NOVA

University of Newcastle Research Online

nova.newcastle.edu.au

McDonald, Skye, Fisher, Alana, Togher, Leanne, Tate, Robyn, Rushby, Jacqueline, English, Therese, Kelly, Michelle, Mathersul, Danielle, Froreich, Francesca, Francis, Heather. "Adolescent performance on The Awareness of Social Inference Test: TASIT" Published in *Brain Impairment*, Vol. 1, Issue 1, Pages 3-18, (2015).

Available from: <http://dx.doi.org/10.1017/BrImp.2015.7>

This article has been published in a revised form in *Brain Impairment* <http://dx.doi.org/10.1017/BrImp.2015.7>. This version is free to view and download for private research and study only. Not for re-distribution, re-sale or use in derivative works. © Cambridge University Press.

Accessed from: <http://hdl.handle.net/1959.13/1324046>

Adolescent performance on The Awareness of Social Inference Test: TASIT

McDonald, Skye¹

Fisher, Alana¹

Togher, Leanne²

Tate, Robyn³

Rushby, Jacqueline¹

English, Therese¹

Kelly, Michelle¹

Mathersul, Danielle¹

Frreich, Francesca¹

Francis, Heather¹

¹School of Psychology, University of New South Wales

²Faculty of Health Sciences, the University of Sydney

³Faculty of Medicine, the University of Sydney Correspondence: Professor Skye

McDonald, School of Psychology, University of NSW, Sydney, AUSTRALIA, 2052.

ABSTRACT

Primary objective: Social cognition underlies social skills and can be disrupted in numerous developmental and acquired brain disorders during childhood and adolescence. Despite this, there are few tools to assess social cognition clinically in this age group. This study examined adolescent performance on The Awareness of Social Inference Test (TASIT), a valid measure of social cognition in adults. Design: Cross sectional design examining performance on Parts 1, 2 and 3 of TASIT (and alternate forms) in Australian girls and boys with varying levels of English familiarity. Methods: 665 school children from private and government schools were administered TASIT subtests. Of these, 464 students aged 13-15 were selected to provide normative data. Scores from a further 97 provided information about the effects of lack of English familiarity. Results: The two Forms of TASIT were statistically equivalent for two of the three parts. Adolescents performed lower than adults although the differences were not large. Some incremental effects were seen for chronological age. Gender effects were apparent on all subtests. Lack of English familiarity (i.e. English not spoken at home) reduced scores a further 6-13% relative to high English proficiency. Conclusions: TASIT appears suitable for adolescents. Norms are best aggregated across ages in adolescence and stratified according to gender.

INTRODUCTION

It is well established that brain disorders affect cognition in pediatric populations. The tradition of performance based assessment to elucidate disorders in attention, language, memory, perception and praxis is firmly in place. This is not the case in the field of social cognition. Indeed, it is only relatively recently that social cognition has been considered an area of competence to be assessed in its own right. While it has long been recognised that many brain disorders, both developmental (e.g., autism spectrum disorders) and acquired (e.g., traumatic injuries, stroke) can impede the normal development of social skills, the assessment of such difficulties was either subsumed under assessment of generic cognitive abilities or overlooked entirely. However, increasingly, researchers have identified social cognition as a distinct area of skill that is crucial to the development of normal social skills during adolescence and which can be differentially disrupted by brain disorders. It is therefore one that needs careful clinical examination.

Social cognition refers to the ability to attend to and identify social cues (such as facial and vocal expressions, gestures and ‘body language’) and to interpret these, in context, to make judgments concerning the emotions, beliefs and attitudes of others and the intentions behind their actions, an ability known as Theory of Mind (ToM). Such abilities are critical not only to understand social behaviour, but also to understand everyday conversation. Communication is a complex phenomenon. There are usually layers of meaning in the simplest conversational exchange, where it is not only what is said, but *how* it is said, that conveys messages between speakers about the nature of the social interaction, be it a formal encounter, social banter, diplomatic side stepping and so on. In particular, speakers often convey messages indirectly, such as when hinting, or being sarcastic or ironic in order to address interpersonal and social goals. The meaning behind such exchanges cannot be understood unless the listener can infer the emotional and mental state of the speaker.

Emotion perception is a skill that appears to develop progressively through childhood and early adolescence (Kolb, Wilson, & Taylor, 1992; McClure, 2000). Similarly, the ability to make ToM judgments develops incrementally, linked to growing sophistication in the understanding of indirect language (Winner & Leekam, 1991). Such maturation intensifies during puberty where social cognition development accompanies the growing preoccupation with social acceptance by peers (Blakemore & Choudhury, 2006). There are significant gender differences in social cognitive development. Girls demonstrate a consistent advantage over boys for emotion perception from infancy to adolescence (Lee et al., 2013; McClure, 2000). Young children and pre-pubescent girls also appear to outperform boys on tasks requiring mentalising ability (Bosacki, 2000; Hughes & Dunn, 1998) as do adult women compared to men (Baron-Cohen, Jolliffe, Mortimore, & Robertson, 1997).

Social cognition is disrupted in a wide variety of clinical conditions. Indeed, marked impairment in social interaction is a core characteristic of Autism Spectrum Disorder (ASD: APA, 2000). Consistent with this, children, adolescents and young adults with ASDs and the related disorder of alexithymia generally have poor social cognition (e.g., Baron-Cohen, 1989; Couture et al., 2010; Di Martino et al., 2009; Jellesma, Rieffe, Meerum Terwogt & Westenberg, 2009). Children and adolescents with acquired brain injuries also frequently experience deficits in socio-emotional functioning, peer acceptance and specifically, social cognition (Beauchamp & Anderson, 2010; Yeates et al., 2007). For example, adolescents with traumatic brain injuries are often poor at recognising emotions (especially facial affect), a deficit that is not readily accounted for by cognitive function (Tonks, Williams, Frampton, Yates, & Slater, 2007). They also often do not realise when social communication conventions (e.g., bragging) have been violated and do not recognise implied meanings like sarcasm (Turkstra, McDonald, & DePompei, 2001). Poor social cognition is also associated with less talk about thoughts and feelings in conversations with peers (Stronach & Turkstra,

2008). Individuals with schizophrenia, likewise, experience deficits in social cognition including poor understanding of emotional expressions (Edwards, Jackson, & Pattison, 2002), ToM (Brune, 2005; S. Couture, M., Penn, & Roberts, 2006) and indirect language (Sparks, McDonald, Lino, O'Donnell, & Green, 2010).

Despite growing interest in the assessment of social cognition in children and adolescents, there are few clinical instruments available, especially in the area of emotion perception. One experimental tool was designed by Dennis and Barnes (Dennis, Barnes, Wilkinson, & Humphreys, 1998) to assess ability to identify real and feigned emotions in schematic line drawings of faces. The real and feigned emotions test is performed poorly by children with ASDs and brain injury in some research (Dennis et al., 2013; Dennis et al., 1998; Dennis, Lockyer, & Lazenby, 2000) but not all (Anderson et al., 2013).

There are more tests available to assess ToM either directly or indirectly in children and adolescents. These usually rely upon responses to written and spoken texts and pictures as a means to tap understanding of the feelings and beliefs of the story characters. Experimental tasks include the Strange Stories task (Happe, 1994), the Faux Pas test (Baron-Cohen, O'Riordan, Stone, Jones, & Plaisted, 1999), the social problem solving measure (Pettit, Dodge, & Brown, 1988) and the Interpersonal Negative Strategies task (Yeates, Schultz, & Selman, 1991). A similar approach is found in some clinical instruments, e.g. the Pragmatic judgement subtest of the 'Comprehensive Assessment of Spoken Language' test (Carrow-Woolfolk, 1999), the Social Language Development Test (Adolescent form) (Bowers, Huisinigh, & LoGiudice, 2010) and the making inferences subtest of the Test of Language Competence- Expanded Edition (Anderson et al., 2013; Wiig & Secord, 1989).

As social cognition is, by definition, an ability to interpret social cues, it is reasonable to question the ecological validity of text or pictorially based approaches to its assessment.

One promising approach is the use of virtual reality programs to represent realistic everyday scenarios within which to assess social understanding (Hanten et al., 2011). An alternative is to use conventional visual or audiovisual media to assess social cognition in adolescents. There are relatively few such instruments specifically developed for adolescents. One of the few is the Video Social Inference Test (Turkstra, 2008) which was designed to measure social cognition in adolescents with brain injury. This test uses adolescent actors and has been shown to be sensitive to brain injury but awaits further research as to normative data and standardisation.

As an alternative, some tests that have been developed to assess adult social cognition use audiovisual stimuli and may have applicability to younger people. For example the Florida Affect Battery (FAB; Bowers, Blonder, & Heilman, 1991) is a comprehensive test of emotion perception which examines both (photographed) facial expressions and tone of voice in ten subtests involving face and sound affect discrimination, naming, comprehension and selection, matching, or conflict detection (i.e. whether semantic content and prosody conflict) and cross-modal matching. The FAB has been tested with children and has shown to be sensitive to maturation (Tonks, Williams, Frampton, Yates, & Slater, 2007) and also to the effects of brain injury (Tonks et al., 2007). Its validity in predicting social skills in children and adolescents has not been examined. The Reading the Mind in the Eyes test (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001; Baron-Cohen, Wheelwright, Spong, Scahill, & Lawson, 2001) requires participants to choose a term to match mental state based on the eye region of the face. While performance on this task improves with developing maturity (Tonks et al., 2007) and is sensitive to clinical impairments (Baron-Cohen, Wheelwright, et al., 2001), its construct validity as a measure of ToM has been questioned (Gregory et al., 2002; Johnston, Miles, & McKinlay, 2008).

Another option is the The Awareness of Social Inference Test (TASIT) (McDonald,

Flanagan, & Rollins, 2011) which is an audiovisual based tool that was designed to provide an ecologically valid assessment of social cognition for adults with traumatic brain injury (McDonald, Flanagan, Rollins, & Kinch, 2003). TASIT has three parts (1) Emotion Evaluation Test (EET), (2) Test of Social Inference (Minimal) (SI-M) and (3) Test of Social Inference (Enriched) (SI-E) each with alternate forms for re-testing. Where Part 1 taps understanding of emotional expression from integrated dynamic facial, vocal and gestural cues, Parts 2 and 3 assess recognition of more subtle emotional states as well as theory of mind and understanding of pragmatic inference. Part 2 is 'minimal' because there are no additional cues apart from the dialogue and demeanor of the speakers. Part 3 is 'enriched' due to the provision of a prologue or visual edit that gives extra information regarding the true state of affairs.

TASIT has proven to be sensitive to deficits in social cognition in adults with traumatic brain injuries (McDonald & Flanagan, 2004; McDonald, Flanagan, Martin, & Saunders, 2004; McDonald et al., 2003; McDonald & Saunders, 2005), and predictive of real world social functioning (McDonald et al., 2004). It is also sensitive to social cognition deficits in individuals with schizophrenia (Kern et al., 2009; Kosmidis, Eleni, P., Maria, & Panayiotis, 2008; Sparks et al., 2010), ASDs (Mathersul, McDonald, & Rushby, 2013), fronto-temporal dementia (Kipps, Nestor, Acosta-Cabronero, Arnold, & Hodges, 2009; Rankin et al., 2009) and right hemisphere lesions (Blake, 2009; Fournier, Calverley, Wagner, Poock, & Crossley, 2008).

TASIT has also been used to examine deficits in social cognition in adolescents with TBI (McDonald et al., 2013), where it proved to be sensitive to deficits in social communication as described by both the adolescent with brain injury and their parents. It was also sensitive to IQ and injury severity in the clinical group. Thus, despite the fact that TASIT was designed for adults and uses scenes of adult actors in everyday encounters, it

appears to be suitable as an instrument to measure social cognition in adolescents. The purpose of the following study was to provide more comprehensive normative data for TASIT in typically developing adolescents than available at the time of publishing TASIT manual and to examine the psychometric validity of TASIT for adolescents by comparing performance on Form A and B of the test. It was also designed to provide information regarding the construct validity of TASIT for adolescents, by examining performance patterns with increasing age from 13-15 years and according to gender. Specifically, it was hypothesised that TASIT performance would improve with chronological age across these early adolescent years, although possibly not meet adult levels of performance. It was also hypothesised that girls would out-perform Boys. Finally, given the language demands of TASIT, the effects of familiarity with English were explored.

METHODS

Participants

A total of 665 students were recruited from six high schools in the Sydney metropolitan area between 2002 and 2014. The high schools involved were: two single-sex government schools, one coeducational government school, and three single-sex Catholic schools. Participants were excluded if they or their parents reported any of the following: a history of brain injury or other neurological condition; learning or intellectual disability; developmental delay; an Autism Spectrum Disorder or Asperger's disorder diagnosis; diagnosed psychiatric conditions; uncorrected vision and/or hearing impairments.

Materials

The Awareness of Social Inference Test (TASIT; McDonald et al., 2003) is an ecologically valid, and clinically sensitive and valid, measure of simple emotion perception and complex social cognition with established adult norms in healthy and various clinical populations (e.g., traumatic brain injury, schizophrenia, and dementias). Participants are required to integrate cues from various sources (e.g., facial expressions, prosody, gesture, and social context) to interpret the emotions, beliefs and intentions (i.e., ToM) and conversational meanings imparted by target characters in videotaped conversational interactions. All scripts and response probes in TASIT use simple language (see Appendix A for example script) suitable for school children. Specifically, according to two well-recognised metrics of readability (Kincaid, Fishburne, Rogers & Chissom, 1975), TASIT scripts were rated 97.7/100 (for Flesch Reading Ease) indicating that they could be understood by an average 11 year old native English speaker and 1.2 (using the Flesch–Kincaid Grade Level), which indicates that a native English speaker would need 1.2 years of formal education on average to understand the text. There are two equivalent forms (i.e. versions), A and B, which each consist of three parts:

Part 1: *The Emotion Evaluation Test* comprises 28 short video clips portraying one of six basic emotions (happy, sad, fear, disgust, surprise and anger) or no emotion (neutral); the scripts are all semantically neutral and are enacted by method actors who invoke the genuine emotion in themselves as part of the scene. Participants circle the emotion (from seven randomly-ordered choices) that best corresponds to what the character is feeling.

Part 2: *Social Inference – Minimal* comprises fifteen video clips depicting sincere and sarcastic interactions between two actors; the sarcasm items are further divided into *simple sarcasm*, whereby a neutral dialogue is made sarcastic by the demeanour of the speaker and *paradoxical sarcasm* whereby the script (also enacted sarcastically) makes no sense unless it

is understood to be sarcastic, e.g., Gary says “*Hard at it I see*” and Michael responds “*I am having a break!*”.

Part 3: *Social Inference – Enriched* comprises sixteen scenes of similar type scripts to that in Part 2 but which also contain a visual edit or prologue that conveys additional information about the true state of affairs before or after the dialogue of interest. Eight of the scenes depict sarcastic exchanges and the other eight depict one speaker lying to another, usually in order to be kind. Comprehension in both Parts 2 and 3 is tapped via forced-choice responses (yes/no/don’t know) to 4 response probes per item. These ask questions about what the speakers are: (1) feeling; i.e. their emotional states; (2) thinking; i.e. their beliefs and knowledge about the situation (*first order ToM*); (3) doing; i.e. what they intend their conversational partner to think or feel as a result of the interaction: (*second order ToM*); and (4) saying, i.e., whether they want the literal or non-literal meaning of their message to be believed.

In addition to TASIT, all participants’ parent(s) also completed a screening questionnaire about their child/ren, which asked for information about their child/ren’s age and level of education completed, neurological and mental health history, sensory problems, language and ethnic background. The questionnaire was administered in order to characterise the sample demographics, identify ineligible participants or potential confounding variables.

Procedure

Ethics approval was sought and obtained to conduct research in New South Wales government and Catholic schools. Researchers from the University of New South Wales (UNSW) contacted principals from ethics-approved schools via letter and subsequently by telephone call to introduce the research and discuss their school’s willingness to participant. Mutually convenient testing times were then arranged with principals agreeing to participate.

Information sheets and consent forms were distributed to students by their class teacher; students were asked to complete these and give a copy to their parents to review and sign. Parents also completed a screening questionnaire about their child/ren. Only students who returned signed copies of the consent forms- theirs and their parents'- were allowed to participate.

All testing was done in schools by visiting UNSW researchers who were experienced with administering TASIT. Participants were tested in their class groups during normal lesson times. Depending on timetabling and staff availability, testing lasted between one and two class periods (approximately 40 to 50 minutes each). Thus, participants completed one, two, or three parts, taken from Form A and/or B; there were no instances where participants completed the same part from both forms. TASIT was administered in line with instructions outlined in the revised manual (McDonald, Flanagan, & Rollins, 2011) and videos were played on DVD players. Researchers paused the DVD between clips to allow participants time to indicate their responses on their response sheet; for parts 2 and 3 researchers also read aloud each of the four questions one-by-one for participants. To discourage response sharing, participants were reminded that the research was “not a test” and that their thoughts were the most important part of the research. Upon testing completion, participants were debriefed and thanked for their time and effort.

Data analysis

Statistical analyses were conducted using IBM SPSS Statistics 22. Descriptive and frequency analyses were used to characterise sample demographics. A series of independent sample t-tests compared performance on Parts 1, 2, and 3 of Forms A and B of TASIT. Analyses of Variance (ANOVAs) with planned contrasts (Helmert) were used to compare the

performance of adults and adolescents (13, 14, and 15 year olds) and younger and older adolescents (13 versus 14 and 15 year olds; 14 versus 15 year olds) on Parts 1, 2, and 3 of Forms A and B of TASIT. Finally, Analyses of Covariance (ANCOVAs), controlling for age, were used to test for gender differences and English language familiarity effects (Polynomial contrasts) on performance on Parts 1, 2, and 3 of Forms A and B. Given the large number of planned contrasts, a more conservative p -value of 0.01 was adopted for these analyses. For all other analyses, $p < 0.05$ was treated as statistically significant.

RESULTS

Sample characteristics

Of the total sample ($N=665$), twenty participants (3.0%) met at least one exclusion criteria, resulting in a final sample of 645. Demographic characteristics for the final sample are summarised in Table 1. Recruited participants mainly came from single-sex Catholic schools (boys: $n=145$, 22.5% and $n=142$, 22%; girls: $n=122$, 18.9%), followed by single-sex government schools (boys: $n=144$, 22.3%; girls: $n=78$; 12.1%) and a coeducational government school ($n=9$, 1.4%). The final sample comprised slightly more males ($n=373$, 57.8%) than females, and was mostly aged 13 to 15 years old (13: $n=160$, 24.8; 14: $n=255$, 39.5; 15: $n=152$, 23.6%) with a relatively small number of participants aged younger than 13 ($n=39$, 6.0%) or older than 15 ($n=39$, 6.0%). The majority of the sample spoke only English at home ($n=372$, 58.0%), followed by English and another language ($n=150$, 23.4%), and another language only ($n=117$, 18.3%). Six students did not report their language background. The most commonly-spoken languages at home other than English, included: Arabic ($n=56$, 8.7%), Vietnamese ($n=26$, 4.0%), and Chinese/Cantonese ($n=24$, 3.7%).

TABLE 1 ABOUT HERE

The performance of 464 students (69.8% initial sample) who reported speaking English at home (including those who spoke more than one language at home) and who were aged between 13 and 15 was then examined for age and gender effects and used to create norms for Parts 1–3. Students who were younger than 13 or older than 15 were excluded as there were too few available students in these age groups to provide meaningful normative data.

Performance according to age

Performance on the different parts of TASIT (Forms A and B) for the different age groups is depicted in Tables 2-5. For comparison, mean performance of adults taken from the TASIT manual (McDonald et al., 2011) is also included.

Part 1

Performance on Part 1 of TASIT (Emotion Evaluation Test) is depicted in Table 2. Total scores on Forms A and B were marginally statistically equivalent ($p=0.055$), however, more variation emerged on the parallel subtests. Specifically, in Part 1, Surprise, Angry and Sad items were easier on A than B and Neutral items were easier on B than A. In almost all cases the difference was .5 or less of one point (out of a maximum of 4). Analysis of performance across the four ages (13, 14, 15, adult) for Forms A and B confirmed that, in both cases, performance improved with age. This was, however, primarily to do with the superior performance of the adults compared to the adolescents on both Form A and B (Table 7). Comparisons between individual age groups in adolescence revealed no differences for either Form of Part 1.

TABLE 2 ABOUT HERE

Part 2

Overall performance for Part 2 (sincere items, simple sarcasm and paradoxical sarcasm) is depicted in Table 3 while performance on the four different probe types (collapsed across subtests) is depicted in Table 4. Analyses revealed a significant difference between Forms A and B total scores, such that Form A participants performed slightly better than Form B (2.33 points out of a total of 60: Table 3). Within the different exchange types, simple sarcasm items for Form A appeared slightly easier than Form B. Meanwhile within the response probes, Do and Feel questions of Form A were slightly easier than Form B.

TABLE 3 and 4 ABOUT HERE

Analysis of scores on Forms A and B confirmed that, once again, performance improved with age. Adults performed better than adolescents overall on Forms A and B (Table 7).

Moreover, older adolescents (14 and 15 year olds) were better than younger adolescents (13 year olds) overall, $F(3,231)=13.85$, $\Psi=3.72$, $p=0.003$, at recognising paradoxical sarcasm, $F(3,231)=21.51$, $\Psi=1.92$, $p=0.003$, and at making inferences about what a character was trying to say or make another character do, $F(3,231)=17.97$ and 19.41 , $\Psi=1.54$ and 1.17 , p 's= 0.002 and <0.001 . On Form B, adults were better than adolescents at recognising sincerity, simple and paradoxical sarcasm, $F(3,130)=5.33$, 6.50 , 5.62 , $\Psi=3.31$, 2.86 , and 2.79 , p 's <0.001 , and making inferences in response to all (Do, Say, Think, and Feel) probe types, $F(3,130)=11.62$, 7.90 , 4.28 , and 7.48 , $\Psi=2.85$, 2.16 , 1.66 , and 2.30 , p 's ≤ 0.001 . On Form B, unlike Form A, no differences were observed between the different adolescent age groups. .

Part 3

Overall performance for Part 3 (sarcasm and lies) is depicted in Table 5 while performance on the four different probe types (collapsed across subtests) is depicted in Table 6.

TABLE 5 and 6 ABOUT HERE

Forms A and B were statistically equivalent for total scores as were scores across the two exchange types (lies and sarcasm) and four response probe types (Do, Say, Think, and Feel). Analysis of overall scores for Forms A and B confirmed that performance on Part 3 also improved with age. Adults performed better than adolescents overall on both Forms A and B (Table 7), however, Comparisons of the adolescent age groups revealed no differences in performance on either Form A or B.

Performance according to gender

The overall performance of girls and boys for Parts 1, 2 and 3 of TASIT (Forms A and B) are depicted in Table 7. Once again, only those who spoke English exclusively, or in combination with another language, at home, were included. For comparison, the adult normative data are replicated in the bottom rows.

TABLE 7 ABOUT HERE

The effects of gender on each Part of TASIT were examined entering age as a covariate to control for varying ages of students tested on each Part. Gender effects emerged on all total scores such that, after controlling for age, adolescent girls performed significantly better than adolescent boys on both forms of Part 1, Part 2 and Part 3.

Familiarity with English

In Table 8, scores are presented for Forms A and B of TASIT, according to familiarity with English: English only speakers ($n=342$, age: $M=13.94$, $SD=.744$), speakers of English and another language at home ($n=121$; age: $M=14.09$, $SD=0.764$) and those who speak another language at home exclusively ($n=97$; age: $M=13.87$, $SD=.731$).

Comparisons between students based upon usage of English were conducted using age as a covariate to control for the varying ages in each group. These analyses revealed that scores decreased linearly as English usage decreased on all three parts of Forms A and B.

TABLE 8 ABOUT HERE

DISCUSSION

The Awareness of Social Inference Test (TASIT) is an established test of social cognition with a number of advantages over other existing tools. It uses naturalistic audiovisual scenes with professional actors engaged in every day conversations thus providing dynamic social cues including facial and vocal expressions and gesture. TASIT uses these stimuli to probe core aspects of social cognition including emotion perception and theory of mind judgements ranging from simple mental judgements (such as beliefs and feelings) to more complex mental judgements (such as what one speaker intends towards another). It also examines understanding of irony and deception.

This study aimed to develop normative data for the use of The Awareness of Social Inference Test (TASIT) with adolescents and to examine its construct validity as a measure of social cognition in this age range. The performance of 464 students who spoke English at school and at home, aged between 13-15, were used to create norms while comparison with a further 97 students (aged 13-15) who had limited use of English at home enabled us to examine the effects of English familiarity. These adolescent norms extend those published in TASIT's manual (McDonald et al, 2011) and provide an opportunity to look at adolescent performance on all parts of TASIT.

The first thing of note is that although these adolescents were 15 or younger, they were able to interpret TASIT scenes relatively accurately. On average they achieved a score of 22-23 (Form B and A respectively) out of the maximum of 28 on Part 1. This was only

1.61 (Form A) and 1.71 (Form B) points less than their adult counterparts, a difference of around 6% of total scores. On Part 2 they were also relatively accurate, achieving 49/60 for Form A and 47/60 for Form B. Relative to Part 1, these scores were somewhat lower than that produced by adults, i.e., a difference of 5.17 (Form A) and 6.27 (Form B) points, equating to a drop of roughly 9-10%. Performance levels were similarly good for Part 3, i.e., an overall score of 49/64 (Form A) and 48/64 (Form B) which represented a drop of around seven points relative to adults for both forms or approximately 11%. TASIT was defined as a criterion referenced test such that those with a normal range of social skills are expected to perform highly. The relatively good performances by young adolescents on all three parts of TASIT are consistent with this expectation and suggest that it is a suitable measure for this age range. Furthermore, Form A and B scores were generally equivalent except for some minor differences and produced similar patterns to those seen in the adult normative study (McDonald et al., 2003). Specifically on Part 1, the two forms were equivalent and the same emotion categories were easier on A than B and visa versa in both adolescents and adults. On Part 2, adults produced comparable total scores on both forms whereas adolescents scored higher overall on Form A than B although the magnitude of the difference (2.3/60) was trivial. At a subtest level, the adolescents had higher scores on the simple sarcasm items on Form A relative to Form B whereas for adults the paradoxical sarcasm of Form A was easier. In Part 3 also, both adolescents and adults produced statistically equivalent scores for Forms A and B. Where the adolescents produced comparable scores for both lies and sarcasm across the two forms, in adults, sarcasm was slightly easier on Form A and lies on Form B.

The pattern of performance on TASIT in adolescence also attests to its construct validity. The finding that Part 1 was performed so well by the adolescents suggests that emotion perception ability as measured by TASIT is relatively well developed by early adolescence. This is consistent with other research that has accuracy in judging emotions in

photographs of faces develops incrementally through childhood (Rosenberg-Kima & Sadeh, 2010) reaching adult levels by 13-14 (Kolb et al., 1992), although this is not a universal finding, with other research indicating continuing improvement in emotion perception into the early 20s (Williams et al., 2009). Certainly, the results from this study suggest that, when realistic, multi-modal audiovisual emotions are used as stimuli, there was a great deal of overlap in emotion perception abilities within our group over the ages of 13-15 such that no incremental pattern was seen and such that performance neared (but did not quite reach) adult competencies.

Our results demonstrated that the adolescents were also relatively accurate when assessing social interactions that conveyed indirect meanings such as sarcasm and lies (Parts 2 and 3) although, once again, not as accurate as adults and seemingly not as accurate as when simply judging emotions. Given the rather complex questions asked of each scene in Parts 2 and 3 it is reasonable to question the role of developing language and cognitive abilities in producing these slightly deflated scores. Language ability has a moderate to large effect on theory of mind tasks in younger children (under the age of seven) (Milligan, Astington, & Dack, 2007) and development in executive function has also been speculated to contribute to increasing competence on mentalising tasks (Blakemore, 2008). However, language, along with working memory and processing speed, appears to be less of a contributor by the time children are 12 to 15 years old (Vetter, Leipold, Kliegel, Phillips, & Altgassen, 2013). The alternative possibility is that, consistent with our hypotheses, TASIT performance on Parts 2 and 3 reflected improving competence in theory of mind judgements with age and, relatedly, improving ability to understand non-literal meanings in conversation. Certainly, there is evidence for growing sophistication of mentalising ability through early childhood (Wimmer & Perner, 1983) and into adolescence (Choudhury, Blakemore, & Charman, 2006; Keulers, Evers, Stiers, & Jolles, 2010). This is paralleled by the increasing

competence in this age range for understanding non-literal meanings. Children from 6 to 9 years of age begin to grasp sarcasm and irony (Dews et al., 1996) and also deception (Demorest, Meyer, Phelps, Gardner, & Winner, 1984). There is evidence that this capacity continues to be refined into adulthood (Demorest et al., 1984) although studies vary with respect to when full competence is reached and whether different kinds of mental judgements (e.g., sarcasm versus deception) develop before others. As evidence of the construct validity of TASIT for adolescents, we found an incremental pattern on Part 2 (Form A) such that older adolescents (14 and 15 year olds) out- performed the 13 year olds when understanding paradoxical sarcasm. They were also better than their younger counterparts recognising intention (the ‘do’ and ‘say’ questions tapping second order theory of mind judgements) which is consistent with our prior work demonstrating a strong relation between second order theory of mind and understanding sarcasm (McDonald & Flanagan, 2004). This pattern was not, however, observed in the simple sarcasm scenes of Part 2, or in sarcastic subsets of Form B or Part 3 (both forms). In each of these cases there was overlap between the different age groups such that no differential performance according to age was observed, except for the superior performance of the adults in every case.

To some extent, the overlap in scores between adolescents aged 13-15 on the majority of subtests of TASIT is not surprising. While prior research leads us to anticipate a developmental trajectory in social cognition abilities during adolescence, it is also the case that there are enormous individual differences in the rate and timing of maturity physically, emotionally and cognitively during this time. Some 15 year olds may not have yet achieved the same level of maturity in skills and abilities, including social cognition, as that seen in precocious 13 year olds. A similar lack of age effects on social cognition measures for older adolescents has also been reported (Taylor, Barker, Heavy & McHale, 2013). Some inkling of these individual differences is seen in the variability of the adolescent scores relative to the

adults. Indeed, across all parts of TASIT and alternate forms, the estimated SD of each adolescent age group was larger than the adults by 27-100%. Such variance, combined with the fact that we grouped adolescents loosely according to age in years (rounded) rather than months, will diminish the likelihood of seeing systematic differences according to age groups separated by a single year. Overall, the finding that the adolescents did perform significantly less well than the adults, attests to the construct validity of TASIT as an instrument sensitive to developing levels of competence in social cognition. In view of the significant overlap of scores in the 13-15 year old age group, combined with the inherent variability of maturation during this period of development, TASIT normative scores for adolescents may be most useful as aggregated scores rather than broken down into years.

The uniform reliance on audiovisual scenes in TASIT stimuli provides a unique opportunity to consider the developmental trajectory of different facets social cognition side by side. Studies into the developmental trajectory for both emotion perception and for social cognition vary as to whether they suggest one kind of ability matures before the other. Part of the problem with such research is that there are few if any comparable measures of emotion perception and ToM. Most studies examining these two constructs either individually or combined use disparate means such as written text, cartoons and static photographs (Demorest et al., 1984; Dews et al., 1996; Keulers et al., 2010; Kolb et al., 1992; Vetter et al., 2013; Williams et al., 2009). In contrast, TASIT's use of similar audiovisual scenes to assess both emotion perception and ToM affords an opportunity to observe how these different abilities mature in adolescence relative to adults. Examination of mean differences between adults and adolescents for Part 1 versus Parts 2 and 3 suggests that adolescents aged 13-15 are closer to having mastered simple emotion perception than they are at understanding complex theory of mind and speaker intentions.

As further evidence of its construct validity, we found that girls, regardless of age,

outperformed boys on all parts of TASIT. The advantage afforded by being female was consistent and clear and far outweighed the ambiguous improvements with chronological age. This further attests to the construct validity of the test and is consistent with prior research demonstrating a female advantage in children, adolescents and adults (Baron-Cohen et al., 1997; Bosacki, 2000; Hughes & Dunn, 1998; Lee et al., 2013; McClure, 2000). It does suggest that gender based norms are important to consider with TASIT and other measures of social cognition.

Finally, because TASIT is heavily reliant upon good English comprehension, it was important to examine the effects of familiarity with English. As apparent in Table 8, there was a significant decrement in performance on TASIT as English familiarity decreased. For those who spoke English exclusively at home, the difference between the performance of adolescents and adults was only 1-2 points on Part 1 and 4-6 points on Parts 2 and 3. In contrast, for those adolescents who were much less familiar with English, speaking another language at home exclusively, decrements from the adult scores were much larger, i.e. 4 points on Part 1 and 10-12 points on Parts 2 and 3. This means the adolescents with least familiarity with English trailed behind their native English speaking peers by 6-13% across the various subtests. The performance of these adolescents suggests they were still able to perform the tasks well above chance levels and their scores provide some tangible information regarding the effects of lack of English familiarity.

In conclusion, this study provides a comprehensive normative data set for adolescent performance on TASIT. By sampling a large group of students aged 13-15 it was revealed that TASIT can be suitable for use with adolescents and produces scores that were predictable and that were only 6-11% lower than adult normative data. TASIT yielded patterns of performance in terms of age and gender that attest to its construct validity as a measure of social cognition in adolescence. Given the inherent variability of adolescent development,

few differences were found in the scores of adolescents of 13, 14 and 15, and it is recommended that aggregate scores for this age group are more useful for normative comparisons. On the other hand, gender differences were strongly apparent and consequently, gender based normative comparisons are recommended. Table 8 provides such aggregated, gender based norms. The students who contributed to these normative data were almost exclusively from single-sex schools, so there is the possibility that these gender based disparities might be somewhat different when girls and boys are schooled together. Future research needs to address this question. It is also important to note that there are facets of social cognition that are not assessed by TASIT. Moral reasoning, and empathy for example, are facets of social cognition that undergo critical development in adolescence and can be disrupted in acquired brain injury (Beauchamp, Dooley & Anderson, 2013). Given the wide variety of clinical conditions that can interrupt normal social cognitive development, and the importance of such skills for the development of social competence, it is important that we have valid assessment tools for adolescents. While TASIT does not assess all aspects of social cognition, it has proven validity in numerous adult clinical conditions and also in adolescents with TBI (McDonald et al., 2013) and, in the context of the normative data from this study, may provide a useful adjunct assessment in this age group.

References

- Anderson, V., Beauchamp, M. H., Yeates, K. O., Crossley, L., Hearps, S. J. C., & Catroppa, C. (2013). Social Competence at 6 Months Following Childhood Traumatic Brain Injury. *Journal of the International Neuropsychological Society*, 19(05), 539-550.
- APA. (2000). *Diagnostic and statistical manual of mental disorders* (4th, Text Revision ed.). Washington, D C: Author.
- Baron-Cohen, S. (1989). The autistic child's theory of mind: a case of specific developmental delay. *Journal of Child Psychology & Psychiatry & Allied Disciplines*, 30, 285-297.
- Baron-Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another advanced test of theory of mind: Evidence from very high functioning adults with autism or Asperger Syndrome. *Journal of Child Psychology & Psychiatry & Allied Disciplines*, 38(7), 813-822.
- Baron-Cohen, S., O'Riordan, M., Stone, V., Jones, R., & Plaisted, K. (1999). Recognition of faux pas by normally developing children with asperger syndrome or high-functioning autism. *Journal of Autism & Developmental Disorders*, 29(5), 407-418.
- Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: A study with normal adults and adults with Aspergers Syndrome or high functioning autism. *Journal of Child Psychology and Psychiatry*, 42, 241-251.
- Baron-Cohen, S., Wheelwright, S., Spong, A., Scahill, V., & Lawson, J. (2001). Are intuitive physics and intuitive psychology independent? A test with children with Asperger Syndrome. *Journal of Developmental & Learning Disorders*, 5, 1-58. .
- Beauchamp, M., & Anderson, V. (2010). SOCIAL: An integrative framework for the development of social skills. *Psychological Bulletin*, 136, 39-64.
- Blake, M. L. (2009). Inferencing processes after right hemisphere brain damage: Maintenance of inferences. *Journal of Speech, Language, and Hearing Research*, 52, 359-372.
- Blakemore, S.J., 2008. The social brain in adolescence. *Nature Reviews Neuroscience* 9, 267-277.
- Blakemore, S. J., & Choudhury, S. (2006). Development of the adolescent brain: implications for executive function and social cognition. . *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 47, 296–312.
- Bosacki, S. L. (2000). Theory of mind and self-concept in preadolescents: Links with gender and language. *Journal of Educational Psychology*, 92(4), 709-717.
- Bowers, L., Huisinigh, R., & LoGiudice, C. (2010). *Social Language Development Test - Adolescent Manual*. East Moline, IL LinguiSystems, Inc.
- Brune, M. (2005). Emotion recognition, 'theory of mind,' and social behavior in schizophrenia. *Psychiatry Research*, 133(2-3), 135-147.
- Carrow-Woolfolk, E. (1999). *Comprehensive Assessment of Spoken Language*. Circle Pines, MN: American Guidance Services, Inc.
- Choudhury, S., Blakemore, S.-J., & Charman, T. (2006). Social cognitive development during adolescence. *Social Cognitive and Affective Neuroscience*, 1(3), 165-174.
- Couture, S., M., Penn, D., L., & Roberts, D., L. (2006). The Functional Significance of Social Cognition in Schizophrenia: A Review. *Schizophrenia Bulletin*, 32(Suppl1), S44-S63.
- Couture, S. M., Penn, D. L., Losh, M., Adolphs, R., Hurley, R., & Piven, J. (2010). Comparison of social cognitive functioning in schizophrenia and high functioning autism: more convergence than divergence. *Psychological Medicine*, 40, 569-579.

- Demorest, A., Meyer, C., Phelps, E., Gardner, H., & Winner, E. (1984). Words Speak Louder Than Actions: Understanding Deliberately False Remarks. *Child Development*, 55(4), 1527-1534.
- Dennis, M., Agostino, A., Taylor, H. G., Bigler, E. D., Rubin, K., Vannatta, K., . . . Yeates, K. O. (2013). Emotional Expression and Socially Modulated Emotive Communication in Children with Traumatic Brain Injury. *Journal of the International Neuropsychological Society*, 19(01), 34-43.
- Dennis, M., Barnes, M. A., Wilkinson, M., & Humphreys, R. P. (1998). How children with head injury represent real and deceptive emotion in short narratives. *Brain and Language*, 61(3), 450-483.
- Dennis, M., Lockyer, L., & Lazenby, A. L. (2000). How High-Functioning Children with Autism Understand Real and Deceptive Emotion. *Autism*, 4(4), 370-381.
- Dews, S., Winner, E., Kaplan, J., Rosenblatt, E., Hunt, M., Lim, K., . . . Smarsh, B. (1996). Children's Understanding of the Meaning and Functions of Verbal Irony. *Child Development*, 67(6), 3071-3085.
- Di Martino, A., Ross, K., Uddin, L. Q., Sklar, A. B., Castellanos, F. X., & Milham, M. P. (2009). Functional brain correlates of social and nonsocial processes in autism spectrum disorders: an activation likelihood estimation meta-analysis. *Biological Psychiatry*, 65, 63-74.
- Edwards, J., Jackson, H. J., & Pattison, P. E. (2002). Emotion recognition via facial expression and affective prosody in schizophrenia: A methodological review. *Clinical Psychology Review*, 22(6), 789-832.
- Fournier, N. M., Calverley, K. L., Wagner, J. P., Poock, J. L., & Crossley, M. (2008). Impaired social cognition 30 years after hemispherectomy for intractable epilepsy: The importance of the right hemisphere in complex social functioning. *Epilepsy & Behavior*, 12, 460-471.
- Gregory, C., Lough, S., Stone, V., Erzinclioglu, S., Martin, L., Baron-Cohen, S., & Hodges, J. R. (2002). Theory of mind in patients with frontal variant frontotemporal dementia and Alzheimer's disease: Theoretical and practical implications. *Brain*, 125(4), 752-764.
- Hanten, G., Cook, L., Orsten, K., Chapman, S. B., Li, X., Wilde, E. A., . . . Levin, H. S. (2011). Effects of traumatic brain injury on a virtual reality social problem solving task and relations to cortical thickness in adolescence. *Neuropsychologia*, 49(3), 486-497.
- Happe, F. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism & Developmental Disorders*, 24(2), 129-154.
- Hughes, C., & Dunn, J. (1998). Understanding mind and emotion: Longitudinal associations with mental state talk between young friends. *Developmental Psychology*, 34(5), 1026-1037.
- Jellesma, F.C., Rieffe, C., Meerum Terwogt, M., Westenberg, M., 2009. Do I feel sadness, fear or both? Comparing self-reported alexithymia and emotional task performance in children with many or few somatic complaints. *Psychology and Health*, 24, 881-893.
- Johnston, L., Miles, L., & McKinlay, A. (2008). A critical review of the Eyes Test as a measure of social-cognitive impairment. *Australian Journal of Psychology*, 60, 135-141.
- Kern, R. S., Green, M. F., Fiske, A. P., Kee, K. S., Lee, L., Sergi, M. J., . . . Nuechterlein, K. H. (2009). Theory of mind deficits for processing counterfactual information in persons with chronic schizophrenia. *Psychological Medicine*, 39, 645-654.

- Keulers, E. H. H., Evers, E. A. T., Stiers, P., & Jolles, J. (2010). Age, Sex, and Pubertal Phase Influence Mentalizing About Emotions and Actions in Adolescents. *Developmental Neuropsychology*, 35(5), 555-569.
- Kincaid, J.P., Fishburne, R.P., Rogers, R.L., & Chissom, B.S. (1975). Derivation of New Readability Formulas (Automated Readability Index, Fog Count, and Flesch Reading Ease formula) for Navy Enlisted Personnel. Research Branch Report 8-75. Chief of Naval Technical Training: Naval Air Station Memphis.
- Kipps, C. M., Nestor, P. J., Acosta-Cabronero, J., Arnold, R., & Hodges, J. R. (2009). Understanding social dysfunction in the behavioural variant of frontotemporal dementia: the role of emotion and sarcasm processing. *Brain*, 132, 592-603.
- Kolb, B., Wilson, B., & Taylor, L. (1992). Developmental changes in the recognition and comprehension of facial expression: Implications for frontal lobe function. *Brain and Cognition*, 20(1), 74-84.
- Kosmidis, M. H., Eleni, A., P., B. V., Maria, G., & Panayiotis, I. (2008). Studying social cognition in patients with schizophrenia and patients with frontotemporal dementia: Theory of mind and the perception of sarcasm. *Behavioural Neurology*, 19(Print), 65-69.
- Lee, N. C., Krabbendam, L., White, T. P., Meeter, M., Banaschewski, T., Barker, G. J., . . . Shergill, S. S. (2013). Do you see what I see? Sex differences in the discrimination of facial emotions during adolescence. *Emotion*, 13(6), 1030-1040.
- Mathersul, D., McDonald, S., & Rushby, J. A. (2013). Understanding advanced theory of mind and empathy in high-functioning adults with autism spectrum disorder. *Journal of Clinical and Experimental Neuropsychology*, 35(6), 655-668.
- McClure, E. (2000). A Meta-Analytic Review of Sex Differences in Facial Expression Processing and Their Development in Infants, Children, and Adolescents.[Article]. *Psychological Bulletin*, 126(3), 424-453.
- McDonald, S., English, T., Randall, R., Longman, T., Togher, L., & Tate, R. L. (2013). Assessing social cognition and pragmatic language in adolescents with brain injuries. *Journal of the International Neuropsychological Society*, 19, 1-11.
- McDonald, S., & Flanagan, S. (2004). Social Perception Deficits After Traumatic Brain Injury: Interaction Between Emotion Recognition, Mentalizing Ability, and Social Communication. *Neuropsychology*, 18(3), 572-579.
- McDonald, S., Flanagan, S., Martin, I., & Saunders, C. (2004). The ecological validity of TASIT: A test of social perception. *Neuropsychological Rehabilitation*, 14, 285-302.
- McDonald, S., Flanagan, S., & Rollins, J. (2011). *The Awareness of Social Inference Test (Revised)*. Sydney, Australia: Pearson Assessment.
- McDonald, S., Flanagan, S., Rollins, J., & Kinch, J. (2003). TASIT: A New Clinical Tool for Assessing Social Perception after traumatic brain injury. *Journal of Head Trauma Rehabilitation*, 18, 219-238.
- McDonald, S., & Saunders, J. C. (2005). Differential impairment in recognition of emotion across different media in people with severe traumatic brain injury. *Journal of the International Neuropsychological Society*, 11(4), 392-399.
- Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and Theory of Mind: Meta-Analysis of the Relation between Language Ability and False-Belief Understanding. *Child Development*, 78(2), 622-646.
- Pettit, G. S., Dodge, K. A., & Brown, M. M. (1988). Early Family Experience, Social Problem Solving Patterns, and Children's Social Competence. *Child Development*, 59(1), 107-120.

- Rankin, K. P., Salazara, A., Gorno-Tempin, M. L., Sollberger, M., Wilson, S. M., Pavlica, D., . . . Miller, B. L. (2009). Detecting sarcasm from paralinguistic cues: Anatomic and cognitive correlates in neurodegenerative disease *Neuroimage*, 47, 2005-2015.
- Rosenberg-Kima, R. B., & Sadeh, A. (2010). Attention, Response Inhibition, and Face-Information Processing in Children: The Role of Task Characteristics, Age, and Gender. *Child Neuropsychology*, 16(4), 388-404.
- Sparks, A., McDonald, S., Lino, B., O'Donnell, M., & Green, M. J. (2010). Social cognition, empathy and functional outcome in schizophrenia. *Schizophrenia Research*, 122, 172-178.
- Stronach, S. T., & Turkstra, L. S. (2008). Theory of mind and use of cognitive state terms by adolescents with traumatic brain injury. *Aphasiology*, 22(10), 1054-1070.
- Taylor, S.J., Barker, L.A., Heavey, L., McHale, S., 2013. The typical developmental trajectory of social and executive functions in late adolescence and early adulthood. *Developmental Psychology* 49, 1253-1265.
- Tonks, J., Williams, H. W., Frampton, I., Yates, P., & Slater, A. (2007). Assessing emotion recognition in 9–15-years olds: Preliminary analysis of abilities in reading emotion from faces, voices and eyes. *Brain Injury*, 21(6), 623-629.
- Tonks, J., Williams, W. H., Frampton, I., Yates, P., & Slater, A. (2007). Reading emotions after child brain injury: A comparison between children with brain injury and non-injured controls. *Brain Injury*, 21(7), 731-739.
- Turkstra, L. S. (2008). Conversation-based assessment of social cognition in adults with traumatic brain injury. *Brain Injury*, 22(5), 397-409.
- Turkstra, L. S., McDonald, S., & DePompei, R. (2001). Social information processing in adolescents: Data from normally developing adolescents and preliminary data from their peers with traumatic brain injury. *Journal of Head Trauma Rehabilitation*, 16(5), 469-483.
- Vetter, N. C., Leipold, K., Kliegel, M., Phillips, L. H., & Altgassen, M. (2013). Ongoing development of social cognition in adolescence. *Child Neuropsychology*, 19(6), 615-629.
- Wiig, E. H., & Secord, W. (1989). *Test of Language Competence-Expanded Edition*. San Antonio, TX: Psychological Corporation.
- Williams, L. M., Mathersul, D., Palmer, D. M., Gur, R. C., Gur, R. E., & Gordon, E. (2009). Explicit identification and implicit recognition of facial emotions: I. Age effects in males and females across 10 decades. *Journal of Clinical and Experimental Neuropsychology*, 31(3), 257-277.
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13(1), 103-128.
- Winner, E., & Leekam, S. (1991). Distinguishing irony from deception: understanding the speaker's second-order intention. *British Journal of Developmental Psychology*, 9, 257-270.
- Yeates, K. O., Bigler, E. D., Dennis, M., Gerhardt, C. A., Rubin, K. H., Stancin, T., . . . Vannatta, K. (2007). Social Outcomes in Childhood Brain Disorder: A Heuristic Integration of Social Neuroscience and Developmental Psychology. *Psychological Bulletin*, 133(3), 535–556.
- Yeates, K. O., Schultz, L. H., & Selman, R. L. (1991). The Development of Interpersonal Negotiation Strategies in Thought and Action: A Social-Cognitive Link to Behavioral Adjustment and Social Status. *Merrill-Palmer Quarterly* (1982-), 37(3), 369-405.

		Males n (%)	Females n (%)
Age	12	11 (2.9)	28 (10.8)
	13	95 (25.5)	65 (23.9)
	14	154 (41.3)	101 (37.1)
	15	84 (22.5)	68 (25.0)
	16	25 (6.7)	9 (3.3)
	17	3 (0.8)	1 (0.4)
	18	1 (0.3)	0 (0)
English usage at home^a	English only	190 (70.1)	182 (49.2)
	English and other	58 (21.4)	92 (24.9)
	Other only	23 (8.5)	94 (25.4)
LOTE^a	Arabic	34 (18.0)	22 (27.2)
	Vietnamese	13 (6.9)	13 (16.0)
	Chinese	11 (5.9)	1 (1.2)
	Hindi	11 (5.8)	3 (3.7)
	Cantonese	10 (5.3)	2 (2.5)
	Tagalog	10 (5.3)	2 (2.5)
	Tongan	9 (4.8)	5 (6.2)
	Bengali	10 (5.3)	0 (0)

Table 1: Demographic characteristics of the final sample ($N=645$) for males and females separately. ^a = percentages based on valid percentages; LOTE = Languages Other Than English spoken by $\geq 5\%$ of male and/or female samples.

Part 1	Form A						Form B					A - B difference: adolescents (all)			
	Age	13	14	15	All	Adults	13	14	15	All	Adults	Mean	CI	<i>t</i> (259)	<i>d</i>
	<i>N</i>	36	32	50	118	88	37	83	23	143	46				
Happy	Mean	3.64	3.31	3.46	3.47	3.72	3.59	3.47	3.30	3.48	3.50	-0.001	-.18-0.17	-.01	.01
	<i>SD</i>	0.54	0.69	0.68	0.65	0.52	0.64	0.75	0.93	0.76	0.66				
Surprised	Mean	3.53	3.75	3.56	3.60	3.76	3.59	3.27	3.26	3.35	3.57	0.25	.08-0.43	2.83*	.34
	<i>SD</i>	0.65	0.44	0.68	0.62	0.53	0.60	0.86	0.96	0.82	0.65				
Neutral	Mean	2.94	3.03	2.96	2.97	2.96	3.35	3.42	3.30	3.38	3.07	-0.40	0.61-0.19	3.71*	.47
	<i>SD</i>	0.83	0.93	0.95	0.90	0.90	0.86	0.87	0.70	0.84	0.57				
Sad	Mean	3.11	3.28	3.40	3.28	3.60	2.73	2.89	2.61	2.80	3.07	0.48	.29-.67	4.96*	.61
	<i>SD</i>	0.79	0.73	0.61	0.70	0.62	0.80	0.83	1.12	0.87	0.57				
Angry	Mean	3.06	3.34	3.40	3.28	3.68	2.84	2.84	2.22	2.74	3.15	0.55	0.33-0.77	4.92*	.58
	<i>SD</i>	0.72	0.65	0.73	0.72	0.60	0.80	1.08	1.41	1.09	0.84				
Anxious	Mean	3.25	3.25	3.42	3.32	3.61	3.27	3.16	3.00	3.16	3.50	0.15	-.08-.39	1.31	.17
	<i>SD</i>	1.03	0.92	0.76	0.89	0.75	0.87	0.97	1.24	0.99	0.81				
Revolted	Mean	3.50	3.31	3.12	3.29	3.63	3.57	3.54	3.39	3.52	3.78	-0.23	-.46-.00	-1.96	.25
	<i>SD</i>	0.85	1.03	1.17	1.05	0.65	0.56	0.85	0.94	0.80	0.51				
Total (28)	Mean	23.11	23.28	23.32	23.25	24.86	22.95	22.59	21.09	22.44	24.15	0.83	-.02-1.69	1.93	.23
	<i>SD</i>	2.734	3.072	2.684	2.79	2.11	3.06	4.27	5.20	4.180	2.53				

Table 2: Means and standard deviations by age for each emotion, Part 1: Emotional Evaluation Test, Forms A and B, 95% CI = confidence interval; *t* = independent sample t-statistic, * = $p < 0.01$

Exchange type		Form A					Form B					A-B difference: adolescents (all)			
	Age	13	14	15	All	Adults	13	14	15	All	Adults	Mean	CI	<i>t</i> (223)	<i>d</i>
	<i>N</i>	36	33	48	117	98	37	46	25	108	73				
Sincere (20)	Mean	15.39	15.85	15.60	15.61	16.51	14.97	14.52	14.72	14.72	15.67	.88	-.07-1.84	1.83	.25
	<i>SD</i>	3.02	3.30	3.11	3.12	3.25	3.73	4.24	4.19	4.03	3.41				
Simple sarcasm (20)	Mean	16.81	18.06	17.75	17.55	18.56	16.57	16.33	16.40	16.43	18.67	1.12	.35-1.89	2.87*	.38
	<i>SD</i>	2.95	2.56	2.27	2.61	1.94	3.22	3.37	2.92	3.19	1.89				
Paradoxical Sarcasm (20)	Mean	14.92	17.03	15.58	15.79	19.04	16.30	15.20	16.32	15.83	18.40	-.05	-1.04-.94	-.09	.01
	<i>SD</i>	3.89	3.15	4.12	3.86	1.44	3.49	4.00	3.33	3.69	2.14				
Total (60)	Mean	47.11	50.94	48.94	48.94	54.11	47.30	45.61	47.44	46.61	52.88	2.33	.28-4.38	2.24**	.31
	<i>SD</i>	6.76	7.05	6.70	6.12	4.29	7.83	9.00	8.64	8.50	5.30				

Table 3: Part 2: Social Inference (Minimal): Means and standard deviations by age for each exchange type, Forms A and B, CI = 95%

confidence intervals; *t* = independent sample *t*-statistic, * = $p < 0.01$, ** = $p = 0.026$

Response Probe		Form A						Form B						A -B difference: adolescents (all)			
	Age	13	14	15	All	Adults	13	14	15	All	Adults	Mean	CI	<i>t</i> (223)	<i>d</i>		
	<i>N</i>	36	33	48	117	98	37	46	25	108	73						
Do (15)	Mean	11.25	12.58	11.77	11.84	13.72	11.41	10.93	11.24	11.17	13.19	.67	0.09-1.24	2.29**	.30		
	SD	1.98	2.05	1.93	2.03	1.19	2.22	2.54	2.33	2.37	1.79						
Say (15)	Mean	11.53	12.76	12.44	12.25	13.62	12.00	11.65	12.52	11.97	13.49	.28	-.30-.85	.95	.13		
	SD	1.82	1.79	2.18	2.02	1.45	2.27	2.45	2.22	2.34	1.36						
Think (15)	Mean	11.56	12.42	11.87	11.93	13.00	12.30	11.76	12.04	12.01	13.21	-.08	-.64-.49	-.27	.04		
	SD	1.87	2.22	1.92	2.01	1.42	2.05	2.44	2.62	2.27	1.44						
Feel (15)	Mean	12.72	13.18	12.85	12.91	13.77	12.14	11.74	11.52	11.82	13.10	1.08	0.52-1.64	3.80*	.51		
	SD	1.78	2.04	1.86	1.88	1.16	2.11	2.37	2.66	2.35	1.50						

Table 4: Part 2: Social Inference (Minimal): Means and standard deviations by age for each probe for Forms A and B, CI = 95% confidence intervals; *t* = independent sample *t*-statistic, * = $p < 0.01$, ** = $p = 0.023$.

Exchange type		Form A						Form B					A -B difference: adolescents (all)			
	Age	13	14	15	All	Adults	13	14	15	All	Adults	Mean	CI	<i>t</i> (249)	<i>d</i>	
	<i>N</i>	40	49	46	135	123	32	66	17	115	63					
Lies (32)	Mean	25.83	26.69	25.91	26.11	27.77	25.59	26.74	26.52	26.39	28.59	-.28	-1.26-.70	-.56	.07	
	SD	3.63	5.46	3.37	4.34	2.89	3.76	3.20	3.43	3.40	2.71					
Sarcasm (32)	Mean	23.65	22.45	22.46	22.80	27.87	24.41	21.47	21.29	22.26	26.52	.54	-.65-1.73	.89	.11	
	SD	4.03	4.32	4.50	4.29	3.17	5.07	5.06	4.83	5.16	4.05					
Total (64)	Mean	49.22	48.59	48.37	48.64	55.64	50.00	47.56	47.82	48.28	55.11	.36	-1.41-2.14	.40	.05	
	SD	6.59	6.68	6.46	6.56	4.82	7.66	7.78	7.54	7.72	5.28					

Table 5: Part 3: Social Inference (Enriched): Means and standard deviations by age for each exchange type, Forms A and B, CI = 95%

confidence intervals; *t* = independent sample *t*-statistic, * = $p < 0.01$

Response Probe		Form A					Form B					A -B difference: adolescents (all)			
	Age	13	14	15	All	Adults	13	14	15	All	Adults	Mean	CI	<i>t</i> (249)	<i>d</i>
	<i>N</i>	40	49	46	136	123	32	66	17	115	63				
Do (16)	Mean	12.10	11.84	11.91	11.93	14.33	12.00	11.76	11.06	11.72	13.29	.21	-.60-1.02	.52	.07
	<i>SD</i>	2.07	2.04	2.15	2.07	1.56	2.71	4.82	2.05	3.99	1.95				
Say (16)	Mean	11.88	11.49	11.24	11.49	13.72	11.81	11.45	11.65	11.58	13.49	-.09	-.70-.52	-.29	.04
	<i>SD</i>	2.41	2.33	2.15	2.30	1.84	2.74	2.48	2.37	2.52	2.73				
Think (16)	Mean	13.08	13.00	12.78	12.94	13.63	13.50	12.80	12.59	12.97	14.16	-.02	-.48-.43	-.10	.02
	<i>SD</i>	1.66	1.62	2.19	1.83	1.31	1.34	1.92	2.03	1.81	1.22				
Feel (16)	Mean	12.45	12.27	12.39	12.34	13.98	12.69	12.39	12.41	12.48	14.03	-.14	-.68-.40	-.51	.06
	<i>SD</i>	2.23	2.27	2.06	2.19	1.50	2.15	2.17	2.29	2.17	1.67				

Table 6: Part 3: Social Inference (Enriched): Means and standard deviations by age for each probe for Forms A and B, CI = 95% confidence

intervals; *t* = independent sample *t*-statistic, * = $p < 0.01$

		Form A			Form B		
		Part 1	Part 2	Part 3	Part 1	Part 2	Part 3
Max Score		28	60	64	28	60	64
Female students	<i>N</i>	49	61	51	68	53	75
	<i>Age</i>	13.71	14.02	13.92	13.96	14.04	13.92
	<i>M</i>	24.27	50.93	51.04	23.94	48.51	50.04
	<i>SD</i>	2.85	6.26	5.37	2.83	7.87	6.83
Male students	<i>N</i>	69	56	85	75	55	40
	<i>Age</i>	14.41	14.20	14.12	13.85	13.75	13.78
	<i>M</i>	22.52	46.77	47.20	21.08	44.78	44.98
	<i>SD</i>	2.518	6.99	6.81	4.72	8.74	8.29
Female-Male difference	<i>M</i>	1.74	4.17	3.84	2.90	3.38	5.07
	<i>CI¹</i>	.76-2.72	1.74-6.59	1.75-5.92	1.62-4.18	1.79-15.73	2.21-7.92
	<i>F</i>	16.29*	12.64*	10.84*	21.38*	5.65**	10.84*
	<i>df</i>	1, 115	1,114	1,132	1,140	1,105	1,112
Adults	<i>N</i>	88	98	98	98	123	123
	<i>M</i>	24.86	54.11	55.64	24.15	52.88	55.11
	<i>SD</i>	2.11	4.29	4.82	2.53	5.30	5.28
Adult-Adolescent (all) difference	<i>Ψ</i>	1.47*	4.59*	6.70*	3.59*	9.23*	10.61*
	<i>F</i>	7.00	13.86	30.84	7.23	10.03	22.99
	<i>df</i>	3,237	3,231	3,260	3,169	3,130	3,149

Table 7. Means and standard deviations for adolescents by gender, *F* = ANCOVA *F* statistic, *Ψ* = contrast estimate (Helmert), CI = 95%

confidence intervals; *df* = degrees of freedom * = $p < 0.01$; ** = $p = 0.019$. ¹ Mean differences and CIs are estimated based on raw scores. *F* values are estimated on covariate adjusted means.

		Form A			Form B		
		Part 1	Part 2	Part 3	Part 1	Part 2	Part 3
Max Score		28	60	64	28	60	64
English only at home	<i>N</i>	72	77	99	106	89	87
	<i>Age</i>	14.06	14.17	13.93	13.85	13.88	13.84
	<i>M</i>	23.32	49.49	49.33	22.74	47.04	49.03
	<i>SD</i>	2.55	7.30	6.62	4.06	8.70	8.13
English plus another	<i>N</i>	46	40	37	37	19	28
	<i>Age</i>	14.22	13.98	14.35	14.05	13.95	13.96
	<i>M</i>	23.13	47.88	46.78	21.49	44.58	45.92
	<i>SD</i>	3.15	6.07	6.11	4.40	7.31	5.79
Another language only	<i>N</i>	20	16	13	44	31	32
	<i>Age</i>	14.35	14.19	14.15	13.77	13.48	13.56
	<i>M</i>	21.05	42.31	44.08	19.82	42.16	45.19
	<i>SD</i>	3.62	6.60	7.90	4.12	9.01	6.80
Language familiarity effect	Ψ	-1.62*	-5.09*	-3.77*	-2.12*	-14.32*	-9.57*
	<i>F</i>	4.86	7.25	4.81	8.37	5.73	4.62
	<i>df</i>	2,134	2,129	2,144	2,183	2,136	2,144

Table 8: Overall performance on Parts 1, 2 and 3 of TASIT (Forms A and B) broken down into groups based on familiarity with English. Ψ = contrast estimate (linear) df = degrees of freedom * = $p < 0.01$

Appendix A

Example script

(Husband and wife couple, Ruth and Gary are in a room, Gary is reading a book. Ruth places a washing basket full of clothes on the table beside him, which she starts folding.)

Ruth Are you busy?

 Can you help me fold the washing?

Gary No, I'm not busy. I'm just here twiddling my thumbs here.

Ruth Oh, sorry. I know, you've got a lot on.

 I'll do it myself.

Gary No, really. I've got plenty of time.

 Is there something else you'd like me to do?

Ruth (sighs)