

# Do animated triangles reveal a marked difficulty among autistic people with reading minds?

Autism  
2021, Vol. 25(5) 1175–1186  
© The Author(s) 2021



Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/1362361321989152  
journals.sagepub.com/home/aut



Alexander C Wilson 

## Abstract

This meta-analysis tested whether autistic people show a marked, isolated difficulty with mentalising when assessed using the Frith-Happé Animations, an advanced test of mentalising (or ‘theory of mind’). Effect sizes were aggregated in multivariate meta-analysis from 33 papers reporting data for over 3000 autistic and non-autistic people. Relative to non-autistic individuals, autistic people underperformed, with a small effect size on the non-mentalising control conditions and a medium effect size on the mentalising condition. This indicates that studies have reliably found mentalising to be an area of challenge for autistic people, although the group differences were not large. It remains to be seen how important mentalising difficulties are in accounting for the social difficulties diagnostic of autism. As autistic people underperformed on the control conditions as well as the mentalising condition, it is likely that group differences on the test are partly due to domain-general information processing differences. Finally, there was evidence of publication bias, suggesting that true effects on the Frith-Happé Animations may be somewhat smaller than reported in the literature.

## Lay abstract

Autistic people are thought to have difficulty with mentalising (our drive to track and understand the minds of other people). Mentalising is often measured by the Frith-Happé Animations task, where individuals need to interpret the interactions of abstract shapes. This review article collated results from over 3000 people to assess how autistic people performed on the task. Analysis showed that autistic people tended to underperform compared to non-autistic people on the task, although the scale of the difference was moderate rather than large. Also, autistic people showed some difficulty with the non-mentalising as well as mentalising aspects of the task. These results raise questions about the scale and specificity of mentalising difficulties in autism. It also remains unclear how well mentalising difficulties account for the social challenges diagnostic of autism.

## Keywords

animated triangles, autism, mentalising, meta-analysis, theory of mind

Autism has often been associated with difficulties in mentalising, that is, our drive to track and understand the minds of other people (Frith, 2001). Several tests have been devised to measure mentalising skills; the most fundamental is the false-belief task, which requires individuals to track a character’s false belief to predict their behaviour. While such tasks can capture mentalising difficulties in younger children, these are not sensitive to the real-world social difficulties of older autistic children and adults (Frith, 1994). This has led to the development of ‘advanced’ mentalising tests, including those which *explicitly* ask test-takers to interpret mental states, such as the Strange Stories (Happé, 1994) and the Reading the Mind in the Eyes Test

(Baron-Cohen et al., 2001). Other advanced mentalising tests are more *implicit* and measure a person’s tendency to attribute mental states without an explicit requirement for the interpretation of behaviour. The most commonly used of these tests is the Frith-Happé Animations (Abell et al., 2000; Castelli et al., 2002). This test has advantages over

University of Oxford, UK

### Corresponding author:

Alexander C Wilson, Department of Experimental Psychology, Anna Watts Building, Woodstock Road, Oxford, OX2 6GG, UK.  
Email: alexander.wilson2@psy.ox.ac.uk

other tests, as it includes well-matched control conditions that help rule out non-mentalising accounts of performance on the test. In addition, the test may be a relatively pure measure of mentalising, as it does not make the same demands as the Reading the Mind in the Eyes Test on emotion perception, a rather different skill to mentalising (Oakley et al., 2016), and it involves nonverbal materials that may somewhat reduce the high verbal requirements of other tests such as the Strange Stories (which shows a moderate relationship with verbal ability; e.g. Devine & Hughes, 2016).

In the Frith-Happé Animations, several cartoon clips of moving triangles are presented, and the test-taker is prompted to consider what is happening in each clip, responding either through verbal descriptions or a multiple-choice format (White et al., 2011). There are three conditions, one of which targets mentalising, while two conditions control for more general skills in perceiving and interpreting movement and action. In these control conditions, the triangles either move *randomly* (e.g. swirling around the screen) or in a simple *goal-directed* fashion (e.g. chasing each other across the screen). In the critical mentalising condition, the triangles interact as if they are trying to influence each other's mental states (e.g. by persuasion or deception). Given the theory that autism involves a core impairment in mentalising (Frith, 2001), autistic people might be expected to underperform specifically in the mentalising condition. This view was supported by the original studies using the animations with adults diagnosed with Asperger's (Castelli et al., 2002) and less cognitively able autistic children (Abell et al., 2000). Both groups were less likely to provide 'appropriate' descriptions, specifically, of the mentalising clips. Example descriptions from the Castelli et al. (2002) study include: 'The big triangle was trying to make the little one go out, but he doesn't want to' (appropriate description of the 'coaxing' mentalising clip) and 'The two triangles are obviously angry with each other – they are fighting' (inappropriate description of the same clip). In later research, an objective multiple-choice version of the task was found to be 'as sensitive as the traditional subjective method in demonstrating the well-established mentalizing impairment in autism' (White et al., 2011, p. 152) and this version has since been developed as a 'fast and straight forward measure of ToM [theory of mind] in autistic and neurotypical adults, to be used in future research and clinical settings' (Livingston et al., 2019). However, recent work sounds a cautionary note about the sensitivity of the Frith-Happé Animations, as a large sample of autistic adults scored lower, but only slightly, on both mentalising *and* control animations compared to non-autistic adults (Wilson & Bishop, 2020). This is more consistent with a subtle domain-general difficulty in inferring meaning rather than a marked, but isolated, difficulty with mentalising.

Given these inconsistencies, this study presents a meta-analysis to evaluate the extent to which other studies

support, or undermine, the hypothesis that the Frith-Happé Animations test reveals a specific mentalising difficulty in autistic people.

## Methods

A meta-analysis was carried out by screening the citation lists in Web of Science for the original studies that developed the Frith-Happé Animations (Abell et al., 2000; Castelli et al., 2002) as well as the citation list for a more recent multiple-choice version of the test (White et al., 2011). After removing duplicates, 1781 titles/abstracts were screened, and of these, 121 papers were read to determine eligibility. Papers were excluded because they were theoretical/review papers ( $n = 15$ ), did not present the Frith-Happé Animations ( $n = 53$ ), did not include autistic and control groups ( $n = 9$ ), only presented fMRI data ( $n = 3$ ), only presented protocols ( $n = 2$ ), had overlapping samples ( $n = 4$ ), did not supply relevant data ( $n = 1$ ) or predominantly included less verbally able individuals ( $n = 1$ ). 33 papers were included in the meta-analysis.

Studies were included if they presented the Frith-Happé Animations to a group of autistic children or adults and to a control group. Participants needed either (1) to produce verbal descriptions of the animations that were rated according to the method of Abell et al. (2000) or similar criteria set out by Castelli et al. (2002) or (2) to complete the multiple-choice questions (MCQs) devised by White et al. (2011). Under the criteria of Castelli et al. (2002), verbal descriptions are rated for appropriateness and use of mental state language (although the latter is only meaningful for the mentalising animations). The MCQ version includes two question types: classification (of each animation as showing random, physical or mental interaction) and identification of feelings (represented in the mentalising animations). As this review compared performance *across* animation types, the relevant performance indices were appropriateness ratings and accuracy on the classification MCQs. Studies were eligible if the task was presented in full or part (e.g. without the control animations). Only one study has used the Frith-Happé Animations with less cognitively able individuals (Abell et al., 2000; mean full-scale intelligence quotient (FSIQ) of 74). Therefore, to reduce heterogeneity, this study was excluded to focus on samples of individuals with average-range general ability.

The following data were extracted from all papers: sample characteristics (sample size, age, gender, verbal ability and use of autism assessments), task characteristics (number and type of animations presented, scoring criteria used and the presence of inter-rater quality check), and means and SDs for performance on each animation type (in terms of appropriateness ratings and/or the classification MCQs, depending on the procedure used in this paper). Where means and SDs were not available in the paper, effect sizes were calculated on the basis of test statistics or authors were

emailed. Authors were also emailed to confirm that samples were largely independent if there was uncertainty whether samples might have overlapped, for example, if the same authors published more than one paper in quick succession.

A meta-analysis was carried out in R (R Core Team, 2020) using the metafor package (Viechtbauer, 2010). Data and code are available on the Open Science Framework (<https://osf.io/qa8p2/>). For each animation type reported in each paper, the standardised mean difference between autistic and non-autistic groups (Hedges'  $g$ ) was computed; these are all shown in Figure 1. All individual effect sizes were aggregated in a multivariate meta-analysis to test the hypothesis that autistic people have a specific difficulty with processing mentalising animations compared to goal-directed and random animations. Animation type (goal-directed, random, mentalising) was included as a fixed effect in the meta-analysis, and the hypothesis was tested through the significance of the three levels of this fixed effect. The goal-directed condition was taken as the reference level of the fixed effect, as it offers a high-level control for the mentalising condition. Since most samples contributed more than one effect size to the meta-analysis (as most people were presented with more than one animation type), study was included as a random effect to model the dependencies between effect sizes calculated from the same groups of people. For the same reason, standard errors of effect sizes from the same paper are likely to be correlated. As papers typically have no reason to report this information, cluster robust estimation was used to allow for unknown covariances, as suggested by Hedges et al. (2010).

This meta-analysis depends on the close matching of autistic and non-autistic groups across studies. While most studies had attempted to match participants for verbal ability, age and gender, there were some small discrepancies, as can be seen in Table 1, so the robustness of the results to these participant variables was assessed. For all studies, it was possible to determine the standardised mean difference in verbal IQ (or an equivalent measure) between the autistic and non-autistic groups recruited into the study, so this control variable was added to the initial meta-analysis as a fixed effect alongside animation type to control for any group differences in verbal ability. Differences in age and the proportion of males versus females across groups were not always calculable from the studies. Therefore, these factors were assessed in subsequent models re-running the meta-analysis across subsets of studies reporting this information (For age, the standardised mean difference was computed, whereas for gender the proportion of males in the autistic group was divided by proportion of males in the non-autistic group.).

After running the meta-analysis, Cook's distance was used to assess for samples exerting undue influence on the results. Then, three further fixed effects – age-group (adult or child sample), format (verbal description or MCQs) and the inverse of the sample size – were included to

investigate whether these possible moderators accounted for heterogeneity in individual effect sizes. The inverse of the sample size was included as a moderator to assess for publication bias, as a relationship between smaller studies and larger effects might exist if the former were only published if a large effect was found.

## Results

Data from 1530 autistic and 1569 non-autistic people, drawn from 33 papers, were included in the meta-analysis. There were 2138 adults (1067 non-autistic and 1071 non-autistic) and 961 children (463 autistic and 498 non-autistic). Due to incomplete reporting, it was not possible to determine the exact gender distribution, but among adults, approximately 68% of the autistic and 59% of the non-autistic individuals were male, and among children, equivalent percentages were approximately 83% and 74%. Tables 1 and 2 present sample and task characteristics for each paper, and Figure 1 shows forest plots for effect sizes for each animation type from each paper.

Animation type (goal-directed movement, random movement, mentalising) was investigated as a fixed effect predicting autistic participants' interpretation of the clips. Each level of the fixed effect was significant, with autistic people showing less normative interpretations than non-autistic people. Controlling for any group differences in verbal ability, absolute effect sizes [95% confidence intervals (CIs)] were small for the animations with random,  $g = -0.35 [-0.51, -0.19]$ , and goal-directed movement,  $g = -0.35 [-0.48, -0.22]$ , and there was a medium effect size for the mentalising animations,  $g = -0.62 [-0.74, -0.50]$ , all  $p < 0.001$ . To test the hypothesis that autistic people have specific difficulty with the mentalising animations, relative differences in effect size were assessed. Compared to the reference level (goal-directed movement), there was no significant difference for the random condition,  $p = 0.997$ , and a small-sized increase in group difference for mentalising,  $g = -0.27 [-0.41, -0.14]$ ,  $p < 0.001$ . Table 3 shows these effect sizes separately for the child and adult samples. As noted in section 'Methods', it was possible that imperfect matching for age and gender across studies might have influenced results, so the meta-analysis was re-run across the subset of studies that had reported this information. Study-level differences in age,  $p = 0.633$ , and gender,  $p = 0.799$ , did not contribute to the models. The analyses indicate that autistic people had some difficulty with all types of animation, but somewhat more for the mentalising ones, when controlling for group differences in verbal ability, gender and age.

However, there was significant heterogeneity in effect sizes,  $Q(64) = 205.17$ ,  $p < 0.001$ , and so sources of heterogeneity were explored, first through influential case diagnostics and then through moderator analysis. One study (Clemmensen et al., 2016) was associated with a high

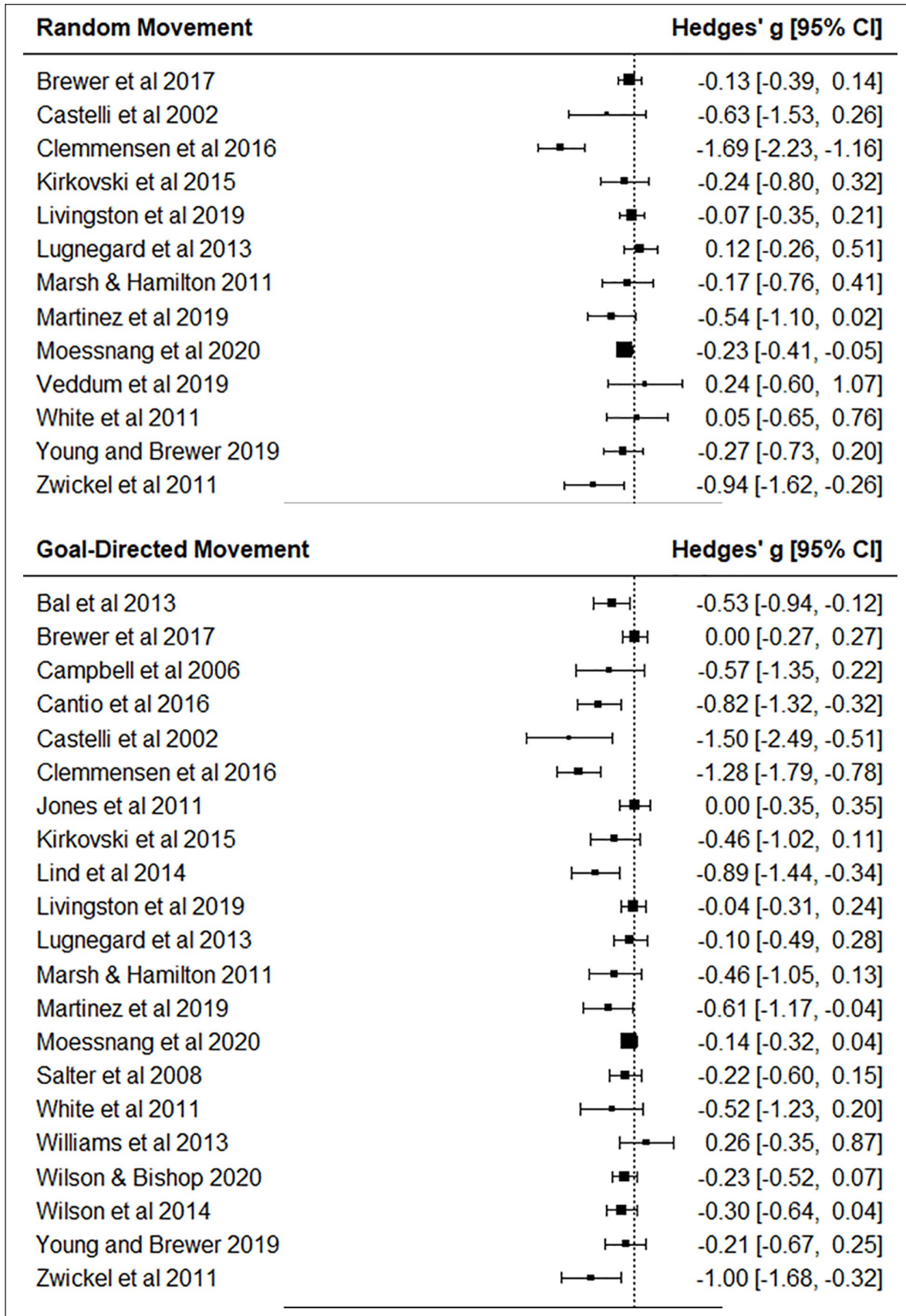


Figure 1. (Continued)



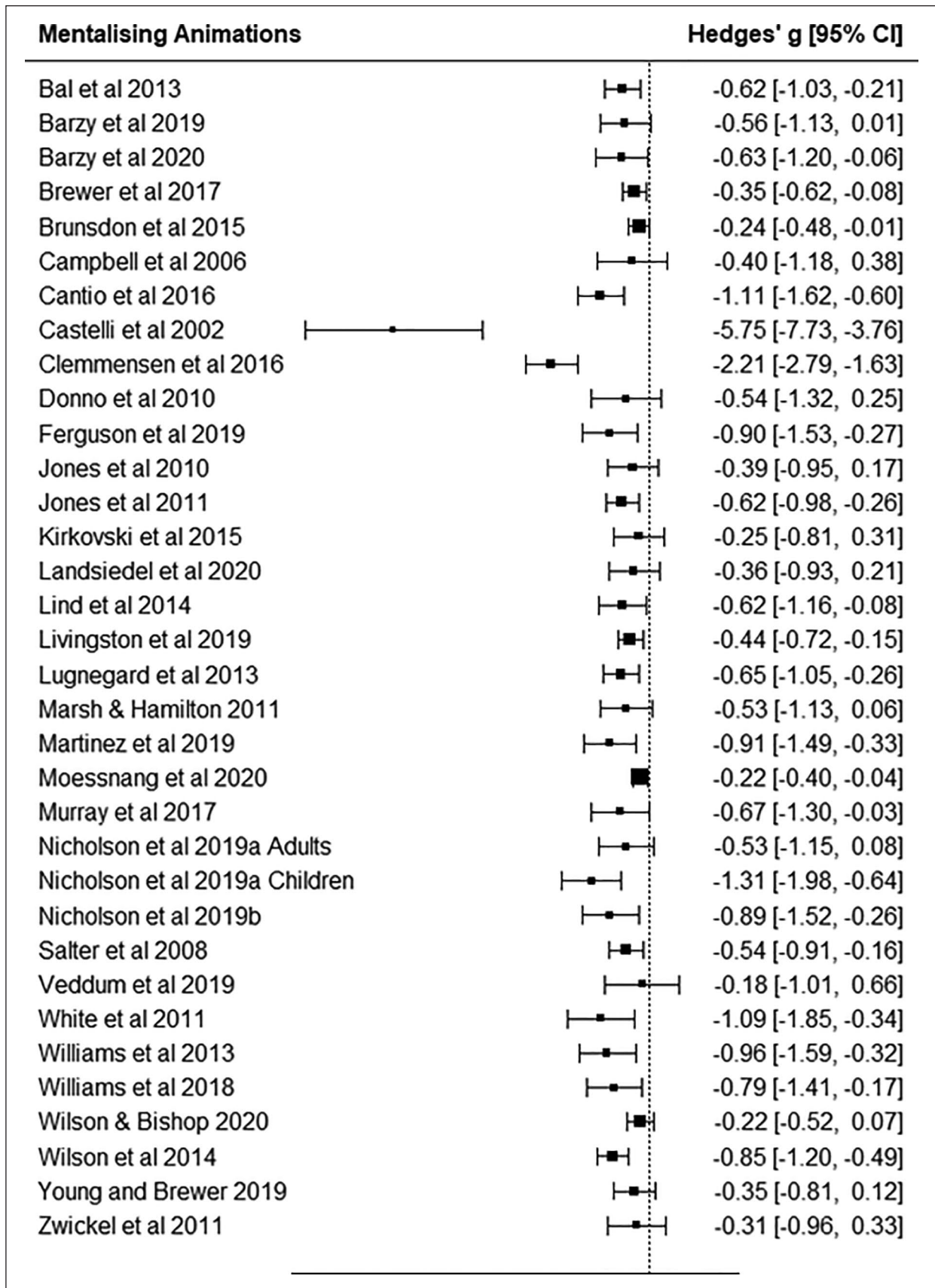


Figure 1. Forest plots for group differences between autistic and non-autistic people on the Frith-Happé Animations by condition.

**Table 1.** Sample characteristics.

| Study   | <i>n</i> | Age          | Percent male | Verbal ability | Diagnostic interview |
|---|----------|--------------|--------------|----------------|----------------------|
| <b>Adult samples</b>                                  |          |              |              |                |                      |
| Barzy et al., 2019                                    | 24       | 33;9 (11;2)  | 71           | 102 (11)       | ADOS                 |
|   | 26       | 34;9 (17;5)  | 69           | 100 (9)        |                      |
| Barzy et al., 2020                                    | 25       | 34;5 (10;9)  | 68           | 104 (12)       | ADOS                 |
|   | 24       | 33;0 (16;11) | 71           | 100 (10)       |                      |
| Brewer et al., 2017                                   | 163      | 27;0 (11;10) | 69           | 103 (15)       |                      |
|   | 80       | 26;1 (10;2)  | 30           | 111 (14)       |                      |
| Castelli et al., 2002                                 | 10       | 33 (7;7)     |              | 104            |                      |
|   | 10       | 25 (4;10)    |              | 111            |                      |
| Ferguson et al., 2019                                 | 22       | 36;2 (8;5)   | 68           | 101 (15)       | ADOS                 |
|   | 21       | 37;7 (9;9)   | 68           | 105 (11)       |                      |
| Kirkovski et al., 2015                                | 27       | 30;7 (8;10)  | 48           | 104 (13)       |                      |
|   | 23       | 30;6 (10;3)  | 48           | 106 (15)       |                      |
| Landsiedel et al., 2020                               | 25       | 34;10 (11;5) | 84           | 105 (14)       | ADOS                 |
|   | 23       | 38;3 (13;2)  | 83           | 104 (10)       |                      |
| Lind et al., 2014                                     | 27       | 35;6 (13;3)  | 78           | 112 (15)       | ADOS                 |
|   | 29       | 33;3 (16;2)  | 76           | 113 (12)       |                      |
| Livingston et al., 2019 <sup>a</sup>                  | 71       | 31;6 (11;8)  | 49           | 80 (22)        |                      |
|   | 160      | 28;10 (12;1) | 50           | 83 (8)         |                      |
| Lugnegard et al., 2013                                | 53       | 27;4 (4;1)   | 49           | 104 (12)       | DISCO                |
|   | 50       | 28;10 (9;4)  | 38           | 99 (11)        |                      |
| Marsh & Hamilton, 2011 <sup>b</sup>                   | 24       | 33 (10;11)   |              | 113 (17)       | ADOS                 |
|   | 21       | 32;2 (10;1)  |              | 113 (14)       |                      |
| Martinez et al., 2019                                 | 28       | 22;7 (3;6)   | 78           | 104 (14)       | ADI-R                |
|   | 23       | 23;3 (3;1)   | 83           | 107 (13)       |                      |
| Moessnang et al., 2020 <sup>b,c</sup>                 | 271      | 17;11 (4;5)  | 74           | 107 (14)       | ADOS, ADI-R          |
|   | 221      | 18 (5;6)     | 65           | 109 (12)       |                      |
| Murray et al., 2017                                   | 20       | 30;7 (6;6)   | 100          | 105 (17)       |                      |
|   | 20       | 30;8 (6;3)   | 95           | 111 (12)       |                      |
| Nicholson, Williams, Carpenter, & Kallitsounaki, 2019 | 21       | 37;3 (11;10) | 62           | 106 (11)       | ADOS                 |
|   | 21       | 41;2 (14;0)  | 71           | 104 (11)       |                      |
| Nicholson, Williams, Grainger, et al., 2019           | 21       | 37;2 (19;11) | 71           | 103 (12)       | ADOS                 |
|   | 22       | 37;3 (12;4)  | 76           | 107 (10)       |                      |
| Veddum et al., 2019 <sup>c</sup>                      | 11       | 25;1         | 36           | 105            |                      |
|   | 11       | 25;5         | 36           | 108            |                      |
| White et al., 2011                                    | 16       | 33 (10;4)    | 75           | 111 (13)       | ADOS                 |
|   | 15       | 36;6 (9;10)  | 73           | 114 (12)       |                      |
| Williams et al., 2018                                 | 22       | 35;10 (11;7) | 82           | 102 (15)       | ADOS                 |
|   | 21       | 36;4 (12;0)  | 76           | 107 (10)       |                      |
| AC Wilson & Bishop, 2020 <sup>d</sup>                 | 71       | 38;0 (14;0)  | 35           | 15.5 (5.3)     | ADOS                 |
|   | 120      | 30;11 (11;3) | 45           | 12.1 (4.5)     |                      |
| CE Wilson et al., 2014                                | 64       | 26;0 (7;0)   | 100          | 110 (14)       | ADOS, ADI-R          |
|   | 71       | 28;0 (6;0)   | 100          | 109 (13)       |                      |
| Young & Brewer, 2019                                  | 32       | 33;4 (13;10) | 63           | 104 (14)       |                      |
|   | 41       | 21;8 (5;11)  | 37           | 103 (10)       |                      |
| Zwikel et al., 2011                                   | 19       | 37;0         |              | 117            | ADOS                 |
|   | 18       | 39;0         |              | 115            |                      |
| <b>Child samples</b>                                  |          |              |              |                |                      |
| Bal et al., 2013                                      | 41       | 10;6 (2;3)   | 78           | 115 (19)       |                      |
|   | 58       | 10;9 (2;5)   | 71           | 115 (14)       |                      |
| Brunsdon et al., 2015 <sup>c</sup>                    | 138      | 13;6 (0;8)   | 83           | 94 (17)        | ADOS, ADI-R, DAWBA   |
|   | 148      | 12;9 (1;1)   | 69           | 102 (15)       |                      |
| Campbell et al., 2006                                 | 13       | 13;2 (1;9)   |              | 96 (18)        | 3Di                  |
|   | 13       | 13;4 (2;1)   |              | 96 (16)        |                      |
| Cantio et al., 2016                                   | 31       | 11;0 (1;4)   | 81           | 105 (14)       | ADOS, ADI-R          |
|   | 37       | 10;10 (1;4)  | 70           | 103 (16)       |                      |

(Continued)

**Table 1.** (Continued)

| Study   | <i>n</i> | Age         | Percent male | Verbal ability | Diagnostic interview |
|---|----------|-------------|--------------|----------------|----------------------|
| Clemmensen et al., 2016                               | 33       | 11;4 (1;11) | 73           | 92 (16)        |                      |
|   | 41       | 10;11 (1;9) | 41           | 106 (16)       |                      |
| Donno et al., 2010                                    | 9        | 10;4 (1;9)  | 89           | 106 (19)       | 3Di                  |
|   | 22       | 9;5 (1;8)   | 82           | 106 (16)       |                      |
| Jones et al., 2010                                    | 21       | 13;1 (1;8)  | 100          | 94 (12)        |                      |
|   | 31       | 11;6 (1;5)  | 100          | 100 (13)       |                      |
| Jones et al., 2011                                    | 79       | 15;6 (0;6)  | 91           | 84 (16)        | ADOS, ADI-R          |
|   | 50       | 15;6 (0;6)  | 96           | 88 (20)        |                      |
| Nicholson, Williams, Carpenter, & Kallitsounaki, 2019 | 21       | 12;11 (1;6) | 76           | 108 (9)        |                      |
|   | 21       | 12;8 (1;2)  | 71           | 109 (11)       |                      |
| Salter et al., 2008                                   | 56       | 10;4 (3;2)  | 84           | 99 (19)        | 3Di, ADOS            |
|   | 56       | 10;8 (3;6)  | 84           | 98 (15)        |                      |
| Williams et al., 2013                                 | 21       | 10;7 (2;0)  |              | 104 (18)       |                      |
|   | 21       | 10;7 (1;4)  |              | 106 (14)       |                      |

ADOS: Autism Diagnostic Observation Schedule; DISCO: Diagnostic Interview for Social and Communication Disorders; ADI-R: Autism Diagnostic Interview-Revised; DAWBA: Development and Wellbeing Assessment; 3Di: Developmental, Diagnostic and Dimensional Interview.

For each study, the first row relates to the autistic group and the second row to the non-autistic group. All studies required autistic participants to have a clinical diagnosis of autism based on Diagnostic and Statistical Manual of Mental Disorders (DSM)/International Classification of Diseases (ICD) criteria. Some studies gave participants or their families a diagnostic interview to confirm the diagnosis; this information is shown in the far column. For age, means (SDs) are given in years; months. For verbal ability, standard scores on norm-referenced tests of verbal intelligence are given.

<sup>a</sup>Verbal ability for this paper was measured using raw scores from the Spot the Word Task.

<sup>b</sup>Authors kindly supplied data for the Frith-Happé Animations relating to a larger sample size than reported in the paper (exclusions in these papers were made based on fMRI criteria). Descriptive statistics given here for age, sex and verbal ability relate to the slightly smaller samples reported in the paper.

<sup>c</sup>These studies did not provide descriptive statistics for verbal ability, so full-scale IQ scores have been reported here.

<sup>d</sup>For this paper, verbal ability was measured using a non-standardised vocabulary test.

Cook's distance. Exclusion of this study made limited difference to the effect sizes; random,  $g = -0.24$  [-0.38, -0.11], goal-directed,  $g = -0.34$  [-0.48, -0.20], and mentalising,  $g = -0.58$  [-0.70, -0.46]; difference in effect size between the goal-directed and mentalising animations remained significant,  $p = 0.002$ . With this study still excluded, three moderators were added to the model. These moderators were age-group (child, adult), task format (verbal description, MCQ) and inverse of the sample size. As the inverse of the sample size approaches zero, we have a hypothetical sample of unlimited size and precision, and so coefficients for other fixed effects reflect predictions under this situation. Table 4 shows the results of this moderator analysis. Inverse of the sample size was a significant predictor in the model, and coefficients for the various conditions of the Frith-Happé Animations were somewhat smaller, indicating evidence for publication bias. The moderator analysis also demonstrated a trend-level link between administration of the multiple-choice (rather than verbal) task format and slightly smaller group differences between autistic and non-autistic samples. There were no differences in effect size when comparing children and adults.

## Discussion

Across the studies collated in this review, autistic people experienced a gradient of difficulty on the Frith-Happé Animations, with a small effect size difference between

autistic and non-autistic people on the control conditions and an additional small increase in effect size on the mentalising animations. Analysis indicated that similar effects were found across children and adults on the spectrum and that there was evidence of publication bias slightly inflating these effects. It has been claimed that 'impairments in individuals with autism can be revealed in characteristic inaccuracies in mental state attribution to animated shapes' (Castelli et al., 2002, p. 1845). On the one hand, this meta-analysis did find a reliable difference between autistic and non-autistic people in mentalising skills as measured by the task. On the other hand, there are questions about the scale and specificity of the difference.

In the first study using the Frith-Happé Animations with autistic adults (Castelli et al., 2002), there was a very substantial difference between autistic and non-autistic people on the mentalising animations ( $g = -5.75$ ), but widescale replication across many studies has not shown such a stark difference. After accounting for differences in verbal ability and performance on the control animations (including the higher-level 'goal-directed' control condition), there was a small but reliable difference between autistic and non-autistic groups. Other meta-analyses have shown much larger group differences on mentalising tasks (e.g. Chung et al., 2014; Velikonja et al., 2019), but crucially these reviews have not controlled for cognitive processes other than mentalising, as this review has, so we should be sceptical about the size and specificity of these group differences.

**Table 2.** Study characteristics.

|   | Behavioural session | fMRI session | Mentalising | Goal-directed | Random | Castelli rating | Abell rating | Inter-rater check | White MCQs |
|---|---------------------|--------------|-------------|---------------|--------|-----------------|--------------|-------------------|------------|
| Adult samples   |                     |              |             |               |        |                 |              |                   |            |
| Barzy et al., 2019 <sup>a</sup>                       | X                   |              | 4           |               |        |                 | X            | X                 |            |
| Barzy et al., 2020 <sup>a</sup>                       | X                   |              | 4           |               |        |                 | X            | X                 |            |
| Brewer et al., 2017                                   | X                   |              | 4           | 4             | 4      |                 |              | NA                | X          |
| Castelli et al., 2002                                 |                     | X            | 4           | 4             | 4      | X               |              | X                 |            |
| Ferguson et al., 2019 <sup>a</sup>                    | X                   |              | 4           |               |        |                 | X            | X                 |            |
| Kirkovski et al., 2015                                |                     | X            | 4           | 4             | 4      |                 |              | NA                | X          |
| Landsiedel et al., 2020                               | X                   |              | 4           |               |        |                 | X            | X                 |            |
| Lind et al., 2014                                     | X                   |              | 4           | 4             |        |                 | X            | X                 |            |
| Livingston et al., 2019                               | X                   |              | 4           | 4             | 4      |                 |              | NA                | X          |
| Lugnegard et al., 2013                                | X                   |              | 4           | 4             | 4      | X               |              |                   |            |
| Marsh & Hamilton, 2011                                | X                   |              | 4           | 4             | 4      |                 |              | NA                | X          |
| Martinez et al., 2019                                 | X                   |              | 4           | 4             | 4      | X               |              |                   |            |
| Moessnang et al., 2020 <sup>b</sup>                   | X                   | X            | 4           | 4             | 4      | X               |              | NA                | X          |
| Murray et al., 2017                                   | X                   |              | 4           |               |        | X               |              |                   |            |
| Nicholson, Williams, Carpenter, & Kallitsounaki, 2019 | X                   |              | 4           |               |        |                 | X            | X                 |            |
| Nicholson, Williams, Grainger, et al., 2019           | X                   |              | 4           |               |        |                 | X            | X                 |            |
| Veddum et al., 2019                                   | X                   |              | 4           |               | 4      | X               |              | X                 |            |
| White et al., 2011 <sup>c</sup>                       | X                   |              | 4           | 4             | 4      | X               |              | X                 |            |
| Williams et al., 2018                                 | X                   |              | 4           |               |        | X               |              | X                 |            |
| AC Wilson & Bishop, 2020                              | X                   |              | 4           | 4             |        | X               |              | X                 |            |
| CE Wilson et al., 2014                                | X                   |              | 4           | 2             |        | X               |              |                   |            |
| Young & Brewer, 2019                                  | X                   |              | 4           | 4             | 4      |                 |              | NA                | X          |
| Zwikel et al., 2011                                   | X                   |              | 4           | 4             | 4      | X               |              | X                 |            |
| Child samples   |                     |              |             |               |        |                 |              |                   |            |
| Bal et al., 2013 <sup>d</sup>                         | X                   |              | 4           | 4             |        | X               |              | X                 |            |
| Brunsdon et al., 2015                                 | X                   |              | 4           |               |        |                 | X            |                   |            |
| Campbell et al., 2006                                 | X                   |              | 4           | 4             |        | X               |              | X                 |            |
| Cantio et al., 2016 <sup>d</sup>                      | X                   |              | 4           | 4             |        | X               |              | X                 |            |
| Clemmensen et al., 2016 <sup>c</sup>                  | X                   |              | 4           | 4             | 4      | X               |              | X                 | X          |
| Donno et al., 2010                                    | X                   |              | 4           |               |        |                 |              | X                 |            |
| Jones et al., 2010                                    | X                   |              | 4           |               |        | X               |              | X                 |            |
| Jones et al., 2011                                    | X                   |              | 4           | 2             |        | X               |              | X                 |            |
| Nicholson, Williams, Carpenter, & Kallitsounaki, 2019 | X                   |              | 4           |               |        |                 | X            | X                 |            |
| Salter et al., 2008                                   | X                   |              | 4           | 4             |        | X               |              | X                 |            |
| Williams et al., 2013                                 | X                   |              | 2           | 2             |        |                 | X            | X                 |            |

This table shows the type of session used for data collection; the number of animations used; the rating criteria (Abell et al., 2000; Castelli et al., 2002) used for verbal descriptions (with or without inter-rater checks); and use of the multiple-choice questions (MCQs) of White et al. (2011).

<sup>a</sup>Whereas effect sizes in the meta-analysis were generally calculated on the basis of means and SDs reported in papers, these studies did not present SDs but did give effect sizes. Therefore, these effect sizes were directly aggregated in the meta-analysis, with sampling variances calculated on the basis of the effect size and sample size.

<sup>b</sup>Although behavioural and fMRI sessions were used in this study, data on the verbal descriptions have yet to be fully published, so only effect sizes based on data collected on the MCQs in the fMRI session have been included in the meta-analysis. Note, however, that the authors reported no group difference on total scores on the verbal descriptions, so results are likely to be similar to the null results collected in the fMRI session.

<sup>c</sup>Although the MCQs were administered in both these studies alongside the verbal description paradigm, the papers did not report scores broken down by animation type, so only data on verbal descriptions are analysed in this review.

<sup>d</sup>These studies did not present means or SDs, so effect sizes were calculated on the basis of test statistics.

There are a couple of ways of interpreting the small effect established in this study. It may represent a modest difference between the *average* autistic and non-autistic person

in mentalising, which, along with other cognitive factors, may have real-world impacts on social learning and behaviour. However, it is worth noting that there is limited



**Table 3.** Effect sizes across the child and adult samples.

|   | Hedges' <i>g</i> | Lower 95% CI | Upper 95% CI |
|---|------------------|--------------|--------------|
| CHILDREN (autistic <i>n</i> = 430; non-autistic <i>n</i> = 457) |                  |              |              |
| Goal-directed   | -0.25            | -0.62        | 0.11         |
| Mentalising   | -0.62            | -0.88        | -0.37        |
| ADULTS (autistic <i>n</i> = 1067; non-autistic <i>n</i> = 1071) |                  |              |              |
| Goal-directed   | -0.37            | -0.54        | -0.19        |
| Mentalising   | -0.56            | -0.70        | -0.42        |
| Random  | -0.25            | -0.38        | -0.12        |

The outlying study (Clemmensen et al., 2016) involved children and has been removed from analysis. The random movement condition was only presented in studies involving adults.

**Table 4.** Results of the moderator analysis.

|                     | Hedges' <i>g</i> | Lower 95% CI | Upper 95% CI | <i>p</i> value |
|---------------------|------------------|--------------|--------------|----------------|
| Intercept           | -0.20            | -0.40        | 0.01         | 0.065          |
| Random              | 0.05             | -0.10        | 0.19         | 0.514          |
| Mentalising         | -0.23            | -0.37        | -0.09        | 0.002          |
| Age-group           | 0.03             | -0.22        | 0.27         | 0.832          |
| Format              | 0.15             | -0.03        | 0.32         | 0.094          |
| Verbal ability      | 0.01             | -0.31        | 0.34         | 0.945          |
| Inverse sample size |                  |              |              | 0.045          |

The intercept reflects performance on the goal-directed movement condition of the Frith-Happé Animations. Hedges' *g* values for the other conditions represent relative differences compared to the intercept rather than the absolute effect size.

evidence that mentalising tests predict autistic features or day-to-day social skills (as reviewed by Gernsbacher & Yergeau, 2019), so the real-world significance of a modest difference in mentalising remains difficult to assess. Alternatively, the effect size at the group level may disguise heterogeneity, with a substantial difficulty in mentalising only experienced by a subgroup of autistic people. In the case of the present analysis, this view is speculative, but it has received empirical support in a large-scale heterogeneity analysis that found mentalising differences only in some autistic subgroups (Lombardo et al., 2016). This is in line with the theory that there are multiple cognitive influences on autism that vary in their impact for any particular individual (e.g. Brunsdon & Happe, 2014).

A different interpretation of the results would revolve around the task: that the Frith-Happé Animations might not be sensitive to individual differences in mentalising. First, it should be noted that restricted variance is not a problem for the task when presented to general population or clinical groups, as no study included in this review reported a ceiling effect. Therefore, the question is not about the presence of individual differences on the task, but what these represent. On the one hand, the Frith-Happé Animations and other mentalising tasks do not tend to correlate highly, if at all, suggesting that we cannot be confident in precisely what accounts for variance in performance on the tests. Gernsbacher and Yergeau (2019) collate evidence for the poor convergence between different mentalising tests, and

conclude that mentalising lacks construct validity. On the other hand, we could equally argue that mentalising is not a single ability but a set of multiple, specific skills – a view that is supported by neural accounts of the ‘social brain’ (Schaafsma et al., 2015). Indeed, it is within the social neuroscience literature that we find the strongest argument for the validity of the Frith-Happé Animations as a test of mentalising. Studies show that the mentalising condition reliably activates social-cognitive networks in the brain that partially overlap with activation patterns observed for other mentalising tasks (see Schurz et al., 2014 for a meta-analysis). In the largest neuroimaging study of the Frith-Happé Animations to date (*n* = 394; Moessnang et al., 2020), there was robust activation across the mentalising network bilaterally in the pSTS and surrounding regions, as well as in precuneus, inferior frontal gyrus (IFG), medial prefrontal cortex (mPFC) and temporal poles. Furthermore, this neuroimaging study found no mean differences between autistic and non-autistic people in task-related activation in the mentalising network (Moessnang et al., 2020). This is consistent with the idea that the social difficulties diagnostic of autism might not be well-accounted for by mentalising difficulties.

As noted above in the meta-analysis, autistic people tended to underperform on the control animations. This raises the question whether difficulties on the Frith-Happé Animations can be explained to some extent by more general difficulties with interpreting motion, whether at a perceptual or higher cognitive level, which manifest across all

animation types. Perceptually, autistic people have been found to perform differently on tasks involving detection and discrimination of global and biological motion without mentalising demands (e.g. Klin et al., 2009; Milne et al., 2002; Robertson et al., 2014). Interestingly, in their meta-analysis of global and biological motion tasks, Van der Hallen et al. (2019) found that autistic people underperformed to a very similar degree as they did on the control animations in the present study,  $g = -0.30$  ( $-0.17, -0.44$ ), with no differences by task or motion type (Van der Hallen et al., 2019). These sensory-perceptual differences may impact on performance of a more complex task involving moving stimuli such as the Frith-Happé Animations. Alternatively, difficulties could emerge at a higher cognitive level, as the task involves assigning a 'narrative' to moving stimuli. Consistent with the central coherence hypothesis (Happé & Frith, 2006), autistic individuals may show less tendency to integrate the actions and interactions of animated shapes into a central narrative, and similar domain-general processing differences may apply across conditions. In line with this view, supplemental material to this article shows continuity in performance across the control and mentalising conditions ( $r = 0.66$ , 95% CI: 0.43, 0.89) in almost 200 autistic and non-autistic adults, suggesting that the conditions share overlapping cognitive demands.

This meta-analysis presented evidence that autistic people show less difficulty in understanding social narratives in abstract animations than early reports indicated. This suggests that we should be cautious about suggesting that mentalising is necessarily an area of marked difficulty for autistic people, although we should also not underplay the subtle but reliable difference in mentalising that did emerge, which may impact on the behavioural phenotype in autism. Given that group differences between autistic and non-autistic people also emerged on the control conditions of the Frith-Happé Animations, it is possible that performance across the task is influenced, in addition to the mentalising demands, by domain-general abilities in perceiving and assigning meaning to motion.

### Acknowledgements

The author would like to thank Professor Dorothy Bishop for comments on this article.

### Funding

The author received no financial support for the research, authorship and/or publication of this article.

### Declaration of conflicting interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### ORCID iD

Alexander C Wilson  <https://orcid.org/0000-0001-7077-1618>

### Supplemental material

Supplemental material for this article is available online.

### References

- Abell, F., Happé, F., & Frith, U. (2000). Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development. *Cognitive Development, 15*(1), 1–16. [https://doi.org/10.1016/S0885-2014\(00\)00014-9](https://doi.org/10.1016/S0885-2014(00)00014-9)
- 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 Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 42*(2), 241–251. <https://doi.org/10.1111/1469-7610.00715>
- Brunsdon, V. E., & Happe, F. (2014). Exploring the 'fractionation' of autism at the cognitive level. *Autism, 18*(1), 17–30. <https://doi.org/10.1177/1362361313499456>
- Castelli, F., Frith, C., Happé, F., & Frith, U. (2002). Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain, 125*(8), 1839–1849. <https://doi.org/10.1093/brain/awf189>
- Chung, Y. S., Barch, D., & Strube, M. (2014). A meta-analysis of mentalizing impairments in adults with schizophrenia and autism spectrum disorder. *Schizophrenia Bulletin, 40*(3), 602–616. <https://doi.org/10.1093/schbul/sbt048>
- Clemmensen, L., Bartels-Velthuis, A. A., Jespersen, R. A., van Os, J., Blijd-Hoogewys, E. M., Ankerstrøm, L., . . . Jepsen, J. R. (2016). A psychometric evaluation of the Danish version of the Theory of Mind Storybook for 8-14 year-old children. *Frontiers in Psychology, 7*, 330. <https://doi.org/10.3389/fpsyg.2016.00330>
- Devine, R. T., & Hughes, C. (2016). Measuring theory of mind across middle childhood: Reliability and validity of the Silent Films and Strange Stories tasks. *Journal of Experimental Child Psychology, 149*, 23–40. <https://doi.org/10.1016/j.jecp.2015.07.011>
- Frith, U. (1994). Autism and theory of mind in everyday life. *Social Development, 3*(2), 108–124. <https://doi.org/10.1111/j.1467-9507.1994.tb00031.x>
- Frith, U. (2001). Mind blindness and the brain in autism. *Neuron, 32*(6), 969–979. [https://doi.org/10.1016/s0896-6273\(01\)00552-9](https://doi.org/10.1016/s0896-6273(01)00552-9)
- Gernsbacher, M. A., & Yergeau, M. (2019). Empirical failures of the claim that autistic people lack a theory of mind. *Archives of Scientific Psychology, 7*(1), 102–118. <https://doi.org/10.1037/arc0000067>
- Happé, F. G. (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 and Developmental Disorders, 24*(2), 129–154. <https://doi.org/10.1007/BF02172093>
- Happé, F. G., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 36*(1), 5–25. <https://doi.org/10.1007/s10803-005-0039-0>
- Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods, 1*(1), 39–65. <https://doi.org/10.1002/jrsm.5>

- Klin, A., Lin, D. J., Gorrindo, P., Ramsay, G., & Jones, W. (2009). Two-year-olds with autism orient to non-social contingencies rather than biological motion. *Nature*, *459*(7244), 257–261. <https://doi.org/10.1038/nature07868>
- Livingston, L. A., Shah, P., & Happé, F. (2019, November 17). Further developing the Frith-Happé animations: A quicker, more objective, and web-based test of theory of mind for autistic and neurotypical adults. <https://doi.org/10.31234/osf.io/grwd6>
- Lombardo, M. V., Lai, M., Auyeung, B., Holt, R. J., Allison, C., & Baron-Cohen, S. (2016). Unsupervised data-driven stratification of mentalizing heterogeneity in autism. *Scientific Reports*, *6*, 35333. <https://doi.org/10.1038/srep35333>
- Milne, E., Swettenham, J., Hansen, P., Campbell, R., Jeffries, H., & Plaisted, K. (2002). High motion coherence thresholds in children with autism. *Journal of Child Psychology and Psychiatry*, *43*(2), 255–263. <https://doi.org/10.1111/1469-7610.00018>
- Moessnang, C., Baumeister, S., Tillmann, J., Goyard, D., Charman, T., Ambrosino, S., . . . EU-AIMS LEAP Group. (2020). Social brain activation during mentalizing in a large autism cohort: The Longitudinal European Autism Project. *Molecular Autism*, *11*(1), 17. <https://doi.org/10.1186/s13229-020-0317-x>
- Oakley, B., Brewer, R., Bird, G., & Catmur, C. (2016). Theory of mind is not theory of emotion: A cautionary note on the Reading the Mind in the Eyes Test. *Journal of Abnormal Psychology*, *125*(6), 818–823. <https://doi.org/10.1037/abn0000182>
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Robertson, C. E., Thomas, C., Kravitz, D. J., Wallace, G. L., Baron-Cohen, S., Martin, A., & Baker, C. I. (2014). Global motion perception deficits in autism are reflected as early as primary visual cortex. *Brain*, *137*(9), 2588–2599. <https://doi.org/10.1093/brain/awu189>
- Schaafsma, S. M., Pfaff, D. W., Spunt, R. P., & Adolphs, R. (2015). Deconstructing and reconstructing theory of mind. *Trends in Cognitive Sciences*, *19*(2), 65–72. <https://doi.org/10.1016/j.tics.2014.11.007>
- Schurz, M., Radua, J., Aichhorn, M., Richlan, F., & Perner, J. (2014). Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neuroscience and Biobehavioral Reviews*, *42*, 9–34. <https://doi.org/10.1016/j.neubiorev.2014.01.009>
- Van der Hallen, R., Manning, C., Evers, K., & Wagemans, J. (2019). Global motion perception in autism spectrum disorder: A meta-analysis. *Journal of Autism and Developmental Disorders*, *49*(12), 4901–4918. <https://doi.org/10.1007/s10803-019-04194-8>
- Velikonja, T., Fett, A. K., & Velthorst, E. (2019). Patterns of nonsocial and social cognitive functioning in adults with autism spectrum disorder: A systematic review and meta-analysis. *JAMA Psychiatry*, *76*(2), 135–151. <https://doi.org/10.1001/jamapsychiatry.2018.3645>
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48. <https://www.jstatsoft.org/v36/i03/>
- White, S. J., Coniston, D., Rogers, R., & Frith, U. (2011). Developing the Frith-Happé Animations: A quick and objective test of theory of mind for adults with autism. *Autism Research*, *4*(2), 149–154. <https://doi.org/10.1002/aur.174>

## References included in meta-analysis

- Bal, E., Yerys, B. E., Sokoloff, J. L., Celano, M. J., Kenworthy, L., Giedd, J. N., & Wallace, G. L. (2013). Do social attribution skills improve with age in children with high functioning autism spectrum disorders? *Research in Autism Spectrum Disorders*, *7*(1), 9–16. <https://doi.org/10.1016/j.rasd.2012.07.004>
- Barzy, M., Ferguson, H. J., & Williams, D. (2019). Perspective influences eye movements during real-life conversation: Mentalising about self vs. others in autism. <https://doi.org/10.31234/osf.io/7druh>
- Barzy, M., Filik, R., Williams, D., & Ferguson, H. J. (2020). Emotional processing of ironic versus literal criticism in autistic and nonautistic adults: Evidence from eye-tracking. *Autism Research*, *13*(4), 563–578. <https://doi.org/10.1002/aur.2272>
- Brewer, N., Young, R. L., & Barnett, E. (2017). Measuring theory of mind in adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *47*(7), 1927–1941. <https://doi.org/10.1007/s10803-017-3080-x>
- Brunsdon, V. E., Colvert, E., Ames, C., Garnett, T., Gillan, N., Hallett, V., . . . Happé, F. (2015). Exploring the cognitive features in children with autism spectrum disorder, their co-twins, and typically developing children within a population-based sample. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *56*(8), 893–902. <https://doi.org/10.1111/jcpp.12362>
- Campbell, R., Lawrence, K., Mandy, W., Mitra, C., Jeyakuma, L., & Skuse, D. (2006). Meanings in motion and faces: Developmental associations between the processing of intention from geometrical animations and gaze detection accuracy. *Development and Psychopathology*, *18*(1), 99–118. <https://doi.org/10.1017/S0954579406060068>
- Cantio, C., Jepsen, J. R., Madsen, G. F., Bilenberg, N., & White, S. J. (2016). Exploring ‘the autisms’ at a cognitive level. *Autism Research*, *9*(12), 1328–1339. <https://doi.org/10.1002/aur.1630>
- Castelli, F., Frith, C., Happé, F., & Frith, U. (2002). Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain*, *125*(8), 1839–1849. <https://doi.org/10.1093/brain/awf189>
- Clemmensen, L., Bartels-Velthuis, A. A., Jespersen, R. A., van Os, J., Blijd-Hoogewys, E. M., Ankerstrøm, L., . . . Jepsen, J. R. (2016). A psychometric evaluation of the Danish version of the Theory of Mind Storybook for 8-14 year-old children. *Frontiers in Psychology*, *7*, 330. <https://doi.org/10.3389/fpsyg.2016.00330>
- Donno, R., Parker, G., Gilmour, J., & Skuse, D. H. (2010). Social communication deficits in disruptive primary-school children. *The British Journal of Psychiatry*, *196*(4), 282–289. <https://doi.org/10.1192/bjp.bp.108.061341>
- Ferguson, H. J., Black, J., & Williams, D. (2019). Distinguishing reality from fantasy in adults with autism spectrum disorder: Evidence from eye movements and reading. *Journal*

- of *Memory and Language*, 106, 95–107. <https://doi.org/10.1016/j.jml.2019.03.001>
- Jones, A. P., Happé, F. G., Gilbert, F., Burnett, S., & Viding, E. (2010). Feeling, caring, knowing: Different types of empathy deficit in boys with psychopathic tendencies and autism spectrum disorder. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 51(11), 1188–1197. <https://doi.org/10.1111/j.1469-7610.2010.02280.x>
- Jones, C. R., Swettenham, J., Charman, T., Marsden, A. J., Tregay, J., Baird, G., . . . Happé, F. (2011). No evidence for a fundamental visual motion processing deficit in adolescents with autism spectrum disorders. *Autism Research*, 4(5), 347–357. <https://doi.org/10.1002/aur.209>
- Kirkovski, M., Enticott, P. G., Hughes, M. E., Rossell, S. L., & Fitzgerald, P. B. (2016). Atypical neural activity in males but not females with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 46(3), 954–963. <https://doi.org/10.1007/s10803-015-2639-7>
- Landsiedel, J., & Williams, D. M. (2020). Increasing extrinsic motivation improves time-based prospective memory in adults with autism: Relations with executive functioning and mentalizing. *Journal of Autism and Developmental Disorders*, 50(4), 1133–1146. <https://doi.org/10.1007/s10803-019-04340-2>
- Lind, S. E., Williams, D. M., Bowler, D. M., & Peel, A. (2014). Episodic memory and episodic future thinking impairments in high-functioning autism spectrum disorder: An underlying difficulty with scene construction or self-projection? *Neuropsychology*, 28(1), 55–67. <https://doi.org/10.1037/neu0000005>
- Livingston, L. A., Shah, P., & Happé, F. (2019, November 17). Further developing the Frith-Happé animations: A quicker, more objective, and web-based test of theory of mind for autistic and neurotypical adults. <https://doi.org/10.31234/osf.io/grwd6>
- Lugnegård, T., Unenge Hallerbäck, M., Hjärthag, F., & Gillberg, C. (2013). Social cognition impairments in Asperger syndrome and schizophrenia. *Schizophrenia Research*, 143(2–3), 277–284. <https://doi.org/10.1016/j.schres.2012.12.001>
- Marsh, L. E., & Hamilton, A. F. (2011). Dissociation of mirroring and mentalising systems in autism. *NeuroImage*, 56(3), 1511–1519. <https://doi.org/10.1016/j.neuroimage.2011.02.003>
- Martinez, G., Mosconi, E., Daban-Huard, C., Parellada, M., Fananas, L., Gaillard, R., . . . Amado, I. (2019). ‘A circle and a triangle dancing together’: Alteration of social cognition in schizophrenia compared to autism spectrum disorders. *Schizophrenia Research*, 210, 94–100. <https://doi.org/10.1016/j.schres.2019.05.043>
- Moessnang, C., Baumeister, S., Tillmann, J., Goyard, D., Charman, T., Ambrosino, S., . . . EU-AIMS LEAP Group. (2020). Social brain activation during mentalizing in a large autism cohort: The Longitudinal European Autism Project. *Molecular Autism*, 11(1), 17. <https://doi.org/10.1186/s13229-020-0317-x>
- Murray, K., Johnston, K., Cunnane, H., Kerr, C., Spain, D., Gillan, N., . . . Happé, F. (2017). A new test of advanced theory of mind: The ‘Strange Stories Film Task’ captures social processing differences in adults with autism spectrum disorders. *Autism Research*, 10(6), 1120–1132. <https://doi.org/10.1002/aur.1744>
- Nicholson, T., Williams, D., Carpenter, K., & Kallitsounaki, A. (2019). Interoception is impaired in children, but not adults, with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 49(9), 3625–3637. <https://doi.org/10.1007/s10803-019-04079-w>
- Nicholson, T., Williams, D. M., Grainger, C., Lind, S. E., & Carruthers, P. (2019). Relationships between implicit and explicit uncertainty monitoring and mindreading: Evidence from autism spectrum disorder. *Consciousness and Cognition*, 70, 11–24. <https://doi.org/10.1016/j.concog.2019.01.013>
- Salter, G., Seigal, A., Claxton, M., Lawrence, K., & Skuse, D. (2008). Can autistic children read the mind of an animated triangle? *Autism*, 12(4), 349–371. <https://doi.org/10.1177/1362361308091654>
- Veddem, L., Pedersen, H. L., Landert, A. L., & Bliksted, V. (2019). Do patients with high-functioning autism have similar social cognitive deficits as patients with a chronic cause of schizophrenia? *Nordic Journal of Psychiatry*, 73(1), 44–50. <https://doi.org/10.1080/08039488.2018.1554697>
- White, S. J., Coniston, D., Rogers, R., & Frith, U. (2011). Developing the Frith-Happé Animations: A quick and objective test of theory of mind for adults with autism. *Autism Research*, 4(2), 149–154. <https://doi.org/10.1002/aur.174>
- Williams, D. M., Nicholson, T., & Grainger, C. (2018). The self-reference effect on perception: Undiminished in adults with autism and no relation to autism traits. *Autism Research*, 11(2), 331–341. <https://doi.org/10.1002/aur.1891>
- Williams, D., Boucher, J., Lind, S., & Jarrold, C. (2013). Time-based and event-based prospective memory in autism spectrum disorder: The roles of executive function and theory of mind, and time-estimation. *Journal of Autism and Developmental Disorders*, 43(7), 1555–1567. <https://doi.org/10.1007/s10803-012-1703-9>
- Wilson, C. E., Happé, F., Wheelwright, S. J., Ecker, C., Lombardo, M. V., Johnston, P., . . . Murphy, D. G. (2014). The neuropsychology of male adults with high-functioning autism or Asperger syndrome. *Autism Research*, 7(5), 568–581. <https://doi.org/10.1002/aur.1394>
- Wilson, A. C., & Bishop, D. (2020). Judging meaning: A domain-level difference between autistic and non-autistic adults. *Royal Society Open Science*, 7(11), 200845. <https://doi.org/10.1098/rsos.200845>
- Young, R. L., & Brewer, N. (2020). Brief report: Perspective taking deficits, autism spectrum disorder, and allaying police officers’ suspicions about criminal involvement. *Journal of Autism and Developmental Disorders*, 50(6), 2234–2239. <https://doi.org/10.1007/s10803-019-03968-4>
- Zwicker, J., White, S. J., Coniston, D., Senju, A., & Frith, U. (2011). Exploring the building blocks of social cognition: Spontaneous agency perception and visual perspective taking in autism. *Social Cognitive and Affective Neuroscience*, 6(5), 564–571. <https://doi.org/10.1093/scan/nsq088>