symptoms. If this is genuinely the case (see the misconception theory described next), then the explanation for schizophrenia is ultimately not cognitive at all, despite the suggestive pattern of cognitive disabilities. Most importantly from our point of view, mechanical failures of this sort dramatically undermine straightforward reverse engineering theorizing. The reason is this: Given the number of systems that are likely to be involved in any such mechanical failure, reasoning backward from a pattern of cognitive and other behavioral disabilities to the functional architecture of the system is an Herculean task. In particular, there is no good reason to assume that seemingly discrete domains at the cognitive level will correspond to actual domains in the architecture of the system. As we put it earlier, these may not be real, but only architecturally virtual cognitive domains.

**Misconnection and Dysmetria**

Motor control theories of misattributed agency undermine the strong ToMM hypothesis for schizophrenia by accounting for a significant—some say “core”—subset of symptoms in terms of a malfunctioning lower level sensorimotor subsystem, rather than one that is devoted to higher level cognitive processes. A second way of opposing the strong ToMM hypothesis would be more radical still, by seeking to explain schizophrenia as an essentially mechanical disorder. In this case, even if the sensorimotor system is affected, it is not a problem specific to the sensorimotor system per se—that is, a cognitive problem relating
to the generation and use of motor predictions. It is, rather, a noncognitive problem that affects, among a variety of functionally unrelated things, the generation and use of motor predictions.

Some recent theories of schizophrenia take this approach. Although differing in many details, both Nancy Andreasen and Peter Liddle share the view that schizophrenia involves disruption to a circuit integrating functionally diverse cognitive processes whose neural substrates are located in different areas of the brain. As Liddle put it, “the evidence indicates that the abnormality characteristic of schizophrenia is impaired coordination of activity at the diverse sites engaged during mental processing rather than a static loss of function at any site” (Liddle, 2001, p. 72).

For Liddle, the relevant circuit involves a series of linked cortico-striato-thalamo-cortical feedback loops whose activation affects the diverse regions implicated by neuroimaging studies of schizophrenia. He suggests that the cause of this misconnection is most likely dopaminergic regulation of this circuit.

For Andreasen, the explanatory strategy is to identify the neural substrates of cognitive malfunction in schizophrenia and investigate possible linking circuitry. The circuit she identifies is a cortico-cerebellar-thalamo-cerebellar-cortical loop. For Andreasen, however, the misconnection is explained as a lack of temporal integration of activation across distributed brain regions. Hence she dubs the misconnection syndrome a case of dysmetria (Andreasen, Paradiso, & O’Leary, 1998; Andreasen et al., 1999).

Patients suffering from schizophrenia have a misconnection syndrome that leads them to make abnormal associations between mental representations; they lack the ability to distinguish between self and nonself, and, due to an incapacity to suppress the multiple stimuli normally bombarding human consciousness, they are further unable to distinguish between the important and the trivial. In light of these difficulties, a number of problems follow: Internal representations may be attributed to the external world, leading to hallucinations. Perceptions or other information will be misconstrued with inappropriate associations leading to delusional misinterpretations. Online monitoring of language or thoughts will be impaired or mistimed, leading to disorganized speech and disorganized thinking. Behavior will not be adequately monitored, leading to social awkwardness, excessive aggressiveness or shyness, or other abnormalities in behavioral activities. The inability to monitor may also lead to the “freezing” characteristic of catatonic immobility or a failure to inhibit expressed as catatonic excitement. Difficulties in inhibiting or prioritizing may also lead to the various negative symptoms such as avolition or anhedonia, much as a computer locks when it cannot match signals sent at an incorrect rate or to an incorrect place (Andreasen et al., 1999, p. 4).

Andreasen calls her theory “cognitive” rather than “mechanical.” But we should note that exactly the same explanation could be offered by someone who focuses on finding the mechanical link between diffusely affected systems. The only thing missing would be the coda of cognitive unification provided by suggesting a functional role for the linking.

Why, then, are Andreasen acco...
The mind is, rather, a noncognitive system. Its functional role is not explained by the interaction of functionally unrelated things, the genetic predispositions. It takes this approach. Although different areas of the brain, they share the view that the circuit integrating functionally diverse substrates are located in different areas and that the abnormality of coordination of activity at the diverse levels rather than a static loss of function at a single level.

Theory of Mind in Autism and Schizophrenia

Theories of the mind are of three types: (1) structural-functional; (2) cognitive; and (3) computational. Structural-functional theories hold that the mind is a collection of specialized processing units, each of which is responsible for a specific function. Cognitive theories hold that the mind is a collection of mental processes, each of which is responsible for a specific function. Computational theories hold that the mind is a collection of algorithms, each of which is responsible for a specific function.

The cognitive approach to the mind is the most recent and the most influential. It is based on the idea that the mind is a collection of mental processes, each of which is responsible for a specific function. This approach has been successful in explaining many of the phenomena that are of interest to psychologists, such as memory, attention, language, and thinking.

The computational approach to the mind is the oldest and the least influential. It is based on the idea that the mind is a collection of algorithms, each of which is responsible for a specific function. This approach has been unsuccessful in explaining many of the phenomena that are of interest to psychologists, such as memory, attention, language, and thinking.

The structural-functional approach to the mind is the most recent and the least influential. It is based on the idea that the mind is a collection of specialized processing units, each of which is responsible for a specific function. This approach has been successful in explaining many of the phenomena that are of interest to psychologists, such as memory, attention, language, and thinking.

Both autism and schizophrenia present a distinctive profile of spared abilities and disabilities, with “core” symptoms that invite unification in terms of discretely impaired cognitive function. By the principles of reverse engineering,
such unification suggests there may be a single underlying cognitive entity that is differentially affected. However, although such reasoning is not always fallacious, it may inhibit research into more basic underlying cognitive and/or mechanical explanations of “core” symptoms. In our view, this research becomes especially important if there is no explanation of a range of further symptoms with which these core symptoms are characteristically comorbid. ToMM theorists have generally recognized the need to widen the explanatory scope of their accounts, but have failed to do so in any convincing way. Admittedly, their proposals are often speculative, awaiting further empirical investigation. But it may be that they persist in this approach partly because it seems unlikely that such distinctive higher order cognitive abnormalities could be explained except by way of selective impairment to dedicated higher order cognitive systems. In this chapter, we have argued that such cognitive abnormalities may in fact be caused—developmentally or, indeed, architecturally—by damage to lower order cognitive mechanisms and/or mechanical malfunction, and there are good reasons to prefer such explanations for both autism and schizophrenia. In any case, we propose that theorists should only retain high-level cognitive hypotheses where these function as theoretically ineliminable constraints on implementation theories for cognitive capacities. In the case of both autism and schizophrenia, continuing research suggests that, even though ToM deficits constitute a clinically interesting set of related symptoms, their explanation will not involve a dedicated ToM module. In fact, our bet is that this module will prove to be eliminable as the deeper cognitive and noncognitive structure of both disorders becomes more fully understood.

NOTES

1. Mirroring our earlier point, because mechanical malfunction inevitably produces cognitive malfunction, some might doubt the viability of distinguishing between mechanical and cognitive explanations for various disorders. Although we think the distinction between PKU and, say, prosopagnosia is clear enough, we reinforce the idea by considering another computer analogy. Suppose some instability in the microcircuitry of a PC causes a particular, resource-hungry, program to crash first. The program is indeed a discrete computational entity identified by its selective malfunction, but the ultimate cause is not computational but mechanical. The reason is simply this: The solution is not to reinstall the program, but to fix the microcircuitry.

2. Some of the material in this section draws heavily on ideas developed in McGeer (2001) and McGeer (forthcoming). For additional arguments against the ToM approach to autism, see Gerrans (1998).

3. There are a number of variations of this task, but one simple version that has been used on autistic populations is the so-called “Sally–Ann” task (Baron-Cohen, Leslie, & Frith, 1985; cf. Wimmer & Perner, 1983): Children are shown two dolls, “Sally” and “Ann.” Sally has a basket in which she places a marble. Then, she goes away leaving her basket behind. Ann takes Sally’s basket and hides the marble in it. Sally returns, and the children are asked to predict what Sally will do. To pass, children must correctly predict that Sally has a belief and failing false photograph: Like recent studies suggest this may be an area where children do equally well on both tasks once differences between them have been eliminated to continue to show the dramatic physical symptoms (Peterson & Siegel, 1998).

5. For a defence of this perspective and for the following limited summary is culled from Tager-Flusberg, & Cohen (2000). See also (1999).

6. The theoretical perspective advocated by the following work is the perspective advocated by these authors. Although their social deficits, Uta Frith in particular (WCC) view of autism in contrast to the WCC cannot account for autistic abnormal case (for instance, autistic savant talents). And process incoming information in context with higher level meanings. For an explanation (1989) and Happé (1999). At one time, on the grounds that autistic assets are not recent, she seems to have leaned more WCC, ToM, and executive dysfunction, pulse control, shifting tasks, and the initiative, Happé herself has also advocated some instance in her writing: “To date, the experimental evidence is that these two aspects of autism interact in context might contribute to everyday social play a part in certain social impairments... example... could hamper emotion recog- nition, with respect to ToM, and finally, settling views as yet, we’re日晚间, highlighting in this chapter.

7. This is not to suggest that Gopnik endorses the capacities discussed in this article. In Gopnik, theory of mind is— as the term suggests is to be explained in terms of a child’s work on progressively elaborating an idea, and new evidence (see, e.g., Gopnik, Capps, 1998).

8. As yet ToMM theorists give no explanation for which the explanation should not work as reliably as one.
A single underlying cognitive entity that
though such reasoning is not always falla-
basic underlying cognitive and/or me-
chisms. In our view, this research becomes
vation of a range of further symptoms
characteristically comorbid. ToMM theo-
widens the explanatory scope of their
convinced way. Admittedly, their pro-
other empirical investigation. But it may
nty because it seems unlikely that such
malities could be explained except by
ed higher order cognitive systems. In
cognitive abnormalities may in fact be
fectively—by damage to lower or-
nal malfunction, and there are good
both autism and schizophrenia. In any
only retain high-level cognitive hypoth-
ly ineliminable constraints on implica-
ties. In the case of both autism and sch
states that, even though ToM deficits
related symptoms, their explanation will
fact, our bet is that this module will
cognitive and noncognitive structure of
nderstood.

NOTES

4. Mechanical malfunction inevitably produces
the viability of distinguishing between me-
various disorders. Although we think the
agnosia is clear enough, we reinforce the
alogy. Suppose some instability in the mi-
source-hungry, program to crash first. The
onal entity identified by its selective mal-
putational but mechanical. The reason is
it the program, but to fix the microcircuit.
y heavily on ideas developed in McGee
additonal arguments against the ToM ap-
task, but one simple version that has been
"Sally–Ann" task (Baron-Cohen, Leslie,
1983): Children are shown two dolls, "Sally"
he places a marble. Then, she goes away
leaving her basket behind. Ann takes Sally’s marble out of the basket and puts it in a
box. Sally returns, and the children are asked "where will Sally look for her marble?"
To pass, children must correctly predict that Sally will look in the basket where she
believes the marble to be, as opposed to the box where they know the marble is
themselves.

5. Although some studies indicate normal children may show weak superiority on false-
photo tasks (Slaughter, 1998) or dissociation in the opposite direction (passing false
belief and failing false photograph; Leslie & Thaiss, 1992; Zaitchik, 1990), more
recent studies suggest this may be an artefact of experimental design. Normal 4-year
olds do equally well on both tasks once incidental conversational and linguistic dif-
fferences between them have been eliminated. Autistic subjects, on the other hand,
continue to show the dramatic physical–mental dissociation seen in earlier studies
(Peterson & Siegal, 1998).

6. For a defence of this perspective and for details of autistic abnormalities from which
the following limited summary is culled, see the collected papers in Baron-Cohen,

7. The theoretical perspective advocated in this paper stands in some contrast to other
work done by these authors. Although both have been friendly to a ToM account of
social deficits, Uta Frith in particular has advocated a "weak central coherence"
(WCC) view of autism in contrast to the ToM approach on grounds that the latter
cannot account for autistic abnormal capacities in areas unrelated to social cognition
(for instance, autistic savant talents). According to WCC, autistics lack a capacity to
process incoming information in context, thereby using contextual cues to deliver
higher level meanings. For an explanation and defense of this approach, see Frith
(1989) and Happé (1999). At one time, Frith seemed to reject the ToM approach on
the grounds that autistic assets and deficits are likely to have the same origin.
Recently, she seems to have leaned more toward a multiple deficit approach involving
WCC, ToM, and executive dysfunction (affecting planning, working memory, im-
pulse control, shifting tasks, and the initiation and monitoring of action; Frith, 1997).
Happe herself has also advocated something like a weak ToM (multiple dysfunction)
stance in her writing: "To date, the experimental findings suggest that weak central
coherence and theory of mind are somewhat independent... However, it is likely
that these two aspects of autism interact, and failure to integrate information in
context might contribute to everyday social difficulties. Featural processing might
play a part in certain social impairments. Piecemeal processing of faces, for
example... could hamper emotion recognition" (Happe, 1999, p. 220). This theo-
retical vacillation with respect to ToM may well indicate that the authors have no
settled views as yet, due precisely to the kinds of explanatory difficulties we are
highlighting in this chapter.

8. This is not to suggest that Gopnik endorses the kind of modularity approach to ToM
capacities discussed in this article. In Gopnik’s view (and the view of her collabora-
tors), theory of mind is—as the term suggests—a theoretical competence. Its acquisi-
tion is to be explained in terms of a child’s native capacity for theory building put to
work on progressively elaborating an innate proto-theory of agency in the face of
new evidence (see, e.g., Gopnik, Capps, & Meltzoff, 2000).

9. As yet ToMM theorists give no explanation for why a late-developing ToM mecha-
nism should not work as reliably as one that develops on schedule, other than its
lateness. Perhaps to avoid this problem, some suggest that high-functioning autistic kids that are able to pass theory-of-mind tests do so by means of "compensatory mechanisms"—general problem-solving capacities deployed in the social realm. However, it is hard to see how this use of "general problem-solving capacities" would not require some specialized knowledge about psychological matters, in which case the expertise of normal subjects might well be explained by the vast difference in their knowledge base acquired through years of training, rather than by the presence of any specialized mechanism for psychological reasoning.

It is interesting to compare these various clinical populations with ‘Down’s syndrome children who do pass false-belief tasks at the same mental age as normal children. Hence, Down’s syndrome children are developmentally retarded, but they seem to follow a normal developmental trajectory (see, e.g., Baron-Cohen, Leslie, & Frith, 1985). By contrast, congenitally blind children and deaf children of hearing parents follow the same abnormal developmental trajectory as autistic children: They are unusually delayed in passing theory-of-mind tasks compared with non-social reasoning tasks. They also show autistic-like abnormalities in social, communicative, and imaginative abilities; but, as is not the case with autistic children, these abnormalities tend to disappear as they become more able to relate to others through developing skills that overcome their handicaps in a context of able and responsive others (Brown et al., 1997; Peterson & Siegal, 1998). It is also notable that deaf children whose parents are native signers, and who therefore have rich protoco-conversational and conversational interactions with others from an early age, do not show any “autistic” social or cognitive abnormalities in the nature of their conversational behavior (Meadow, Greenberg, Erting, & Carmichael, 1981). In particular, they do not have any difficulty passing “theory-of-mind” reasoning tasks. These comparative results are discussed in Peterson and Siegal (1999).

REFERENCES


some suggest that high-functioning autistic
is do so by means of “compensatory mecha-
ies deployed in the social realm. However,
pologies” would not re-
not psychological matters, in which case the
be explained by the vast difference in their
of training, rather than by the presence of
gical reasoning.
clinical populations with Down’s syndrome
the same mental age as normal children.
mentally retarded, but they seem to
ly (see, e.g., Baron-Cohen, Leslie, & Frith,
children and deaf children of hearing parents
al trajectory as autistic children. They are
mind tasks compared with non-social read-
obnormalities in social, communicative, and
as with autistic children, these abnor-
are able to relate to others through develop-
has taken a context of able and responsible others
1998). It is also notable that deaf children
the age, do not show any “aut-
the nature of their conversational behav-
ich, 1981). In particular, they do not
“reasoning tasks. These comparative re-
1999).

REFERENCES

ess on autism and theory of mind. Cam-
bridge, MA: MIT Press.

International Review of Research in Mental

Baron-Cohen, S., Leslie, A. M., & Frith, C. D.
(1985). Does the autistic child have a theory of

Blakemore, S.-J., Goodbody, S. J., & Wolpert,
D. M. (1998a). Predicting the consequences of
our own actions: The role of sensorimo-
tor context estimation. Journal of Neuro-
science, 18, 7511–7518.

delusions, a phenomenological approach.

Brown, R., Hobson, R. F., Lee, A., & Steven-
son, J. (1997). Are there “autistic-like” fea-
tures in congenitally blind children? Journal
of Child Psychology and Psychiatry, 38(6),
693–703.

Campbell, J. (1999). Schizophrenia, the space
of reasons and thinking as a motor process.
Monist, 8, 609–625.

Carruthers, P. (1996). Autism as mind-blind-
ness: An elaboration and partial defense. In
P. Carruthers & P. Smith (Eds.), Theories of
theories of mind (pp. 22–38). Cambridge:
Cambridge University Press.

Currie, G. (1996). Simulation theory, theory-
tory, and the evidence from autism. In P.
Carruthers & P. Smith (Eds.), Theories of
theories of mind (pp. 242–256). Cambridge:
Cambridge University Press.

Dapretti, E., Franck, N., Georgieff, N., Proust,
J., Pacherie, E., Dalery, J., & Jeannerod, M.
(1997). Looking for the agent: An investiga-
tion into consciousness of action and self-
consciousness in schizophrenic patients.
Cognition, 63, 71–86.

Decety, J., Perani, D., Jeannerod, M.,
Bettinardi, V., Tadary, B., Woods, R.,
motor representations with PET. Nature, 371,
600–602.

essays on mind and psychology. Monto-
gomery, VT: Bradford Books.

Williams syndrome: From genotype through
to the cognitive phenotype. American Journal
of Medical Genetics: Seminars in Medical
Genetics, 97(2), 164–171.

Elman, J., Bates, E., Johnson, M., Karmiloff-
Rethinking innateness: A connectionist perspec-
tive on development. Cambridge, MA: MIT
Press.

symptoms of schizophrenia reflect impair-
ments in the perception and initiation of

Frith C. D. (1992). The cognitive neuropsy-
chology of schizophrenia. Hove: Lawrence
Erlbaum.

Frith, C. D., Blakemore, S. J., & Wolpert, D.
M. (2000). Abnormalities in the awareness and
control of action. Philosophical Transac-
tions of the Royal Society of London B, 355,
1771–1788.

Firth, C. D., & Done, D. J. (1989). Experiences
of alien control in schizophrenia reflect a
disorder in the central monitoring of action.
Psychological Medicine, 19, 359–363.

Oxford: Blackwell.

Firth, U. (1998). Cognitive explanations of

Firth, U. (1997). The neurocognitive basis of
autism. Trends in the Cognitive Sciences,
1(2), 73–77.

and self-consciousness: What is it like to be

Gallagher, S. (2000). Philosophical concep-
tions of the self: Implications for cognitive
21.

phrenia: A cognitive model of immunity to
error through misidentification. In D. Zahavi
(Ed.), Exploring the self: Philosophical and
psychopathological perspectives on self-
experience (pp. 203–239). Philadelphia: John
Benjamins.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti,
G. (1996). Action recognition in the pre-

Gallese, V., Fadiga, L., Fogassi, L., Luppino,
circuit for hand grasping movements in the
monkey: Evidence from reversible inactiva-
tion experiments. In P. Thier & H. O.
Karnath (Eds.), Parietal lobe contributions to
orientation in 3-D space (pp. 255–269).
Heidelberg: Springer-Verlag.

consciousness of external reality: A “who”
system for consciousness of action and self-
consciousness. Consciousness and Cogni-

Gergely, G., & Watson, J. (1996). The social
biofeedback theory of parental affect mir-
roring: The development of emotional self-
awareness and self-control in infancy.
International Journal of Psychoanalysis, 77,
1181–1212.

development. Mind and Language, 13, 56–75.

Gerrans, P. (2002). Theory of mind in evolu-
tionary psychology. Biology and Philosophy,
17, 305–321.


British Journal of Developmental Psychology, 16, 301–320.


Wing, L., & Gould, J. (1979). Severe impairments of social interactions and associated abnormalities in children: Epidemiology and
Suppose we use the term *theory of mind* to mean that having a theory of mind is simply part of everyday folk psychological practice, with the term being most useful for prediction. This use of the term is noncommittal and is not meant to explain the fact of a capacity to identify with others in our lives; or perhaps in terms of a mixture of capacities. At least in its current use of *theory of mind* is also noncommittal. It is, after all, our psychological abilities resides in a spectrum.

Once we abstract away from questions about modularity, it can be said that having a theory of mind—being able to engage in everyday folk psychological understanding—is fundamental to our understanding of the human condition. As theory of mind is . . . fundamental to an understanding of mental development, it is natural to suppose that, just as a naïve,