Humour Comprehension and Use of Mental State Language in Williams Syndrome and Down Syndrome

Smriti Krishan¹, Jennifer Batchelor¹ and Melanie Porter¹,²*

¹Department of Psychology, Macquarie University, Australia
²Centre for Atypical Neurodevelopment, Macquarie University, Australia

Abstract

Good humor comprehension and mental state language use is reported to be associated with good theory of mind (ToM) capabilities. There is limited research on humor comprehension in Williams Syndrome (WS) and Down Syndrome (DS), but related research in non-literal language and humor comprehension in people with intellectual disabilities suggest humor comprehension may be impaired in these populations. Furthermore, there is mixed evidence for ToM in WS and DS and limited research exists on the use of mental state language in either WS or DS. So far, the literature suggests that people with WS have are good at inferring emotions, however both WS and DS individuals have difficulties with producing cognitive language (e.g., thinks, believes). The current study sought to examine humour comprehension and use of mental state language in WS and DS relative to one another and relative to neuro typical controls, as well as the relationship between humour comprehension and mental state language use. Participants (30 WS; 18 DS; 27 CA; 30 MA controls-17 MA matched to WS; 13 MA matched to DS) were shown 23 cartoon jokes and were asked to explain what was funny. The results revealed poor humour comprehension in WS and DS relative to CA matched controls, but WS and DS individuals showed a comparable performance to each other and to MA matched controls. Furthermore, the groups only differed on the use of physical emotion words (e.g., laughing, crying, screaming) and cognition words (e.g., thinks). WS used fewer physical emotion words than DS and CA matched controls, but their frequency of use was equal to MA matched controls. The DS group did not differ relative to either CA or MA matched controls. Regarding use of cognition words, both WS and DS groups used fewer than CA controls, but performed similarly to each other and to their relative MA matched controls. Lastly, humour comprehension was not associated with humour comprehension for WS or DS individuals, only CA and DS MA matched controls. The study provides unique findings and evidence for humour comprehension in WS and DS that is similar to MA levels, but below CA levels. Results highlight the importance of intellectual ability in processing humour. Social implications and avenues for future research are discussed.

Humour is an important aspect of daily human life across the world. Laughter is a pattern of vocalizations in response to a humorous stimulus, and is one of the first social vocalizations produced by an infant, after crying McGhee [1]. We encounter humour in many different forms in our day to day lives, including interpersonal joking and banter; comedy films and TV shows and cartoons in newspapers, books, or on the internet Martin [2]. Humour is considered an essential phenomenon for successful social interaction Yip & Martin [3], as it aids in communicating ideas, feelings and opinions Nezlek & Derks [4].

To produce humour, one needs to be able to process information from the environment or from memory and manipulate ideas, words and actions to generate a funny action or comment. Understanding humour involves processing the meaning of the humorous stimuli that has been heard or viewed, and appraising it as humorous and non-serious Martin [2].

Although there are many theories of humour comprehension, one of the most widely accepted is the incongruity-resolution theory Suls [5]. The theory states that in order to understand humour, one first needs to detect an incongruity, which is usually an outcome that does not conform to the prediction the recipient made. Then the next step is to resolve the incongruity by finding a rule that makes the ending follow from the rest of the joke. Accordingly, understanding humour is viewed as a problem solving task.

Keywords: Theory of mind; Emotion

Development of Humour

Children and adults laugh at objects and events that do not conform to their existing schemas or perception of the world. That is, they laugh at incongruity. Schemas are constantly building as children develop, thus the things that may have been funny at an earlier age often become non-humorous at a later age, and other things that may not have been funny at a younger age, are humorous when older due to more sophisticated schemas.
allowing newer incongruities to be more comprehensible Martin [2].

McGhee [1] proposed a four-stage model of humour development. In stage 1 (18-24 months) children deliberately play with objects in ways that are incongruous to their normal functional use. For example, using a bucket as a hat. In stage 2 (24-36 months), children alter the relationship between verbal labels and objects, that is, incongruous labelling of objects and events such as inventing nonsense words and using wrong names for objects or body parts. Stage 3, conceptual incongruity (3-7 years) involves children deliberately altering one or more properties or attributes of events, objects or words instead of just mislabelling it. An example of jokes that children may find funny at this stage is a picture of a cat with two heads saying moo instead of meow. In stage 4, multiple meanings (7-11 years), children develop awareness of ambiguity and words with multiple meanings and that is regarded as the onset of adult humour.

Humour and Theory of Mind

Theory of Mind (ToM) is the ability to recognize and attribute mental states such as beliefs, intentions, knowledge, desires and emotions to self and others in order to predict behavior and outcome Premack & Woodruff [6]. ToM starts developing from infancy and early childhood (ages 2-3 years), when children develop skills such as joint attention, recognizing and displaying emotions, imitating expressions, pretending to be someone else (e.g. pretend play), recognizing that different people have different desires and that their actions are determined by these Westby & Robinson [7]. Between the age of 4 and 5 years, children are able to predict what others are thinking and, in turn, predict other people’s actions based on their beliefs and thoughts (first order ToM) Westby & Robinson [7]. By ages 6-8 years, children can predict what someone is thinking (or believes) about what someone else is thinking (or believes) (second order ToM) Westby & Robinson [7]. Around 8-10 years, children develop higher order ToM, that is, they are able to understand, detect and use deception, lies and figurative language Westby & Robinson [7].

Studies investigating humour comprehension in neuro developmental populations that have reported ToM deficits, such as autism spectrum disorder or schizophrenia, have provided evidence that ToM deficits impact on humour comprehension abilities Emerich et al. [8] and Wu et al. [9]. Studies have shown that both ToM and humour comprehension are related to activation in the prefrontal cortex, which is known to mediate abstraction, reasoning and problem solving (executive function) Aboulafia-Brakha et a. [10], Samson et al. [11].

A recent study by Wu et al. [9] investigated humour comprehension and appreciation of incongruity-resolution jokes and nonsense jokes in high school students with ASD. The results showed that the individuals with ASD understood fewer of the incongruity-resolution jokes compared to typical age, sex and intelligence matched controls and they also found nonsense jokes funnier than incongruity resolution jokes compared to typically developing peers. The authors suggest that the lack of ToM may interfere with understanding the punch line. However, the authors measured comprehension of jokes by asking participants to rate how funny each joke was on a 9-point scale. This is problematic, because how funny a joke seems is subjective and depends on one’s sense of humour. That is, even if the individual comprehends the joke, they may rate it as not being funny, because they did not find it humorous, thus the comprehension measure may have poor construct validity. Furthermore, the study utilized purposive sampling, that is, ASD participants were selected by the researcher based on notable impairments in communication and social interaction, as well as an overall IQ greater than 70. This may limit the generalisibility of the results, especially because ASD is a very heterogeneous disorder Lenroot et al. [12].

Another study examined humour in ASD Emerich et al. [8] asked participants to select the correct ending to a selection of cartoon jokes and found that clinical groups made more errors in selecting the correct ending compared to typically developing controls. However, the study had a small sample which consisted of only eight ASD participants and eight typically developing children, thus generalizability of results to the clinical population may be low. Additionally, although the control subjects were age matched and sex matched to the clinical subjects, IQ was not controlled, which may be a confounding variable to humour comprehension Greengross & Miller [13].

Schizophrenia is also a disorder with known deficits in ToM abilities Brüne [14]. Varga et al. [15] investigated verbal humour comprehension in adults with schizophrenia. They presented all participants with story form verbal jokes and semantic scenarios to measure literal interpretation abilities. They also measured ToM abilities using first order (knowledge about characters’ beliefs) and second order ToM tasks (what one character thinks about another character’s thoughts). The results showed that individuals with schizophrenia performed significantly lower on the humour task compared to healthy controls, which involved participants rating a humorous scenario as funny and justifying their response. Individuals with schizophrenia in their study also performed lower on second order ToM tasks compared to healthy controls, but no group differences were obtained on first order ToM tasks. Varga et al. [15] did not find any significant associations of comprehension of verbal humour with ToM abilities.

Similarly, Marjoram et al. [16] conducted a study investigating appreciation of picture jokes in schizophrenia patients. The participants were presented with 63 single image cartoons of which 31 were ToM cartoons, that is, they required attribution...
Humour in Williams syndrome

Williams’ syndrome (WS) is a rare neuro developmental disorder involving a deletion on chromosome 7, which, in turn, affects 26 to 28 genes Schubert [19]. It is characterized by cardiac defects, atypical facial features and mild to moderated intellectual disability (IQ usually ranges from 60-70, on average). Individuals with WS tend to have good receptive vocabularies, but poor performance on measures of nonverbal ability, such as drawing, visuospatial, planning and number processing Bellugi et al. [20], Brock [21]. Individuals with WS are typically extremely friendly and outgoing Doyle et al. [22], however, even with excessive desire for social contact, they have difficulty maintaining friendships Laws & Bishop [23]. Individuals with WS often have difficulties with pragmatics, such as turn taking during conversations, conversational coherence and appreciation of conversational context Laws & Bishop [23].

No studies to date have examined comprehension of picture jokes in WS. There are, however, studies that have looked into non-literal language comprehension in this population. Verbal humour can be considered a non-literal form of language, that is, its understanding requires pragmatic competence, such as drawing inferences about the speaker’s utterance Varga et al. [15]. Sullivan et al. [24] investigated whether adolescents with WS, Prader-Willi syndrome and individuals’ with non-specific intellectual deficits were able to distinguish between lies and jokes. Prader-Willi Syndrome is another neuro developmental disorder associated with intellectual disability Cassidy et al. [25]. Sullivan et al. [24] presented individuals with short stories that ended with either a lie or a joke and asked them to classify the story as a lie or a joke and give a reason for their response. The results indicated no differences between the groups, with all participants finding it difficult to identify ironic jokes and classified them as lies, as they did not correspond to reality. These errors are similar to errors made by younger typically developing children. The study did not employ a healthy control group and it is, therefore, difficult to judge the degree of impairment relative to typical humour performance by age matched typically developing children.

Moreover, a study by Godbee & Porter [26] explored comprehension of sarcasm, metaphor and simile in WS. The results showed that the WS children were poorer at comprehending non-literal language compared to chronologically age matched typically developing children, but performance of WS was not significantly different to mental age matched control children.

Humour in Down Syndrome (DS)

Down Syndrome (DS) is a genetically based disorder caused by the presence of an extra chromosome 21. Apart from the physical characteristics, it is most commonly associated with varying degrees of intellectual disability and delayed language ability Abbeduto et al. [27]. There’s been limited research on humour comprehension in Reddy V et al. [28] conducted a cross sectional study on humour and laughter of DS and ASD children matched on nonverbal IQ. They asked parents to video record or report/recall instances of humorous interactions (eliciting laughter, clowning, teasing) of their children. Results indicated that DS children tended to laugh more at or with other’s laughter, had more instances of clowning (making others laugh) and engaged in teasing more than ASD children. St James and Tager-Flusberg [29] conducted a similar observational study in age and language matched DS and ASD groups, in which they observed and recorded incidences of laughter during mother-child play interactions over the course of one year. Results indicated that the DS group overall had more incidences of laughter than the ASD group. However, it is difficult to judge if humour in DS is impaired relative to typically developing children from these results and instances of laughter do not necessarily reflect humour comprehension.

Evidence from other intellectual disorders

Several studies on humour have been conducted with children with non-specific intellectual disabilities. Bruno et al. [30] conducted a study on verbal humour comparing children with learning (academic) disabilities, children with intellectual disabilities and typically developing children. Children were asked to complete the jokes by selecting the correct joke ending. They were then asked to also explain the joke. The authors had 3 categories of jokes corresponding to McGhee’s proposed stages of humour development. The three categories were phonological (humour due to changes in sounds, thus creating multiple meanings), lexical humour (same word used in different meanings) and incongruity based humour (based on illogical resolution). They found that typically developing children scored higher than the other two groups overall and that children with an intellectual disability found it harder to explain phonological humour than lexical or cognitive humour.

Similarly, Degabriele & Walsh [31] conducted a study exploring humour comprehension in children with mild to
moderate intellectual disabilities using an appreciation task, as well as a comprehension task. In the appreciation task children watched episodes of SpongeBob SquarePants and assigned ratings of funniness after each scene. They were also presented with 12 question and answer type jokes/ridicules, with or without visual aids (pictures, gestures, acting), and then given a multiple choice question about why the joke was funny.

Results showed that the participants rated physical jokes (funny action) and visual jokes (something looks funny) as the funniest compared to the verbal jokes (funny spoken comment or joke). They also found that when a joke is told with supporting gestures or pictures, comprehension of the joke is higher in children with intellectual disabilities compared to when the jokes are not supported. It is important to note that the study had a very small sample size of nine, which limits the generalizability and did not include a typically developing control group.

**Theory of mind and Williams Syndrome**

Karmiloff Smith et al. [32] conducted one of the first studies on ToM in WS. Karmiloff-Smith et al. [32] gave participants ToM tasks measuring first and second order abilities. The authors found that individuals with WS performed similar to typically developing children and performance was significantly better than individuals with ASD. However, the authors only used ToM measures that were language based, which is a potential confound since verbal abilities in WS are a relative strength and could have aided their performance on the ToM measures. Another study assessed ToM in WS adults using a visual task in which participants were presented with photographs of eyes and asked to name the mental state Tager-Flusberg et al. [33]. Results showed that individuals with WS performed better than those with Prader-Willi Syndrome, but that only 50% of the WS participants performed as well as CA matched typical adults. This suggests that there are some potential ToM difficulties in WS. However, the authors did not match for IQ, which was significantly different in WS and Prader-Willi participants compared to typical adults.

Furthermore, Porter, Coltheart & Langdon [34] conducted a ToM investigation in WS individuals using a non-verbal picture sequencing task and found that individuals with WS performed significantly lower than CA matched typically developing participants on false belief tasks, but performance was similar on other aspects of ToM (intention, pretence, social script knowledge, cause and effect reasoning).

Santos & Deruelle [35] conducted a study examining attribution of intentions in WS compared to typically developing children via a visual measure, as well as a verbal measure (designed to be equivalent) of attributing intentions. They found that there was no difference across the groups in task performance when material was presented verbally, however, the WS individuals tended to make more errors in selecting the correct response for the visual task compared to the controls. Furthermore, the WS participants made more errors in attributing intentions than controls in the visual condition. The two studies mentioned above both individually matched WS participants to MA matched controls.

Tager-Flusberg & Sullivan [36] proposed that there are two components of ToM and that WS individuals have relatively more impairments in one of these components. The two components outlined are the social-cognitive component and a social-perceptual component. The social-cognitive component refers to the traditional ToM definition. That is, being able to represent mental states (false beliefs, intentions), typically accessed via traditional false belief ToM tasks (e.g. Sally-Anne). The social perceptual component, on the other hand, is the ability to distinguish between people and objects and also make implicit judgments about mental states through facial expressions and body language. Tager-Flusberg and Sullivan suggest that WS individuals have relatively less difficulty with social-perceptual ToM, but more difficulties with social-cognitive ToM Tager-Flusberg & Sullivan [36].

**Theory of mind and Down Syndrome**

ToM research in DS is limited, but generally suggests that ToM capabilities are commensurate with IQ levels in DS. Baron-Cohen et al. [37], for example, were one of the first to investigate ToM in DS and compared performance on false belief tasks (Sally-Anne task) in ASD. DS and typically developing controls. The results indicated that DS and typically developing children performed equally well on the false belief tasks. However, the authors did not match participants on MA or CA, both of which could potentially confound ToM results due to the different developmental trajectories Abbeduto et al. [38]. Some studies do not suggest intact ToM in DS, for example, Giavouri et al. [39] compared ToM performance in DS individuals, children with Intellectual disabilities (IQ ranged from 55-70) and typically developing controls, with all participants being matched on overall MA, but not CA (intellectual disabilities group had higher CA). The results signified the DS group performed significantly lower on the false belief tasks than typically developing and intellectual disability groups, implying that ToM is not intact in DS of relevance in this context is that most of the literature on ToM in DS assesses traditional false belief tasks.

**Mental state language in WS and DS**

Mental state language is the term given to the linguistic expression of a person’s thoughts, emotions and beliefs and some suggest is an indicator of ToM abilities Longobardi et al. [40]. Indeed, Mental State language use and ToM abilities are reported to be associated in typically developing children Dunn et al. [41].

There is limited research on mental state language use in WS and DS populations. Tager-Flusberg & Sullivan [36] explored the ability to explain an action through the use of mental states in...
A few studies have explored mental state language use in WS and/or DS populations using a narrative task Jones et al. [42], Van Herwegen et al. [43]. Van Herwegen et al. [43], for example, examined the use of emotions in narratives in WS compared to CA matched typically developing children. Results revealed that both WS and CA controls produced similar proportions of emotion words, but WS produced more emotion words related to sadness than CA controls. A similar study employing narratives found contradictory evidence Jones et al. [42], Reilly et al. [44]. The Jones et al. [42] study revealed typically developing controls used more cognitive inferences (motivation, causality and mental states) than WS group, but WS used more social engagement language (character speech, sound effects) than age matched normal controls.

Reilly et al. [44] examined the use of mental state language in WS, DS and MA matched controls. Reilly et al. coded each participant’s narrative in response to a picture for affective enhancers (affective states, character speech, sound effects) and social cognitive enhancers (mental states, inferences, causality and negation). Results indicated that WS used significantly more affective language relative to DS and MA matched controls. WS group also used more cognitive language than DS, but did not differ from MA controls. Furthermore, DS group used less of each category relative to MA controls. There are no studies examining the relationship between mental state language use and humour comprehension in WS and DS, perhaps surprising given both mental state language use and humour comprehension are reported to be associated with ToM Varga et al. [15], Marjoram et al. [16], Bihrlé et al. [17], Dunn et al. [41].

**Aims**

The first aim of the study was to examine humour comprehension (understanding of picture jokes) in WS and DS both relative to one another and relative to typically developing CA and MA matched controls. The second aim was to compare use of mental state language across the four groups. The final aim was to investigate how use of mental state language related to the understanding of humour.

**Hypotheses**

It was hypothesized that once IQ was controlled WS and DS groups would display similar levels of humour understanding and significantly lower humour comprehension scores than the CA matched control group. No significant difference in performance was expected between WS, DS and their MA matched controls.

Regarding the use of mental state language, it was expected that: WS participants would not differ significantly from CA matched or MA matched controls in their use of mental state language; WS would produce more mental state language compared to DS participants; DS participants would provide less mental state language compared to CA and MA matched controls. Thirdly, it was predicted that a higher score on humour comprehension would be associated with higher usage of mental state words in the explanation.

**Method**

**Participants**

5.1.1. Williams Syndrome: Twenty-seven English speaking Williams Syndrome (WS) individuals (16 males) aged 5-31 years (M=14.89, SD=6.96) were recruited from Williams Syndrome Australia Limited to participate in the study. All participants had a formal clinical diagnosis of WS confirmed genetically using a FISH (Fluorescence in situ hybridization) test to look for the elastin gene deletion. Mental age (MA) ranged from 3.58-9.33 years (M=5.57, SD=1.35) as assessed using the Woodcock Johnson psycho-educational Battery, Revised (WJ-R).

5.1.2. Down Syndrome: Sixteen English speaking individuals with Down Syndrome (10 males) aged 6 to 31 years (M=14.46, SD=6.36) were recruited through the NSW Departments of Community Services and Ageing, Disability and Home Care and the Down Syndrome Association NSW. Mental age was assessed via the Wj-R, ranged from 4-6.16 years (M=4.87, SD=0.75).

**Neuro typical Controls**

All controls were screened for intellectual disability Woodcock Johnson [45], neurological and psychological impairment, learning difficulties and developmental disability.

**Chronological age matched controls**

Participants were 27 (13 females) English speaking neuro typical adults. Thirteen participants were first year psychology students recruited from Macquarie University and 14 were children recruited from Neuronauts, a children’s science club at Macquarie University. Their chronological age ranged from 5 to 31 years and mean age was 15.15 (SD=6.81).

**Mental age matched controls**

Thirty MA control subjects were recruited through Neuronauts. Seventeen (9 females) participants were individually matched to WS participants on mentally age using the WJ-R, which formed the MA comparison group for WS (MAWS). The mean MA of this group was 6.11 (SD=1.46), ranging from 3.67 to 9.5. An additional 13 (5 males) participants were individually matched to 16 DS participants on MA to form the MA control group for DS (MADS). MA of the latter group ranged from 4 to 6, with a mean MA of 5.17(SD=0.61).
Materials and Apparatus

Picture jokes

A total of 23 humorous cartoons, selected from children’s picture books, were used to measure humour comprehension. Durkin & Ferguson [46], Factor & Marshall [47,48], Factor & Viska [49,50], Rosenbloom & Hoffman [51]. The jokes also measured ToM indirectly, as the jokes often required the individual to be able infer beliefs, emotions or desires of the characters in the cartoons and thus could be explained using mental state terms, similar to those used in Marjoram et al. [16] and Bihrle et al. [17]. For example, in Figure 1, to understand the first joke, individuals would need to know that the bird hunter thinks it’s a real bird, a false belief that the explorer character holds. To understand the second, participants could interpret the joke by inferring that the character is happy and unware of the dog licking her because she thinks it’s the boy kissing her. The joke responses were scored on a scale of 0 to 2. A score of 0 reflected an irrelevant or incorrect understanding of the joke, a score of 1 was awarded if the response was partially correct and was identifying parts of the humorous aspects and a score of 2 for a response if it was comprehensively identifying the core humorous aspects. Each joke had specific scoring criteria (available from the third author on request) Figure 1.

Woodcock-Johnson Psycho-educational Battery-Revised
Woodcock & Mather [52] the Woodcock Johnson psycho-educational battery revised (WJ-R) Woodcock & Johnson [52], a test of cognitive ability, is a revised version of the test first published in 1977. It is based on the Horn-Catell theory of intelligence. The standard battery of tests 1-7 (memory for names, memory for sentences, visual matching, incomplete words, visual closure, and picture vocabulary and analysis synthesis) was administered to each participant to measure their cognitive ability and calculate MA. The subtests measured each of the different abilities that compromise intellectual ability outlined by Horn-Catell theory including: long term retrieval; short term memory; processing speed; auditory processing; visual processing; comprehension and fluid reasoning. The WJ-R has sound psychometric properties, including reliability and validity Woodcock & Mather [52].

Procedure

The cartoon jokes were presented to the participants in a randomised order and participants were asked to explain what was funny/what the joke was for each. Participants were prompted if they gave no or vague response or said “I don’t know”. The responses of each participant were recorded verbatim. All participants were then administered the WJ-R to calculate their MA.

Analysis/scoring

All the jokes were scored by two independent raters and the correlation between the scores yielded 95% agreement. The remaining 5% with disagreement were scored by a third
blind rater to get a final joke score. The final score given to each participant per joke was the majority score among the 3 raters, and if there was no majority agreement, the three scores were averaged (three joke responses overall had no majority agreement). Total joke score (sum of score on all 23 jokes) was calculated for each participant and was used to measure humour comprehension.

Use of mental state language was analyzed by counting the number of mental state words or phrases the participants used in their response. Mental state words were categorized as: emotion words (e.g., sad); physical emotion words (e.g., laughing, crying); desires (e.g., wants); cognition words (e.g., thinks); pain (e.g., hurt) and character speech (Full list of words under each category is available from the third author on request). These categories were similar to those employed in Tager-Flusberg & Sullivan [36] and Reilly et al. [44], encapsulating both affective and cognitive language. The proportion of each category words to total number of words used (for all 23 jokes) was calculated before prompt and after prompt.

Results

Demographics

Table 1: Means, Standard Deviations and Range of CA and MA for each Group.

<table>
<thead>
<tr>
<th></th>
<th>WS</th>
<th>DS</th>
<th>CA controls</th>
<th>MAWS</th>
<th>MADS</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA</td>
<td>Mean(SD)</td>
<td>14.89 (6.96)</td>
<td>14.46 (6.36)</td>
<td>15.15 (6.81)</td>
<td>6.15 (1.54)</td>
</tr>
<tr>
<td>Range</td>
<td>6.00-31.00</td>
<td>6.75-31.00</td>
<td>5.17-30.25</td>
<td>3.67-10.00</td>
<td>4.24-7.33</td>
</tr>
<tr>
<td>MA</td>
<td>Mean(SD)</td>
<td>5.58 (1.35)</td>
<td>4.87 (0.75)</td>
<td>16.66 (9.29)</td>
<td>6.11 (1.46)</td>
</tr>
<tr>
<td>Range</td>
<td>3.58-9.33</td>
<td>4.00-6.16</td>
<td>4.17-32.00</td>
<td>3.67-9.50</td>
<td>4.00-6.00</td>
</tr>
</tbody>
</table>

Note: CA is in years and MA was measured via the WJ-R is also in years; MAWS= MA matched controls to WS; MADS= MA matched controls to DS

Refer to Table 1 for means and standard deviations of CA and MA for each group. WS, DS and CA matched controls did not differ significantly on chronological age (F (2,67)=0.05, p=.95). There was no significant difference in mental age between the WS and the corresponding MA matched control group (MAWS) (t(45)=-1.04, p=.304) or between the DS and their MA matched control group (MADS) (t(29)=-1.47, p=.154) Table 1.

Humour Comprehension

7.2.1. WS vs DS vs CA controls: Correlation analysis revealed that MA was significantly related to joke scores both before prompt (r=.63, p<0.001) and after prompt (r=.66, p<0.001) (Prompts were given if participants did not give an adequate response to the jokes). Therefore, MA was entered as a covariate in all following analysis on humour comprehension.

Repeated measures univariate analysis of covariance (ANCOVA) with group as a between subjects’ factor (three levels: NCA, WS and DS), prompt as a within subjects’ factor (two levels: before prompt and after prompt) and mental age as a covariate was conducted to test for any differences between the groups on the joke task. The results revealed a significant main effect for group (F (2,66) =16.86, p<0.001). However, there was no significant difference overall in scores before or after prompt (F(1,66)=.52, p=.473) and no significant prompt by group interaction (F (2, 66) =2.93, p=.061) thus there were no differences in scores before prompt and after prompt across the three groups.

Follow up pair wise comparisons for the significant main effect for group revealed that the WS group scored significantly lower on the jokes task than the CA control group (F(1, 66)=21.62, p<0.001) and that the DS group scored significantly lower than the NCA control group (F(1,66)=32.24, p<.001). However, there was no significant difference in joke scores between the WS and DS groups (F(1,66)=3.75, p=0.172) (See Table 2 for mean score differences). All p values were Bonferroni adjusted Table 2.

Table 2: Descriptive Statistics of Humour Comprehension Score for WS, DS and CA and MA matched controls.

<table>
<thead>
<tr>
<th></th>
<th>WS</th>
<th>DS</th>
<th>CA Controls</th>
<th>MAWS</th>
<th>MADS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prompt</td>
<td>23.48 (7.2)</td>
<td>19.56 (5.82)</td>
<td>36.63 (5.28)</td>
<td>24.53 (7.18)</td>
<td>22.08 (9.03)</td>
</tr>
<tr>
<td>After</td>
<td>23.96 (7.13)</td>
<td>20.13 (6.06)</td>
<td>37.11 (5.22)</td>
<td>24.58 (7.12)</td>
<td>22.62 (7.30)</td>
</tr>
</tbody>
</table>

Note: Mean (SD); MAWS= MA matched controls to WS; MADS= MA matched controls to DS
Clinical groups vs MA matched groups

See Table 1 and Table 2 for means and standard deviations for MA and humour comprehension, respectively. Correlation analysis revealed that MA was significantly associated with jokes score before prompt (r=.52, p<.005) and joke score after prompts (r=0.51, p<.005) for WS group. Similarly, MA was significantly and positively associated before prompt joke score (r=0.56, p<.05) and after prompt joke score (r=.59, p<.05) for the DS group and the WSMA controls (before prompt r=0.64, p<0.05; after prompt r=0.64, p<0.05) but not for DSM controls before prompt (r=0.14, p=0.645) or after prompts (r=0.05, p=0.862).

WS vs MA controls

The mean and standard deviations for WS and MA matched controls are presented in Table 2. A repeated measures ANCOVA comparing the WS group to the WSMA control group revealed no significant main effect for prompt (F (1,44)=1.85, p=0.181) and no significant prompt by group interaction (F(1,44)=1.80, p=0.187). There was also no significant difference in joke score between the WS group and the WSMA matched control group (F (1,44)=0.33, p=0.567).

DS vs MA controls

The results were similar for the DS group compared to their DSM matched controls. There was no significant main effect for prompt (F(1,28)=1.22, p=0.279) and no significant prompt by group interaction (F(1,28)=0.83, p=0.316). Furthermore, there were no significant difference in total joke score between the DS group and the DSMA matched control group (F (1,28)=0.27, p=0.607). Table 2 for means and standard deviations.

Use of mental state language

A repeated measures ANCOVA with group as a between subjects’ factor (five levels; WS, DS, CA controls, MAWS and MADS) and prompt as a within subjects’ factor (two levels; before prompt and after prompt) revealed no significant differences in total use of mental state language see Figure 2 between groups before and after prompt (F (1,26)=0.46, p=0.503). Therefore, only the before prompt data was used for the subsequent analyses Figure 2.

Figure 2: Mean Total mental state language use for each group (%).

WS vs DS vs CA Controls

Multivariate analysis of covariance (MANCOVA) with the six different categories of mental state language (emotion, physical emotion, desire, cognition, pain and character speech) as the dependent variables and MA as a covariate was conducted to determine if there were any differences in the proportion of mental state words used by the three groups (WS, DS and CA matched controls) (Table 3 for descriptive information for each group for each category). The overall test revealed a significant difference between groups in the use of mental state language (F(12,122)=2.13, p<0.05). Further univariate analysis revealed a significant difference between the three groups in the use of physical emotion words (F (2,66)=6.02, p<0.005) and cognition words (F(2,66)=5.28, p<0.005). There were no significant differences between the three groups in the use of desire words (F(2,66)=0.55, p=0.581), total emotion words (F(2,66)=3.8, p=0.689), pain related words (F(2,66)=1.59, p=0.211) or character speech phrases (F(2,66)=0.29, p=0.749). All p values for the univariate analysis were compared to the critical α=0.01, adjusted for the multiple dependent variables.

Follow up pair wise comparisons of the three groups for physical emotion words revealed that there was a significant difference in use of physical emotion words between the WS and DS groups (F(1,66)=11.79, p<0.005), with the WS group using significantly fewer physical emotion words compared to DS. There were no significant differences in the use of physical emotion words between the WS and CA matched control groups (F(1,66)=2.27, p=0.141) or between the DS and CA matched control groups (F (1,66)=1.87, p=.528). Comparisons for the cognition words indicated that WS used significantly fewer cognition words compared to the CA control group (F(1,66)=6.46, p<0.05). The DS group also used significantly fewer cognition words compared to the CA control group (F(1,66)=10.21, p<0.05). However, there was no significant differences in number of cognition words used by the DS and WS groups (F (1,66)=1.38, p=.735). All p values for the pair wise comparisons were bonferroni adjusted. Refer to Table 3 for descriptive statistics of mental state language for each group.

WS vs MA matched controls

The MANCOVA revealed a significant difference between the WS and MA matched control groups on the use of mental state language (F(6,39)=3.79, p<0.05). However, further analysis revealed that there was no significant difference between groups for desire words (F (1,44)=1.46, p=0.233), pain (F (1,44)=1.94, p=0.170), emotion words (F (1,44)=0.01, p=0.918), physical emotion words (F(1,44)=2.56, p=0.117) cognition words (F(1,44)=2.15, p=0.150) or character speech phrases (F (1,44)=3.29, p=0.077). All p values for univariate analysis compared to critical alpha of 0.01, adjusted for mulitple dependent variables. See Table 3 for means and standard deviations.
DS vs MA matched controls

A similar analysis was conducted to compare the DS and their MA control group. No significant differences were observed in overall mental state language use between the two groups (F (5, 24)=0.722, p =0.614). See Table 3 means and standard deviations of mental state language use Table 3.

**Table 3:** Mean Number percentage (%) of mental state words used and Standard Deviations for WS, DS and CA matched and MA matched controls.

<table>
<thead>
<tr>
<th>GROUP</th>
<th>Emotion</th>
<th>Physical Emotion</th>
<th>Cognition</th>
<th>Desire</th>
<th>Pain</th>
<th>Character Speech</th>
</tr>
</thead>
<tbody>
<tr>
<td>WS</td>
<td>2.42 (2.03)</td>
<td>0.47(0.40)</td>
<td>0.81 (0.49)</td>
<td>0.25 (0.44)</td>
<td>0.08 (0.18)</td>
<td>0.19 (0.37)</td>
</tr>
<tr>
<td>DS</td>
<td>1.97 (1.61)</td>
<td>1.02 (0.75)</td>
<td>0.44 (0.58)</td>
<td>0.30 (0.53)</td>
<td>0.20 (0.27)</td>
<td>0.20 (0.40)</td>
</tr>
<tr>
<td>CA Controls</td>
<td>2.24 (1.05)</td>
<td>0.64 (0.43)</td>
<td>2.26 (1.31)</td>
<td>0.32 (0.42)</td>
<td>0.09 (0.15)</td>
<td>0.13 (0.28)</td>
</tr>
<tr>
<td>MAWS</td>
<td>2.05 (1.26)</td>
<td>0.35 (0.27)</td>
<td>0.61 (0.55)</td>
<td>0.39 (0.42)</td>
<td>0.04 (0.11)</td>
<td>0.05 (0.09)</td>
</tr>
<tr>
<td>MADS</td>
<td>2.26 (1.39)</td>
<td>0.60 (0.60)</td>
<td>0.44 (0.54)</td>
<td>0.65 (0.71)</td>
<td>0.12 (0.19)</td>
<td>0.07 (0.75)</td>
</tr>
</tbody>
</table>

Note: Mean (SD); MAWS: MA Matched Controls to WS; MADS: MA Matched Controls to DS.

**Mental State language and humour comprehension**

Since there were no significant differences between groups on either humour comprehension or total mental state language use, before prompt or after prompt (see above), the following correlation analysis was conducted using only the before prompt data. Correlation analysis was conducted to examine whether any associations existed in the total use of mental state language and the score on the picture joke task for each group. There was a significant, positive correlation between total mental state language use and joke scores for CA matched controls (r=0.44, p<0.05) and MADS controls (r=0.58, p<0.05). No significant associations between total mental state language use and scores on the joke task were found for the DS group (r=0.13, p=0.628), WS group (r = -0.19, p=0.356) or MAWS group (r=0.35, p=0.163).

**Discussion**

The aim of the present study was to examine differences in humour comprehension and the use of mental state language in individuals with WS, DS and typically developing CA and MA matched controls and to explore whether humour comprehension and mental state language use was related.

**Humour comprehension**

It was hypothesized that the WS and DS groups would have lower humour comprehension than CA matched controls, but that no such differences would exist relative to the MA matched controls. In accordance with these predictions, WS and DS groups scored significantly lower than the CA matched control group on the joke task, but there were no significant differences in humour comprehension between the WS or DS and MA matched control groups. These findings may be interpreted to suggest that WS and DS groups understand humour at a level commensurate with their mental age, but below the level of their chronological age.

Findings are consistent with Godbee & Porter [26], which showed that WS children had more difficulty than CA matched controls, but performed similarly to MA matched children in understanding non literal language, such as sarcasm, metaphors and simile. Although, Godbee & Porter [26] did not directly examine visual humour comprehension, some similar processes are likely to underlie the comprehension of visual humorous stimuli.

Both WS and DS are associated with intellectual disability Wang & Bellugi [53], and these conditions are also both associated with executive functioning impairments Carney et al. [54], Zelazo et al. [55], Lew et al. [56]. A possible next step for future research is to explore the relationship between humour comprehension and executive functions in these populations.

**Use of mental state language**

Results revealed no significant difference in the use of mental state words (in any category) between WS and CA matched controls; except for the use of cognition words (WS produced fewer). Results also revealed that the WS group did not differ significantly in their use of mental state language compared to MA matched controls, which was in accordance with predictions. These findings suggest that WS are able to infer mental states from cartoons at the level of CA and MA matched controls, but had more difficulties with producing cognition words. Furthermore, there were no significant differences in the use of mental state language between WS and DS groups, apart from on physical emotion words (e.g., laughing, crying, screaming). WS used fewer physical emotion words than DS, which was not predicted.

The DS group also used fewer cognition words than CA matched controls, which is in accordance with expectations. On the other hand, the DS group did not differ significantly from MA matched controls. These findings were consistent with Reilly et al. [44] who showed that DS used fewer emotional and cognitive language enhancers whilst telling a story. This finding is also consistent with Baron-Cohen et al. [37] who found DS individuals performed at the level of typical controls on false belief tasks. However, it is not consistent with another study looking at ToM in DS, which also controlled for MA, and found that DS performed lower than MA matched typically developing children Giaouris et al. [57]. The discrepancy is likely due to the type of task used, as Tager-Flusberg & Sullivan [36] used traditional false belief tasks.
where they are explicitly asked what character thinking is. The present task measured implicit ability to infer mental state and may have been more subtle.

**Humour and mental state language use**

Total use of mental state language only significantly correlated with joke score for CA matched and DSMA matched controls (for DS), suggesting that humour comprehension was not associated with ToM abilities in WS or DS. This finding is inconsistent with suggestions that ToM and humour appreciation are necessarily related. Some jokes in our set could be described as physical (e.g., girl sitting on a tack), which may, at least partially account for our findings. Further research is warranted.

**Strengths and Limitations**

No previous studies have specifically explored humour comprehension (in the form of understanding picture jokes) in WS or DSA. A particular strength of this study was that each control participant was individually matched to either WS or DS individuals on the basis of mental age. DS and WS participants were also matched on chronological age (albeit at a group level). That allowed us to control for the influence of intellectual level and the amount of social experience gained with age on joke task performance. However, due to WS and DS often having an uneven cognitive profile, with strengths in some intellectual domains and weaknesses in others, it is possible that there still may be differences in some aspects of cognitive ability Jarrold & Brock [58]. In the present study, there was too much inter-subtest scatter (in the WJ-R) to be able to individually match on all the different subtests of the WJ-R and, in turn, the different aspects of intellectual ability (e.g., verbal comprehension, working memory, and reasoning) were not controlled. Finally, the sample size of each group was small (albeit relatively large relative to other studies in the area), so caution needs to be exercised when generalizing the results to the wider population [59-65].

**Implications and Future research**

The results of the current research contribute to the limited literature on humour comprehension in WS and DS. Humour is essential for successful social interactions and is considered a positive attribute Martin [2]. An impairment in humour could lead to poor peer relationships and in turn, social isolation. Techniques to enhance humour comprehension in people with WS and DS may be able to aid them in using humour more effectively during socialization and, in turn, enhance social interactions. Furthermore, these findings also provide an insight into the mental processes involved in humour processing and comprehension, and the results suggest that intellectual ability plays a role.

The results also contribute to the expanding literature on ToM abilities in WS and DS, specifically about use of mental state language in WS and DS and the relationship between ToM and humour understanding. Since WS and DS had more difficulties with cognition words (e.g., thinks), perhaps these findings could be used to form a training or therapy program to enhance their ability of inferring these mental states, for example, via more exposure to this vocabulary during early childhood, when ToM is still developing. This may also enhance WS and DS individuals’ conversation and social skills.

However, there are still questions about the topic at hand. Future research should be conducted with a greater number of participants in each group, with the aim of matching participants on each dimension of cognitive ability. That would allow use of regression analyses to examine the main cognitive contributors to the understanding of humour and, in particular, the role of language and executive functioning in humour comprehension. Related to this, future research should include independent measures of language and executive function, specifically measures of expressive and receptive language, working memory, reasoning and problem solving abilities.

Future research should also try and examine whether mode of presentation of humour (e.g., verbally or visually) affects humour comprehension in these (and other) clinical groups. Furthermore, data on humour comprehension could be gathered from multiple sources, such as measuring latencies of time taken to respond, which would reveal if there are any differences in WS, DS and CA and MA matched controls in time taken to process the stimuli and the jokes. Additionally, more research is required to examine how humour comprehension relates to mental state language in WS and DS, or perhaps whether WS and DS use different mechanisms to neuro typical individuals for humour comprehension [66-69].

**Conclusion**

In conclusion, the results reveal impaired humour comprehension relative to chronological age matched controls in both WS and DS, but both clinical groups seem to understand humour equally well to one another and to typically developing mental-age matched peers. Furthermore, WS and DS were found to use similar levels of mental state language, except for cognition words and physical emotion words (e.g., laughing). Findings highlight the importance of intellectual ability in humour comprehension and have implications relating to quality of life and social experience in WS and DS. Impairments in humour comprehension can cause difficulties in socializing and, thus, awareness of this deficit in WS and DS can help enhance their social interactions and conversational skills. Further research needs to be conducted with stricter controlling and matching on potentially confounding variables, such as verbal language abilities and executive functioning to yield a further understanding.

**Acknowledgement**

Thank you to participants and to Williams Syndrome Australia Limited and Australian Rotary Health who funded this work.
References

An empirical study with some educational implications. Procedia-Social and Behavioral Sciences 2: 3883-3887.


