Measuring Social-Cognitive Functions in Children with Somatotropic Axis Dysfunction

David Skuse  Kate Lawrence  Joey Tang
Behavioural and Brain Sciences Unit, Institute of Child Health, London, UK

Key Words
Brain  ·  Children  ·  Growth hormone receptors  ·  Insulin growth factor I receptors  ·  Social cognition

Abstract
Growth hormone (GH) and insulin-like growth factor I (IGF-I) are expressed in specific regions of the central nervous system during early human development. They may consequently influence aspects of cognition, or emotional and behavioural adjustment from childhood to adulthood, in conditions associated with abnormalities of the somatotropic axis. GH receptors are relatively common within hippocampal and perihippocampal regions that are primarily involved in declarative memory for facts and events. They are also located in structures (e.g., the putamen) that are involved in the processing of social perceptions. IGF-I receptors have been discovered in the amygdala and prefrontal cortex, which contribute to the neural circuits known as the ‘social brain’. The evaluation of emotional, social and behavioural adjustment among children who have deficiencies in GH or IGF-I functional integrity requires the objective assessment of their social-cognitive competence. We describe a computerized test battery, the Schedules for the Assessment of Social Intelligence (SASI), which has been shown to possess excellent psychometric properties in terms of reliability and validity. The SASI, which can be used by both children and adults, may provide new evidence for deficits and treatment effects of GH/IGF-I on emotional, behavioural and cognitive functions.

Introduction
Growth hormone (GH) and insulin-like growth factor I (IGF-I) have recently been implicated in terms of potential effects on the central nervous system (CNS), although it is not clear whether such effects are mediated directly by GH, by IGF-I or by the hypothalamic neuropeptide GH-releasing hormone (GHRH) [1]. GH binding sites have been reported in highest concentration in the choroid plexus, the hippocampus, putamen, thalamus, pituitary and hypothalamus. There is evidence that GH messenger RNA exists in the human brain, but this does not imply that the hormone is produced in the CNS. It could be acting via autocrine or paracrine mechanisms, and be actively transported across the blood-brain barrier [2]. IGF-I receptors have been discovered in the hippocampus and perihippocampal regions, as well as in the amygdala, cerebellum and frontal cortex [3]. We may deduce, from these patterns of receptor expression, that specific neurocognitive functions could be influenced by the somatotropic axis, bearing in mind that we do not know whether such effects – if any – are mediated directly or indi-
rectly. The highest concentration of IGF-I receptors in the human brain is found in hippocampal regions, although the precise localization is not as well understood as it is in rodent models such as the rat [4]. There are also very high numbers of GH receptors in the hippocampus, where they decline with age [5]. Hippocampal and perihippocampal areas are primarily involved in memory systems, especially declarative memory in verbal and visuospatial domains. Declarative memory refers to the everyday memories that we use when consciously recalling facts such as the capital of a country, or events such as the last time we went to the cinema [6]. Declarative memory allows remembered facts or events to be evaluated and compared, and it essentially represents the external world. We might, therefore, anticipate an association between declining activity of the somatotropic axis in the ageing brain and impairments in memory-related performance on standardized tasks. To a degree, this association has indeed been observed [5], but it has not been possible to demonstrate a causal link.

Children who have been deficient in GH could, theoretically, have long-term deficits in memory as a consequence of that deficiency during a critical period of brain growth and organization, even if later GH replacement therapy has been instituted. There is no unequivocal evidence to support this hypothesis, but that could be because the association is masked by potential confounding variables. These include coincident endocrinological deficits and indirect effects of hormonal deficiency on brain growth and maturation, which might be mediated via social and educational experiences. There may also have been a failure to measure a sufficiently wide spectrum of cognitive or other abilities that are linked to brain functions putatively influenced by GH/IGF-I [7]. Accordingly, the use of more subtle measures of dysfunction among individuals with a history of GH deficiency (GHD) is to be commended, such as event-related potentials that reflect attentional processing [8].

GH binding sites have been reported in the putamen, which is one component of a collection of nuclei sometimes known as the striatum (caudate, putamen and nucleus accumbens). Whereas the putamen is part of the basal ganglia, it has functions that are not directly associated with movement, one of which seems to be the interpretation of negative facial expressions. In this role it is associated with the caudate nucleus, as well as the amygdala and anterior insula [9]. IGF-I binding sites in the CNS also include the amygdala and the caudate nucleus [10].

Taken together, these observations raise the possibility that the functional integrity of the so-called ‘social brain’, which is specialized for the analysis of social perceptions (including facial expressions), could be influenced by the somatotropic axis. In cases where there are abnormalities in the functioning of that axis, particularly if they occur during early development, consequent deficits may be subtle, and require the employment of novel standardized measures of social cognition to detect them. There is a degree of specificity: for example, neural responses to facial expressions of increasing happiness are normally associated with increased activity in the putamen and in the bilateral fusiform gyri (inferior temporal regions that are specialized for the analysis of facial perceptions) [11]. Other facial expressions that have been associated with putamen activity include disgust, fear and sadness [12]. Involvement of the putamen (as well as other nuclei in the region) in disease processes such as Huntington’s chorea can lead to the loss of ability to discriminate between a limited range of facial expressions (in this example, disgust; see [13]).

In order to make a comprehensive evaluation of the cognitive consequences of GHD, it is necessary to assess a range of cognitive processes. Bearing in mind the distribution of GH/IGF-I receptors in the human brain, tests ought to include declarative memory, attention, executive functions (planning, flexibility) and processing speed. For measurement of many of these skills, the Cambridge Neuropsychological Testing Automated Battery (CAN-TAB) [14] is appropriate. This battery of computer-administered tasks complements standard intelligence quotient (IQ) tests, and published data on developmental trends in the executive functions it measures, from childhood to adulthood, are now available [14, 15]. However, there is also, in theory, a good case for measuring aspects of social cognition, although, until recently, there has been no comprehensive battery of social cognition tasks that are standardized in children or adults to measure objectively the functional integrity of the ‘social brain’ [16].

In order to answer the need for a standardized set of measures of social-cognitive competence, we developed the Schedules for the Assessment of Social Intelligence (SASI). The SASI comprises a set of tasks that are usually administered via computerized presentation, directly by a trained psychologist. We have also developed methods that can be administered over the Internet – a procedure that allows for efficient data collection from large numbers of individuals. The SASI is designed to be administered to children and adults, aged between 6 and
60+ years. It is possible to generate centile scores or z-scores for performance, from the standardized scores according to the individual’s age and sex.

**Components of the SASI**

The SASI comprises the following components:
- facial expression recognition task
- face recognition memory task
- gaze-monitoring task
- Theory of Mind task.

**Facial Expression Recognition Task**

In collaboration with Paul Ekman, we have automated the administration of a series of facial expressions [17] that are widely used in cognitive research. In this task, individuals are presented with images of different facial emotions and are required to select the correct expression from a list of possible answers. A total of 60 faces are presented. There are 10 examples (male and female, balanced) for each of the following emotions: fear, anger, disgust, sadness, happiness and surprise. This task was developed in view of the important role played by the interpretation of facial expressions in the functions of the ‘social brain’ [16].

**Face Recognition Memory Task**

We have computerized the administration of a Face Recognition Memory Test devised by Elizabeth Warrington [18]. There are two phases in this task: learning and testing. During the learning phase, individuals are presented with images of faces and are asked to judge whether they are ‘nice’ or ‘not nice’. In the testing phase, individuals are shown two faces in each trial and are required to choose the one that was presented in the learning phase. This task has been widely used in cognitive research and requires encoding, memorial storage and decoding of facial images over time. There is evidence to suggest that task performance is supported by amygdala function in females [19].

**Gaze-Monitoring Task**

We have devised a novel computerized task to measure accuracy in detection of gaze direction from static photographs, with eyes looking out directly from the photo but deviated between 5 and 20 degrees. In this task, individuals have to decide in each trial whether the person in the photograph is looking directly into his/her eyes or looking to the left/right. In people with autistic traits, an association between performance on this task, fear expression recognition and ability to respond appropriately to a Theory of Mind task has been observed [20].

**Theory of Mind Task**

The ability to detect contingency is an essential component of our understanding of events we can perceive visually. We normally distinguish unconsciously between mechanical and intentional social contingencies, based on simple movement patterns. Accordingly, neural networks that are activated in typical ‘Theory of Mind’ tasks have been shown to respond to simple animated cartoons of simple abstract symbols (such as triangles of different size), the movements of which imply:
- that they are living
- that they have intentions (such as coaxing or tricking).

The movement patterns are designed to evoke from the observer comments which imply that ‘mentalizing’ is occurring [21]. People with early onset damage to the amygdala do not use ‘mentalizing terms’ when describing these movements [22]. These individuals find it difficult to provide accurate responses that capture the appropriate and apparently intentional movements of the abstract cartoon objects engaged in activities such as ‘flirting’, ‘mocking’ or ‘teasing’ one another. The finding implies that damage to the neural integrity of the ‘social brain’ will be associated with loss of Theory of Mind skills. This task has been employed in studies of individuals on the autistic spectrum in both childhood [23] and adulthood [24, 25] who typically have compromised Theory of Mind abilities. Its psychometric properties have been studied in typically developing children aged between 6 and 18 years [26].

**Administration of the SASI**

The administration of the full SASI takes about 30 min. All scoring is automatic, with the exception of the Theory of Mind task, which currently requires an audio-recording of the observer’s spontaneous comments describing the ‘story’ evoked by the movement of the symbolic cartoon figures, which are then rated by trained staff. There is a brief (~5 min) practice session at the start of the test battery. The format of the task is designed to be interesting and engaging to both children and adults. Response times are automatically recorded. A response must be chosen before it is possible to move on to the next item. The sequence of task administration is discretion-
ary. Preliminary studies confirmed that the tasks could be interpreted correctly and could be completed by children as young as 6 years, but younger children often had difficulty when a verbal recognition/response was required.

**Testing the SASI**

*Test Populations*

The SASI has been tested, by different approaches according to age, in individuals recruited from a variety of sources. Children (n = 477; 246 boys and 231 girls, aged between 6.00 and 18.00 years; mean age 11.14 years) were individually tested at school by trained psychologists. An additional sample of 271 adults (97 men and 174 women; mean age 26.00 years) was also tested individually. Furthermore, the test battery was made available on the Internet, with a restricted URL, and was completed by 848 staff and students of University College London (UCL) (329 men and 519 women, aged between 17.90 and 79.60 years; mean age 29.96 years), bringing the total to 1,596 individuals.

*Ethical Permission*

Ethical permission for testing with the SASI in both the child and adult studies was obtained from the Research Ethics Committee of Great Ormond Street Hospital for Children and the Institute of Child Health, University College London, UK.

*Reliability*

The administration of the SASI test battery is completely automated. Data capture does not require the intervention of a tester and there is no possibility of error in the automated scoring by the computerized algorithms (authors’ unpublished data). For school children and clinical groups, explanation is necessary beforehand.

Several reliability measures were examined, namely external reliability, discriminant reliability and test-retest reliability [for descriptions of these measures, see 27].

*External Reliability*

External reliability of the SASI test battery was tested by comparing the mean scores of individuals who performed the tasks unsupervised over the Internet with those of individuals who were supervised by a tester. Each of the groups comprised 189 individuals (59 males and 130 females), matched by age and gender. The mean ages of the supervised and the unsupervised males were 27.25 years (range 16.00–45.00 years) and 27.35 years (range 16.47–44.68 years), respectively. The mean ages of the supervised and the unsupervised females were 26.68 years (range 16.00–41.00 years) and 26.71 years (range 16.23–41.20 years), respectively. All 189 matched case pairs performed the facial expression recognition task. Of the 189 matched case pairs, 160 performed the face recognition memory task and 48 performed the gaze-monitoring task. It was not possible to administer the Theory of Mind task over the Internet, therefore external reliability was not established.

*Discriminant Reliability*

Discriminant reliability of the facial expression recognition task, the face recognition memory task and the gaze-monitoring task was established by comparing mean scores from the lowest 10% and the highest 10% of the two test groups mentioned above for each of the administration methods (supervised and unsupervised). The purpose of discriminant reliability is to evaluate the ability of the test to discriminate significantly between individuals with good and poor performance. Clearly, if the task is not associated with a substantial discrimination, in relatively large proportions of individuals, there will be little or no difference between the mean performance of those in the lowest scoring centiles (conventionally taken to be the lowest 10% of the sample) and those in the highest scoring centiles (conventionally taken to be the highest 10% of the sample).

*Test-Retest Reliability*

Test-retest reliability was established by administering the SASI test battery on serial occasions. The test battery does not provide feedback to the individuals taking the test. Consequently, it is not possible for them to tell whether their responses are correct or not. Test-retest reliability was established by re-contacting those who had completed the web-based tasks on an earlier occasion. Individuals were invited to perform the tasks again in order for test-retest reliability to be established. Of the original sample of 848 staff and students of UCL, 159 (53 men, 106 women; age 17.90–79.60 years; mean age [standard deviation, SD] 32.87 [9.79] years) performed one or more of the three tasks again (89 performed the facial expression recognition task, 120 performed the gaze-monitoring task and 77 performed the face recognition memory task). The mean time elapsed between the two tests was 20 months (range 3–28 months).
Inter-Rater Reliability
The Theory of Mind task was scored independently by two researchers. The inter-rater reliability of the scores was established using intraclass correlations (ICCs).

Validity
Content Validity
Content validity of the SASI was established on the basis that it measures a range of cognitive domains that are closely associated with activity within the ‘social brain’ [28] and with impairments in the functional integrity of neural activity in autistic individuals [29]. The range of social-cognitive tasks that are relevant to these aims includes measures of facial expression recognition, facial recognition memory, gaze direction detection and Theory of Mind.

Known-Group Criterion Validity
Known-group criterion validity of the facial expression recognition task, the face recognition memory task, the gaze-monitoring task and the Theory of Mind task was assessed by comparing the mean scores of patients with clinically diagnosed autism spectrum disorder with those of normal controls, matched by age, gender and IQ score. As discussed previously, individuals with autistic conditions characteristically possess deficits in social-cognitive competence in a range of domains. Each group selected for this analysis comprised 53 children (46 boys and 7 girls). The mean age of the clinical group was 9.46 years (range 6.13–20.93 years); the mean age of the control group was 9.23 years (range 6.00–17.00 years). The mean Wechsler Abbreviated Scale of Intelligence (WASI) [30] full-scale IQ scores were 98.43 and 98.60 for the clinical and the control groups, respectively, a non-significant difference indicating close matching had been achieved.

Results
Reliability

External Reliability (supervised by human researcher vs. unsupervised over the Internet)

Facial Expression Recognition Task. One-way analysis of variance (ANOVA) performed on mean scores of the facial recognition task revealed no significant differences between individuals tested under supervision and those who performed the same task unsupervised on the Internet, with the exception of recognition of the fearful expression, on which the unsupervised group performed better than the supervised group (table 1).

Face Recognition Memory Task. One-way ANOVA revealed a non-significant difference in mean scores of the face recognition memory task between the supervised and unsupervised populations (42.89 [4.00] and 43.70 [3.88], respectively; $F(1,319)<1$, not significant).

Gaze-Monitoring Task. One-way ANOVA of mean scores of the gaze-monitoring task revealed a non-significant difference between the supervised and unsupervised groups (60.05 [8.86] and 60.76 [9.51], respectively; $F(1,94)<1$, n.s.).

Discriminant Reliability

Facial Expression Recognition Task. Mean scores from the lowest 10% and the highest 10% of the test populations were compared in a one-way ANOVA for each facial expression, except happy, for each of the administration methods (supervised and unsupervised). A ceiling effect in the happy expression recognition renders the comparison unfeasible. For all the other expressions, the low and high scorers differed significantly (all $F >400, p<0.001$).

Face Recognition Memory Task. Mean scores from the lowest 10% and the highest 10% of the test populations were compared in a one-way ANOVA for each facial expression. The low and high scorers differed significantly under both supervised and unsupervised conditions ($F(1,40)=618.76, p<0.001$ and $F(1,41)=716.41, p<0.001$, respectively).

Gaze-Monitoring Task. Mean scores from the lowest 10% and the highest 10% of the test populations were compared in a one-way ANOVA for each of the administration methods. The low and high scorers differed significantly under both supervised and unsupervised conditions ($F(1,40)=618.76, p<0.001$ and $F(1,41)=716.41, p<0.001$, respectively).

<table>
<thead>
<tr>
<th>Facial expression</th>
<th>Administration method</th>
<th>$F_{(1,376)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Happy</td>
<td>supervised</td>
<td></td>
</tr>
<tr>
<td></td>
<td>unsupervised</td>
<td></td>
</tr>
<tr>
<td></td>
<td>over the Internet</td>
<td></td>
</tr>
<tr>
<td>Happy</td>
<td>9.96 (0.19)</td>
<td>9.92 (0.30)</td>
</tr>
<tr>
<td>Surprised</td>
<td>8.96 (1.21)</td>
<td>9.01 (1.11)</td>
</tr>
<tr>
<td>Fearful</td>
<td>7.76 (2.13)</td>
<td>8.46 (1.56)</td>
</tr>
<tr>
<td>Sad</td>
<td>8.43 (1.50)</td>
<td>8.27 (1.29)</td>
</tr>
<tr>
<td>Disgusted</td>
<td>8.61 (1.35)</td>
<td>8.54 (1.39)</td>
</tr>
<tr>
<td>Angry</td>
<td>8.63 (1.20)</td>
<td>8.70 (1.35)</td>
</tr>
</tbody>
</table>

* $p<0.001$.
nificantly under both supervised and unsupervised conditions ($F_{(1,10)} = 134.89, p < 0.001$ and $F_{(1,11)} = 205.62, p < 0.001$, respectively).

Test-Retest Reliability

Facial Expression Recognition Task. Paired-samples $t$ tests were performed on mean scores of the facial expression recognition task comparing test and retest scores of individuals who performed the task twice, revealing a non-significant difference in mean score for each of the six facial expressions (table 2).

Face Recognition Memory Task. A paired-samples $t$ test performed on the mean scores of the face recognition memory task comparing individual’s test and retest scores revealed a significant difference (improvement) in mean score for the second administration of the test ($t_{(76)} = 6.29, p < 0.001$).

Gaze-Monitoring Task. A paired-samples $t$ test performed on the mean scores of the gaze-monitoring task comparing test and retest scores of individuals revealed a non-significant difference in mean score ($t_{(119)} = 1.14$, not significant).

Inter-Rater Reliability

Theory of Mind Task. ICCs were performed on each of the six measures of the Theory of Mind task. All measures showed significant correlations between the two raters in both the clinic and the control group (tables 3 and 4, respectively).

Table 2. Mean scores (and SDs) for the facial expression recognition task in the determination of test-retest reliability

<table>
<thead>
<tr>
<th>Facial expression</th>
<th>Test score</th>
<th>Retest score</th>
<th>$t_{(88)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Happy</td>
<td>9.78 (1.12)</td>
<td>9.64 (1.56)</td>
<td>1.05</td>
</tr>
<tr>
<td>Surprised</td>
<td>8.88 (1.56)</td>
<td>8.60 (1.93)</td>
<td>1.77</td>
</tr>
<tr>
<td>Fearful</td>
<td>8.64 (1.57)</td>
<td>8.48 (1.99)</td>
<td>0.66</td>
</tr>
<tr>
<td>Sad</td>
<td>7.99 (1.63)</td>
<td>8.21 (1.83)</td>
<td>–1.29</td>
</tr>
<tr>
<td>Disgusted</td>
<td>8.64 (1.78)</td>
<td>8.79 (1.79)</td>
<td>–0.87</td>
</tr>
<tr>
<td>Angry</td>
<td>8.69 (1.59)</td>
<td>8.85 (1.53)</td>
<td>–0.81</td>
</tr>
</tbody>
</table>

Table 3. Mean scores (and SDs) for the Theory of Mind task in the determination of inter-rater reliability for the clinic group

| Goal-directed: intentionality | 8.63 (2.21) | 9.22 (2.57) | 0.654* |
| Goal-directed: appropriateness | 4.50 (1.80) | 4.88 (2.45) | 0.852** |
| Goal-directed: length | 11.84 (4.27) | 11.81 (4.11) | 0.983** |
| Theory of Mind: intentionality | 9.94 (2.76) | 10.16 (3.25) | 0.761** |
| Theory of Mind: appropriateness | 2.66 (1.96) | 2.84 (2.08) | 0.827** |
| Theory of Mind: length | 13.56 (3.66) | 13.44 (3.53) | 0.988** |

Table 4. Mean scores (and SDs) for the Theory of Mind task in the determination of inter-rater reliability for the control group

<table>
<thead>
<tr>
<th>Control group</th>
<th>Mean</th>
<th>ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goal-directed: intentionality</td>
<td>11.19 (1.48)</td>
<td>11.16 (1.53)</td>
</tr>
<tr>
<td>Goal-directed: appropriateness</td>
<td>5.92 (1.25)</td>
<td>5.81 (1.31)</td>
</tr>
<tr>
<td>Goal-directed: length</td>
<td>14.00 (2.34)</td>
<td>14.06 (2.31)</td>
</tr>
<tr>
<td>Theory of Mind: intentionality</td>
<td>11.89 (2.07)</td>
<td>12.22 (2.14)</td>
</tr>
<tr>
<td>Theory of Mind: appropriateness</td>
<td>3.91 (1.33)</td>
<td>3.95 (1.30)</td>
</tr>
<tr>
<td>Theory of Mind: length</td>
<td>14.84 (1.95)</td>
<td>14.88 (2.00)</td>
</tr>
</tbody>
</table>

** $p < 0.001$.  
* $p < 0.005$; ** $p < 0.001$.  

Horm Res 2005;64(suppl 3):73–82

Skuse/Lawrence/Tang
Validity

Content Validity

Both the facial expression recognition task and the face recognition memory tasks are computerized versions of well-established tasks [17, 18]. The gaze-monitoring task was devised with the aim of discriminating individuals who appeared to have difficulty in detecting direct gaze because of neurodevelopmental anomalies [20]. The Theory of Mind task has been employed in previous studies and discriminates individuals with impaired social cognitive skills due to both acquired brain damage and congenital neurodevelopmental anomalies [23–26]. All these tasks have proved to be sensitive measures for their purposes and are essential for assessing social-cognitive abilities.

Known-Group Criterion Validity

Facial Expression Recognition Task. Of the 53 matched case pairs of children tested, 5 were removed due to missing values. One-way ANOVAs revealed no significant differences in the mean scores of the two groups, with the exception of recognition of the fearful expression, in which the clinical group performed more poorly than the control group, as predicted (table 5).

Face Recognition Memory Task. Of the 53 matched case pairs of children tested, 4 were removed due to missing values. A one-way ANOVA conducted on mean scores revealed a small but significant difference between the clinical and control groups (32.87 [7.07] and 35.63 [5.97], respectively; $F_{(1,96)} = 4.37, p < 0.050$), with a poorer performance in the clinical group.

Gaze-Monitoring Task. Of the 53 matched case pairs, 2 were removed due to missing values. A one-way ANOVA conducted on mean scores revealed a significant difference between the clinical and control groups (44.79 [10.27] and 48.99 [11.00], respectively; $F_{(1,100)} = 3.96, p < 0.050$). As predicted, performance was poorer in the clinical group.

Theory of Mind Task. Of the 53 matched case pairs, 32 performed the Theory of Mind Task. A $3 \times 2$ ANOVA was performed on the average scores between the two raters of each of the two types of measures: Goal-Directedness and Theory of Mind. The three measures for each of them were: intentionality, appropriateness and length. The between-subjects factor was group (clinic and control). Whenever the sphericity assumption was violated, the degrees of freedom were corrected with the Greenhouse-Geisser epsilon.

For goal-directedness, the ANOVA revealed a significant main effect of measure ($F_{(1,88)} = 206.90, p < 0.001$) and a significant main effect of group ($F_{(1,62)} = 21.19, p < 0.001$; the clinic group [mean = 8.48] scored significantly lower than the control group [mean = 10.36]). The two factors did not interact significantly ($F_{(2,124)} = 1.27, n.s.$). In other words, the clinic group scored significantly lower than the control group on all three measures for goal-directedness (fig. 1).

For Theory of Mind, the ANOVA revealed a significant main effect of measure ($F_{(2,97)} = 557.75, p < 0.001$) and a significant main effect of group ($F_{(1,62)} = 11.63, p < 0.001$; the clinic group [mean = 8.77] scored significantly lower than the control group [mean = 10.28]). The two factors did not interact significantly ($F_{(2,124)} < 1, n.s.$). In other words, the clinic group scored significantly lower than the control group on all three measures for Theory of Mind (fig. 1).

Discussion

Reliability and Validity of the SASI

Facial Expression Recognition Task

Performance of this task was compared under supervised and unsupervised (i.e. web-based testing) conditions. For the six facial expressions presented in the test, the mean scores of the two test populations differed significantly only with regard to the fearful faces, with performance being better in unsupervised individuals than in supervised individuals. The absolute difference (approximately 7% of total possible score) was small. Such a difference can be explained by the fact that fear is a more difficult emotion to interpret than the other emotions included in the test and is therefore more susceptible to a
larger variance in scores, reflected by the large SDs in both test populations (the largest recorded for all the facial expressions tested; see table 1). Mean score for this facial expression was the lowest in the supervised group and the second lowest in the unsupervised group. We have previously assessed facial expression recognition abilities in typically developing children (aged 6–16 years) and found that fearful (and disgusted) expressions were significantly less well recognized than other emotions (authors’ unpublished data).

Discriminant reliability was highly significant for all the facial expressions in this task, suggesting that it is suitable for populations that have subtle deficits in facial expression recognition. This is further supported by the known-group criterion validity analyses, in which a clinical sample of autistic children showed a significant deficit in recognizing the fearful facial expressions compared with a matched control group.

Test-retest analysis indicated that an individual’s performance on the face recognition memory task of the SASI did not improve significantly over time.

Face Recognition Memory Task
There was no significant difference between mean scores of supervised and unsupervised individuals for the face recognition memory task, suggesting that human supervision does not facilitate or impede performance of this task in any way. Since this was a between-group design, the non-significant difference in mean score indicates little individual difference in this task across different groups taken from the normal population.

Discriminant reliability was highly significant for this task, suggesting that the task is suitable for populations that have subtle deficits in face recognition memory. This is further supported by the known-group criterion validity analysis in which autistic children showed a significant deficit in face recognition memory compared with a matched control group.

Test-retest analysis indicated that an individual’s performance on the face recognition memory test improved significantly over time. Such a finding is not surprising since people generally improve in memory tasks with practice.

Gaze-Monitoring Task
There was no significant difference in the mean scores of supervised and unsupervised individuals for the gaze-monitoring task, suggesting that human supervision does not facilitate or impede performance of this task in any way. Since this was a between-group design, the non-sig-
significant difference in mean score indicates little individual difference in this task across different groups taken from the normal population.

Discriminant reliability was highly significant for this task, suggesting that the task is suitable for populations that have subtle deficits in gaze direction detection. This is further supported by the known-group criterion validity analysis, in which autistic children showed a significant deficit in gaze direction detection compared with a matched control group.

Test-retest analysis indicated that an individual’s performance on this task of the SASI did not improve significantly over time.

Theory of Mind Task
The significant ICCs indicated a good concordance between the scores from two independent raters on the Theory of Mind task. Furthermore, the known-group criterion validity analysis was highly significant for this task, in which children with autism showed a significant deficit in attributing mental states compared with a matched control group, suggesting that the task is suitable for populations who lack a Theory of Mind.

Summary of the Reliability and Validity of the SASI
The reliability and validity of each of the SASI component tasks have been assessed exhaustively. All four tasks have been shown to be reliable and valid, and the battery has the advantage that it is applicable for administration to adults as well as children from the age of 6 years upwards in an identical form.

Potential Use of the SASI
The SASI has proved to be a sensitive battery for assessing social-cognitive skills and we advocate its use in populations that have subtle deficits in social cognition – for example, those who have impaired functioning of, or impaired mechanisms related to, the functional integrity of the ‘social brain’.

Conclusion
There is a large body of research that indicates GH may exert a direct or indirect impact on the developing brain. As a consequence, deficiencies and functional abnormalities could affect cognitive functions such as attention and memory [3, 30]. GH may influence appetite, energy, mood and well-being, in addition to its effects on cognitive processes [31]. In view of the anatomical distribution of GH/IGF-I receptors in the brain, we predict that individuals who are deficient in the production of GH or IGF-I, who have neurosecretory dysfunction of the somatotropic axis or who lack appropriate mechanisms for translating hormonal activity into the brain could lack the normal integrated functioning of subcortical nuclei and cortical regions that are collectively known as the ‘social brain’. If this prediction is correct, we might expect deficiencies of GH in childhood to result in subtle abnormalities in the development of social perception, and in behavioural difficulties in relation to social interaction skills. We would expect isolated GHD in later life (e.g. among the elderly) to also be associated with deficits in social cognition. However, it is important to bear in mind that such problems in cognitive processing are not necessarily reflected in terms of real-world behaviour, in either children or adults, because the latter is influenced by many other factors including both biological and social mechanisms.

Acknowledgements
We are grateful to John Morris and Lynette Spence, who devised and established the set of tasks used in the SASI. We are indebted to Paul Ekman for his cooperation. Simon Gornall wrote the software that enabled the computerization of these tasks. Others who assisted in the gathering of data relevant to the reliability and validity studies include Ben Corden, Chris Chaddock, Will Mandy and David Spektor. The development work was supported by a grant from the National Alliance for Autism Research (USA).
References