Dissociation of social affect and theory of mind in a case of Asperger syndrome[*]

NEIL SMITH, BEATE HERMELIN and IANTHI TSIMPLI

Abstract

We report on a case of Asperger syndrome (KH), who has considerable linguistic ability with a verbal IQ of 153. Strikingly, he readily and correctly identifies the use of sarcasm, showing mastery of the meta-representation and dissociation characteristic of 'interpretive use'. Current theories of the syndrome largely agree on the constellation of properties which define it, but disagree on whether the basic cause resides in a deficit in Theory of Mind or in a social and emotional deficit. We suggest an account in a 'quasi-modular' version of the modularity hypothesis, as developed by Smith and Tsimpli. We specify a range of quasi-modules (including Theory of Mind and Social Interaction, itself divided into Social Cognition and Social Affect); we outline the structure of the Emotional component, including basic and derived emotions; and we then spell out some of their inter-connections and their relation to the Language Faculty. Our tentative conclusion is that all these components may dissociate, and that high intelligence, combined with linguistic ability, may mask a deficit in Theory of Mind, though not in Social Affect.

1 Introduction

Asperger syndrome is standardly viewed (see e.g. Wing, 1991) as a form of autism, characterised by the constellation of properties in (1):

(1)   A lack of understanding of the mental states of others.
       Ineptitude in social interaction.
       Impairment of emotional responses from the earliest stages of development, both as regards the expression of emotions and, especially, understanding those of other people.

That these properties are typical of Asperger syndrome subjects is not at issue, but it is unclear whether the syndrome is correctly characterised by their individual necessity and joint sufficiency. Ultimately, this doubt follows from the fact that the aetiology of

autism itself is a matter of controversy. There is a consensus that it involves a problem with Theory of Mind (see e.g. Frith, 1989; Baron-Cohen, 1995; Yirmiya et al, 1998), but there is disagreement over whether this is the root of the condition or a side-effect of a more basic emotional or social deficit (e.g. Hobson, 2002; Klin, 2000). We attempt to shed light on the issue by reference to a case (KH) in which these properties appear to dissociate.

We are not the first to highlight such dissociation. Klin (2000), Tager-Flusberg and Sullivan (1999), and Garfield et al (2001) have argued that Theory of Mind may fractionate, with identifiable cognitive and affective sub-parts. We build on this work, and argue further that a cognitive deficit may be compensated for by high linguistic and meta-linguistic ability, whereas an affective deficit may not be. As a result, someone with all the other manifestations of Asperger syndrome may appear to have an intact Theory of Mind in the traditional sense.

2 Background

There has been considerable progress in the controlled experimental study of the nature and origin of autistic children's apparent withdrawal from the world. Following the pioneering work of O'Connor & Hermelin (1970), Premack & Woodruff (1978), Wimmer & Perner (1983), Baron-Cohen et al (1985), Perner et al (1989), (see also Hermelin, 2001), a consensus gradually emerged that the primary cause of autism in all its forms was a defective Theory of Mind. The explanatory force of this aetiology was then generalised beyond the cognitive domain to include the emotions. Thus, it has been suggested that Asperger syndrome subjects' emotional ineptitude is a function of a deficit in theory of mind - e.g. Tantam, (1991: 158) who writes that Asperger subjects characteristically have trouble *interpreting* emotions, especially from facial expressions; and Rieffe & Terwogt (2000:601), who suggest that deaf children's impaired understanding of emotions is attributable to "an impaired theory of mind", specifically: "[d]eaf children explain other people's emotions by means of desires regardless of the typicality of that emotion" (2000:607).

More recently, alternative explanations have been provided by Hobson (2002) and Klin et al (2002), who trace the origin of autistic disorders to social and emotional problems. Hobson writes (2002:5) that the "autistic child's lack of emotional connectedness with others … has quite startling implications for the child's ability to think"; and Klin et al (2002:895) claim that "the core social disorder defines the condition and likely affects the development and expression of these other skills". That is, these researchers reverse the direction of causation between the emotional and the cognitive, attributing the impairments characteristic of autism ("often including abnormalities in language" – Hobson, 2002:6) to social and emotional

impoverishment.

It is clear then that the relation between the cognitive and affective components of Asperger syndrome is not straightforward. We think that the case of KH may contribute to an understanding of the situation, and argue that the characterisation of Asperger subjects needs to be cast in terms of a dissociable combination of deficits in two modules: Theory of Mind and Social Interaction.

The effects of a defective Theory of Mind are seen most clearly in variants of 'false belief' tasks, such as 'Sally-Anne' and 'Smarties', where children below the age of four or so, and autistic subjects regardless of age, are unable to conceive that someone else can have a representation of the world at variance with their own representation of the true state of affairs. Such meta-representations (i.e. representations of representations) emerge at about the age of two, when children also begin to develop pretend play (Leslie, 1987). During normal development, such mental representations develop into the ability to 'read the minds' of others for their intentions, feelings and mental states. There is a certain tension between the early age at which pretend play appears (around 2) and the later age at which success in false belief tasks begins (around 4). However, a more elaborated version of Theory of Mind (see e.g. Baron-Cohen, 1995) breaks it down into a number of precursory sub-abilities, e.g. EDD - Eye-Direction Detector, and SAM - Shared-Attention Mechanism), which come on-line at different stages. For instance, Harris (1989:212-213) observes that while autistic children are incapable of attributing emotions to others they have a Shared-Attention Mechanism, as shown by the fact that they can understand pointing to an object as attracting their visual attention, but fail to deduce other implications of this pointing, such as that the object may cause emotions in the person pointing at it, which he in turn may want the other person to share.

Fruitful as this attempt in terms of a Theory of Mind deficit has been to account for the impairments in social interaction by those with autism, there have also been indications that such an explanation may be necessary but not sufficient (Klin, 2000), as it is well-known that perceptual, affective and cognitive factors interact in a range of other conditions. Thus profoundly deaf children (Peterson & Siegal, 1995, 2000) and mentally retarded children (Yirmiya et al, 1998) show impaired social interaction. Hobson (2002) shows that children with normal mental development, but who are congenitally blind or severely visually impaired, pass fewer theory of mind tests than sighted children, and that this latter group's verbal mental age is highly correlated with performance on theory of mind tasks. Similar correlations between IQ and success with theory of mind problems have been widely reported in relation to autism (e.g. Happé, 1995.). Moreover, Hermelin (2001) reported that in normal development, recognition of emotions involves the integration of different modalities in that it uses both voice cues and facial expression.

A further complicating factor is that Frith et al. (1994) have reported that some autistic children who passed Theory of Mind tests also showed better social adaptation in real life situations than those who were unable to pass these tests. They concluded that there was a minority amongst those with autism who had at least some competence in thinking about the thoughts of others in real life interpersonal context.

In one attempt to tease out the respective roles of social and cognitive factors, Klin (2000) developed a Social Attribution Task (SAT) and tested it on normal, autistic and Asperger Syndrome subjects. Using geometric shapes that "act like people", Klin (2000:833; cf. Heider & Simmel, 1944; also Abell et al, 2000) predicted that AS subjects would base their descriptions more on physical/geometric considerations than social ones. That is, their 'folk physics' should be better preserved than their 'folk psychology', where both of these domains are conceptualised cognitively. It follows that a deficit in Theory of Mind alone is inadequate to explain certain aspects of autistic behaviour, as "having Theory of Mind skills does not guarantee commensurate social adaptation skills" (Klin, 2000:832).

Hobson (2002) highlights a number of interesting difference between autists and normals. In the interpretation of point-light people, autists made no mention of feelings (2002:56); on IQ tests where the vocabulary pertained to emotion (2002:187) there were systematic difference between autists and normals; and in tests where the subjects had to draw a 'house' versus 'my house', as opposed to a 'person' versus 'self' (2002:236f.) there was a marked disparity in ability: autists were as good as or better than normals in drawing different houses, but markedly worse in characterising the difference between self and other. Most strikingly, picture sorting tasks differentiated between normals and autists: normals typically sorted by facial expression, whilst autists never did (Lee & Hobson, 1998; Hobson, 2002:214), yet autists were equally good at identifying the emotions expressed on faces and upside down faces (Hobson, 2002:247), whereas normals were markedly better at interpreting faces the right way up. Similarly, autists fail to develop a concept of 'friend' (Hobson, 2002:228) and seem to have minds that are "thinly populated" (idem 233).

For present purposes, the most important point to be derived from Klin's and Hobson's work is the differentiation into 'cognitive' and 'affective' components of the mind, reminiscent of Tager-Flusberg and Sullivan's (1999) division of social intelligence into 'social-cognitive' and 'social-affective' (cited in Garfield et al, 2001:523). We would highlight Klin's observations that: "…although higher ToM or metalinguistic capacity may have given the HFA [high-functioning autist] and AS [Asperger syndrome] participants a higher vocabulary to use when questioned, it did not necessarily facilitate spontaneous understanding of the social plot depicted in the cartoon" (Klin, 2000:840), and: "If one does not naturally seek social meaning in the environment, one's ToM capacities are of little avail: in other words, there is a need for a theory of how ToM skills are put into action" (Klin, 2000:841).

This leads us to the subject of the present report.

3 The Participant

At the time of testing, KH was an 11 year old boy (date of birth 22.8.89), diagnosed with Asperger syndrome and with a verbal IQ (WISC, administered 13/8/98 by Ms J. Wilson) of 153, a performance IQ of 105 and a full scale IQ of 135. These verbal skills reflect an ability better than 99.9% of his age band. He is also reported (Connell, 1998) as having most success in "non-verbal reasoning tasks" such as Raven's matrices, though he found it "very difficult to cope … socially and emotionally" with (a Rudolf Steiner) school, and had been educated at home by his mother since Easter 1998. He is agile and has a great sense of balance, but is nonetheless mildly apraxic, being unable to "tie his shoe laces or ride a bike" (Moore); he cooperates well, but has "an over-concern with rituals and routines", and displays an array of autistic like behaviours, as revealed in the PIA (Parent Interview for Autism), reported in Moore; and he is also said to live "in a fantasy world of his own" (Connell).

From the beginning his mother noticed that he was in many ways different from other small children. He never smiled back at her and did not seem to respond differently to people and objects. As a toddler he developed various obsessive mannerisms such as rubbing his head to the point of injury, biting his cheek, and making strange repetitive noises. He also hummed incessantly and monotonously, shook his head for long periods of time and ran round and round in circles. Repetitive movements and a need for sameness were also noted as well as a failure to respond to painful stimuli, allied with over-sensitivity to certain noises. He was quite unable to relate to, or play with, other children - when other children chatted, he would cover his ears with his hands; and it seemed to his mother that he was living in a world of his own. However, he showed advanced language development.

He entered primary school when he was 5 years old but, despite the fact that his reading skills were outstanding from the beginning, it soon became obvious that he was anxious and unhappy there, especially during 'playtime'. When he was 6 he transferred to a Rudolf Steiner school, but his social and behavioural problems became even more obvious, his school performance was erratic and unpredictable, and it was decided that he should be taught at home. Accordingly, until the age of 12 he was tutored at home by a professional. He was diagnosed at age 8 as suffering from Asperger Syndrome. The report was based on Rutter's (1978) classical criteria of autism: deficits in reciprocal social and emotional interactions, impaired communication and stereotyped behaviour with an insistence on fixed routines. It was also noted that despite his verbal skill, he did not take the abilities of his interlocutor into account, and tended to use unusually complex language and rare words which

even his 17 year old sister did not understand. His non-verbal communication was impaired in that he often could not understand other people's facial expressions or gestures. He loved word-play such as anagrams, he invented elaborate fantasy games, and he read books voraciously.

At age 11, KH wrote a book *Asperger Syndrome, the Universe and Everything* (Hall, 2000), in which he comments on how happy the diagnosis of Asperger Syndrome had made him, as at last he understood why he had always felt different from other people. In addition to his linguistic ability, which we summarise shortly, KH has other talents. At 11 he passed the GCSE maths exam at grade B, and at age 12 obtained a grade A in the GCSE Information Technology exam. In September 2001, KH started as a boarder at a special school for children with Asperger Syndrome. Lessons are highly structured and there is an excellent staff: pupil ratio. From the beginning, KH has been positive and enthusiastic about the school, and claims to have experienced no homesickness.

When we met KH, he was not unfriendly but rather remote. After he had tested us on our (somewhat minimal) knowledge of Ulster English, he became co-operative and enthusiastic, and provided us with the data reported on below.

4 The Tests

We tested KH on a battery of linguistic tasks which had been previously devised to test the polyglot savant Christopher (Smith & Tsimpli, 1995, Tsimpli & Smith, 1998). The aim had been to establish whether Christopher's knowledge of his first language (English) and various aspects of his use of that language fell within normal limits, and we applied the same reasoning to KH. Accordingly, the tests covered a wide range including examples of the kind illustrated in (2) to (7).

(2) Judgements of grammaticality across a wide range

In each case KH was given a written sentence and asked to mark it right or wrong and, if wrong, to provide a corrected version, as in (c):

    a. Remember Susan to feed the dog   X
    b. Would you remind her to buy some ginger, please.   ok
    c. Alex gets often into trouble with the police.   X
       "Alex often gets into trouble with the police"
    d. The weather today is beautiful, is it?   X

KH's judgements of well- and ill-formedness were normal: that is we have reason to believe that his knowledge of language (his I-language, in Chomsky's, 1986, sense) is not distinct from that of the normal population. These results indicate that KH has control of simple agreement relations, selection (both c-selection and s-selection), phenomena of binding, basic movement and adjacency constraints, etc. More interestingly, when NS pronounced (d) with a sarcastic Fall + Rise intonation, he immediately accepted it, commenting that it was then 'sarcasm'. This mastery of so-called 'interpretive use' is extremely rare in the autistic and Asperger population.

(3) Tests of inferential ability involving the use of modus ponens, modus tollens, etc. in which KH had to read a dialogues and then answer the question in the last line by underlining either 'Yes' or 'No' or 'Don't know'.

Michael said: 'If George comes I shan't be able to play.'
Fred said: 'George is coming.'
Do you think Michael will be able to play?   Yes/<u>No</u>/Don't know

Given that autistic people are usually literal minded and often incapable of lying, it is of considerable interest that he added spontaneously: "Unless he's lying".

(4) Tests of the use of implicated assumptions and conclusions, in which he had to read a dialogue and then answer the question in the last line by underlining either 'Yes' or 'No' or 'Don't know'.

a. John said: 'Would you like some coffee?'
   Mary replied: 'Coffee would keep me awake.'
   Do you think Mary accepted the coffee?   Yes/<u>No</u>/Don't know

He spontaneously asked "What time of day was it?" before answering; again showing appropriate awareness of possible divergence from a default interpretation. This was confirmed by his response to the next example:

b. Mary said: 'I have to work all night tonight.'
   John said: 'Would you like some coffee?'
   Mary replied: 'Coffee would keep me awake.'
   Do you think Mary accepted the coffee?   <u>Yes</u>/No/Don't know

(5) Tests of the use of discourse connectives, such as *after all,* and *you see*, in which he was asked to read an example and fill the gap, marked by ---, with one of the forms: *After all/Anyway/Moreover/So/Therefore/You see*:

a. Jill was waiting for her boyfriend in the park. She was very depressed and miserable. --- she'd just lost the pet dog her boyfriend had given her.
   "You see"

b. John and Bill wanted to catch an elephant. They spent six weeks hunting for one in the jungle, but however hard they looked they couldn't find one. They thought this was very strange: --- elephants are very big and should be easy to find.
   "After all"

c. When I was young I wanted to be a pilot. At that time I used to have all sorts of mad ideas; I even hoped to become chief test pilot. --- I decided to learn to fly.
   "So"

(6) Tests of the construal of discourse referents in ambiguous contexts. Here he was asked: "Who do you think 'he' refers to in sentence (a)? Who does 'he' refer to in sentence (b)? Are sentences (c) and (d) both correct?"

a. <u>John</u> telephoned Bill. <u>He</u> needed to speak to him.
b. John telephoned <u>Bill</u>. <u>He</u> refused to speak to him.

KH underlined as indicated, and added re (b) "or pointless if it was John".

c. John telephoned Mary. She refused to talk to him.   ok
d. John telephoned Mary. She wanted to talk to him.   ok
   [KH added re (d) "He wants to talk to *her*"]

(7) Tests of making anagrams; word-making from a given head-word; etc.

a. He provided 'dawn' for 'wand', and 'anger' for 'range'. He was unable to provide an anagram for 'chesty', but was pleased when 'scythe' was supplied.

b. He was equally proficient at making words from a Head-word: given the examples shown in (c) and (d), he provided those in inverted commas in (c') and (d'):

c. DISASTER: star, aster, tried …

c'. "A, I, AS, DISASTER"

d. BATTLEFIELD: file …

d'. "BATTLE, FIELD, A, I, AT, BAT, FILED"

He was reasonably competent at completing word-ladders (something that Christopher, for instance, never understood at all). Given: 'Hate - Love' in (e) as an example, he provided the examples for 'Heat - Fire', and 'Hand - Foot' in (f) and (g):

e. HATE   have   hove   LOVE
f. HEAT   feat   fear   FIRE
g. HAND   hood   hoot   FOOT

On all of these tests KH performed extremely well. Most strikingly, he performed flawlessly on judgements of irony and sarcasm (8a); metaphor (8b); jokes (8c); the use/mention distinction (8d), and other examples involving what is known as 'interpretive use' (Sperber & Wilson, 1995) or meta-representational ability.

(8) a.   Can you explain what is meant by the following <u>underlined</u> expressions?
     i.   "<u>He's a fine friend</u>". Said of someone who has just kicked you.
           His instantaneous response was: "They're not much. Sarcasm".
     ii.   John and Mary went to a party, where both of them became very sick, and had to go home early, because John gave Mary too much to eat and drink. Mary said: "<u>What a wonderful party!</u>"
           "More sarcasm".
     iii. The judge told the traitor that he was "<u>a credit to his country</u>".
           "More sarcasm; or he was a foreigner".

     b.   <u>No man is an island</u>
           "No-one's alone".

     c.   i.   Diner: "Waiter, what's that fly doing in my soup?"
                 Waiter: "It looks like the breast-stroke, Sir"
                 He responded with "It's called a bad joke".
           ii.   A Russian minister visits a car factory. The manager goes out of his way to show him around and at the end of the tour offers the minister a free car.
                 "Oh no", says the minister, "I can't accept it".
                 "In that case I'll sell it to you for five roubles".
                 The minister hands him a ten rouble bill:
                 "In that case, I'll have two".
                 His response to this one was to laugh - presumably, not such a bad joke - but surprisingly normal.

(8) d.   He found nothing remarkable with the use/mention distinction, illustrated in (i) versus (ii), and (iii) versus (iv) (that is, with and without 'scare quotes'):
    i.   Dogs have four legs
    ii.   Dogs has four letters
    iii. Kenneth is a great writer
    iv.   "Kenneth" is a great name

Most strikingly, he gave perfect judgements on written versions of 'Sally-Anne' (false belief) tasks, as illustrated in (9):

(9) He was asked to answer the questions at the end of the description:
    a.   Three children, Alexia, Maureen and Jill are playing in the kitchen. While the others are watching, Alexia puts a chocolate under a teacup. Maureen then goes out of the room. While she is away, Jill removes the chocolate from under the cup and puts it in a saucepan. When Maureen comes back into the room, where will she look for the chocolate?
    He responded: "Under the cup unless she heard the saucepan".

    b.   John, Clare and Chris were tidying up the house together. They agreed to put the stamps in the top drawer of the desk. When Clare had gone out, John and Chris changed their mind and put the stamps on the bookcase. Where do you think Clare will look for the stamps when she comes back?
    "In the top drawer".

    c.   John, Tony and Paddy agreed to meet in the House of Commons on Wednesday. Later, John and Tony changed the meeting till Thursday, but neither of them remembered to tell Paddy. Which day did Paddy arrive for the meeting?
    "Wednesday, except in [inaudible]"

An interesting partial parallel in the literature is provided by Happé's (1991) subject David, who has "apparently good understanding of others' minds" (1991:218). This conclusion was arrived at on the basis of his use of mental state predicates like 'think'; though even here there was a lot of "I think …" but only one example of "I think he thought …". Given the performance of KH and David, it is of interest to read Happé's (1991:234) remark: "It is widely reported that even the most verbally able autistic people (that is, people with Asperger syndrome) fail to understand non-literal speech such as sarcasm, joking and metaphorical expression". She stresses Sperber & Wilson's (1995) notion of comprehension of intention, and the use of inferential communication, and we should test KH further in this domain.

5 The Framework

In earlier work, we argued that the 'central system' (in either Fodor's, 1983, or Chomsky's, 1975, 1984 sense) was characterised by a Basic Processing Mechanism (à la Anderson, 1992), a general memory structure, a set of Special Processors, and, crucially, a set of (quasi-)modules. Quasi-modules (Tsimpli & Smith, 1998; see also Smith & Tsimpli, 1995, where they are described as 'central modules') have some of the properties characteristic of Fodorian modules (Fodor, 1983): their operations are domain specific, fast, mandatory, general to the species, and subject to idiosyncratic pathological breakdown. Importantly, however, they do not share all of those characteristics: their operations are not informationally encapsulated and the vocabulary over which they operate is conceptual rather than perceptual. These quasi-modules include at least: the language faculty (in part), Theory of Mind, moral judgement, folk physics, folk psychology, number and music, plus Social Interaction.

Before elaborating the details of the framework we presuppose, it is necessary to make a number of conceptual clarifications. The first is the difference between language and communication. The traditional philosophical position which equates the two is untenable: one can communicate without using language and one can use language for purposes other than communication (for discussion, see Smith, 1999:149f.). This is not to deny that communication using language is infinitely more subtle than non-verbal communication, nor that language is used most strikingly for communication, but the two are in principle separable. The need to make such a distinction in the present context arises from the recently expressed view that "communication leads to thought" (Hobson, 2002:106). There is a sense in which this is obviously true, but there is a logically prior need for the mental (linguistic) representation of what the speaker intends to communicate.

Second, it is essential to differentiate competence and performance in Chomsky's sense; (e.g. Chomsky, 1965; Smith, 1999); that is knowledge of Natural Language and the use of that knowledge in the production and comprehension of utterances. In order to distinguish these notions from other aspects of cognition, we designate this Competence$_{NL}$ and Performance$_{NL}$. The development of such knowledge in normal children is standardly seen as a matter of 'growth' (the joint product of innate and environmental factors) rather than the result of explicit teaching. That is, along with the general community of linguists, we do not accept such claims as: "to learn language and to use words correctly, a child needs to receive correction as correction…" (Hobson, 2002:113). For discussion, see Bowerman, 1987; Smith, 1999).

The competence/performance distinction we are appealing to is a classic one and now relatively uncontroversial. We also think it is necessary to draw a third basic distinction which, while parallel to this, is less widely discussed: that between

knowledge of the Language of Thought in the sense of Fodor (1975) and the use of that knowledge in particular acts of thinking. That is, we assume that the thought processes deployed in (e.g.) problem solving and the 'fixation of belief' are in principle distinct from processes which use natural languages such as German, Yoruba or English. We refer to this distinction as Competence$_{LoT}$ and Performance$_{LoT}$. It is necessary to make this distinction explicit, because of the tendency in some of the literature to run them together. For instance, in discussing "the child's ability to think" and "the roots of thought" Hobson (2002:5, 7, see especially p.105.) fails to distinguish the content of thought and the algebra or syntax of thought, which renders that content transparent. Despite this conceptual distinction between competence of two different kinds, it remains the case that the demands of compositionality (the claim that the meaning of the whole is made up of the meaning of its parts) entail that there can't be too much divergence (in the adult) between Competence$_{NL}$ and Competence$_{LoT}$ (see Cormack & Smith, 2002). Indeed, the usual evidence for the nature of both Competence$_{LoT}$ and Competence$_{NL}$ comes from Performance$_{NL}$, but it is generally accepted that one can deploy the resources provided by one's ability to think without using natural language to do so. This holds, even if the link between Competence$_{NL}$ and Competence$_{LoT}$ is typically mediated by conceptual elements represented in a natural language lexicon.

There is then an interesting empirical question whether normal and autistic subjects are differentiated - inter alia - by the former having 'perfect' Competence$_{NL}$ and the latter having a putatively imperfect Competence$_{NL}$, or if the difference resides elsewhere. It is relatively uncontroversial that autistic subjects manifest anomalous Performance$_{NL}$ in their language behaviour, but it is not self-evident that this is a function of the language faculty alone or of its interaction with other modules. To illustrate the kind of interaction we mean, consider the development of Competence$_{NL}$ and Theory of Mind. On the one hand, it has been claimed that the development of '*that*-clauses' in the language faculty is necessary for the emergence of Theory of Mind (De Villiers & de Villiers, 1999; see also Papafragou, 2002; Tsimpli & Smith, 1998); on the other hand, it has been argued that the linguistic ability to give an epistemic rather than a deontic interpretation to modal sentences is dependent on the emergence of the metarepresentational ability of Theory of Mind feeding into the language faculty (see Papafragou, 1998, for discussion). This contrast is exemplified by the difference in interpretation of the examples containing the modal auxiliary *must* in (10), where (10a) is 'epistemic' – reporting an inference - and (10b) is 'deontic' – reporting a requirement:

(10)   a. John's so untidy, he *must* be unmarried
        b. If John is to enter this competition, he *must* be unmarried

These various conceptual distinctions are usually associated with claims about the innateness of the faculties concerned. The issue of innateness is largely irrelevant to our present concerns: that something is innate in each domain seems to us to be undeniable, but we see no reason for the a priori acceptance of innateness in one domain and its rejection in others. Thus Hobson talks about the "infant's innate capacity for social engagement" (2002:28), and appears happy to accept innateness for vision (2002:54), but he rejects "the innateness hypothesis" (2002:257) for language and theory of mind, saying that it is merely "a cloak for ignorance" (2002:29).

We are happy to agree with Hobson's claims about the importance of social interaction, but it seems clear that there are cognitive prerequisites to socialisation (where 'cognitive prerequisites' means in our terms some mastery of Competence$_{LoT}$). The converse view that there are social prerequisites to cognition could only be true on a construal of cognition as Performance$_{LoT}$, rather than Competence$_{LoT}$. What is crucial is the direction (if any) of causation. A similar conceptual issue arises with regard to the emotions and whether there are social prerequisites to emotional development. We believe with Hobson that there must be such prerequisites, though we are unconvinced that "to perceive emotion is also to react to emotion" (Hobson, 2002:39), as we think that this links emotion and thought too closely (cf. p.151 where he says "thinking and feeling [are] integral to each other"). It is anyway noteworthy that emotion and cognition dissociate in moral judgement – see Greene & Haidt, 2002.

It is implausible that the gamut of emotions should be subserved by one or even several quasi-modules of the type postulated to account for social interaction. Following de Sousa's (1987) ground-breaking work, we adopt a (Platonist) tri-partite characterisation of human nature in terms of desire, reason and emotion. Sperber & Wilson (1995:73-74) have argued for the special status of 'belief' and 'desire' vis-à-vis other mental states, and we would add the emotions to these. (See Smith, 2002). That is, these three categories do not themselves define quasi-modules, but underlie some of the quasi-modules that make up the rest of mental structure, specifically Social Interaction and Theory of Mind.

Any theory of the emotions must take account of the distinction between 'basic' and 'derived' emotions. The former include those emotions which have a direct physiological underpinning and are instantly (perhaps universally) identifiable: fear, terror, happiness, disgust, anger, joy - perhaps the six basic emotions of Ekman (1973). These are, we take it, 'decoded' from facial expression together with other bodily manifestations such as trembling. The 'derived' emotions, such as envy, shame, embarrassment, and jealousy, are more complex in a variety of ways. First, they necessitate an understanding of the distinction between Self and Other; second, they rely on what is made mutually manifest – usually on the basis of linguistic evidence, and hence what is inferentially as opposed to directly Interpretable; and third, they presuppose access to other systems, specifically Social Interaction and Theory of

Mind, demonstrating their lack of informational encapsulation. Both kinds of emotion may be the object of conscious introspection: 'being afraid' and 'knowing that one is afraid' are quite different (see e.g. Kenny, 1963).

Even the derived emotions are not of uniform complexity: shame and envy are typically a relation among three entities - two persons and a property: e.g. person A envies person B some property or possession (her good looks or large income); whereas jealousy is a layered relation among three persons: e.g. person A is jealous of person B because of B's relation to person C. Experiencing derived emotions would then necessitate the availability of a Theory of Mind, and the identification of derived emotions in others would require both this and some further interpretive or inferential ability. One implication of this complexity is that experiencing some emotions involves a fully functioning Theory of Mind, and their identification in others presupposes both this and some considerable sophistication in Social Interaction.

As is clear from the earlier discussion, there is widespread agreement that it is necessary to separate the affective and the cognitive. It is less clear how that separation should be effected. It is uncontroversial that Theory of Mind includes a cognitive component that allows an individual to understand the point of view of someone else. It is also clear that one could have a cognitive component of social interaction, which allowed one to understand the role of other individuals (e.g. as authority figures) without being able to ascribe to those individuals a point of view distinct from one's own. Similarly, we think that there is an affective (i.e. non-cognitive) component which interacts with each of these others, but whose status as part of either one is unclear. We assume that KH suffers from (putatively related) deficits in his ability to represent and interpret (derived) emotions and in his Social Interaction quasi-module. His Theory of Mind may either be intact or, if impaired, then his high intelligence enables him to circumvent any resulting cognitive problems, even though it is not similarly effective in the affective and social domains. We discuss both possibilities.

6 The Analysis

KH's linguistic and cognitive abilities indicate either that he has no significant deficit in his Theory of Mind or that any deficit is compensated by his high intelligence. We accept the distinction drawn by many between the cognitive and the affective components of human mentation, and suggest that different components of the mind may be differentially compensated by high intelligence, especially verbal intelligence: that is, intelligence might mask a cognitive deficit but is relatively unhelpful in the presence of affective or emotional disorders.

This claim is clearly in need of justification, as there is a general consensus that high Verbal Mental Age (VMA) in autistic and Asperger Syndrome subjects is not sufficient to guarantee success on (second-order) Theory of Mind tasks, even if they can pass first-order tasks. We accordingly need to spell out the distinction between a cognitive deficit in Theory of Mind and an affective deficit in the Social Interaction quasi-module, assuming that these involve respectively the attribution of beliefs and the attribution of emotions. That is, we take it that KH's affective deficit is attributable to an impairment of his Social Interaction quasi-module which inhibits the formulation and expression of derived emotions.

We hypothesize that KH's Theory of Mind is intact, but even if it is somewhat impaired, we argue that the potential cognitive effects of such a deficit may be masked by the presence of exceptionally good language, including not only a high VMA but also a system which underpins good inferential abilities carried out in the language of thought. Specifically, KH's normal performance in language-based Theory of Mind tasks may be explained by reference to his exceptionally good language and meta-linguistic abilities, where meta-linguistic skills presuppose 'central coherence' in the sense of Frith (1989; see also Happé, 1999), and require inferential processes whose premises involve linguistic and linguistically-based representations. If this is the case, KH's exceptional abilities lie in the central workings of a Basic Processing Mechanism (in the sense of Anderson, 1992; see Smith & Tsimpli, 1995 for discussion), which appears to be within the middle/high range of performance. This shows a marked contrast with the *savant* Christopher, whose performance IQ is drastically lower. Our claim is then that compensatory strategies for a deficit in Theory of Mind representations may be developed on the basis of an intact Basic Processing Mechanism and meta-linguistic ability.

We attribute KH's understanding of the second-order representations involved in sarcasm, irony and standard Theory of Mind tasks to compensatory strategies that are exceptionally developed on the basis of a high IQ and explicit (taught) knowledge. We know, for instance, that his mother has identified instances of sarcasm for him in her attempts to explain the workings of the world to him. The combination of an intact Basic Processing Mechanism, which can efficiently process 'explicit' knowledge of belief concepts, together with his meta-linguistic ability manipulating lexical, syntactic and prosodic information, enable KH to distinguish between literal and non-literal uses of language. It is also significant that KH understands epistemic uses of modal verbs (e.g. "that will be his wife" as an inference about the present rather than a prediction of the future).

The possibility of developing this kind of compensatory strategy for a putatively deficient Theory of Mind is compatible with the claim that it is quasi-modular. Whether Theory of Mind is unitary, or fractionated into the Cognitive and the Affective, we would expect impairment to lead to a deficit in the attribution of

emotions, because emotions – unlike beliefs – are not encoded exclusively linguistically. Indeed, the contrast between the need for language in expressing beliefs *versus* emotions is clear from the pattern that is observed in children's development (e.g. Harris, 1989, and Garfield et al, 2001, on hearing-impaired children with Theory of Mind deficits). In particular, the attribution of emotions which are not shared by the subject seems to precede the attribution of beliefs (by roughly a year and a half). Given KH's problems of social and emotional interaction, it seems natural to assume a deficit in some quasi-modular domain responsible for generating mental representations of the Self/Other type (cf. Smith & Tsimpli, 1995). In fact, if the attribution of emotions to others is as problematic for him as appears to be the case, we would expect the explanation to reside either in his having a deficit in his Theory of Mind or, more plausibly, in an affective module (the Social Interaction quasi-module) which is accessible to the Theory of Mind. As we have argued independently for Christopher, the quasi-modular status of Theory of Mind allows for compensatory strategies to develop. What is exceptional in KH's case is that a high-functioning Basic Processing Mechanism, together with his inferential and meta-linguistic abilities, enables these strategies to mask the effects of a deficit in the cognitive domain, but not in the affective domain.

To make such an analysis plausible, we need to flesh out somewhat the quasi-modular framework we are presupposing. We take it as axiomatic that we all have some mentally represented theory of Social Interaction. In line with Tager-Flusberg and Sullivan (1999), we suggest that Social Interaction is itself divided into Social Cognition, which links most directly to Theory of Mind, and Social Affect, which links most directly to the emotions. Social Cognition must include categories of Authority, Equality, and Propinquity, where Authority, for instance, may in turn be subdivided into the Known (family, teachers, etc.) and the Unknown (police, traffic wardens, God, etc.); Propinquity and Equality are subdivisible in terms of Family *versus* Friends, and so on. Social Affect pertains mainly to the nature of the engagement with the others identified by Social Cognition, and draws heavily on the emotions. The two may dissociate independently, and both may link with other components of the mind. Thus even though Social Cognition is more directly related to Theory of Mind than is Social Affect, the latter crucially shares with the developing Theory of Mind the need to exploit a Shared Attention Mechanism (in the sense of Baron-Cohen, 1995) to provide a foundation for emotional engagement. Whether each of Social Cognition and Social Affect is an autonomous Quasi-module in its own right, or should rather be viewed as sub-structures of a single entity is still moot. The way we have formulated it here leaves both possibilities open.

This account is directly compatible with the analysis we had constructed for Christopher to accommodate the fact that the familiarity of the participants improved his Theory of Mind performance. In Tsimpli & Smith (1998) we used Johnson Laird's

(1983) notion of Mental Model to account for Christopher's improvement when the stories included known protagonists, suggesting that this reduces the amount of processing that the central processing mechanism or some system parasitic on Theory of Mind will use. This indicates that Theory of Mind and Social Interaction are penetrable and have quasi-modular status. The absence of informational encapsulation is likewise implicit in any attempt to derive deficits in social ability from an impaired Theory of Mind, even if this derivation is in fact insufficient to account for the whole of the observed phenomena. It is to this putative inadequacy that we turn next.

Given our framework as modified here, there are several plausible hypotheses which could account for the behaviour of KH and other Asperger syndrome subjects: they could have a defective Social Interaction quasi-module, with repercussions for both Social Cognition and Social Affect; either of the latter two could be intact or impaired; they could have an intact Social Interaction Quasi-module but an impaired Theory of Mind (as perhaps in congenitally blind or profoundly deaf children); or both these could be intact but the interaction between them be impaired. Given the pattern of behaviour shown by KH, we think that the most plausible scenario is that his Theory of Mind is intact, his Social Cognition is largely unimpaired, and the locus of difficulty resides in an impaired Social Affect.

Asperger subjects characteristically have a problem with the imaginative projection of emotions (cf. Harris, 2000); that is, their problem is cognitive AND emotional. These conceptual domains meet in Theory of Mind and in Social Intelligence, in both of which the distinction between Self and Other is represented.

We have argued that Theory of Mind is quasi-modular in virtue both of its dependence on conceptual representation and its lack of informational encapsulation. We wish to make the same claim with regard to Social Intelligence. However, there is a putative distinction between the cognitive and affective domains. Whereas cognitive representations are necessarily propositional, affective representations are in part propositional and in part non-propositional. We take it that the propositional is directly related to the linguistically encoded in a way that the non-propositional is not. It is accordingly plausible that the distinction between the cognitive and the affective may be linguistically manifest in, for instance, the earlier development of basic emotions than beliefs. If this is the case, the correlation between verbal abilities and Theory of Mind in high-functioning autists and Asperger Syndrome subjects reflects the availability of a possible compensatory strategy. Crucially, language is a necessary but not sufficient condition for apparently successful performance in Theory of Mind tasks. That it is not sufficient is clear from Christopher's defective performance; probably attributable to his impaired 'central coherence', which entails that his excellent linguistic abilities cannot mask his Theory of Mind deficit completely.

By contrast, the partly non-propositional nature of an affective deficit means that it cannot be masked by such linguistically-based compensatory strategies. Given that the

development of emotions is deficient in all autistic children, including Asperger Syndrome subjects, the basic question is whether the development of emotions can proceed without a well-functioning Theory of Mind. We suspect that the answer to this is negative but with the following two conditions: the distinction between basic and derived emotions, and the distinction between recognizing basic emotions and processing this information further. More specifically, basic emotions - e.g. anger, disgust, surprise - may have transparent physiological counterparts and be thereby recognisable. For example, even autistic children can perceive an angry face and even experience some sympathy on the basis of their corresponding feelings. The difference between the normal and the autistic child is that the autistic child is incapable of inferring what the implications of this emotion would be for the other person, e.g. that he or she would like the object that caused this emotion to be removed from sight. For this a functioning Theory of Mind is required. Complex emotions by hypothesis require Theory of Mind representations and hence are not within the autistic child's cognitive range. Similarly, if Social intelligence is dependent on the appropriate attribution of concepts such as authority, Theory of Mind is again a prerequisite.

In sum, we suggest that KH's unusual pattern of performance both supports a (quasi‒)modular analysis of human cognition, and simultaneously casts doubt on the two simplest assumptions in the literature: that the aetiology of Asperger syndrome is attributable exclusively to either a deficit in Theory of Mind or to a deficit in Social Interaction.

References

Abell, F., F. Happé & U. Frith (2000) "Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development". *Cognitive Development* 15:1-16.

Anderson, M. (1992) *Intelligence and Development: A Cognitive Theory*. Oxford, Blackwell.

Baron-Cohen, S. (1995) *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge MA, MIT Press.

Baron-Cohen, S., A. Leslie & U. Frith (1985) "Does the autistic child have a Theory of Mind?" *Cognition* 21:37-46.

Bowerman, M. (1987) "The 'no negative evidence' problem. How do children avoid constructing an overly general grammar?". In J. Hawkins (ed) *Explaining Language Universals*. Blackwell.

Chomsky, N. (1975) *Reflections on Language*. New York, Pantheon.

Chomsky, N. (1980) *Rules and Representations*. Oxford, Blackwell.

Chomsky, N. (1984) *Modular Approaches to the Study of Mind*. San Diego, San Diego State University Press.

Chomsky, N. (1986) *Knowledge of Language: Its Nature, Origin and Use*. New York, Praeger.

Connell, J. (1998) Report by Mrs J. Connell on KH: 21.10.98.

Cormack, A. & N. Smith (2002) "Compositionality, Copy theory, and Control". *UCL Working Papers in Linguistics* 14:355-373.

de Sousa, R. (1987) *The Rationality of Emotion*. Cambridge MA, MIT Press.

Ekman, P. (1973) *Darwin and Facial Expression: A Century of Research in Review.* New York, Academic Press.

Fodor, J. (1975) *The Language of Thought.* New York, Crowell.

Fodor, J. (1983) *The Modularity of Mind.* Cambridge MA, MIT Press.

Frith, U. (1989) *Autism: Explaining the Enigma.* Oxford, Blackwell.

Frith, U. (1991) "Asperger and his syndrome". In U. Frith (ed) *Autism and Asperger Syndrome.* Cambridge, CUP; pp. 1-36.

Frith, U., F. Happé, & F. Siddons (1994) "Autism and theory of mind in everyday life". *Social Development*, 3, 108-124.

Garfield, J., C. Peterson & T. Perry (2001) "Social cognition, language acquisition and the development of the Theory of Mind". *Mind & Language* 16:494-541.

Greene, J. & J. Haidt (2002) "How (and where) does moral judgment work?". *Trends in Cognitive Sciences* 6:517-523.

Hall, K. (2000) *Asperger Syndrome, the Universe and Everything.* London, Jessica Kingsley Publishers.

Happé, F. (1991) "The autobiographical writings of three Asperger syndrome adults: problems of interpretation and implications for theory". In U. Frith (ed) *Autism and Asperger Syndrome.* Cambridge, CUP; pp. 207-242.

Happé, F. (1995) "The role of age and verbal ability in the theory of mind task performance of subjects with autism". *Child Development* 66:843-855.

Happé, F. (1999) "Autism: cognitive deficit or cognitive style?" *Trends in Cognitive Sciences* 3:216-222.

Harris, P.L. (1989) *Children and Emotion.* Oxford, Blackwell.

Harris, P. L. (2000) *The Work of the Imagination.* Oxford, Blackwell.

Heider, F. & M. Simmel (1944) "An experimental study of apparent behavior". *The American Journal of Psychology* 57:243-259.

Hermelin, B. (2001) *Bright Splinters of the Mind: A Personal Story of Research with Autistic Savants.* London, Jessica Kingsley Publishers.

Hermelin, B. & N. O'Connor (1970) *Psychological Experiments with Autistic Children.* Oxford, Pergamon Press.

Hobson, P. (2002) *The Cradle of Thought.* London, Macmillan.

Johnson Laird, P. (1983) *Mental Models.* Cambridge, CUP.

Kenny, A. (1963) *Action, Emotion, and Will.* New York, Routledge.

Klin, A. (2000) "Attributing social meaning to ambiguous visual stimuli in higher-functioning autism and Asperger syndrome: the Social Attribution Task". *Journal of Child Psychology and Psychiatry* 7:831-846.

Klin, A., W. Jones, R. Schultz, F. Volkmar & D. Cohen (2002) "Defining and quantifying the social phenotype in autism". *American Journal of Psychiatry* 159:895-908.

Lee, A. & P. Hobson (1998) "Self-description: A controlled study of children and adolescents with autism". *Journal of Child Psychology and Psychiatry* 39:00-00.

Leslie, A. (1987) "Pretense and representation: the origins of 'theory of mind'". *Psychological Review* 94:412-426.

Moore, P. (1998) Report by Dr P.Moore on KH, 31.5.98.

Perner, J., U. Frith, A. Leslie & S. Leekam (1989) "Exploration of the autistic child's Theory of Mind: Knowledge, belief and communication". *Child Development* 60:689-700.

Peterson, C. & M. Siegal (1995) "Deafness, conversation and the Theory of Mind". *Journal of Child Psychology, Child Psychiatry and Allied Disciplines* 36:459-474.

Peterson, C. & M. Siegal (2000) "Insights into Theory of Mind from deafness and autism". *Mind & Language* 15:123-145.

Premack, D. & G. Woodruff (1978) "Does the chimpanzee have a Theory of Mind?" *Behavioral and Brain Sciences* 1:515-526.

Rieffe, C. M.M. Terwogt (2000) "Deaf children's understanding of emotions: desire takes precedence". *Journal of Child Psychology and Psychiatry* 41:601-608.

Rutter, M. (1978) "Diagnosis and definition". In M. Rutter & E. Schopler (eds) *Autism: A Reappraisal of Concepts and Treatment.* New York, Plenum. Pp. 1-25.

Smith, N.V. (1999) *Chomsky: Ideas and Ideals*. Cambridge, CUP.

Smith, N.V. (2002) "Wonder". *Glot International* 6:55-57.

Smith, N.V. & I.-M. Tsimpli (1995) *The Mind of a Savant.* Oxford, Blackwell.

Sperber, D. & D. Wilson (1995) *Relevance: Communication and Cognition.* Oxford, Blackwell.

Tager-Flusberg, H. & K. Sullivan (1999) "Are children with Williams Syndrome spared in the Theory of Mind? Unpublished MS, University of Massachusetts.

Tantam, D. (1991) "Asperger syndrome in adulthood". In U. Frith (ed) *Autism and Asperger Syndrome.* Cambridge, CUP; pp. 147-183.

Tsimpli, I.-M & N.V. Smith (1998) "Modules and Quasi-modules: Language and theory of mind in a polyglot savant". *Learning and Individual Differences.* 10:193-215.

Wimmer, H. & J. Perner (1983) "Belief about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception". *Cognition* 13:103-128.

Wing, L. (1991) "The relationship between Asperger's syndrome and Kanner's autism". In U. Frith (ed) *Autism and Asperger Syndrome.* Cambridge, CUP; pp. 93-121.

Yirmiya, N., O. Erel, M. Shaked & D. Solomonica-Levi (1998) "Meta-analyses comparing theory of mind abilities of individuals with autism, individuals with mental retardation, and normally developing individuals. *Psychological Bulletin* 124:283-307.