A Meta-Analysis of Cognitive Flexibility in Autism Spectrum Disorder

Claudia Lage [a*], Eleanor S. Smith[a], Rebecca P. Lawson[a]

[a] Department of Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, United Kingdom

*Corresponding author at: Department of Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, United Kingdom
Email address: cl821@cam.ac.uk

Abstract

Cognitive flexibility is a fundamental cognitive process that underlies adaptive behaviour in response to behavioural change. Studies examining the profile of cognitive flexibility in autism spectrum disorder (ASD) have reported inconsistent findings. To address whether difficulties with cognitive flexibility are characteristic of autism we conducted a random-effects meta-analysis and employed subgroup analyses and meta-regression to assess the impact of relevant moderator variables such as task, outcomes and age. Fifty-five studies were included and comprised of 1901 autistic individuals without intellectual disabilities and 1846 neurotypical controls, with an age range of 4 to 85 years. The results showed that autistic individuals have greater difficulties with cognitive flexibility, with an overall statistically significant small to moderate effect size. Subgroup analyses revealed a significant difference between different task outcomes, with perseverative errors obtaining the largest effect size. In summary, the present meta-analysis highlights the existence of cognitive flexibility difficulties in ASD, in the absence of learning disabilities, but also that this profile is characterised by substantial heterogeneity. Potential contributing factors are discussed.

Keywords: Autism Spectrum Disorder, Cognitive Flexibility, Set Shifting, Meta-Analysis

1. Introduction

Autism Spectrum Disorder (ASD) is a neurodevelopmental condition defined by difficulties in social communication and interaction, along with restricted and repetitive behaviours (DSM-5, American Psychiatric Association, 2013). Although a clear understanding of the causes of ASD remains elusive, several genetic and neurobiological factors have been identified (Lord et al., 2020). Several cognitive theories have also been proposed to underpin the ASD phenotype (e.g. Baron-Cohen et al., 1985; Frith, 1989). The executive dysfunction account of autism, originally proposed by Damasio and Maurer (1978), drew parallels between the rigid and perseverative behaviours of patients with frontal lobe lesions and those diagnosed with autism. Since then, extensive research over the last four decades has highlighted the existence of executive function (EF) difficulties in ASD (e.g. Rumsey, 1985; Alsaedi et al., 2020) and neuroimaging studies have demonstrated structural and functional alterations in frontal regions (Catani et al., 2016; Ecker, 2017; Libero et al., 2015). Although no consensus exists, EF is often divided into the subdomains of working memory, cognitive flexibility, planning and inhibition, and evidence suggests that EF subdomains are separable, yet correlated (Miyake & Friedman, 2012).

Cognitive flexibility, also referred to as set shifting, is the capacity to intentionally shift between different mental tasks or strategies and adjust responses according to environmental changing contingencies (Dajani & Uddin, 2015). In ASD, difficulties in cognitive flexibility have been associated with increased social difficulties (Berger et al., 2003), increased restricted and repetitive behaviours (RRBs) (Faja & Darling, 2019; Miller et al., 2015) and co-occurring symptoms such as anxiety and low mood (Crawley et al., 2020; Ozsivadjian et al., 2021). Growing evidence also suggests a key role for cognitive flexibility in outcomes such as academic achievement (St John et al., 2018), adaptive behaviour (Bertollo et al., 2020) and

quality of life (de Vries & Geurts, 2015). Given that a substantial proportion of autistic adults without intellectual disabilities report lower rates of employment and independent living (Anderson et al., 2014; Frank et al., 2018), fewer relationships (Farley et al., 2009) and reduced quality of life (Mason et al., 2018), it is vital that we enhance our understanding of cognitive flexibility in autism, across the lifespan.

A clear profile of cognitive flexibility in ASD remains elusive, due to inconsistency across studies. Narrative reviews highlight the disparity of cognitive flexibility difficulties and methodological heterogeneity affecting outcomes and the interpretation of findings (Geurts et al., 2009; Hill, 2004; Russo et al., 2007). The most consistent pattern of results has been obtained with the Wisconsin Card Sorting Test (WCST) (Berg, 1948), with higher perseveration, i.e. the tendency to become stuck in set and persist with the same sorting strategy despite incorrect feedback, thought to specifically reflect cognitive flexibility difficulties in ASD (Landry & Al-Taie, 2016). A meta-analysis of EF among autistic children and adolescents, identified cognitive flexibility as one of the core difficulties, with a moderate effect size that remained significant even after co-morbid ADHD and IQ were controlled for (Lai et al., 2017). Nevertheless, there was substantial variability of effect sizes across studies. Similarly, in autistic adults, despite EF difficulties across all subdomains, cognitive flexibility was predominantly affected (Xie et al., 2020). Furthermore, Demetriou et al. (2018) examined the profile of EF difficulties across the lifespan in ASD and obtained an overall moderate effect size. The only quantitative review to date focused specifically on cognitive flexibility in ASD (Leung & Zakzanis, 2014) reported extensive variation in the magnitude of cognitive flexibility difficulties. However, this prior work is limited by the inclusion of participants with learning disabilities, self and parent-report measures, and also a broad range of tasks (e.g. set shifting, task switching and inhibitory control tasks), which together call into question the conclusions that can be drawn regarding cognitive flexibility in ASD.

Several factors should be considered in light of the heterogeneity across studies, including the influence of task and sample characteristics. The inconsistent operationalisation of cognitive flexibility and the myriad of different paradigms (Leung & Zakzanis, 2014) likely contribute to the observed variability. Also, specific task features such as administration format (Demetriou et al., 2019) and type of instructions (Van Eylen et al., 2015) could moderate performance. The maturation of cognitive flexibility follows an inverted U-shaped curve, with a sharp increase during childhood, reaching a peak in early adulthood and deteriorating later in life (Zelazo et al., 2004). However, in ASD, evidence suggests that this pattern is more complex, with higher interindividual variability found throughout development (Van Eylen et al., 2011). The prefrontal cortex (PFC) plays an important role in the development of cognitive flexibility (Buttelman & Karbach, 2017). For instance, in a near-infrared spectroscopy study, during a set shifting task, neurotypical adults exhibited significant activation in the inferior PFC bilaterally and a similar pattern was observed in 5-year-old children (Moriguchi & Hiraki, 2009). However, 3-year-olds who perseverated did not demonstrate significant activation in the right or left inferior PFC throughout the task. Some evidence suggests that autistic adolescents perform better than autistic children in cognitive flexibility tasks (D'Cruz et al., 2013; Van Eylen et al., 2015), and Lai et al. (2017) found a decrease in effect size by 0.062 for each year of increase in mean age. It is possible that greater maturational differences in ASD, due to a protracted development, could explain some of the variability within and across studies and highlight the importance of examining cognitive flexibility within a developmental framework.

In summary, our understanding of cognitive flexibility in ASD remains limited, despite growing evidence suggesting a key role across several outcomes that are important for educational attainment, quality of life and mental health. Also, there is a lack of quantitative reviews focused on cognitive flexibility across the lifespan in ASD. The present meta-analysis

will address previous limitations, and explore heterogeneity with an assessment of how tasks, outcomes and age act as moderating factors. It is hypothesised that autistic people will exhibit significantly greater difficulties in cognitive flexibility compared to typically developing controls. An exploratory approach regarding the influence of key moderating variables will be taken.

2. Methods

2.1. Eligibility Criteria

This systematic review and meta-analysis followed the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines (Moher et al., 2009). The inclusion criteria were defined a priori using the PICOS components below, as recommended by the PRISMA statement.

2.1.1. Participants

Participants with a diagnosis of ASD based on DSM or ICD criteria and/or a diagnosis using other valid diagnostic instruments, such as the Autism Diagnostic Observation Schedule (Lord et al., 2012) and Autism Diagnostic Interview (Rutter et al., 2003) were included. Subjects with learning disabilities (IQ < 70), related medical conditions (e.g. fragile X syndrome), and neurological disorders (e.g. epilepsy) were excluded. No restrictions regarding age were applied.

2.1.2. Interventions/Outcomes

Studies assessing cognitive flexibility using standard set shifting paradigms were included (see below for details of specific tasks). Considering the heterogeneity across studies, to include other types of paradigms would introduce another source of heterogeneity for the quantitative synthesis of the evidence that could obscure vital distinctions at an outcome level and generate misleading results (Cooper, 2017). Additionally, set shifting paradigms are the most frequently used across the literature, with perseveration, thought to specifically reflect difficulties in cognitive flexibility (Hill, 2004; Landry & Al-Taie, 2016).

2.1.3. Comparators

Participants with no reported history of ASD, neurological or psychiatric conditions, without learning disabilities (IQ ≥ 70), matched on at least one IQ measure (e.g. nonverbal IQ) were included as neurotypical (NT) controls.

2.1.4. Study Design

Both cross-sectional and longitudinal designs were included, and for the latter only the baseline data was considered, to avoid practice effects. No other restrictions were applied.

2.2. Literature Search

Four separate electronic searches were performed using Pubmed, Embase, PsycInfo and Scopus as databases and a combination of the following terms: ‘autism’ and variations thereof, ‘Asperger’, ‘pervasive developmental disorder’ and variations thereof, ‘cognitive flexibility’ and variations thereof, ‘executive functioning’ and variations thereof, ‘set-shift’ and several paradigms. The full search strategy is available in Appendix A. The searches were limited to studies in humans, published in English between 1980 (first inclusion of autism diagnosis in

the DSM-III) and June 2020. Manual searches were performed independently by two authors (CL and ESS) using the reference lists of included studies and previous systematic reviews.

2.3. Study Selection

After de-duplication, all records obtained from the electronic searches were sequentially screened by title and abstract. Subsequently, the full text of the remaining articles was examined independently by two authors (CL and ESS) using a piloted eligibility criteria checklist of the PICOS components.

2.4. Data Extraction

The following data was extracted from each included study: age range, mean age and standard deviation, task, sample size, number of males and females, diagnostic method, IQ measurement tool, matching criteria (sex, age, IQ), and mean full-scale, performance/nonverbal and verbal IQs, where applicable. Outcomes were extracted as means and standard deviations for the ASD and NT control group at a single time point. Studies with more than one experiment with no overlap in participants were included separately. Table 1 has a summary of all included studies.

2.5. Set shifting Paradigms

2.5.1. Intra-Extra Dimensional Set-Shift (IED)

The IED from the Cambridge Neuropsychological Test Automated Battery (CANTAB) is a test involving rule acquisition and reversal with two dimensions, i.e. pink shapes and white lines. Participants must choose the correct stimulus based on feedback, and after six correct trials the rule changes. In the first stages of the test the rule changes are intra-dimensional, and later on extra-dimensional, i.e. the white lines become the correct dimension.

2.5.2. Wisconsin Card Sorting Test (WCST)

In the traditional WCST, participants must sort a total of 128 cards according to three categories, i.e. colour, shape and number. The sorting rule must be inferred based on feedback and after ten consecutive correct trials the rule changes without warning. Variations of the WCST, include reduced number of total cards, two sorting categories instead of three and different number of correct trials until the rule changes. Details of the variations included in each respective study are available in Table 1.

2.5.3. Modified Card Sorting Test (MCST)

Participants must sort a total of 48 cards according to three dimensions, i.e. shape, number and colour. The sorting rule must be inferred based on feedback and after six consecutive correct trials, participants are told that the rule has changed.

2.5.4. Two-Choice Reversal Learning Task (2CRL)

Participants are shown two identical stimuli and must choose which one is in the correct location. Feedback is given indicating whether each choice was right or wrong. Without warning, a reversal of the correct location occurs after four, five or six consecutive correct trials.

2.5.5. Probabilistic Reversal Learning Task (PRL)

Participants are shown two identical stimuli and must choose which one is in the correct location. The feedback given is probabilistic, i.e. 80 percent of correct and 20 percent of incorrect responses are randomly reinforced. In the acquisition phase, after eight out of ten consecutive correct trials, without warning, there is a reversal of the correct location.

2.5.6. The Penn Conditional Exclusion Test (PCET)

This test involves four stimuli, in which three are matched based on either shape, size or line thickness. Participants must infer the sorting rule based on feedback and select the stimulus that does not belong. After ten consecutive correct trials, the sorting rule changes without warning.

2.5.7. Set Shifting Task (SST)

In this task, children must work out which cards are the teddy's favourite and sort them according to one of three dimensions, i.e. colour, shape or size. The sorting rule must be inferred based on feedback and after six consecutive correct trials, the rule changes with the presentation of a new deck of cards and a different teddy.

2.5.8. Card Sorting Task (CST)

In this task, participants must match different target stimuli with reference stimuli according to either colour or shape. Participants are given cues that indicate if the matching rules must be repeated or changed.

2.5.9. Computerised Sequencing Game (CSG)

Eight stimuli are presented that differ on two colours and two shapes in each trial. Participants must identify the correct sequence and the rules must be inferred based on feedback. In each trial there is a reversal halfway through, i.e. the first four stimuli are sequenced based on one rule and the subsequent four based on another.

2.5.10. Probabilistic Selection Task (PST)

Three pairs of stimuli are presented in a randomised order and participants must choose which character is most likely to be correct. Probabilistic feedback is given, i.e. for the AB pair, A receives correct feedback 80 percent of the time and B is incorrectly reinforced 20

percent of the time. For the pairs CD and EF, the probabilistic reinforcement schedules are 70:30 and 60:40, respectively.

2.6. Data Analysis

All analyses were performed in R (version 4.0.2) and the packages 'meta' and 'metafor' were used. The standardised mean difference (SMD) between the ASD and NT control group was calculated as Hedges' g (Hedges, 1980). Effect sizes were pooled together using a random-effects model (DerSimonian-Laird estimator for Tau^2) due to heterogeneity between studies. A random-effects model takes into account sampling error and between-study variance (Tau^2) when assigning weights, thus assuming a distribution of effect sizes (Cooper, 2017). A positive Hedges' g indicates that the NT group performed better than the ASD group and the same effect size convention as with Cohen's d was applied, namely g = 0.20 is small, g = 0.50 is medium, g = 0.80 is large (Cohen, 1988). Heterogeneity was assessed using the Q-statistic, with a p-value ≤ 0.10 indicating significant heterogeneity, i.e. variance in effect sizes is not due to sampling error (chance) alone (Higgins et al., 2019). Additionally, the I^2 statistic gives the percentage of total variance in effect sizes that is due to between-study heterogeneity (rather than sampling error), interpreted as follows: I^2 = 25% is low, I^2 = 50% is moderate, and I^2 = 75% is high (Higgins & Thompson, 2002). Heterogeneity was explored and subgroup analyses of task, outcome and age were performed. For the latter, studies were categorised based on mean age reported: 'children $\leq$ 12', 'adolescents >12 <18', and 'adults $\geq$ 18'. A meta-regression with mean age as a moderator was also performed.

2.6.1. Publication Bias and Quality Assessment

Publication bias was assessed with a funnel plot visually inspected for asymmetry and formally evaluated with Egger's test (Egger et al., 1997). The Egger's test quantifies the asymmetry with $p \leq 0.05$ showing significant asymmetry. The Newcastle-Ottawa Scale for

non-randomised case-control studies (Wells et al., 2000) was used to assess the quality of each study. This scale contains a total of eight items covering selection, comparability, and exposure. A star rating system is used, with the highest quality studies receiving one star per item and two stars in the comparability category and up to a total of nine stars.

3. Results

3.1. Study Selection and Characteristics

A total of 6028 records were identified through database searching and an additional 13 through manual searching. The PRISMA flow diagram presented in Figure 1 provides details of the study selection process. Fifty-five studies were included in the quantitative synthesis and comprised of 3747 participants in total, of which 1901 were diagnosed with ASD and 1846 were NT controls. Ages ranged from 4 to 85 years, with a mean age of 17.8 for the ASD group and 18.1 for the neurotypical group.

3.2. Primary Meta-analysis

The pooled SMD between ASD and NT was small to moderate and statistically significant ($g = 0.45$, 95% CI 0.33-0.57, $p < 0.0001$), i.e. autistic participants had significantly more difficulties in cognitive flexibility compared to controls. However, there was significant moderate heterogeneity between studies ($I^2 = 66\%$, $p < 0.01$). As shown by the forest plot in Figure 2, there is substantial variance in effect sizes, ranging from $g = -0.99$ to $g = 2.22$.

3.3. Subgroup Analyses

3.3.1. Effect of Task

There was no significant subgroup difference between tasks ($p = 0.19$), however the WCST had the largest effect size ($g = 0.57$, 95% CI 0.36-0.79). Eleven studies were excluded because there were four or less of each task to form a subgroup. Results are summarised in Table 2.

3.3.2. Effect of Outcome Measure

A significant subgroup difference between outcomes ($p = 0.04$) was found, with perseverative errors showing the largest effect size ($g = 0.55$, 95% CI 0.40-0.70). Perseverative errors and extra-dimensional shift errors formed a subgroup each; all other outcomes due to insufficient numbers were grouped together. Results are summarised in Table 3.

3.3.3. Effect of Age

No significant difference was found between subgroups of children (≤ 12), adolescents (>12 <18) and adults (≥ 18) ($p = 0.27$). The adult subgroup had the largest effect size ($g = 0.53$, 95% CI 0.33-0.73). Results are summarised in Table 4.

3.4. Meta-regression

Age was not a significant moderator ($B = 0.003$, $Se = 0.007$, $p = 0.63$) and did not explain any of the variance in effect sizes ($R^2 = 0\%$). One study was not included in the analysis because there was a significant age difference between the ASD and NT group (Geurts et al., 2020). A meta-regression plot is presented in Figure 3.

3.5. Publication Bias

The funnel plot is presented in Figure 4. The majority of studies is scattered relatively evenly around the pooled effect size, although some asymmetry can be observed. However, the Egger's test result ($p = 0.06$) showed that there is no significant asymmetry, i.e. it can be assumed that there is minimal publication bias and that the pooled effect size is representative.

3.6. Quality Assessment

The majority of studies were rated as having overall adequate quality. Results are summarised in Table 5.

4. Discussion

The present meta-analysis showed that autistic people, without learning disabilities, had significantly more difficulties in cognitive flexibility compared to neurotypical controls. However, there was also significant moderate heterogeneity between studies and it is possible that this variance could at least be partially explained by a broad profile of cognitive flexibility difficulties in ASD. Subgroup analyses revealed a significant difference between outcomes, with perseverative errors obtaining the largest effect size, consistent with previous results (e.g. Landry & Al-Taie, 2016). Although a large body of research has demonstrated that perseveration constitutes a difficulty in ASD, it would be misleading to conclude that cognitive flexibility difficulties are due to preservation alone, as more studies are needed to compare the sensitivity of other outcome measures. For instance, in D'Cruz et al. (2013), there was no significant difference in perseverative errors, however the autistic group made significantly more regressive errors, which were positively correlated with behavioural rigidity, indexed by clinical ratings of RRBs.

Additionally, in the subgroup analyses, no significant differences between tasks were found, nonetheless, the WCST obtained the largest effect size, consistent with previous results (e.g. Leung & Zakzanis, 2014). The specificity of the WCST is often debated (e.g. Nyhus & Barcelo, 2009) as it relies on multiple EF domains, including working memory and inhibition (Russo et al., 2007). This multifactorial EF demand has been highlighted in neuroimaging research, demonstrating fractionation of cognitive components that are integrated to perform the task (Buchsbaum et al., 2005). Notwithstanding the criticism, as evidence suggests that EF is divided into separable, yet correlated component processes (Miyake & Friedman, 2012), it may not be feasible to develop a completely 'pure' cognitive flexibility task. Nonetheless, Van Eylen et al. (2015), found that even after working memory and inhibition difficulties were controlled for, autistic people still showed significantly more difficulties in cognitive flexibility, with higher perseveration.

As in other meta-analyses (e.g. Lai et al., 2017), the IED task obtained the smallest effect size. It has been suggested that the differences captured by the IED could be partially due to difficulties in sustaining attention, as cognitive flexibility is only assessed at the end of the task when the extra-dimensional shift occurs (Geurts et al., 2009). Consistently, Sinzig et al. (2008), compared the performance of autistic children with and without co-occurring ADHD and found that only the former showed difficulties in the task. In contrast with other tasks such as the WCST that assess cognitive flexibility throughout, it is possible that the IED's stepwise design with the extra-dimensional shift at the end, is not able to fully capture the extent of cognitive difficulties in ASD. However, it is also important to consider that the IED task terminates after 50 trials on any stage, if the learning criterion of six consecutive correct responses is not achieved (Downes et al., 1989). In this sense, it is possible that participants with greater EF difficulties do not reach stage eight, where the extra-dimensional shift occurs. An adjustment is therefore described in the CANTAB administration guide, in which 25 errors should be

added for each failed stage of the task. If studies failed to adjust errors in this manner, it would lead to differences between participants being obscured.

The measurement of cognitive flexibility is challenging due to its inherent complexity, intricate relationships with other EF domains and changeable nature throughout the lifespan. In addition to the myriad of different paradigms, specific administration factors, such as type of instructions given can affect cognitive flexibility (Van Eylen et al., 2011), however this level of detail is rarely reported in studies. Furthermore, the ecological validity of neuropsychological tasks is often debated, as they do not always converge with self-report measures (Toplak et al., 2013). Despite the apparent face-validity, there is a discrepancy between cognitive flexibility difficulties and the prominent 'real-world' behavioural flexibility challenges in ASD, referred to as the 'paradox of cognitive flexibility in autism' (Geurts et al., 2009). Moving forward, more research is needed to address this 'paradox', using cognitive tasks that confer greater ecological validity in combination with self-report measures to enable a more comprehensive investigation of flexibility in ASD.

Reversal learning (RL) tasks are a widely used translational paradigm to index flexibility (Uddin et al., 2021) and may offer a better trade-off between construct and ecological validity. For instance, in probabilistic RL (D'Cruz et al., 2013; Weiss et al., 2021) the uncertainty in the task is captured by the probabilistic reinforcement schedules, and thus the ability to learn about this uncertainty and flexibly respond to variable contingencies, more closely resembles the 'real-world' flexibility demands of continuously changing environments. Additionally, RL task responses are amenable to computational modeling, which can reveal the latent mechanisms that drive behavioural differences, however there is a paucity of studies adopting computational modeling in autism at present (though see Crawley et al., 2020; Lawson et al., 2017; Manning et al., 2016 for notable exceptions). Cognitive flexibility difficulties in probabilistic RL tasks

have been associated with RRBs in autism (Crawley et al., 2020; D'Cruz et al., 2013). One possibility is that measured difficulties in probabilistic RL tasks could be due to higher response monitoring requirements, i.e. the ability to evaluate behavioural consequences and adjust accordingly to optimise outcomes (Thakkar et al., 2008). It has been suggested that in autistic people, structural and functional alterations of the anterior cingulate cortex (ACC) might underlie response monitoring difficulties and thus contribute to behavioural rigidity (Thakkar et al., 2008). Future neuroimaging studies should explore the role of the ACC during RL tasks and the link with RRBs among autistic individuals.

The meta-regression showed that age did not account for any of the variance between studies and in the subgroup analysis there was no significant difference between the three age groups. It has been proposed that due to a protracted development of cognitive flexibility in ASD, initial differences in cognitive difficulties might be more pronounced in childhood but gradually reduce with increasing age (Van Eylen et al., 2015). Also, with time autistic individuals may become more adept at employing compensatory strategies (Demetriou et al., 2018). Some neuroimaging studies have demonstrated altered brain activity among autistic individuals in the absence of significant differences in task performance (D'Cruz et al., 2016; Schmitz et al., 2006), which might suggest that distinct neurobiological strategies were being employed to achieve the same behavioural outcome. Additionally, cognitive flexibility difficulties in adolescents showed the lowest effect size in our meta-analysis, which is consistent with the findings of Lai et al. (2017). Adolescence is a developmental period of substantial neural changes, including synaptic reorganisation, that could contribute to the reduced differences between autistic and typically developing adolescents (Blakemore & Choudhury, 2006). However, little is known about the developmental trajectory of cognitive flexibility in ASD and more longitudinal studies are needed to address this fully.

In contrast with previous findings (Demetriou et al., 2018), we observed that the greatest effect size for cognitive flexibility difficulties was found in the adult group, although we caution that there was no statistically significant difference across age groups. Buczylowska and Petermann (2016) found among neurotypical adults that the dispersion of EF scores across several domains increased over time due to age-related cognitive decline. Some evidence suggests that this decline could be more accentuated in ASD, for example in Koolschijn et al. (2017), compared to matched typically developing controls, autistic adults showed greater age-related reduction in white matter microstructure. Consistently, Powell et al. (2017) found that aging had a greater impact in adults with ASD, however not all areas of cognitive functioning were equally affected. Whilst some (e.g. category learning) were relatively preserved or showed a similar age-related decline as in typical aging (e.g. free recall), other domains including cognitive flexibility were disproportionately affected. Nevertheless, there was substantial interindividual variability, in other words some older autistic adults showed pronounced cognitive flexibility difficulties whereas others were similarly affected compared to neurotypical controls. Future studies should address this heterogeneity with an investigation of factors, such as education (Correia et al., 2018) and lifestyle (Weng et al., 2018), known to moderate age-related cognitive decline.

4.1. Limitations and Future Directions

One important caveat to consider is that several studies had mixed samples with wide age ranges, some spanning across childhood and into adulthood (e.g. 6 to 44 years - Miller et al., 2015). To perform the age subgroup analysis and meta-regression we had to rely on the study means and it is therefore likely that this led to age-related differences in cognitive flexibility being obscured. Future empirical studies of cognitive flexibility in ASD should endeavour to employ a stratified or longitudinal approach, to enable more precise estimates of cognitive flexibility difficulties across the lifespan. Alternatively, an individual patient data

(IPD) meta-analysis might help to more precisely capture the effects of age on cognitive flexibility. Additionally, in the present meta-analysis the adolescent group included only nine studies, so we caution against drawing any firm conclusions.

Although participants were matched on at least one IQ measure, matching criteria differed substantially across studies, thus not permitting subgroup analyses to be performed, however it is possible that this contributed to the heterogeneity observed. Finding appropriate matching strategies in ASD studies can be a challenge due to the distinctive profiles of cognitive strengths and weaknesses and possible limitations of standard assessment tools to capture these. It has been proposed that one possible avenue is to match on an area of functioning upon which the task heavily relies on, such as verbal abilities, thus allowing for differences to be controlled for (Burack et al., 2004). In the present study we carefully excluded for the presence of learning disabilities as this would confound the profile of cognitive flexibility difficulties, however more research is needed to explore EF across a wide range of abilities within the autistic spectrum, to enable greater generalisability of findings.

The majority of studies did not report the severity of ASD symptoms (e.g. ADOS scores) and this might represent another source of heterogeneity in the observed effects across studies. Cognitive flexibility difficulties have been linked with more pronounced RRBs (e.g. Lopez et al., 2005), therefore it is important for future studies with complete outcome reporting to explore the association between autistic symptom profiles and cognitive flexibility. Furthermore, due to lack of reporting across studies, the impact of co-occurring conditions could not be considered. Anxiety disorders are estimated to affect around 40 percent of autistic individuals (Hollocks et al., 2019; van Steensel et al., 2011). Given this high prevalence rate and that anxiety is known to have a deleterious effect on cognitive flexibility (Park & Moghaddam, 2017; Wilson et al., 2018), it is vital for future studies to take this comorbidity

into consideration. ASD is also highly co-ocurring with ADHD (Hofvander et al., 2009), however before the DSM-5 (American Psychiatric Association, 2013) these two diagnoses were mutually exclusive, thus limiting research studying these two conditions together. The EF difficulties in ADHD are well-documented (Craig et al., 2016; Happe et al., 2006; Sinzig et al., 2008) and growing evidence suggests that autistic individuals with comorbid ADHD have more pronounced cognitive difficulties (Craig et al., 2016; Dajani et al., 2016), thus emphasising the need to take co-occurring conditions into account when considering cognitive flexibility in ASD.

In summary, this meta-analysis has highlighted the existence of cognitive flexibility difficulties among autistic individuals in the absence of intellectual disabilities, but also that this profile is characterised by extensive heterogeneity. While several potential contributing factors have been discussed in the preceding sections, it is essential to consider the possibility that this heterogeneity could represent the broad range of cognitive flexibility profiles across the autistic spectrum. The present findings have therefore important ramifications on a therapeutic level. Considering the prevalence of cognitive flexibility difficulties throughout the lifespan in ASD, links with poorer outcomes (e.g. Bertollo et al., 2020) and the known plasticity of cognitive flexibility particularly during childhood (Buttelmann & Karbach, 2017), very little research has focused on interventions that might support better cognitive flexibility abilities, or strategies to manage everyday situations that require substantial cognitive flexibility burden. Evidence to date from randomised controlled trials on cognitive remediation strategies remains inconsistent (Pugliese et al., 2020), with cognitive enhancement therapy among autistic adults (Eack et al., 2018) and a cognitive behavioural intervention designed for children with ASD (Kenworthy et al., 2014) showing some promising results.

Moving forward, it is vital to explore the heterogeneity within the autistic spectrum and investigate whether there are subgroups with more homogenous cognitive flexibility profiles, to enable the progression from a 'one size fits all' approach towards the development of targeted, autism-specific, cognitive interventions that consider individual profiles of strengths and weaknesses. The emerging field of computational psychiatry, particularly unsupervised machine learning techniques, allow the discovery of hidden structures in data, without the assumption of prior knowledge or labels, and have therefore been used in the identification of previously undetected subtypes within the autistic spectrum (Stevens et al., 2019; Zheng et al., 2020). Furthermore, these data-driven approaches can be combined with theory-driven models of behaviour, such as in probabilistic RL tasks. The present findings should pave the way for research to parse out the heterogeneity of cognitive flexibility, in order to inform interventions and ultimately improve individual support and outcomes for autistic people.

Funding

This work was supported by Wellcome Trust Royal Society Henry Dale Fellowship awarded to RPL [206691/Z/17/Z]. RPL is also a Lister Institute Prize Fellow and supported by an Autistica Future Leaders Award [ID: 7265]. CL is supported by the Portuguese Foundation for Science and Technology [SFRH/BD/144811/2019].

Declaration of Competing Interest

The authors report no declarations of interest.

Appendix A Search Strategy

Database: Embase <1980 to 2020 Week 23>

1 autis*.mp. (73907)
2 "autism spectrum disorder*".mp. (30791)
3 "ASD".mp. (33305)
4 Asperger*.mp. (5443)
5 "pervasive developmental disorder not otherwise specified".mp. (1219)
6 "pervasive developmental disorder*".mp. (3692)
7 1 or 2 or 3 or 4 or 5 or 6 (86750)
8 "cognitive flexibility".mp. (4173)
9 "mental flexibility".mp. (895)
10 "cognitive rigidity".mp. (177)
11 "mental rigidity".mp. (42)
12 "cognitive inflexibility".mp. (297)
13 "mental inflexibility".mp. (13)
14 "executive function*".mp. (56375)
15 "executive dysfunction*".mp. (4941)
16 "set shift*".mp. (2689)
17 "Intra-Extra Dimensional Set Shift*".mp. (100)
18 "Dimensional Change Card Sort*".mp. (146)
19 "Flexible Item Selection".mp. (14)
20 "Modified Card Sort*".mp. (93)
21 "Wisconsin Card Sort*".mp. (4493)
22 "reversal learning".mp. (2417)
23 "probabilistic learning".mp. (254)
24 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 (66062)
25 7 and 24 (2058)
26 limit 25 to (human and english language and yr="1980 -Current") (1764)

References

Alsaedi, R. H., Carrington, S., & Watters, J. J. (2020). Behavioral and neuropsychological evaluation of executive functions in children with autism spectrum disorder in the Gulf region. *Brain Sci, 10*(2).

Ambery, F. Z., Russell, A. J., Perry, K., Morris, R., & Murphy, D. G. M. (2006). Neuropsychological functioning in adults with Asperger syndrome. *Autism, 10*(6), 551-564.

Anderson, K. A., Shattuck, P. T., Cooper, B. P., Roux, A. M., & Wagner, M. (2014). Prevalence and correlates of postsecondary residential status among young adults with an autism spectrum disorder. *Autism, 18*(5), 562-570.

Association, A. P. (2013). *Diagnostic and statistical manual of mental disorders (Fifth edition)*: Arlington: American Psychiatric Publishing.

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

Berg, E. A. (1948). A simple objective technique for measuring flexibility in thinking. *The Journal of General Psychology 39*, 15-22.

Berger, H. J., Aerts, F. H., van Spaendonck, K. P., Cools, A. R., & Teunisse, J. P. (2003). Central coherence and cognitive shifting in relation to social improvement in high-functioning young adults with autism. *J Clin Exp Neuropsychol, 25*(4), 502-511.

Bertollo, J. R., Strang, J. F., Anthony, L. G., Kenworthy, L., Wallace, G. L., & Yerys, B. E. (2020). Adaptive behavior in youth with autism spectrum disorder: The role of flexibility. *J Autism Dev Disord, 50*(1), 42-50.

Blakemore, S. J., & Choudhury, S. (2006). Development of the adolescent brain: implications for executive function and social cognition. *J Child Psychol Psychiatry, 47*(3-4), 296-312.

Braden, B. B., Smith, C. J., Thompson, A., Glaspy, T. K., Wood, E., Vatsa, D., . . . Baxter, L. C. (2017). Executive function and functional and structural brain differences in middle-age adults with autism spectrum disorder. *Autism Research, 10*(12), 1945-1959.

Brady, D. I., Schwean, V. L., Saklofske, D. H., McCrimmon, A. W., Montgomery, J. M., & Thorne, K. J. (2013). Conceptual and perceptual set-shifting executive abilities in young adults with Asperger's syndrome. *Research in Autism Spectrum Disorders, 7*(12), 1631-1637.

Buchsbaum, B. R., Greer, S., Chang, W. L., & Berman, K. F. (2005). Meta-analysis of neuroimaging studies of the Wisconsin card-sorting task and component processes. *Hum Brain Mapp, 25*(1), 35-45.

Buczylowska, D., & Petermann, F. (2016). Age-related differences and heterogeneity in executive functions: Analysis of NAB executive functions module scores. *Arch Clin Neuropsychol, 31*(3), 254-262.

Burack, J. A., Iarocci, G., Flanagan, T. D., & Bowler, D. M. (2004). On mosaics and melting pots: Conceptual considerations of comparison and matching strategies. *Journal of Autism and Developmental Disorders, 34*, 65-73.

Buttelmann, F., & Karbach, J. (2017). Development and plasticity of cognitive flexibility in early and middle childhood. *Front Psychol, 8*, 1040.

Catani, M., Dell'Acqua, F., Budisavljevic, S., Howells, H., Thiebaut de Schotten, M., Froudist-Walsh, S., . . . Murphy, D. G. (2016). Frontal networks in adults with autism spectrum disorder. *Brain, 139*(Pt 2), 616-630.

Chan, R. C. K., Hu, Z.-y., Cui, J.-f., Wang, Y., & McAlonan, G. M. (2011). Social attribution in children with high functioning autism and Asperger syndrome: An exploratory study in the Chinese setting. *Research in Autism Spectrum Disorders, 5*(4), 1538-1548.

Chen, S. F., Chien, Y. L., Wu, C. T., Shang, C. Y., Wu, Y. Y., & Gau, S. S. (2016). Deficits in executive functions among youths with autism spectrum disorders: an age-stratified analysis. *Psychological medicine, 46*(8), 1625-1638.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cooper, H. (2017). *Research Synthesis and Meta-analysis*. USA: Sage.

Craig, F., Margari, F., Legrottaglie, A. R., Palumbi, R., de Giambattista, C., & Margari, L. (2016). A review of executive function deficits in autism spectrum disorder and attention-deficit/hyperactivity disorder. *Neuropsychiatr Dis Treat, 12*, 1191-1202.

Crawley, D., Zhang, L., Jones, E. J. H., Ahmad, J., Oakley, B., San Jose Caceres, A., . . . group, E.-A. L. (2020). Modeling flexible behavior in childhood to adulthood shows age-dependent learning mechanisms and less optimal learning in autism in each age group. *PLoS Biol, 18*(10), e3000908.

Dajani, D. R., Llabre, M. M., Nebel, M. B., Mostofsky, S. H., & Uddin, L. Q. (2016). Heterogeneity of executive functions among comorbid neurodevelopmental disorders. *Sci Rep, 6*, 36566.

Dajani, D. R., & Uddin, L. Q. (2015). Demystifying cognitive flexibility: Implications for clinical and developmental neuroscience. *Trends Neurosci, 38*(9), 571-578.

Damasio, A. R., & Maurer, R. G. (1978). A neurological model for childhood autism. *Archives of Neurology 35*, 777-786.

D'Cruz, A. M., Mosconi, M. W., Ragozzino, M. E., Cook, E. H., & Sweeney, J. A. (2016). Alterations in the functional neural circuitry supporting flexible choice behavior in autism spectrum disorders. *Transl Psychiatry, 6*(10), e916.

D'Cruz, A. M., Ragozzino, M. E., Mosconi, M. W., Shrestha, S., Cook, E. H., & Sweeney, J. A. (2013). Reduced behavioral flexibility in autism spectrum disorders. *Neuropsychology, 27*(2), 152-160.

de Vries, M., & Geurts, H. (2015). Influence of autism traits and executive functioning on quality of life in children with an autism spectrum disorder. *J Autism Dev Disord, 45*(9), 2734-2743.

Demetriou, E. A., DeMayo, M. M., & Guastella, A. J. (2019). Executive function in autism spectrum disorder: History, theoretical models, empirical findings, and potential as an endophenotype. *Front Psychiatry, 10*, 753.

Demetriou, E. A., Lampit, A., Quintana, D. S., Naismith, S. L., Song, Y. J. C., Pye, J. E., . . . Guastella, A. J. (2018). Autism spectrum disorders: a meta-analysis of executive function. *Mol Psychiatry, 23*(5), 1198-1204.

Dichter, G. S., & Belger, A. (2007). Social stimuli interfere with cognitive control in autism. *Neuroimage 35*(3), 1219–1230.

Dichter, G. S., Felder, J. N., & Bodfish, J. W. (2009). Autism is characterized by dorsal anterior cingulate hyperactivation during social target detection. *Soc Cogn Affect Neurosci, 4*(3), 215-226.

Downes, J. J., Roberts, A. C., Sahakian, B. J., Evenden, J. L., Morris, R. G., & Robbins, T. W. (1989). Impaired extra-dimensional shift performance in medicated and unmedicated Parkinson's disease: Evidence for a specific attentional dysfunction. *Neuropsychologia, 27*(11-12), 1329–1343.

Eack, S. M., Hogarty, S. S., Greenwald, D. P., Litschge, M. Y., Porton, S. A., Mazefsky, C. A., & Minshew, N. J. (2018). Cognitive enhancement therapy for adult autism spectrum disorder: Results of an 18-month randomized clinical trial. *Autism Res, 11*(3), 519-530.

Ecker, C. (2017). The neuroanatomy of autism spectrum disorder: An overview of structural neuroimaging findings and their translatability to the clinical setting. *Autism, 21*(1), 18-28.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ, 315*, 629-634.

Faja, S., & Nelson Darling, L. (2019). Variation in restricted and repetitive behaviors and interests relates to inhibitory control and shifting in children with autism spectrum disorder. *Autism, 23*(5), 1262-1272.

Farley, M. A., McMahon, W. M., Fombonne, E., Jenson, W. R., Miller, J., Gardner, M., . . . Coon, H. (2009). Twenty-year outcome for individuals with autism and average or near-average cognitive abilities. *Autism Res, 2*(2), 109-118.

Frank, F., Jablotschkin, M., Arthen, T., Riedel, A., Fangmeier, T., Holzel, L. P., & Tebartz van Elst, L. (2018). Education and employment status of adults with autism spectrum disorders in Germany - a cross-sectional-survey. *BMC Psychiatry, 18*(1), 75.

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

Garcia-Villamisar, D., & Della Sala, S. (2002). Dual-task performance in adults with autism. *Cognitive Neuropsychiatry, 7*(1), 63-74.

Geurts, H. M., Corbett, B., & Solomon, M. (2009). The paradox of cognitive flexibility in autism. *Trends Cogn Sci, 13*(2), 74-82.

Geurts, H. M., Pol, S. E., Lobbestael, J., & Simons, C. J. P. (2020). Executive functioning in 60+ autistic males: The discrepancy between experienced challenges and cognitive performance. *Journal of Autism and Developmental Disorders, 50*(4), 1380-1390.

Goddard, L., Dritschel, B., Robinson, S., & Howlin, P. (2014). Development of autobiographical memory in children with autism spectrum disorders: deficits, gains, and predictors of performance. *Dev Psychopathol, 26*(1), 215-228.

Goldstein, G., Johnson, C. R., & Minshew, N. J. (2001). Attentional processes in autism. *Journal of Autism and Developmental Disorders, 31*(4), 433-440

Gomez-Perez, M. M., Calero, M. D., Mata, S., & Molinero, C. (2016). Discrepancies between direct and indirect measures of interpersonal and neurocognitive skills in autism spectrum disorder children. *Journal of Clinical and Experimental Neuropsychology, 38*(8), 875-886.

Gomez-Perez, M. M., Mata, S., Serrano, F., & Calero, M. D. (2020). Wisconsin Card Sorting Test-learning potential: Usefulness for assessing children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*.

Griebling, J., Minshew, N. J., Bodner, K., Libove, R., Bansal, R., Konasale, P., . . . Hardan, A. (2010). Dorsolateral prefrontal cortex magnetic resonance imaging measurements and cognitive performance in autism. *Journal of Child Neurology, 25*(7), 856-863.

Happe, F., Booth, R., Charlton, R., & Hughes, C. (2006). Executive function deficits in autism spectrum disorders and attention-deficit/hyperactivity disorder: examining profiles across domains and ages. *Brain Cogn, 61*(1), 25-39.

Hedges, L. V. (1980). Unbiased estimation of effect size. *Evaluation in Education: An International Review Series, 4*, 25-27.

Higgins, J. P., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21 (11)* (Wiley Online Library), 1539–1558.

Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (2019). *Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019)*.

Hill, E. L. (2004). Evaluating the theory of executive dysfunction in autism *Developmental Review 24*, 189-233.

Hill, E. L., & Bird, C. M. (2006). Executive processes in Asperger syndrome: Patterns of performance in a multiple case series. *Neuropsychologia, 44*(14), 2822-2835.

Hofvander, B., Delorme, R., Chaste, P., Nyden, A., Wentz, E., Stahlberg, O., . . . Leboyer, M. (2009). Psychiatric and psychosocial problems in adults with normal-intelligence autism spectrum disorders. *BMC Psychiatry, 9*, 35.

Hollocks, M. J., Lerh, J. W., Magiati, I., Meiser-Stedman, R., & Brugha, T. S. (2019). Anxiety and depression in adults with autism spectrum disorder: a systematic review and meta-analysis. *Psychol Med, 49*(4), 559-572.

Kado, Y., Sanada, S., Oono, S., Ogino, T., & Nouno, S. (2020). Children with autism spectrum disorder comorbid with attention-deficit/hyperactivity disorder examined by the Wisconsin card sorting test: Analysis by age-related differences. *Brain and Development, 42*(2), 113-120.

Kaland, N., Smith, L., & Mortensen, E. L. (2008). Brief report: Cognitive flexibility and focused attention in children and adolescents with Asperger syndrome or high-functioning autism as measured on the computerized version of the Wisconsin card sorting test. *Journal of Autism and Developmental Disorders, 38*(6), 1161-1165.

Kaufmann, L., Zotter, S., Pixner, S., Starke, M., Haberlandt, E., Steinmayr-Gensluckner, M., . . . Marksteiner, J. (2013). Brief report: CANTAB performance and brain structure in pediatric patients with Asperger syndrome. *Journal of Autism and Developmental Disorders, 43*(6), 1483-1490.

Keary, C. J., Minshew, N. J., Bansal, R., Goradia, D., Fedorov, S., Keshavan, M. S., & Hardan, A. Y. (2009). Corpus callosum volume and neurocognition in autism. *Journal of Autism and Developmental Disorders, 39*(6), 834-841.

Kenworthy, L., Anthony, L. G., Naiman, D. Q., Cannon, L., Wills, M. C., Luong-Tran, C., . . . Wallace, G. L. (2014). Randomized controlled effectiveness trial of executive function intervention for children on the autism spectrum. *J Child Psychol Psychiatry, 55*(4), 374-383.

Kiep, M., & Spek, A. A. (2017). Executive functioning in men and women with an autism spectrum disorder. *Autism Research, 10*(5), 940-948.

Kilincaslan, A., Mukaddes, N. M., Kucukyazici, G. S., & Gurvit, H. (2010). Assessment of executive/attentional performance in Asperger's disorder. *Turkish journal of psychiatry, 21* (4), 289-299.

Koolschijn, P. C., Caan, M. W., Teeuw, J., Olabarriaga, S. D., & Geurts, H. M. (2017). Age-related differences in autism: The case of white matter microstructure. *Hum Brain Mapp, 38*(1), 82-96.

Lai, C. L. E., Lau, Z., Lui, S. S. Y., Lok, E., Tam, V., Chan, Q., . . . Cheung, E. F. C. (2017). Meta-analysis of neuropsychological measures of executive functioning in children and adolescents with high-functioning autism spectrum disorder. *Autism Res, 10*(5), 911-939.

Landa, R. J., & Goldberg, M. C. (2005). Language, social, and executive functions in high functioning autism: a continuum of performance. *Journal of Autism and Developmental Disorders, 35*(5), 557-573.

Landry, O., & Al-Taie, S. (2016). A meta-analysis of the Wisconsin Card Sort Task in autism. *J Autism Dev Disord, 46*(4), 1220-1235.

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.

Lawson, R. P., Mathys, C., & Rees, G. (2017). Adults with autism overestimate the volatility of the sensory environment. *Nat Neurosci, 20*(9), 1293-1299.

Leung, R. C., & Zakzanis, K. K. (2014). Brief report: cognitive flexibility in autism spectrum disorders: a quantitative review. *J Autism Dev Disord, 44*(10), 2628-2645.

Li, J., Zhu, L., Liu, J., & Li, X. (2014). Social and non-social deficits in children with high-functioning autism and their cooperative behaviors. *Research in Autism Spectrum Disorders, 8*(12), 1657-1671.

Libero, L. E., DeRamus, T. P., Lahti, A. C., Deshpande, G., & Kana, R. K. (2015). Multimodal neuroimaging based classification of autism spectrum disorder using anatomical, neurochemical, and white matter correlates. *Cortex, 66*, 46-59.

Lopez, B. R., Lincoln, A. J., Ozonoff, S., & Lai, Z. (2005). Examining the relationship between executive functions and restricted, repetitive symptoms of Autistic Disorder. *J Autism Dev Disord, 35*(4), 445-460.

Lord, C., Brugha, T. S., Charman, T., Cusack, J., Dumas, G., Frazier, T., . . . Veenstra-VanderWeele, J. (2020). Autism spectrum disorder. *Nat Rev Dis Primers, 6*(1), 5.

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. L. (2012). *Autism diagnostic observation schedule–Second edition (ADOS-2)*. Los Angeles: Western Psychological Services.

Maister, L., Simons, J. S., & Plaisted-Grant, K. (2013). Executive functions are employed to process episodic and relational memories in children with autism spectrum disorders. *Neuropsychology, 27*(6), 615-627.

Manning, C., Kilner, J., Neil, L., Karaminis, T., & Pellicano, E. (2017). Children on the autism spectrum update their behaviour in response to a volatile environment. *Dev Sci, 20*(5).

Mason, D., McConachie, H., Garland, D., Petrou, A., Rodgers, J., & Parr, J. R. (2018). Predictors of quality of life for autistic adults. *Autism Res, 11*(8), 1138-1147.

Merchán-Naranjo, J., Boada, L., del Rey-Mejías, Á., Mayoral, M., Llorente, C., Arango, C., & Parellada, M. (2016). Executive function is affected in autism spectrum disorder, but does not correlate with intelligence. *Rev Psiquiatr Salud Ment, 9*(1), 39-50.

Miller, H. L., Ragozzino, M. E., Cook, E. H., Sweeney, J. A., & Mosconi, M. W. (2015). Cognitive set shifting deficits and their relationship to repetitive behaviors in autism spectrum disorder. *J Autism Dev Disord, 45*(3), 805-815.

Minshew, N. J., Goldstein, G., Muenz, L. R., & Payton, J. B. (1992). Neuropsychological functioning in nonmentally retarded autistic individuals. *Journal of Clinical and Experimental Neuropsychology, 14*(5), 749-761.

Minshew, N. J., Goldstein, G., & Siegel, D. J. (1997). Neuropsychologic functioning in autism: Profile of a complex information processing disorder. *Journal of the International Neuropsychological Society, 3*(4), 303-316.

Minshew, N. J., Meyer, J., & Goldstein, G. (2002). Abstract reasoning in autism: A dissociation between concept formation and concept identification. *Neuropsychology, 16*(3), 327-334.

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Curr Dir Psychol Sci, 21*(1), 8-14.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med, 6*(7), e1000097.

Moriguchi, Y., & Hiraki, K. (2009). Neural origin of cognitive shifting in young children. *PNAS 106*(14), 6017– 6021.

Nyhus, E., & Barcelo, F. (2009). The Wisconsin Card Sorting Test and the cognitive assessment of prefrontal executive functions: A critical update. *Brain Cogn, 71*(3), 437-451.

Ozonoff, S. (1995). Reliability and validity of the Wisconsin Card Sorting Test in studies of autism. *Neuropsychology, 9*(4), 491-500.

Ozonoff, S., Cook, I., Coon, H., Dawson, G., Joseph, R. M., Klin, A., . . . Wrathall, D. (2004). Performance on Cambridge neuropsychological test automated battery subtests sensitive to frontal lobe function in people with autistic disorder: Evidence from the Collaborative Programs of Excellence in Autism network. *Journal of Autism and Developmental Disorders, 34*(2), 139-150.

Ozonoff, S., South, M., & Miller, J. N. (2000). DSM-IV-defined Asperger syndrome: Cognitive, behavioral and early history differentiation from highfunctioning autism. *Autism, 4*(1), 29-46.

Ozsivadjian, A., Hollocks, M. J., Magiati, I., Happe, F., Baird, G., & Absoud, M. (2021). Is cognitive inflexibility a missing link? The role of cognitive inflexibility, alexithymia and intolerance of uncertainty in externalising and internalising behaviours in young people with autism spectrum disorder. *J Child Psychol Psychiatry, 62*(6), 715-724.

Panerai, S., Tasca, D., Ferri, R., Genitori D'Arrigo, V., & Elia, M. (2014). Executive functions and adaptive behaviour in autism spectrum disorders with and without intellectual disability. *Psychiatry Journal, 2014*, 1-11.

Park, J., & Moghaddam, B. (2017). Impact of anxiety on prefrontal cortex encoding of cognitive flexibility. *Neuroscience, 345*, 193-202.

Pellicano, E., Murray, M., Durkin, K., & Maley, A. (2006). Multiple cognitive capabilities/deficits in children with an autism spectrum disorder: "Weak" central coherence and its relationship to theory of mind and executive control. *Development and Psychopathology, 18*(1), 77-98.

Perez Velazquez, J. L., Barcelo, F., Hung, Y., Leshchenko, Y., Nenadovic, V., Belkas, J., . . . Garcia Dominguez, L. (2009). Decreased brain coordinated activity in autism spectrum disorders during executive tasks: Reduced long-range synchronization in the fronto-parietal networks. *International Journal of Psychophysiology, 73*(3), 341-349.

Powell, P. S., Klinger, L. G., & Klinger, M. R. (2017). Patterns of age-related cognitive differences in adults with autism spectrum disorder. *J Autism Dev Disord, 47*(10), 3204-3219.

Pugliese, C. E., Wallace, G. L., Anthony, L. G., & Kenworthy, L. (2020). *Understanding Executive Function Challenges in Autism Spectrum Disorder*. In White, S., Maddox, B., & Mazefsky, C. (Eds.), *The Oxford handbook of autism and co-occurring psychiatric conditions* (Oxford library of psychology).

Robinson, S., Goddard, L., Dritschel, B., Wisley, M., & Howlin, P. (2009). Executive functions in children with autism spectrum disorders. *Brain and Cognition, 71*(3), 362-368.

Rumsey, J. M. (1985). Conceptual problem-solving in highly verbal, nonretarded autistic men. *Journal of Autism and Developmental Disorders, 15*(1), 23-36.

Russo, N., Flanagan, T., Iarocci, G., Berringer, D., Zelazo, P. D., & Burack, J. A. (2007). Deconstructing executive deficits among persons with autism: Implications for cognitive neuroscience. *Brain Cogn, 65*(1), 77-86.

Rut, C., Jose, B., & Antonieta, N. (2018). Age-related cognitive changes: The importance of modulating factors. *Journal of Geriatric Medicine and Gerontology, 4*(2).

Rutter, M., Le Couteur, A., & Lord, C. (2003). *Autism diagnostic interview-revised*. Los Angeles: Western Psychological Services.

Sachse, M., Schlitt, S., Hainz, D., Ciaramidaro, A., Schirman, S., Walter, H., . . . Freitag, C. M. (2013). Executive and visuo-motor function in adolescents and adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 43*(5), 1222-1235.

Sawa, T., Kodaira, M., Oiji, A., Sasayama, D., Iwadare, Y., Ushijima, H., . . . Saito, K. (2013). Dysfunction of orbitofrontal and dorsolateral prefrontal cortices in children and adolescents with high-functioning pervasive developmental disorders. *Ann Gen Psychiatry, 12*(1), 31.

Sawaya, H., McGonigle-Chalmers, M., & Kusel, I. (2019). Identifying set-switching difficulties in autism spectrum disorder using a rule following task. *International Journal of Developmental Disabilities*.

Schmitz, N., Rubia, K., Daly, E., Smith, A., Williams, S., & Murphy, D. G. (2006). Neural correlates of executive function in autistic spectrum disorders. *Biol Psychiatry, 59*(1), 7-16.

Sinzig, J., Morsch, D., Bruning, N., Schmidt, M. H., & Lehmkuhl, G. (2008). Inhibition, flexibility, working memory and planning in autism spectrum disorders with and without comorbid ADHD-symptoms. *Child Adolesc Psychiatry Ment Health, 2*(1), 4.

Solomon, M., Frank, M. J., Ragland, J. D., Smith, A. C., Niendam, T. A., Lesh, T. A., . . . Carter, C. S. (2015). Feedback-driven trial-by-trial learning in autism spectrum disorders. *Am J Psychiatry, 172*(2), 173-181.

St John, T., Dawson, G., & Estes, A. (2018). Brief report: Executive function as a predictor of academic achievement in school-aged children with ASD. *J Autism Dev Disord, 48*(1), 276-283.

Stevens, E., Dixon, D. R., Novack, M. N., Granpeesheh, D., Smith, T., & Linstead, E. (2019). Identification and analysis of behavioral phenotypes in autism spectrum disorder via unsupervised machine learning. *Int J Med Inform, 129*, 29-36.

Thakkar, K. N., Polli, F. E., Joseph, R. M., Tuch, D. S., Hadjikhani, N., Barton, J. J., & Manoach, D. S. (2008). Response monitoring, repetitive behaviour and anterior cingulate abnormalities in autism spectrum disorders (ASD). *Brain, 131*(Pt 9), 2464-2478.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2013). Practitioner review: Do performance-based measures and ratings of executive function assess the same construct? *J Child Psychol Psychiatry, 54*(2), 131-143.

Uddin, L. Q. (2021). Cognitive and behavioural flexibility: Neural mechanisms and clinical considerations. *Nat Rev Neurosci, 22*(3), 167-179.

Van Eylen, L., Boets, B., Steyaert, J., Evers, K., Wagemans, J., & Noens, I. (2011). Cognitive flexibility in autism spectrum disorder: Explaining the inconsistencies? *Research in Autism Spectrum Disorders, 5*(4), 1390-1401.

Van Eylen, L., Boets, B., Steyaert, J., Wagemans, J., & Noens, I. (2015). Executive functioning in autism spectrum disorders: Influence of task and sample characteristics and relation to symptom severity. *Eur Child Adolesc Psychiatry, 24*(11), 1399-1417.

van Steensel, F. J., Bogels, S. M., & Perrin, S. (2011). Anxiety disorders in children and adolescents with autistic spectrum disorders: A meta-analysis. *Clin Child Fam Psychol Rev, 14*(3), 302-317.

Wang, Z., Jing, J., Igarashi, K., Fan, L., Yang, S., Li, Y., & Jin, Y. (2018). Executive function predicts the visuospatial working memory in autism spectrum disorder and attention-deficit/hyperactivity disorder. *Autism Research, 11*(8), 1148-1156.

Weiss, E. O., Kruppa, J. A., Fink, G. R., Herpertz-Dahlmann, B., Konrad, K., & Schulte-Ruther, M. (2021). Developmental differences in probabilistic reversal learning: A computational modeling approach. *Front Neurosci, 14*, 536596.

Wells G., Shea B., O'Connell D., Peterson J., Welch V., Losos M., & Tugwell P. (2000). The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. https://www.ohri.ca/programs/clinical_epidemiology/oxford.asp

Weng, P. H., Chen, J. H., Chiou, J. M., Tu, Y. K., Chen, T. F., Chiu, M. J., . . . Chen, Y. C. (2018). The effect of lifestyle on late-life cognitive change under different socioeconomic status. *PLoS One, 13*(6), e0197676.

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.

Wilson, C. G., Nusbaum, A. T., Whitney, P., & Hinson, J. M. (2018). Trait anxiety impairs cognitive flexibility when overcoming a task acquired response and a preexisting bias. *PLoS One, 13*(9), e0204694.

Xie, R., Sun, X., Yang, L., & Guo, Y. (2020). Characteristic executive dysfunction for high-functioning autism sustained to adulthood. *Autism Res, 13*(12), 2102-2121.

Yerys, B. E., Wallace, G. L., Harrison, B., Celano, M. J., Giedd, J. N., & Kenworthy, L. E. (2009). Set-shifting in children with autism spectrum disorders: Reversal shifting deficits on the Intradimensional/Extradimensional Shift Test correlate with repetitive behaviors. *Autism, 13*(5), 523-538.

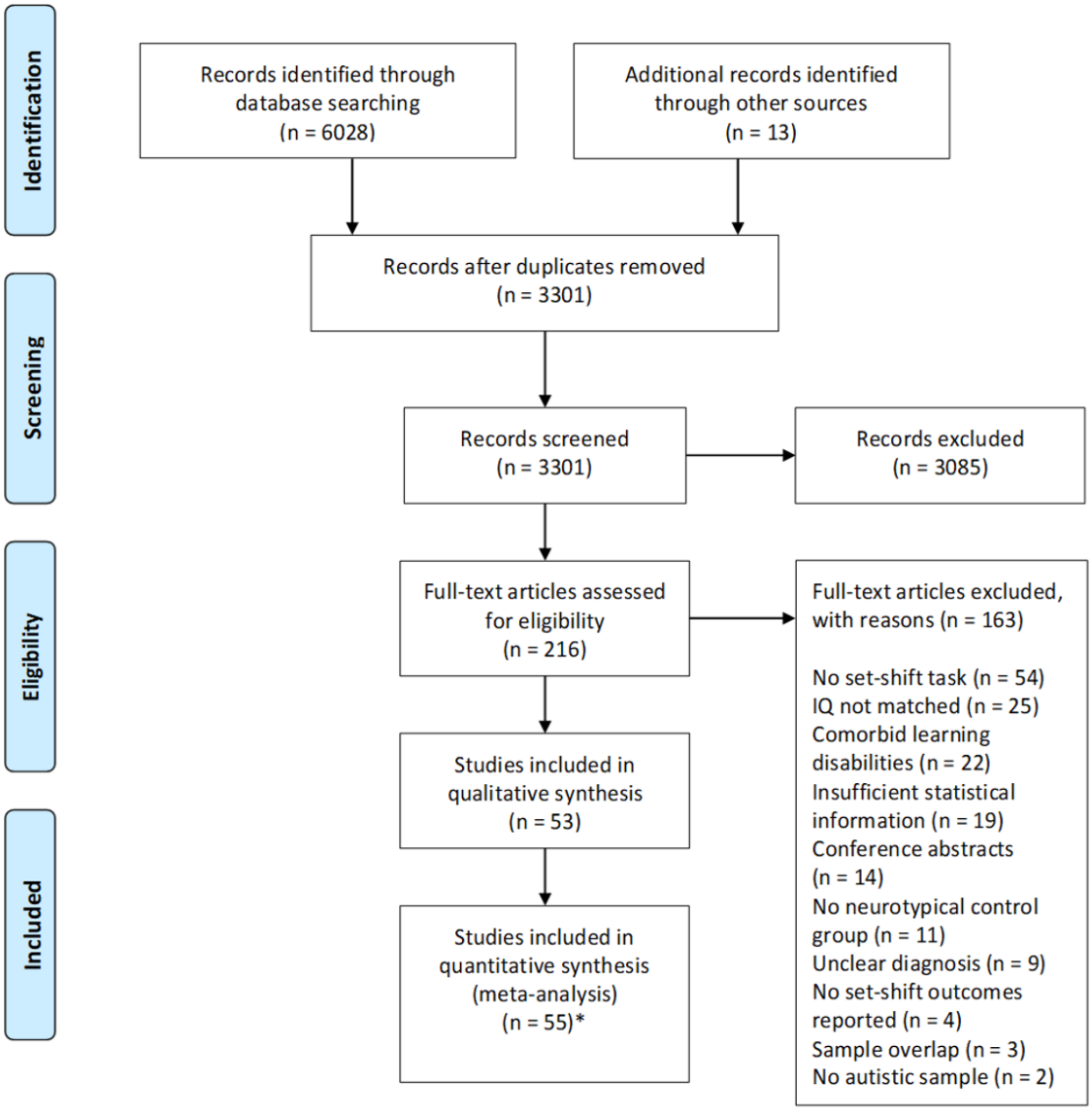
Yeung, M. K., Han, Y. M. Y., Sze, S. L., & Chan, A. S. (2016). Abnormal frontal theta oscillations underlie the cognitive flexibility deficits in children with high-functioning autism spectrum disorders. *Neuropsychology, 30*(3), 281-295.

Zelazo, P. D., Craik, F. I., & Booth, L. (2004). Executive function across the life span. *Acta Psychol (Amst), 115*(2-3), 167-183.

Zhang, L., Tang, J., Dong, Y., Ji, Y., Tao, R., Liang, Z., . . . Wang, K. (2015). Similarities and differences in decision-making impairments between autism spectrum disorder and schizophrenia. *Front Behav Neurosci, 9*, 259.

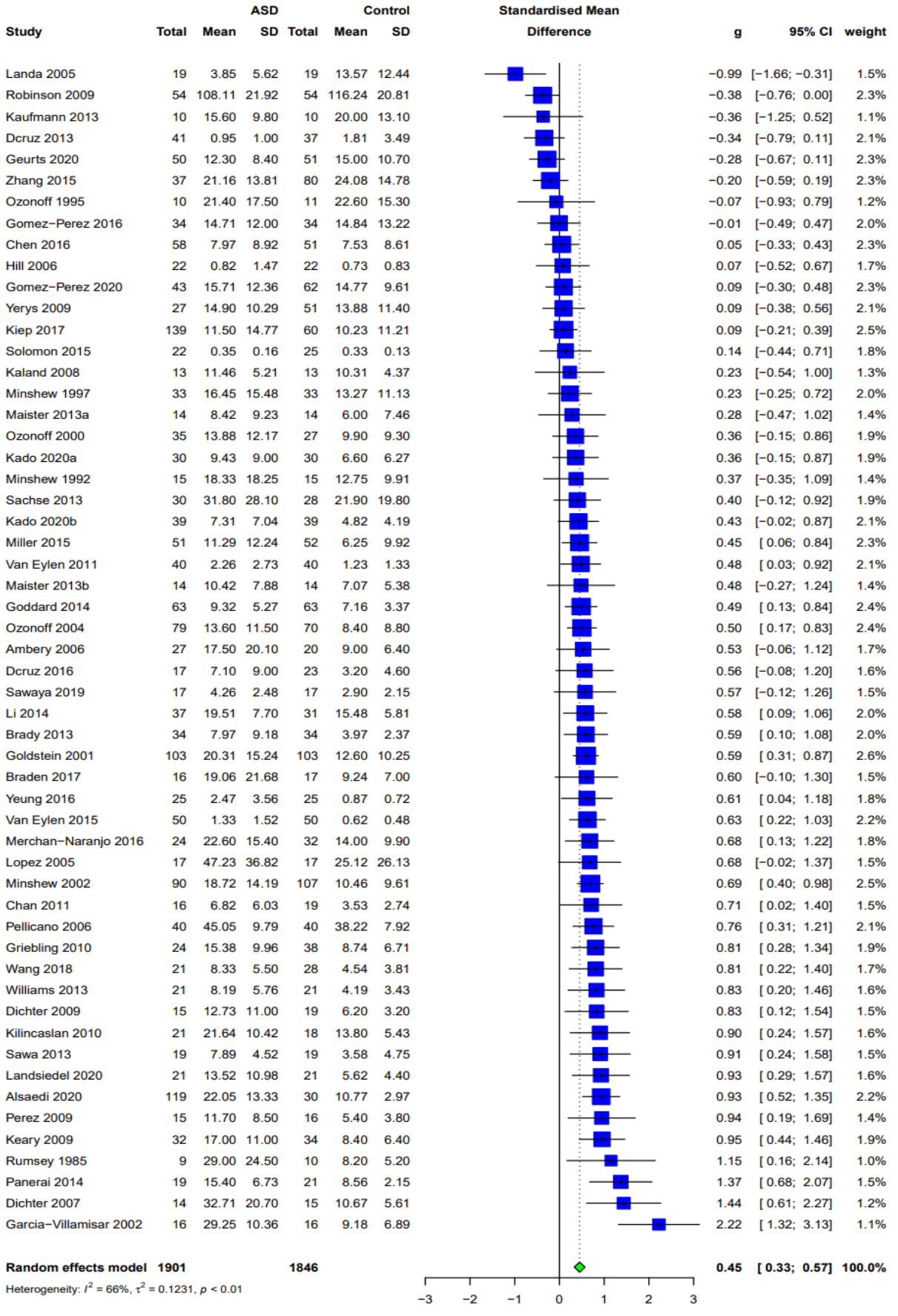
Zheng, S., Hume, K. A., Able, H., Bishop, S. L., & Boyd, B. A. (2020). Exploring developmental and behavioral heterogeneity among preschoolers with ASD: A cluster analysis on principal components. *Autism Res, 13*(5), 796-809.

Figure 1. PRISMA Flow Diagram of Study Selection



*Two studies had more than one experiment with no overlap in participants and were included separately.

Figure 2. Forest Plot of Included Studies



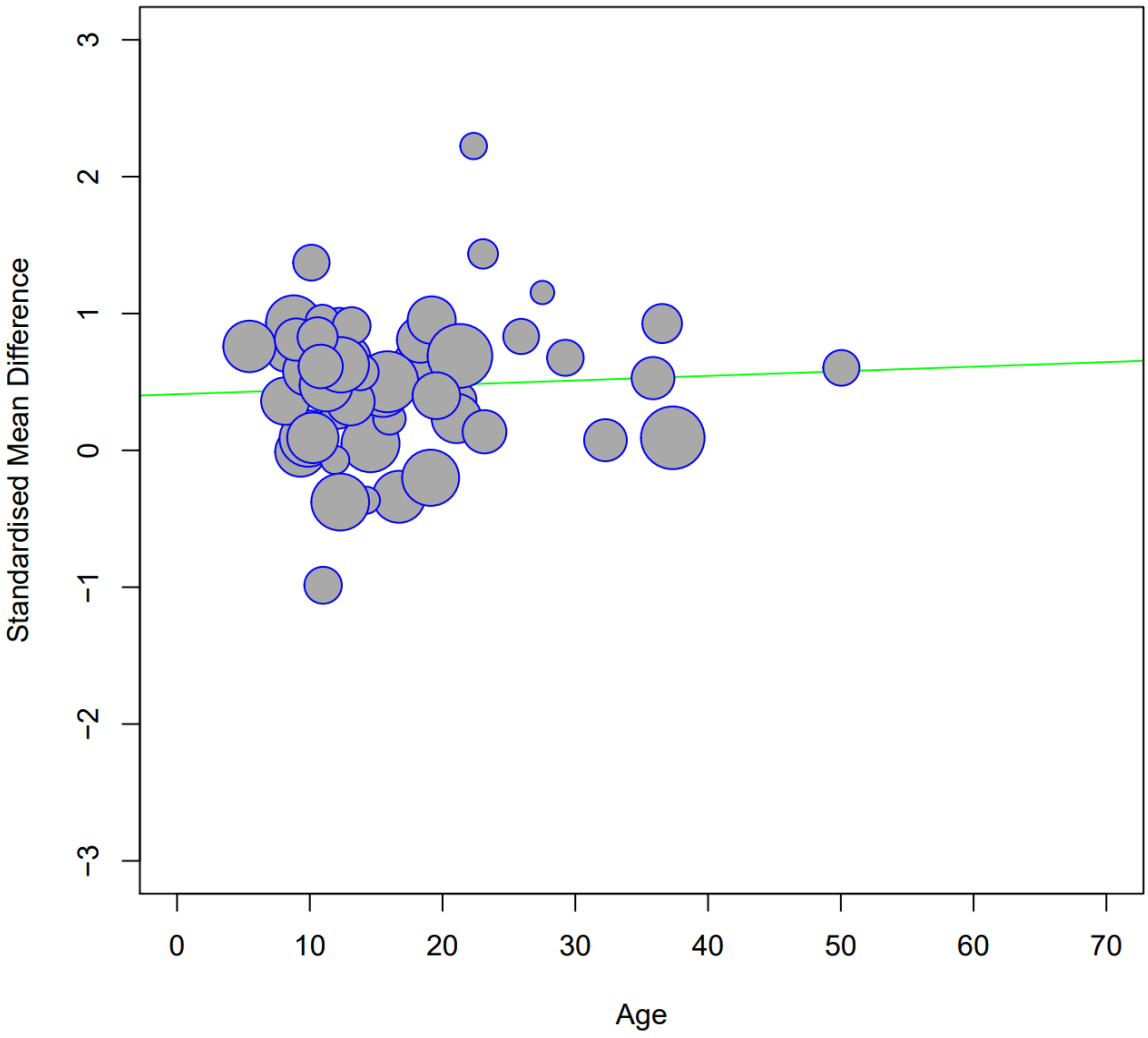


Figure 3. Meta-regression Bubble Plot with Age as a Moderator

Each bubble shows an individual study and the size varies according to the weight assigned under the random-effects models (i.e. studies with larger samples are assigned higher weights as displayed by the larger bubbles). The green line represents the regression line of best fit. Although, the SMD appears to increase slightly with age, the dispersion shows that age is not a significant moderator.

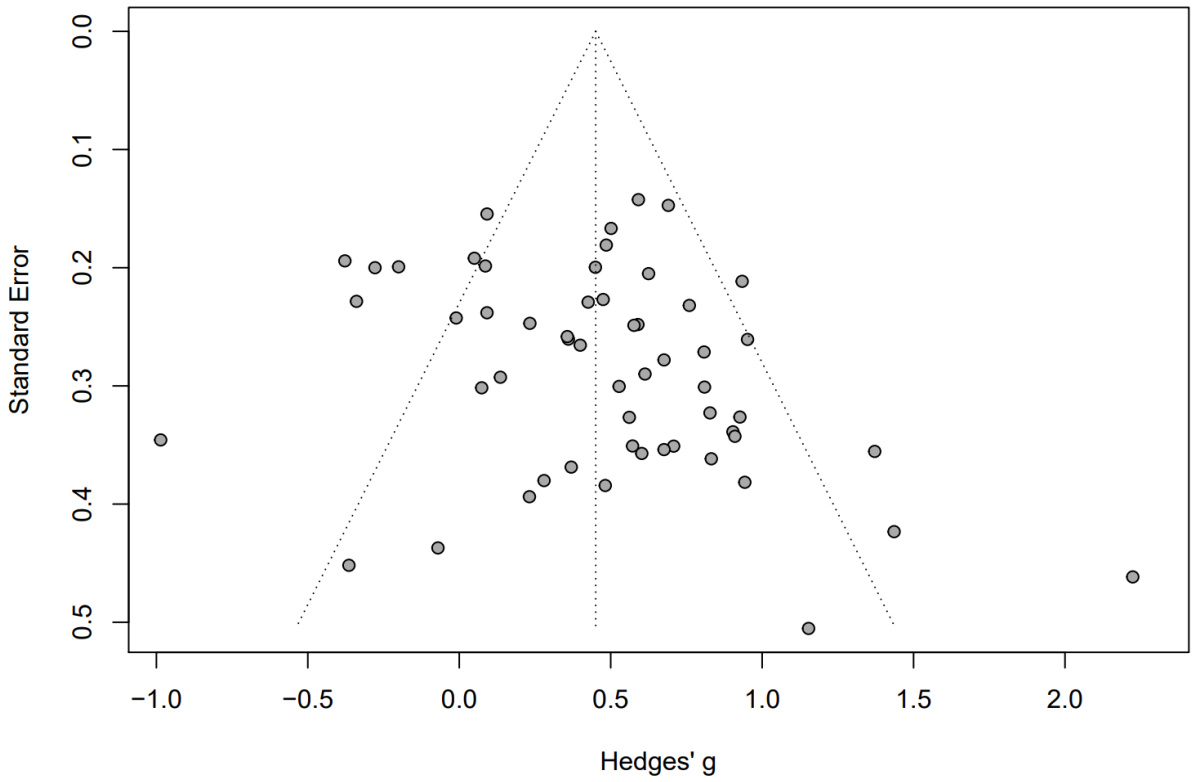


Figure 4. Funnel Plot of Included Studies

Each dot shows an individual study (studies with larger samples have smaller standard errors). The vertical line represents the pooled effect size (g = 0.45) and the diagonal lines (funnel) represent the 95% confidence interval. Although some asymmetry can be observed, most studies are spread relatively evenly around the pooled effect size, i.e. there appears to be minimal publication bias.

Table 1. Data Extracted from Included Studies

			ASD		NT							
Source	**Age Range Mean (SD)**	**Task**	**Sample Size**	**Mean (SD)**	**Sample Size**	**Mean (SD)**	**Diagnostic Criteria**	**IQ Tool**	**Matched**	**FSIQ Mean**	**NVIQ/PIQ Mean**	**VIQ Mean**
Alsaedi 2020	6–12 ASD 8.72 (1.96) NT 9.06 (1.42)	IED	119 M=95 F=24	TE 22.05 (13.33)	30 M=24 F=6	TE 10.77 (2.97)	DSM-IV-TR	RCPM	Age Sex NVIQ	--	ASD 29.76 NT 29.80	--
Ambery 2006	19-67 ASD 37.6 (14.6) NT 33.5 (12)	WCST	27 M=22 F=5	PE 17.5 (20.1)	20 M=16 F=4	PE 9.0 (6.4)	ICD-10 ADI (N=12) ADOS (N=4)	WAIS	Age Sex PIQ VIQ	--	ASD 103.7 NT 109.4	ASD 106.1 NT 107.05
Braden 2017	40–64 ASD 50.1 (--) NT 50.0 (--)	WCST	16 M=16	PE 19.06 (21.68)	17 M=17	PE 9.24 (7)	DSM IV DSM 5 ADOS	KBIT	Age FSIQ	ASD 108.9 NT 110.2	--	--
Brady 2013	16–21 ASD 18.86 (--) NT 18.90 (--)	IED	34 M= 76.5%	ESE 7.97 (9.18)	34 M= -- F= --	ESE 3.97 (2.37)	DSM-IV-TR	WASI	Age Sex FSIQ PIQ VIQ	ASD 112.76 NT 110.44	ASD 108.03 NT 109.03	ASD 114.29 NT 109.32
Chan 2011	6-14 ASD 7.98 (1.9) NT 8.30 (1.98)	MCST	16 M=16	PE 6.82 (6.03)	19 M=19	PE 3.53 (2.74)	DSM-IV	C-WISC	Age FSIQ VIQ	ASD 89.50 NT 101.00	ASD 98.90 NT 99.55	ASD 78.40 NT 105.26
Chen 2016	13-18 ASD 14.72 (1.53) NT 14.41 (1.42)	IED	58 M=57 F=1	ESE 7.97 (8.92)	51 M=50 F=1	ESE 7.53 (8.61)	DSM-IV ADI-R	--	Age Sex FSIQ VIQ PIQ	ASD 107.07 NT 109.92	ASD 107.14 NT 110.80	ASD 107.32 NT 107.53

Abbreviations: IED= Intra-extra dimensional set-shifting task, TE= Total errors, RCMP= Raven's coloured progressive matrices, WCST= Wisconsin card sorting test, PE= Perseverative errors, WAIS= Wechsler adult intelligence scale, KBIT= Kaufman brief intelligence test, WASI= Wechsler abbreviated scale of intelligence, ESE= Extradimensional shift errors, MCST= Modified card sorting test, C-WISC= Chinese version wechsler intelligence scale for children

			ASD		NT							
Source	**Age Range Mean (SD)**	**Task**	**Sample Size**	**Mean (SD)**	**Sample Size**	**Mean (SD)**	**Diagnostic Criteria**	**IQ Tool**	**Matched**	**FSIQ Mean**	**NVIQ/PIQ Mean**	**VIQ Mean**
D'Cruz 2016	7-44 ASD 17.4 (8.6) NT 18.6 (8.4)	2CRL	17 M=12 F=5	PE 7.1 (9.0)	23 M=18 F=5	PE 3.2 (4.6)	DSM-IV-TR ADOS ADI (N=15)	--	Age Sex FSIQ PIQ	ASD 103.9 NT 110.9	ASD 106.7 NT 107.5	ASD 100.4 NT 113.0
D'Cruz 2013	8-44 ASD 15.34 (7.75) NT 18.24 (8.12)	PRL	41 M=33 F=8	PE 0.95 (1.00)	37 M=31 F=6	PE 1.81 (3.49)	DSM-IV-TR ADOS ADI-R (N= 37)	--	Age Sex FSIQ PIQ	ASD 103.9 NT 108.7	ASD 104.73 NT 107.59	ASD 102.00 NT 109.00
Dichter 2007	-- ASD 22.9 (5.2) NT 23.2 (5.7)	WCST-64	14 M=13 F=1	PE 32.71 (20.7)	15 M=14 F=1	PE 10.67 (5.61)	DSM-IV ADI-R ADOS	WASI	Age Sex FSIQ PIQ VIQ	ASD 105.0 NT 105.7	ASD 104.1 NT 103.7	ASD 105.1 NT 106.3
Dichter 2009	-- ASD 23.3 (11.1) NT 28 (7.9)	WCST-C	15 M=14 F=1	PE 12.73 (11)	19 M=18 F=1	PE 6.2 (3.2)	ADI-R (N=5) ADOS (N=10)	WASI (N ASD =9)*	Age Sex FSIQ* PIQ* VIQ*	ASD 102.8 NT 114	ASD 104 NT 113	ASD 101.7 NT 113
Garcia-Villamisar 2002	-- ASD 23.5 (4.31) NT 21.19 (2.51)	WCST	16 M=8 F=8	PE 29.25 (10.36)	16 M=8 F=8	PE 9.18 (6.89)	DSM-IV	SPM	Age Sex NVIQ	--	ASD 42.75 NT 43.69	--
Geurts 2020	60-85 ASD 65.8 (5.6) NT 69.7 (5.6)	WCST-C	50 M=50	PR 12.3 (8.4)	51 M=51	PR 15 (10.7)	DSM-IV	WAIS	FSIQ	ASD 110.7 NT 110.7	--	--
Goddard 2014	8-17 ASD 12.5 (2.8) NT 12.1 (2.2)	WCST-64-C	63 M=51 F=12	PE 9.32 (5.27)	63 M=51 F=12	PE 7.16 (3.37)	DSM-IV-TR	WASI	Age Sex FSIQ	ASD 103.6 NT 104.7	--	--

Abbreviations: WCST= Wisconsin card sorting test, PE= Perseverative errors, WAIS= Wechsler adult intelligence scale, WASI= Wechsler abbreviated scale of intelligence, 2CRL= Two-choice reversal learning task, PRL= Probabilistic reversal learning task, WCST-64= 64 cards total, WCST-C= Computerised version, PR= Perseverative responses, SPM= Standard progressive matrices

Table 1. Data Extracted from Included Studies (cont.)

			ASD		NT							
Source	**Age Range Mean (SD)**	**Task**	**Sample Size**	**Mean (SD)**	**Sample Size**	**Mean (SD)**	**Diagnostic Criteria**	**IQ Tool**	**Matched**	**FSIQ Mean**	**NVIQ/PIQ Mean**	**VIQ Mean**
Goldstein 2001	-- ASD 18.15 (10.14) NT 18.96 (10.1)	WCST	103 F= 13.6%	PE 20.31 (15.24)	103 F= 10.7%	PE 12.6 (10.25)	ADI ADOS	WISC (N=114) WAIS (N= 92)	Age Sex FSIQ PIQ VIQ	ASD 97.57 NT 99.12	ASD 95.65 NT 98.6	ASD 99.6 NT 99.94
Gomez-Perez 2016	7-12 ASD 9.35 (1.28) NT 9.26 (1.46)	WCST-64	34 M=30 F=4	PR 14.71 (12)	34 M=20 F=14	PR 14.84 (13.22)	DSM-IV DSM-5	WISC	Age FSIQ	ASD 106.65 NT 108.12	--	--
Gomez-Perez 2020	7-13 ASD 10.07 (1.65) NT 9.74 (1.56)	WCST-64	43 M=38 F=5	PE 15.71 (12.36)	62 M=33 F=29	PE 14.77 (9.61)	ADOS ADI-R	WISC	Age FSIQ	ASD 94.16 NT 91.56	--	--
Griebling 2010	-- ASD 17.9 (10) NT 18.6 (9)	WCST	24 M=22 F=2	PE 15.38 (9.96)	38 M=36 F=2	PE 8.74 (6.71)	ADI-R ADOS	WAIS WISC	Age Sex FSIQ	ASD 104 NT 104	--	--
Hill 2006	16-64 ASD 31.09 (13.14) NT 33.45 (14.54)	MCST	22 M= 16 F=6	PE 0.82 (1.47)	22 M=14 F=8	PE 0.73 (0.83)	DSM-IV	WAIS	Age FSIQ	ASD 110.5 NT 107.91	--	--
Kado 2020a	5-9 ASD 8.2 (1) NT 8.1 (1)	KWCST	30 M=22 F=8	PE 9.43 (9)	30 M=22 F=8	PE 6.6 (6.27)	DSM-IV-TR DSM-5	WISC	Age Sex FSIQ	ASD 94 NT --	--	--
Kaddo 2020b	10-15 ASD 12.2 (1.3) NT 12.2 (1.4)	KWCST	39 M=34 F=5	PE 7.31 (7.04)	39 M=34 F=5	PE 4.82 (4.19)	DSM-IV-TR DSM-5	WISC	Age Sex FSIQ	ASD 96.9 NT --	--	--

Abbreviations: WCST= Wisconsin card sorting test, PE= Perseverative errors, WAIS= Wechsler adult intelligence scale, MCST= Modified card sorting test, WISC= Wechsler intelligence scale for children, WCST-64= 64 cards total, PR= Perseverative responses, KWCST= Keio version WCST 48 cards+ sorting criterion change after 6 correct trials

			ASD		NT							
Source	**Age Range Mean (SD)**	**Task**	**Sample Size**	**Mean (SD)**	**Sample Size**	**Mean (SD)**	**Diagnostic Criteria**	**IQ Tool**	**Matched**	**FSIQ Mean**	**NVIQ/PIQ Mean**	**VIQ Mean**
Kaland 2008	-- ASD 16.4 (2.84) NT 15.6 (3.07)	WCST-C	13 M=13	PE 11.46 (5.21)	13 M=13	PE 10.31 (4.37)	ICD-10 ADI-R ADOS	WISC	Age FSIQ PIQ VIQ	ASD 109 NT 109.62	ASD 107 NT 106.92	ASD 108.92 NT 110.15
Kaufmann 2013	-- ASD 14.7 (5) NT 13.8 (5.3)	IED	10 M=8 F= 2	ESE 15.6 (9.8)	10 M=8 F=2	ESE 20 (13.1)	DSM-IV-TR ADOS ADI-R	WISC WAIS	Age Sex FSIQ PIQ VIQ	ASD 102.3 NT 109.5	ASD 95.8 NT 106	ASD 107.6 NT 114
Keary 2009	8-45 ASD 19.8 (10.2) NT 18.6 (9.06)	WCST	32 M= -- F= --	PE 17 (11)	34 M= -- F= --	PE 8.4 (6.4)	ADI-R ADOS	WAIS WISC	Age FSIQ PIQ VIQ	ASD 102.9 NT 104	ASD 97.8 NT 102.6	ASD 106.9 NT 104.7
Kiep 2017	19-60 ASD 37.1 (9.8) NT 37.9 (11.1)	WCST-C	139 M=99 F=40	PE 11.5 (14.77)	60 M=35 F=25	PE 10.23 (11.21)	ADI-R	WAIS	Age FSIQ	ASD 109.11 NT 110.71	--	--
Kilincaslan 2010	7-16 ASD 12.44 (2.87) NT 11.96 (2.36)	WCST-C	21 M=18 F=3	PE 21.64 (10.42)	18 M=15 F=3	PE 13.8 (5.43)	DSM-IV	WISC	Age Sex FSIQ PIQ VIQ	ASD 105.52 NT 107.27	ASD 98.35 NT 107.44	ASD 111.17 NT 106
Landa 2005	7-17 ASD 11.01 (2.89) NT 11 (2.85)	IED	19 M= -- F= --	ESE 3.85 (5.62)	19 M= -- F= --	ESE 13.57 (12.44)	ADI-R ADOS	WISC WAIS	Age Sex FSIQ PIQ VIQ	ASD 109.7 NT 113.4	ASD 104.6 NT 108.5	ASD 113.5 NT 115.6
Landsiedel 2020	-- ASD 34.84 (11.42) NT 38.24 (13.19)	WCST-C	21 M=-- F=--	PE 13.52 (10.98)	21 M=-- F=--	PE 5.62 (4.4)	ICD-10 DSM-IV DSM-5	WASI (ASD/NT N= 25/23) *	Age Sex FSIQ PIQ VIQ	ASD 104.32 NT 104.87	ASD 102.92 NT 104.65	ASD 105.2 NT 104.35

Abbreviations: IED= Intra-extra dimensional set-shifting task, WCST= Wisconsin card sorting test, PE= Perseverative errors, WAIS= Wechsler adult intelligence scale, WASI= Wechsler abbreviated scale of intelligence, ESE= Extradimensional shift errors, WISC= Wechsler intelligence scale for children, WCST-C= Computerised version, *= IQ full sample reported/sub-sample remained matched

Table 1. Data Extracted from Included Studies (cont.)

			ASD		NT							
Source	**Age Range Mean (SD)**	**Task**	**Sample Size**	**Mean (SD)**	**Sample Size**	**Mean (SD)**	**Diagnostic Criteria**	**IQ Tool**	**Matched**	**FSIQ Mean**	**NVIQ/PIQ Mean**	**VIQ Mean**
Li 2014	6-12 ASD 9.59 (2.29) NT 10.2 (1.53)	WCST-48-C	37 M= -- F=--	PE 19.51 (7.7)	31 M=-- F=--	PE 15.48 (5.81)	DSM-IV	RSPM	Age Sex NVIQ	--	ASD 109.76 NT 113	--
Lopez 2005	18-45 ASD 29.1 (8) NT 29.4 (11.4)	WCST	17 M=14 F=3	PE 47.23 (36.82)	17 M=11 F=6	PE 25.12 (26.13)	ADI-R ADOS	WAIS	Age Sex PIQ	ASD 77 NT 89	ASD 84.1 NT 87.6	ASD 73 NT 92
Maister 2013a	11-13 ASD 12.2 (0.6) NT 12.1 (0.2)	IED	14 M=14	ESE 8.42 (9.23)	14 M=13 F=1	ESE 6 (7.46)	ADI-R	RSPM	Age NVIQ	--	ASD 43.9 NT 46.4	--
Maister 2013b	9-14 ASD 11.8 (1.4) NT 11.8 (1.1)	IED	14 M=13 F= 1	ESE 10.42 (7.88)	14 M=11 F=3	ESE 7.07 (5.38)	ADI-R	RSPM	Age NVIQ	--	ASD 41.6 NT 44.3	--
Merchan-Naranjo 2016	8-18 ASD 12.8 (2.5) NT 12.9 (2.7)	WCST	24 M=23 F= 1	PE 22.6 (15.4)	32 M=30 F=2	PE 14 (9.9)	DSM-IV ADOS (N=10)	WAIS WISC (NT 2 subtests)	Age Sex FSIQ	ASD 99.2 NT 106.81	--	--
Miller 2015	6-44 ASD 15.1 (8.02) NT 15.9 (7.5)	PCET	51 M=-- F= --	RE 11.29 (12.24)	52 M=-- F=--	RE 6.25 (9.92)	DSM-IV ADI-R ADOS	DAS WASI (ASD/NT N=60/55) *	Age Sex NVIQ	ASD 100.1 NT 108.9	ASD 101.1 NT 106.6	ASD 100.1 NT 110.2
Minshew 1992	15-40 ASD 21.13 (8.02) NT 21.33 (8.3)	WCST	15 M=15	PE 18.33 (18.25)	15 M=-- F= --	PE 12.75 (9.91)	DSM-III-R ADI ADOS	WAIS	Age Sex FSIQ PIQ VIQ	ASD 95.73 NT 96.47	ASD 92.87 NT 93.27	ASD 98.53 NT 99.07

Abbreviations: IED= Intra-extra dimensional set-shifting task, WCST= Wisconsin card sorting test, PE= Perseverative errors, WAIS= Wechsler adult intelligence scale, WASI= Wechsler abbreviated scales of intelligence, ESE= Extradimensional shift errors, WISC= Wechsler intelligence scale for children, WCST-48-C= 48 cards total computerised version, PR= Perseverative responses, PCET= Penn conditional exclusion test, RE= Regressive errors, RSPM= Raven standard progressive matrices, DAS= Differential ability scales, *= IQ full sample reported

			ASD		NT							
Source	**Age Range Mean (SD)**	**Task**	**Sample Size**	**Mean (SD)**	**Sample Size**	**Mean (SD)**	**Diagnostic Criteria**	**IQ Tool**	**Matched**	**FSIQ Mean**	**NVIQ/PIQ Mean**	**VIQ Mean**
Minshew 1997	12-40 ASD 20.91 (9.69) NT 21.21 (9.99)	WCST	33 M=29 F=4	PE 16.45 (15.48)	33 M=29 F=4	PE 13.27 (11.13)	ADI ADOS	WAIS	Age Sex FSIQ	ASD 100.09 NT 100.48	ASD 97.45 NT 99.09	ASD 102.48 NT 101.30
Minshew 2002	-- ASD 21.41 (9.68) NT 21.23 (9.81)	WCST	90 M=-- F=--	PE 18.72 (14.19)	107 M=-- F=--	PE 10.46 (9.61)	ADI-R ADOS	WAIS	Age Sex FSIQ PIQ VIQ	ASD 97.95 NT 100.9	ASD 95.51 NT 99.95	ASD 100.11 NT 101.50
Ozonoff 1995	8-15 ASD 11.9 (2.7) NT 11.9 (1.3)	WCST-C	10 M=9 F=1	PR 21.4 (17.5)	11 M=8 F=3	PR 22.6 (15.3)	DSM-IV	WISC	Age FSIQ PIQ VIQ	ASD 98.1 NT 99.1	ASD 101 NT 98.3	ASD 95.8 NT 100
Ozonoff 2004	6-47 ASD 15.7 (8.7) NT 16 (7.6)	IED	79 M=91%	ESE 13.6 (11.5)	70 M=83%	ESE 8.4 (8.8)	ADI-R ADOS	WISC WAIS	Age Sex FSIQ	ASD 106.3 NT 106	ASD 106 NT 105	ASD 104.9 NT 106.1
Ozonoff 2000	6-20 ASD 13.5 (4.05) NT 12.5 (3.2)	IED	35 M=31 F=4	ESE 13.88 (12.17)	27 M=-- F=--	ESE 9.9 (9.3)	DSM-IV ADI-R ADOS	WISC	Age Sex FSIQ PIQ VIQ	ASD 111.19 NT 111	ASD 105.36 NT 110.6	ASD 114.02 NT 109.9
Panerai 2014	-- ASD 9.23 (3.31) NT 10.94 (2.87)	WCST	19 M=15 F=4	PE 15.4 (6.73)	21 M=14 F=7	PE 8.56 (2.15)	DSM-IV-TR	RCPM	Age Sex NVIQ	--	ASD 23.58 NT 23.38	--
Pellicano 2006	4-7 ASD 5.5 (0.9) NT 5.4 (0.9)	SST	40 M=35 F=5	TNTC 45.05 (9.79)	40 M=31 F=9	TNTC 38.22 (7.92)	DSM-IV ADI-R	PPVT LIPS	Age Sex NVIQ VIQ	--	ASD 113.58 NT 112.52	ASD 101.15 NT 103.25

Abbreviations: IED= Intra-extra dimensional set-shifting task, RCPM= Raven's coloured progressive matrices, WCST= Wisconsin card sorting test, PE= Perseverative errors, WAIS= Wechsler adult intelligence scale, ESE= Extradimensional shift errors, WISC= Wechsler intelligence scale for children, WCST-C= Computerised version, PR= Perseverative responses, SST= Set shifting task, TNTC= Total number of trial to criterion, PPVT= Peabody picture vocabulary test, LIPS= Leiter international performance scale

Table 1. Data Extracted from Included Studies (cont.)

			ASD		NT							
Source	**Age Range Mean (SD)**	**Task**	**Sample Size**	**Mean (SD)**	**Sample Size**	**Mean (SD)**	**Diagnostic Criteria**	**IQ Tool**	**Matched**	**FSIQ Mean**	**NVIQ/PIQ Mean**	**VIQ Mean**
Perez 2009	7-16 ASD 10.8 (3.4) NT 11.1 (2.6)	CST-C	15 M=14 F=1	PE 11.7 (8.5)	16 M=9 F=7	PE 5.4 (3.8)	DSM-IV	WASI	Age VIQ	ASD 111.2 NT 123	--	ASD 119.1 NT 119.9
Robinson 2009	-- ASD 12.5 (2.7) NT 12.08 (2.3)	WCST-64-C	54 M=42 F= 12	PE 108.11 (21.92)	54 M=42 F=12	PE 116.24 (20.81)	DSM-IV	WASI (2 subtests)	Age Sex FSIQ	ASD 103.53 NT 104.80	--	--
Rumsey 1985	18-39 ASD 27 (7) NT 28 (5)	WCST	9 M=9	PE 29 (24.5)	10 M=10	PE 8.2 (5.2)	DSM-III	WAIS	Age FSIQ PIQ	ASD 104 NT 113	ASD 104 NT 111	ASD 103 NT 113
Sachse 2013	14-33 ASD 19.2 (5.1) NT 19.9 (3.6)	IED	30 M=27 F=3	TE 31.8 (28.1)	28 M=24 F=4	TE 21.9 (19.8)	DSM-IV-TR ADI-R ADOS	RSPM	Age Sex NVIQ	--	ASD 105.3 NT 109.3	--
Sawa 2013	10-15 ASD 13.2 (1.7) NT 13.1 (1.9)	KWCST	19 M=17 F=2	PE 7.89 (4.52)	19 M=17 F=2	PE 3.58 (4.75)	DSM-IV-TR	WISC	Age Sex FSIQ PIQ VIQ	ASD 95.95 NT 97.32	ASD 95.26 NT 97.84	ASD 92.96 NT 97.21
Sawaya 2019	12-16 ASD 14.6 (15.5) NT 13 (13.5)	CSG	17 M=16 F=1	RSE 4.26 (2.48)	17 M=11 F=6	RSE 2.9 (2.15)	ADI-R ADOS	WISC subtests	Age NVIQ (1)* VIQ (2)*	--	ASD 98 NT 95	ASD 94 NT 104
Solomon 2015	18-40 ASD 22.95 (5.11) NT 23.36 (4.15)	PST	22 M=18 F=4	LSR 0.35 (0.16)	25 M=21 F=4	LSR 0.33 (0.13)	DSM-IV-TR ADOS	WASI	Age Sex FSIQ PIQ VIQ	ASD 112.64 NT 114.17	ASD 112.18 NT 112.04	ASD 110.82 NT 112.54

Abbreviations: IED= Intra-extra dimensional set-shifting task, TE= Total errors, WCST= Wisconsin card sorting test, PE= Perseverative errors, WAIS= Wechsler adult intelligence scale, WASI= Wechsler abbreviated scale of intelligence, WISC= Wechsler intelligence scale for children, WCST-64-C= 64 cards total computerised version, KWCST= Keio version 48 cards+ sorting criterion change after 6 correct trials, CST-C= Card sorting task computerised, CSG= Computerised sequencing game, RSE= Rule shift errors, PST=Probabilistic selection task, LSR= Lose shift responses, RSPM= Raven standard progressive matrices, *= ASD missing data

			ASD		NT							
Source	**Age Range Mean (SD)**	**Task**	**Sample Size**	**Mean (SD)**	**Sample Size**	**Mean (SD)**	**Diagnostic Criteria**	**IQ Tool**	**Matched**	**FSIQ Mean**	**NVIQ/PIQ Mean**	**VIQ Mean**
Van Eylen 2011	8-18 ASD 11.33 (2.18) NT 11.13 (2.22)	WCST-CTS	40 M=36 F=4	TE 2.26 (2.73)	40 M=36 F=4	TE 1.23 (1.33)	DSM-IV-TR	WISC WAIS subtests	Age Sex FSIQ PIQ VIQ	ASD 105.45 NT 106.76	ASD 104.25 NT 103.88	ASD 106.68 NT 109.65
Van Eylen 2015	8-18 ASD 12.21 (2.58) NT 12.48 (2.72)	WCST-CTS	50 M=30 F=20	PE 1.33 (1.52)	50 M=30 F=20	PE 0.62 (0.48)	DSM-IV-TR	WISC WAIS subtests	Age Sex FSIQ PIQ	ASD 104.32 NT 107.72	ASD 104.32 NT 103.84	ASD 104.32 NT 111.60
Wang 2018	-- ASD 9.05 (2.38) NT 8.92 (1.68)	MCST	21 M=20 F=1	PE 8.33 (5.5)	28 M=19 F=9	PE 4.54 (3.81)	DSM-IV	C-WISC	Age PIQ	ASD 99.53 NT 110.14	ASD 93.32 NT 101.36	ASD 106.21 NT 116.04
Williams 2013	-- ASD 10.6 (2.01) NT 10.59 (1.31)	MCST	21 M=-- F=--	PE 8.19 (5.76)	21 M=-- F=--	PE 4.19 (3.43)	DSM-IV	WASI	Age PIQ VIQ	--	ASD 110.19 NT 107.48	ASD 103.57 NT 106.48
Yerys 2009	6-14 ASD 10.19 (2) NT 10.26 (2.08)	IED	27 M=-- F=--	ESE 14.9 (10.29)	51 M=-- F=--	ESE 13.88 (11.4)	DSM-IV-TR ADI/-R ADOS	WASI WISC (3)# (ASD/NT N=42/84) *	Age Sex FSIQ	ASD 111.95 NT 113.18	--	--
Yeung 2016	6-17 ASD 10.09 (2.58) NT 11.55 (3.53)	WCST-CTS	25 M=19 F=6	PE 2.47 (3.56)	25 M=14 F=11	PE 0.87 (0.72)	DSM-IV-TR ADI-R	WISC subtests	Age Sex FSIQ	ASD 105.24 NT 110.48	--	--
Zhang 2015	-- ASD 18.9 (3.64) NT 19.2 (2.96)	WCST	37 M=31 F=6	PE 21.16 (13.81)	80 M=67 F=13	PE 24.08 (14.78)	DSM-IV-TR	RSPM	Age Sex NVIQ	--	ASD 103.2 NT 108.1	--

Abbreviations: IED= Intra-Extra dimensional set-shifting task, TE= Total errors, WCST= Wisconsin Card Sorting Test, PE= Perseverative errors, WISC= Wechsler Intelligence Scale for children, WAIS= Wechsler Adult Intelligence Scale, WASI= Wechsler Abbreviated Scale of Intelligence, ESE= Extradimensional Shift Errors, MCST= Modified Card Sorting Test, C-WISC= Chinese version Wechsler Intelligence Scale for children, WCST-CTS= Controlled task switching (shape & colour – sorting criterion changes randomly after 7, 8 or 9 corrects trials), RSPM= Raven standard progressive matrices, *= IQ full sample reported/sub-sample remained matched, #= ASD missing data

Table 2. Task Subgroup Results

	N	SMD	95% CI
WCST	21	0.57	0.36- 0.79
WCST-V	12	0.40	0.17- 0.63
IED	11	0.26	-0.009-0.53

SMD= Standardised Mean Difference (Hedges' g), CI= Confidence Interval, WCST= Wisconsin Card Sorting Test, WCST-V= Variations of WCST, IED= Intra-Extra Dimensional Set Shift Task, Subgroup differences *p < .05

Table 3. Outcome Subgroup Results

	N	SMD	95% CI
PE	36	0.55*	0.40-0.70
ESE	9	0.16*	-0.12-0.44
OO	10	0.35*	0.09-0.61

SMD= Standardised Mean Difference (Hedges' g), CI= Confidence Interval, PE= Perseverative Errors, ESE= Extra-Dimensional Shift Errors, OO= Other Outcomes, Subgroup differences *p < .05

Table 4. Age Subgroup Results

	N	SMD	95% CI
Adults	22	0.53	0.33-0.73
Adolescents	9	0.27	0.01-0.52
Children	24	0.44	0.26-0.63

SMD= Standardised Mean Difference (Hedges' g), CI= Confidence Interval, Subgroup differences *p < .05

Table 5. Quality Assessment

Study	S1	S2	S3	S4	C1	E1	E2	E3	Total
Alsaedi 2020	✸	✸	✸	✸	✸✸	✸	✸	✸	9
Ambery 2006	-	-	-	-	-	✸	✸	✸	3
Braden 2017	✸	✸	✸	✸	✸✸	✸	✸	✸	9
Brady 2013	✸	-	-	-	✸	✸	✸	✸	5
Chan 2011	✸	-	✸	-	✸✸	✸	✸	-	6
Chen 2016	✸	-	✸	-	✸	✸	✸	✸	6
D'Cruz 2016	-	✸	✸	✸	✸✸	✸	✸	✸	8
D'Cruz 2013	-	✸	✸	✸	✸✸	✸	✸	✸	8
Dichter 2007	-	-	✸	-	✸✸	✸	✸	✸	6
Dichter 2009	-	-	✸	-	✸✸	✸	✸	✸	6
Garcia-Villamisar 2002	-	-	✸	✸	✸✸	✸	✸	✸	7
Geurts 2020	✸	✸	✸	✸	✸	✸	✸	✸	8
Goddard 2014	✸	✸	✸	-	✸✸	✸	✸	✸	8
Goldstein 2001	-	-	-	✸	✸✸	✸	✸	✸	6
Gomez-Perez 2016	✸	✸	✸	-	✸✸	✸	✸	✸	8
Gomez-Perez 2020	✸	✸	✸	-	✸✸	✸	✸	✸	8
Griebling 2010	-	✸	✸	✸	✸✸	✸	✸	-	7
Hill 2006	✸	✸	✸	-	✸✸	✸	✸	✸	8
Kado 2020 a	✸	-	-	-	✸	✸	✸	✸	5
Kado 2020 b	✸	-	-	-	✸	✸	✸	✸	5
Kaland 2008	✸	-	✸	-	✸✸	✸	✸	✸	7
Kaufmann 2013	-	-	-	✸	✸✸	✸	✸	✸	6
Keary 2009	-	✸	✸	✸	✸✸	✸	✸	✸	8
Kiep 2017	✸	-	-	-	✸✸	✸	✸	✸	6
Kilincaslan 2010	-	-	✸	-	✸✸	✸	✸	✸	6
Landa 2005	-	-	✸	✸	✸✸	✸	✸	✸	7
Landsiedel 2020	✸	-	-	-	✸✸	✸	✸	-	5
Li 2014	✸	-	-	-	✸	✸	✸	-	4
Lopez 2005	✸	-	-	✸	✸✸	✸	✸	✸	7
Maister 2013 a	✸	✸	✸	-	✸✸	✸	✸	✸	8
Maister 2013 b	✸	✸	✸	-	✸✸	✸	✸	✸	8

Table 5. Quality Assessment (cont.)

Study	S1	S2	S3	S4	C1	E1	E2	E3	Total
Merchan-Naranjo 2016	✸	-	✸	✸	✸✸	✸	✸	✸	8
Miller 2015	✸	✸	✸	✸	✸✸	✸	✸	✸	9
Minshew 1992	✸	-	-	✸	✸✸	✸	✸	✸	7
Minshew 1997	✸	-	✸	✸	✸✸	✸	✸	✸	8
Minshew 2002	-	-	✸	✸	✸✸	✸	✸	✸	7
Ozonoff 1995	-	-	-	✸	✸✸	✸	✸	✸	6
Ozonoff 2004	-	-	-	-	✸✸	✸	✸	✸	5
Ozonoff 2000	✸	✸	-	✸	✸✸	✸	✸	✸	8
Panerai 2014	✸	✸	✸	-	✸✸	✸	✸	✸	8
Pellicano 2006	✸	✸	✸	✸	✸✸	✸	✸	✸	9
Perez 2009	-	-	✸	-	-	✸	✸	✸	4
Robinson 2009	✸	✸	✸	-	✸✸	✸	✸	✸	8
Rumsey 1985	✸	-	-	-	-	✸	✸	✸	4
Sachse 2013	✸	-	✸	✸	✸✸	✸	✸	✸	8
Sawa 2013	✸	-	✸	✸	✸✸	✸	✸	✸	8
Sawaya 2019	✸	-	✸	-	-	✸	✸	✸	5
Solomon 2015	-	-	-	-	✸✸	✸	✸	✸	5
Van Eylen 2011	✸	-	✸	-	✸✸	✸	✸	✸	7
Van Eylen 2015	✸	-	✸	✸	✸✸	✸	✸	✸	8
Wang 2018	-	-	✸	-	-	✸	✸	✸	4
Williams 2013	✸	-	-	✸	✸✸	✸	✸	✸	7
Yerys 2009	✸	-	✸	-	✸✸	✸	✸	✸	7
Yeung 2016	✸	-	✸	✸	✸✸	✸	✸	✸	8
Zhang 2015	✸	-	✸	✸	-	✸	✸	✸	6

Selection: S1 - Adequate case definition, S2 - Representativeness of cases, S3 - Selection of controls, S4 - Definition of controls

Comparability: C1 - Comparability of cases and controls on the basis of the design or analysis

Exposure: E1 - Ascertainment of exposure, E2 - Same method of ascertainment for cases and controls, E3 - Non-response rate