Longitudinal studies of challenging behaviours in autistic children and adults: A systematic review and meta-analysis

Clinical Psychology Review 104 (2023) 102320

Contents lists available at ScienceDirect

Clinical Psychology Review

journal homepage: www.elsevier.com/locate/clinpsychrev

Review

Longitudinal studies of challenging behaviours in autistic children and adults: A systematic review and meta-analysis

Dawn Adams [a,b,*], Nicole Dargue [a,b], Jessica Paynter [b,c]

[a] Autism Centre of Excellence, Griffith University, Messines Ridge Road, Mt Gravatt, Brisbane, QLD 4122, Australia
[b] Griffith Institute for Educational Research, Griffith University, Messines Ridge Road, Mt Gravatt, Brisbane, QLD 4122, Australia
[c] School of Applied Psychology and Hopkins Centre, Griffith University, Gold Coast, QLD 4222, Australia

ARTICLE INFO

Keywords:
Autism
Behaviours
Trajectory
Longitudinal
Externalising

ABSTRACT

Autistic children and adults are at increased risk of showing behaviours that may be described as challenging, however, little is known about whether or how these behaviours may change over time. Given the profound impact that challenging behaviour can have on both the autistic individual and their support network, it is critical that the trajectory of challenging behaviours be better understood. This systematic review and meta-analysis identified and synthesised observational longitudinal studies of challenging behaviour in autistic individuals. Fifty-six studies were included in the systematic review, and the effect sizes of 37 independent samples arising from 34 of these reports were examined through meta-analysis. Crucially, across the 37 samples, scores on the measures of challenging behaviour reduced by a small, yet significant, extent over time. Although age of the sample at baseline assessment did not moderate the effect, the time between the baseline assessment and final follow-up and age at final follow-up both moderated the magnitude of the effect, with challenging behaviour scores reducing to a larger extent in (a) studies with longer intervals between baseline and final follow-up and (b) studies with older samples at follow-up. The results from the current systematic review and meta-analysis have both theoretical and practical implications for understanding challenging behaviour over time in autistic individuals. Avenues for future research are also highlighted that may allow better understanding, and therefore support of, challenging behaviour in autistic individuals.

1. Introduction

Autism spectrum disorder (ASD, hereafter autism) is a lifelong neurodevelopmental condition characterised by differences or difficulties in social communication as well as the presence of restricted and repetitive behaviours and interests (American Psychiatric Association, 2013). Population-based prevalence studies suggest that one in 54–100 individuals have a diagnosis of autism[1] (see review by Chiarotti & Venerosi, 2020). The prevalence of autism is higher in males than females (ratio male:females diagnosed is 3.8:1; Loomes, Hull, & Mandy, 2017) and in those with a co-occurring intellectual disability (ID; Tonnsen et al., 2016). One-third of autistic individuals have a co-occurring diagnosis of an ID (defined by intelligence quotient [IQ] $\leq$ 70 and difficulties with adaptive functioning), with a higher rate in females than males (39% vs. 32%; Maenner et al., 2020). Autistic individuals are reported to show elevated rates of behaviours that may be described as challenging, hereafter referred to as challenging behaviours (McClintock, Hall, & Oliver, 2003; Simó-Pinatella, Mumbardó-Adam, Alomar-Kurz, Sugai, & Simonsen, 2019). The focus of this review and meta-analysis is on understanding the level of challenging behaviour over time and any moderators of change over time.

1.1. Challenging behaviour in autistic individuals

Autistic children and adults have an increased likelihood of showing challenging behaviour, or behaviours that challenge, compared to their neurotypical peers (McClintock et al., 2003) and individuals with other neurodevelopmental disorders (Simó-Pinatella et al., 2019). Emerson (1995, p. 44) defined challenging behaviour as culturally abnormal behaviour(s) of such an intensity, frequency or duration that the physical safety of the person or others is likely to be placed in serious jeopardy, or behaviour which is likely to seriously limit use of, or result in

* Corresponding author at: Autism Centre of Excellence, Griffith University, Messines Ridge Road, Mt Gravatt, Brisbane, QLD 4122, Australia.
E-mail address: Dawn.Adams@Griffith.edu.au (D. Adams).

[1] The term autistic has been used in line with the preferred term of the community (Kenny et al., 2016), but the range of preferences for terms is acknowledged.

https://doi.org/10.1016/j.cpr.2023.102320
Received 16 May 2022; Received in revised form 23 February 2023; Accepted 12 July 2023
Available online 20 July 2023

the person being denied access to, ordinary community facilities. Challenging behaviour is therefore an umbrella term for behaviours which include self-injurious behaviour, aggression, and destruction of the environment, amongst other behavioural topographies. Whilst a single term is used within this document to describe this range of behaviours, this does not mean to imply that all topographies of challenging behaviour should be considered as a single entity. Indeed, the aetiology of the differing forms of challenging behaviour places varying emphasis on specific neurobiological processes (e.g., pain) as well as social and non-social reinforcement at different stages of onset and maintenance for different topographies (see review by Vandewalle & Melia, 2021). Supports and strategies for different topographies also differ (Oliver et al., 2022).

It is critical that the trajectory of challenging behaviours are understood, as challenging behaviour of increasing severity can cause long-term damage or injury, result in residential placement, reduce quality of life, and impact the system of supports. Challenging behaviour is associated with poorer physical health outcomes (Oliver et al., 2003) and is a barrier to accessing healthcare to prevent or reduce these poor outcomes (Walsh, Lydon, O'Dowd, & O'Connor, 2020). Higher levels of challenging behaviour are associated with reduced physical, social, and psychological quality of life in both autistic children and adults (Chiang & Wineman, 2014) as well as in their parents (e.g., Estes et al., 2009) and other family members (Yacoub, Dowd, McCann, & Burke, 2018). Challenging behaviour can also impact upon the well-being and stability of paid support professionals (Nevill, Istvan, & Havercamp, 2021).

1.2. Prevalence of challenging behaviour in autistic individuals

A large body of cross-sectional research has explored the prevalence of challenging behaviour in autistic individuals. Accurate prevalence estimates are difficult to determine due to methodological differences, variation in the definition of the behaviours being measured, and variability in participant characteristics (Mazurek, Kanne, & Wodka, 2013; Summers et al., 2017). Further, population-based studies of challenging behaviour in autistic individuals are lacking (Soke, Rosenberg, Hamman, et al., 2016). Large samples ($N > 13{,}000$) from national registers suggest that self-injurious behaviour occurs in approximately 30% of autistic individuals (Soke, Rosenberg, Hamman, et al., 2017); a meta-analysis of 37 papers ($n = 14{,}379$ participants) suggested a prevalence rate of 42% (Steenfeldt-Kristensen, Jones, & Richards, 2020). The prevalence of aggression (from samples of 1380 and 1584 children) is suggested to be 53–56% (Kanne & Mazurek, 2011; Mazurek et al., 2013). The prevalence of some, but not all, challenging behaviours can vary based on the presence and level of co-occurring ID. For example, autistic children with a severe-profound ID were 1.84 times more likely to show self-injurious behaviour than those with an average or above-average IQ but ID did not impact prevalence of aggression (Kurzius-Spencer et al., 2018). Prevalence rates may also be impacted by an interaction between ID and the presence/absence of autism characteristics (Esteves, Perry, Spiegel, & Weiss, 2021). These findings suggest that IQ may be a factor contributing to some of the variability in prevalence rates of differing challenging behaviours.

While research has used a variety of methods to document prevalence of challenging behaviours in autistic individuals, less is known about whether, or how, the level or severity of challenging behaviour changes over time or with age. This means that many important questions about the trajectory of challenging behaviours in autistic individuals remain unanswered. The trajectory of challenging behaviour across age for autistic individuals is very unlikely to be discovered or described by any single study but can be better addressed with meta-analytic methods that aim not only to document the levels of challenging behaviour reported over time, but also to explore the factors which may impact any change reported across studies. In this study, we therefore aim to synthesise the longitudinal data on challenging behaviour in autistic individuals across the life course in order to identify stability or change in the level of such behaviours over time.

Establishing the longitudinal trajectory of challenging behaviours is important in order to inform models of the development of these behaviours as well as to inform systems and supports for the behaviour. The presence of challenging behaviour impacts the quality of life and well-being of the autistic individual (Chiang & Wineman, 2014) and those who support them (e.g., Totsika, Hastings, Emerson, Berridge, & Lancaster, 2011), and results in increased access to services (Adams et al., 2016). These relationships are stronger for chronic and persistent challenging behaviour (e.g., Adams et al., 2018). From a clinical perspective, being aware of the trajectory of challenging behaviours across different age groups will help services to plan effectively for the future needs of autistic individuals (Davies & Oliver, 2013). Such data could also help services to identify critical timepoints at which to intervene to prevent any escalation or increase in challenging behaviour from occurring and to assist services in planning for ongoing need for supports, particularly in light of an ageing population.

1.3. Factors moderating the prevalence and severity of challenging behaviour in autistic individuals

1.3.1. Age and time between assessments

Few cross-sectional studies have reported on the level or severity of challenging behaviour in autistic individuals across the lifespan, with studies tending to focus on either younger or adult populations. Using a broad measure of challenging behaviour, Matson, Mahan, Hess, Fodstad, and Neal (2010) did not identify any significant differences in the level of challenging behaviour between young children (aged 3-6 years), children (7–10 years), or young adolescents (11–14 years). Aligning with this finding, based on 3317 parents or caregivers of autistic individuals aged 2–61, a similar number of behavioural concerns were found across the school years (1.23 for age 5–11 years; 1.34 for age 12–17 years). The number of behavioural concerns remained stable across the two adulthood groups (18–24 years, 25–61 years; both 1.19 concerns), yet was significantly lower than that reported for the school-aged groups (Lai & Weiss, 2017). However, when different adulthood age grouping were used, a decline in challenging behaviour is reported in those aged 50 or above (Totsika, Felce, Kerr, & Hastings, 2010).

Cross-sectional data suggests differing trajectories for different behaviours which can be considered under the umbrella term of challenging behaviour. Fodstad, Rojahn, and Matson (2012) reported a significant increase in the level of aggressive/destructive behaviours from 12 to 25 months ($n = 147$) to 26–39 months ($n = 150$), but no change with age for self-injurious behaviour. Given this potential for topography-specific developmental trajectories of challenging behaviour in autistic individuals, some studies have focussed solely upon on specific behaviour. Duerden et al. (2012) reported no statistical change in severity of self-injurious behaviour from pre- to post-puberty. However, it may be that changes occur later in life; based on a sample of 712 autistic individuals, aged 2–62, Esbensen, Seltzer, Lam, and Bodfish (2009) reported a significant decrease in the severity of self-injurious behaviour with age. Similar to Totsika et al. (2010) finding reporting on a broader measure of challenging behaviour, the most notable decline was in those in middle adulthood and beyond (i.e., aged 40 onwards).

The potential importance of considering a broad age range is also apparent in literature focusing upon aggression. A number of studies using the Child Behaviour Checklist (CBCL) to measure aggression in samples of 147–169 1–5-year-old children did not find any association between level of aggression and age (Hartley, Sikora, & McCoy, 2008; Sikora, Hall, Hartley, Gerrard-Morris, & Cagle, 2008), which would suggest that the level of aggression remains stable over this developmental period. However, using the same measure, but with a larger sample size ($n = 414$) and a broader age range (1–21 years), Farmer et al. (2015) did report a decline in aggressive behaviour with age. Thus, proposing trajectories of challenging behaviour based on cross-sectional

data is problematic due to between-participant variability plus the inherent limitations in difficulties in separating out age-related versus cohort effects. This has led to a recognised need for more longitudinal studies documenting the presentation and level of challenging behaviour for autistic individuals (Vandewalle & Melia, 2021). However, similar to the cross-sectional studies, due to differing samples and measurement approaches there is variability in the results of the longitudinal studies, with some suggesting an improvement (reduction) of challenging behaviours with age (e.g., Gray et al., 2012) while others (e.g., Rattaz, Michelon, Munir, & Baghdadli, 2018) suggest a more stable continuation and persistence of behaviours over time. Any analysis must therefore consider the age of the sample at onset as well as time between assessments, as age at follow-up may also impact the level of challenging behaviour documented.

1.3.2. Sex

Meta-analyses note an unequal ratio of males and females in autism, with a pooled estimate of fewer than four males to each female (Loomes et al., 2017). The impact of sex on the behavioural presentation of autistic males and females may be influenced by age (Fulton, Paynter, & Trembath, 2017), camouflaging (Tubio-Fungueirino, Cruz, Sampaio, Carracedo, & Fernandez-Prieto, 2021), and the socio-construction of sex in relation to both autism and behaviour (Ferri, Abel, & Brodkin, 2018). Given this, and the finding that sex is a risk factor for challenging behaviour in individuals with intellectual disabilities (McClintock et al., 2003), it is important to consider the potential role of sex on the presentation and severity of challenging behaviour in autistic individuals.

Limited studies have explored sex-based differences in prevalence or severity of challenging behaviour in autistic individuals (Hartley & Sikora, 2009; Margari et al., 2019) and of those that have, most have not controlled for the other potential moderators discussed in this review. In general, studies have tended to find little difference between males and females pre-schoolers (Hartley & Sikora, 2009; Postorino, Fatta, De Peppo, & …& Mazzone, L., 2015), primary school aged autistic children (Nasca, Lopata, Donnelly, Rodgers, & Thomeer, 2020) or into adolescence (Brereton, Tonge, & Einfeld, 2006). In contrast, using data on 2418 participants aged 4–18 from the Simons Simplex Collection, Frazier, Georgiades, Bishop, and Hardan (2014) reported more behavioural challenges in autistic females than autistic males and a meta-analysis of the prevalence of self-injurious behaviour in autism identifies a higher prevalence of self-injury in males than females (Steenfeldt-Kristensen et al., 2020). There is therefore a clear need for a meta-analysis to combine samples and potentially overcome some of the previous limitations and explore the potential of sex as a moderator on the severity of challenging behaviour in autistic individuals.

1.3.3. Intellectual disability

There is some cross-sectional evidence to suggest that level of ID may have a differing impact on different topographies of challenging behaviour (McTiernan, Leader, Healy, & Mannion, 2011). Using questions from standardised measures, a series of studies reported that lower levels of IQ, the presence of an ID, or lower levels of adaptive behaviour are associated with higher levels of aggression in cross-sectional studies of pre-schoolers (Hartley et al., 2008), school-aged children (Dominick, Davis, Lainhart, Tager-Flusberg, & Folstein, 2007), early childhood into adolescence (Hill et al., 2014), and adulthood (Rattaz et al., 2018). However, other studies (e.g., Sikora et al., 2008), including a large study from the Simons Simplex registry which reported on 1038 individuals aged 4–17 years (Kanne & Mazurek, 2011), reported that IQ is not correlated with, nor a significant predictor of, challenging behaviour in regression models. A recent meta-analysis of the literature reporting the prevalence of self-injurious behaviour (Steenfeldt-Kristensen et al., 2020) suggested that level of ID is not significantly associated with the prevalence of self-injurious behaviour. However, a systematic review of the factors associated with self-injurious behaviour (Vandewalle & Melia, 2021) concluded that lower IQ and adaptive behaviour may be associated with greater severity or frequency of self-injurious behaviour. The heterogeneity of findings reported in the literature highlights the need to carefully differentiate between factors associated with prevalence and those associated with severity or frequency of challenging behaviour, and longitudinal data are needed to explore whether such factors are impacting challenging behaviour over time.

1.3.4. Measurement of challenging behaviour

In longitudinal studies, measurement must be accurate to effectively assess progress (Reed and Osborne, 2013), to have sufficient reliability to not be impacted by changes in informants or staff, and to be sensitive enough to detect change over time. Alongside considering variability in measures used across studies, there is also notable variability in the score reported from specific measures. Many of the frequently used measures of behaviour can be reported as a subscale score or a total score, which may be calculated by summing scores from multiple subscales such as externalising behaviour, internalising behaviours, or somatic problems. This means that while some studies report on a "broad" measure of behaviour, typically the measure's total score, others report on a "narrow" measure of challenging behaviour, such as the Externalising subscale or the Aggression Behaviour subscale of the Child Behaviour Checklist (CBCL; Achenbach & Rescorla, 2001). The potentially differing impact of exploring a concept with broad or narrow measures is highlighted by the contrasting results between meta-analyses on the social interaction and communication profiles in autistic males and females by Mahendiran et al. (2019), which used broad measures (e.g., total scores), and Wood-Downie, Wong, Kovshoff, Cortese, and Hadwin (2021), which focussed upon narrow measures.

Combining internalising with externalising behaviours (alongside other areas) into a total score may have a particular impact on studies exploring the longitudinal course of challenging behaviour given the differing trajectories of internalising and externalising behaviours reported in autistic individuals with age. Whilst internalising behaviours are reported to increase from young childhood into older childhood (Varela, DuPont, Kamps, et al., 2020) and further increase from adolescence into adulthood before declining into older adulthood (Uljarevic et al., 2020), the studies described above suggest that externalising behaviours may show the opposite pattern of decreasing into adolescence and into adulthood. When the scores reflecting an increase are combined with scores reflecting a decrease, the total score (a broad measure) may be a numerically similar score across ages or over time but may actually reflect a differing experience and presentation in different age groups. Exploring whether challenging behaviour was measured using a total score or with a subscale (or measure) specific to challenging behaviour impacts upon or moderates the course of challenging behaviour documented will therefore be an important element of understanding the differing and often conflicting results reported across studies.

1.4. The current study

With the limited number of longitudinal studies reporting on challenging behaviour in autistic individuals having variability in methodology, measures, samples, and findings, it is difficult to interpret results and, as yet, no single review has brought together these findings to identify commonalities or differences. The lack of consistent research identifying potential risk factors for challenging behaviour also limits the ability to improve supports, strategies, and management of such behaviour (Soke et al., 2017). Therefore, this review will identify studies reporting longitudinal data on challenging behaviour in autistic children, adolescents, and adults using a meta-analytic approach in order to identify any change in levels of challenging behaviour across the lifespan for autistic individuals. In addition to looking at whether the level of challenging behaviours changes over time, the review will also identify which factors have sufficient data to explore potential moderators of changes to levels of challenging behaviour over time and which

require further focus within future research.

To the authors' knowledge, the present study is the first to identify and synthesise the literature which reports on longitudinal data to describe or examine the level of challenging behaviour (or specific challenging behaviours) over time in autistic individuals. As the focus is on levels of challenging behaviour over time, this study will only focus upon studies which are observational (i.e., studies which do not implement an intervention as part of the study design). The aim of the present study was therefore to use meta-analytic techniques to aggregate data from across multiple studies identified through a systematic review to address three unanswered questions about the longitudinal course of challenging behaviour in autistic individuals:

1. How do the levels of challenging behaviour(s) change over time for autistic individuals?
2. Does the extent of change reported in challenging behaviour(s) change depending on the measure used?
3. Do individual or research design factors moderate the level of change in challenging behaviour reported, specifically, age at baseline, sex, presence of ID, time between baseline and final follow-up, drop-out rate, study quality, and measure of challenging behaviour used?

1.5. Ethics

The present meta-analysis used anonymized data and therefore was exempt from approval by the Ethics Committee of the authors' institution.

2. Method

The protocol for this systematic review was registered online with PROSPERO, an international register for systematic reviews with health-related outcomes (Registration number: CRD42020215938; uploaded as supplementary material). Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) standards were followed for all stages of this systematic review.

2.1. Eligibility criteria

Studies were included if they met the following inclusion criteria: (a) study must report data from a measure of challenging behaviour for autistic individuals (including questionnaires and observations, validated or unvalidated, may or may not be focus of study) at multiple time points, with a minimum of 6 months between time points; (b) study is purely observational (i.e., does not report on an intervention implemented as part of the study); (c) study must involve a sample of humans with a diagnosis of autism (i.e., with diagnosis confirmed by educational/clinical report or a standardised measure) and/or their caregivers; and (d) if there is a mixed-diagnosis sample, the data for the autistic participants are reported separately. Reviews, opinions, editorials, book chapters, and descriptive pieces were not included. No limit was placed on the age range of the autistic individuals. This a priori decision was made so that levels of challenging behaviour could be explored across the lifespan.

2.2. Search strategy

The electronic databases ERIC, MEDLINE, and PsycINFO were searched in April 2021 and updated in April 2022 and January 2023. The Cochrane library and PROSPERO database were searched to confirm that no other systematic reviews of the current study topic existed or were registered as being in progress. The search was restricted to reports published after 1994, aligning with publication of DSM-IV (American Psychiatric Association, 1994). Given that quality of each study would be reviewed and that the quality score is entered as moderator into the meta-analysis, gray and non-peer reviewed literature were included as an effective and transparent way to reduce the risk of publication or reporting bias and is in-line with recent recommendations (e.g. Johnson, 2021). The search terms were: autism (autis* OR asperger* OR asd OR pdd OR pervasive developmental disorder), challenging behaviour (challenging behavi* OR aggressi* OR self injur* OR self-injur* OR maladaptive behavio* OR behavio* problem OR problem behavio* OR aberrant behavio* OR self-harm OR aggressive behavio* OR disruptive behavio* OR destructive behavio* OR destruction of property OR behavio* difficulties OR behavio* of concern OR behavio* challenge) AND (longitudinal* OR year* OR trajectory OR follow up OR cohort OR prospective).

2.3. Review strategy

When combined, electronic searches resulted in 5515 records being identified (see Fig. 1). Following removal of duplicates, the titles and abstracts of 3417 reports were assessed against inclusion criteria by the primary rater. A random sample of 20% was blind-double reviewed by a second rater, resulting in 97.5% agreement (Kappa = 0.90). Although inter-rater reliability at both stages of the screening process is necessary for clarity and consistency, it is conducted by <15% of published reviews (Belur, Tompson, Thornton, & Simon, 2021). Conducting inter-rater reliability on a subsample is accepted practice when a review is large and resources are limited (Petticrew & Roberts, 2008). The reports where ratings differed were all included by the primary rater and excluded by the second rater, meaning that the reports that went through to full-text review were more likely to be overly inclusive. Where inclusion or exclusion could not be determined by the title and abstract alone, the full text of 525 reports was assessed by a primary rater. A random sample of 20% of the full-text reports were also double reviewed by a second rater who was blind to the other researcher's ratings, resulting in a 97.9% agreement (Kappa = 0.9). The two reports that differed in ratings were rated by a third rater. Following this, 56 studies met full inclusion for the review.

2.4. Data extraction and synthesis

For each paper, information extracted included study details (including geographical location of study), sample characteristics (sex, prevalence of ID, age), and study design (measure of CB used, time between data points). Data from the measure of challenging behaviour used in the study, including effect sizes for change over time, were also extracted. Data extraction was undertaken for all included papers independently by two of the authors; 100% agreement was reached. A few studies reported exact mean time between assessments, but as per Sowislo and Orth (2013) where valid indicators of time between assessments were given, this information was used to estimate the variable.

2.5. Meta-analytic procedure

Meta-analysis data were analysed using Stata v. 17 (StataCorp, 2021). A study was included in the meta-analysis if it contained sufficient data to calculate Cohen's *d* effect size, using the hierarchy of procedures described below. If a cohort was analysed by more than one study (either in its entirety or a subsample of the larger cohort), only one study (the one with the largest sample size reporting on a standardised measure of challenging behaviour or where that was equal, the one with the longest follow-up) was included in the meta-analysis to ensure independence of effect sizes. This resulted in 16 studies not being included (see Research Team and Study Design subsection of Results for overlap in study cohorts). Where a study reported sufficient data to calculate an effect size on multiple subscales from a measure of behaviour (e.g., the total and externalising subscale), the a priori decision was taken to enter data from the externalising subscale into the meta-analysis. If the study did not report on, or did not report sufficient data on the externalising

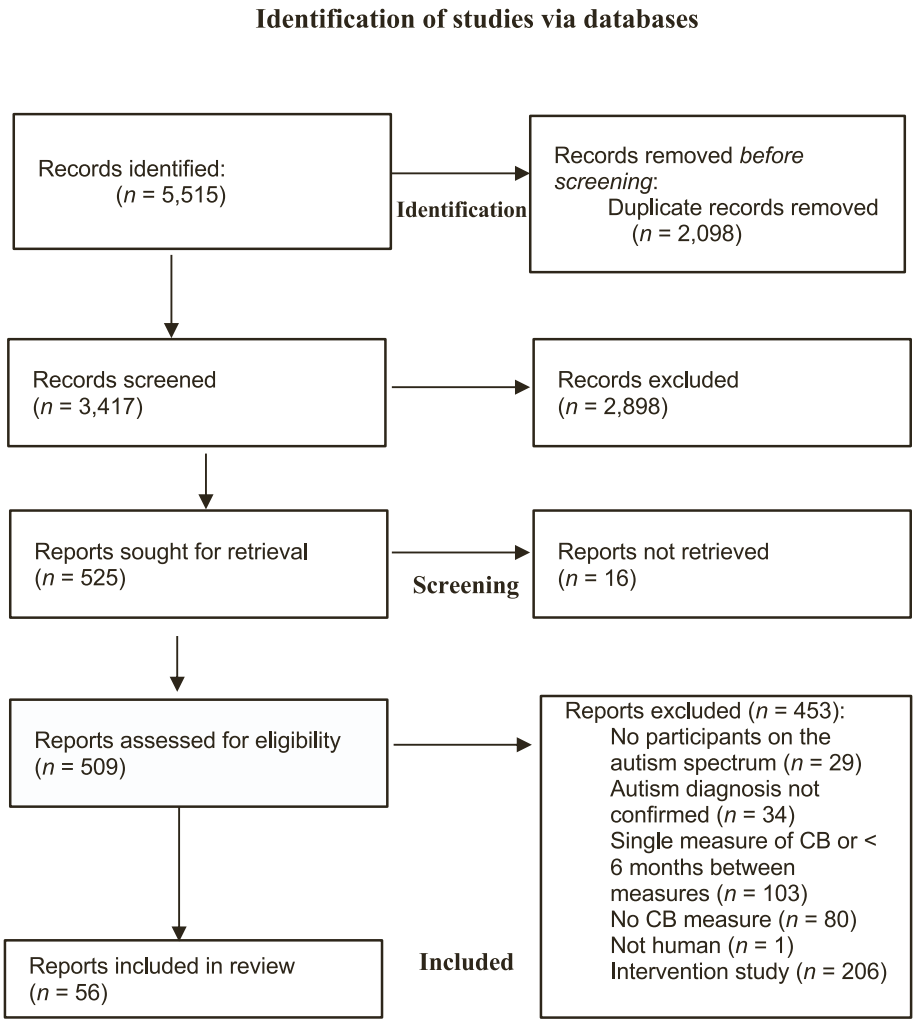

Fig. 1. PRISMA diagram.

subscale, then data from the total score was entered. Studies included in meta-analysis are marked with an asterisk in Table 1.

The effect size used (Cohen's *d*) was interpreted using Cohen (1988) guidelines, with $d = 0.2$ a small effect size, $d = 0.5$ a medium effect size, and $d = 0.8$ a large effect size. Cohen's *d* was calculated so that a positive effect size represents improvement in challenging behaviours over time. When not reported in the article, Cohen's *d* was calculated based on the difference between means (i.e., initial vs. final follow-up) divided by pooled standard deviation. If standard errors were given, these were transformed into standard deviations and entered into the above-mentioned formula. If relevant standard deviation or standard error values were missing, associated F- or t-values were used in conjunction with applicable means to calculate pooled standard deviation (Johnson & Eagly, 2014). Where the relevant F- or t-value was missing, associated *p*-value was used to calculate the t-value, which was then used to obtain the pooled standard deviation (Johnson & Eagly, 2014). Finally, Hedges' (1981) adjustment was applied to calculate unbiased estimates of Cohen's *d*. Unbiased estimates of Cohen's *d*, on average, neither over- nor under-estimate effect size, so the use of such an adjustment is particularly important when conducting a meta-analysis incorporating small sample sizes due to the associated risk of overestimating effect size (Cumming, Fidler, Kalinowski, & Lai, 2012). Estimates of standard error around unbiased Cohen's *d* were calculated using Lipsey and Wilson (2001) formula recommended for within-subjects designs.

2.6. Assessment of study quality

Included studies were evaluated for methodological quality using the National Institute for Health Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies (available from https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools). This 14-item questionnaire includes questions on multiple aspects of study design including study objective, study sample, and population, and the independent and dependent variable selection and measurement. Of the 14 items, one was considered to not be applicable to the current study (i.e., Item 8; "For exposures that can vary in amount or level, did the study examine different levels of the exposure as related to the outcome?"). As a result, the included studies were rated against 13 items. A global score of 26 could be allocated, with 2 points being allocated for each item addressed by the article, 1 point for items that might have been addressed (but were unclear), and 0 points for each item not addressed.

3. Results

3.1. Included reports

Overall, 56 reports were included in the systematic review, and these are summarised in Table 1. Fifty-four reports were peer-reviewed articles, and two were dissertations. Using the process documented in the data analysis section, 32 articles had sufficient and independent data to be entered into the meta-analysis (marked with ^ in Table 1).

Table 1
Studies Included in the Review (ˆ indicates papers included in meta-analysis).

Study	T1 *N*	T1 Age in years (mean, SD, range)	Time from T1 to final timepoint	Final timepoint age (mean, SD, range)	T1% male	T1% ID	Sampling technique, country	Consider interventions?	CB measure	Number of time points behaviour reported	Sig. change over time	*d* [d]
ˆ Anderson, Maye, and Lord (2011)	65	9.75 (2.22)	9 years	18.1 (1.86)	90	NS[b]	Agencies, USA	Medication, therapy	ABC Irritability, Agitation, & Crying	*M* = 14 (4.56)	↓ *	0.91
Baghdadli et al. (2012)	152	4.9 (1.3)	10 years	15 (1.6)	82.2	NS[b]	Epi-TED cohort, France	Yes – therapy	T1 Self-injury absent/mild/ mod/ severe rating, T3 ABC Item 52	3	NS	NA
Baker, Seltzer, and Greenberg (2011)	149	14.77 (1.89)	3 years	17.82 (2.06)	74	56	AAA study, USA	No	SIB-R MBP	2	↓*	NA
Baker, Smith, et al. (2011)	117	19.77 (6.72) 11.3–48.2	7 years	NS	77	57	AAA study, USA	No	SIB-R MBP	4	↓*	NA
Ballaban-Gil, Rapin, Tuchman, and Shinnar (1996)	99	6.8 0.8–20.3	11.3 years	18.1 (NS) 12–29.5	75	61	Clinic, USA	Medication	Parent report during interview	2	NS[c]	NA
ˆ Berkovits et al. (2017)	108	5.7 (1.1)	10 months	NS	82.4	12	Agencies, schools, support groups, USA	No	CBCL Total	2	↔	−0.16
									CBCL EXT	2	↑*	NA
ˆ Bos, Diamantopoulou, Stockmann, Begeer, and Rieffe (2018)	60	11.65 (1.3), 9–15	1.50 years	NS	100	0	From larger study that recruited from clinic, Netherlands	No	CSI, Disruptive Behaviour Scale	3	↓*	0.18
ˆ Breeman et al. (2018)	122	10.2[a] (1.15) 6–13	NS (< 1 year)	NS	100	NS[b]	Schools, Netherlands	Yes	Peer nominations of disobedience	3	↑*	−0.46
ˆ Capal et al. (2020) – Seizure group	22	6.3 (3.51)	3.93 years	10.2 (3.62)	73	53	ATN-RCBA, USA	Medication	ABC Irritability, Agitation, & Crying	2	↓*	NA
									CBCL EXT	2	↓*	0.11
ˆ Capal et al. (2020) – Non-seizure group	450	5.88 (3.13)	3.73 years	9.62 (3.17)	83	43	ATN-RCBA, USA	Medication	ABC Irritability, Agitation, & Crying	2	↔	NA
									CBCL EXT	2	↔	0.27
ˆ Carr and Lord (2013) – Caucasian	80	8.96 (1.28)	5 years	14.13 (0.21)	90	NS[b]	Clinic, USA	Therapy	ABC TBP	2	NS[c]	0.16
ˆ Carr and Lord (2013) – African American	68	10.12 (1.06)	5 years	14.21 (0.26)	87	NS[b]	Clinic, USA	Therapy	ABC TBP	2	NS[c]	0.12
ˆ Chandler et al. (2016)	93	6.83 (1.2)	1.17 years	NS	77	NS[b]	Clinic, UK	No	DBC TBP	2	↓*	0.63
										2	↓*	NA
ˆChowdhury, 2012	143	5.22	5 years	10.39 (2.67)	76	NS[b]	Clinic, USA	Medication, Therapy	NCBRF	2	↓*	0.23
ˆ Dellapiazza, Michelon, Picot, and Baghdadli (2021)	51	5.5 (2.1)	2.90 years	8.6 (2.1)	85	42	From ELENA cohort, France	No	ABC Irritability, Agitation, & Crying	2	↔	0.22
ˆ Dieleman, De Pauw, Soenens, Beyers, and Prinzie (2017)	139	10.2 (2.4) 5.1–6.2	9 years	16.0 (2.3) 11.6–22.6	83	11	Services, Belgium	Medication, other supports	CBCL EXT	3	↓*	1.03
ˆ Eisenhower et al. (2015)	166	5.75 4–7	1.50 years	NS	82	NS	Schools, services, clinics, support groups, USA	No	TRF/CTRF EXT Problems	3	↓*	0.29
ˆ Estes, Rivera, Bryan, Cali, and Dawson (2011)	30	6	3 years	9 (NS)	83	0	Subset from ongoing study, USA	No	ABC Irritability, Agitation, & Crying	2	NS[c]	NA

(*continued on next page*)

Table 1 (*continued*)

Study	T1 *N*	T1 Age in years (mean, SD, range)	Time from T1 to final timepoint	Final timepoint age (mean, SD, range)	T1% male	T1% ID	Sampling technique, country	Consider interventions?	CB measure	Number of time points behaviour reported	Sig. change over time	*d* [d]
^ Gallagher, Pilch, and Hannigan (2018)	34	9	4 years	13 (NS)	NS	0	Growing up in Ireland epidemiological study, Ireland	No	SDQ Total Difficulties	2	↔	0.20
^ Gray et al. (2012)	119	8.7 (4.3) 2.8–19.8	16 years	24.8 (4.7) 17.1–35.2	82	82	ACAD study, Australia	Noted at T1	DBC-MIS	5	↓*	0.65
Greenberg et al. (2006)	149	19.9 (7.84) 11.3–48.9	1.50 years	NS	76	57	AAA study, USA	No	SIB-R EXT	2	NS	NA
Hartley et al. (2012)	199	20.2 (67.6)	6.98 years	NS	NS	59	AAA study, USA	No	SIB-R MBP	4	NS[c]	NA
^ Hastings (2007)	24	9.8 (4.04), 3-19[a]	2 years	11.72 (3.92) 6–20	49[a]	37[a]	NS, UK	No	SDQ Total Difficulties	2	NS[c]	0.36
Helverschou et al. (2021)	132	28.6 (10.6) 16–66	2 years	NS	67	100	Norway	No	ABC Total	3	↓*	NA
^ Herring et al. (2006)	84	3.1 (0.59)	1 year	4.2 (0.6)	89	89	Clinic, Australia	No	DBC TBP	2	↔	−0.11
^ Hickey, Bolt, Rodriguez, and Hartley (2020)	166	9.1 (2.3)	2 years	NS	86	35	Mailings to schools, childcare centres, clinics, community settings, stores, USA	No	TRF Total	3	NS[c]	−0.01
Hsiao et al. (2022)	121	10.3 (3.3)	3.2 years	13.3 (3.5)	88	NS	Clinic, Taiwan		CBCL Aggression	2	NS	NA
Kring et al. (2009)	275	24.8 (9.4) 12.7–51.8	1.50 years	NS	76	68	AAA study, USA	No	SIB-R BPS	2	↓*	NA
^ Kuo, Magill-Evans, and Zwaigenbaum (2015)	29	15.1 (2.3)	1 year	NS	90	NS	Services, support organisations, Canada	No	CBCL EXT	2	NS[c]	0.26
^ Laverty, Oliver, Moss, Nelson, and Richards (2020)	67	13.4 (7.7)	10 years	23.9 (7.7)	81	NS	Autism parent group, UK	Access to paediatrician	(CBQ) SIB present/absent	3	↓*	0.37
^ Li, Bos, Stockmann, and Rieffe (2020)	57	4.6 (1.0) 1.8–5.9	2.10 years	NS	100	0	Autism clinic, Netherlands	No	ECI-4 EXT	3	NS[c]	0.18
^ Libove et al. (2017)	13	11.0 (1.3) 8–12	2.38 years	13.6 (1.7) 9–16	100	0	Recruited for longitudinal neuroimaging study, USA	Medication, Therapy	CBCL EXT	2	↓*	0.64
^ Lin, Iao, Lee, and Wu (2021)	75	2.1 (0.4)	1.50 years	3.6 (0.9)	83	NS	Clinic, Taiwan	Therapy	CBCL EXT	2	NS[c]	0.23
Martin-Borreguero et al. (2021) – Regression	19	3.6 (± 0.9)	2 years	NS	85	NS	NS, Spain	Parents agreed to make no changes to any psycho educational treatment. Children on medications excluded	SDQ Total Difficulties	2	↔	NA
Martin-Borreguero et al. (2021) – No regression	33	3.6 (± 0.9)	2 years	NS	82	NS			SDQ Total Difficulties	2	↔	NA
^ May, Cornish, and Rinehart (2014)	56	Females: 9.7 (1.86). Males 10.5 (1.93), All: 7–12	1 year	NS	50	0	Clinics, autism register, Australia	Medication	Conners Aggression	2	NS[c]	−0.13–0.02
Mazurek et al. (2019)	437	5.07 (2.14), 2–10	3.93 years	8.90 (2.23)	83	NS[b]	ATN-RCBA, USA	Medication	CBCL Aggressive Behaviour Scale	2	NS[c]	NA
Melville et al. (2008)	50	37.8 (14.1)	2 years	NS	77	100	Services, UK	Medication	C21st Health Check Problem Behaviours score	2	NS[c]	NA
Orm, Øie, Fossum, Andersen, and Skogli (2021)	26	12.0 (2.3)	2 years	14.2 (2.4)	64	0	Hospital clinics, Norway	No	CBCL Externalising	2	↓*	NA

(*continued on next page*)

Table 1 (*continued*)

Study	T1 *N*	T1 Age in years (mean, SD, range)	Time from T1 to final timepoint	Final timepoint age (mean, SD, range)	T1% male	T1% ID	Sampling technique, country	Consider interventions?	CB measure	Number of time points behaviour reported	Sig. change over time	*d* [d]
^ Osborne, McHugh, Saunders, and Reed (2008)	83	8.8	10 months	NS	97	NS	Local education authorities, UK	No	SDQ Total Difficulties	2	↓*	0.31
Peters-Scheffer, Didden, and Korzilius (2012)	102	1.93 (0.65)	2 years	NS	75	100	Previous/ ongoing study, Netherlands	No	CBCL Total	3	NS[c]	NA
Peverill et al. (2019)	363	3.23 (0.73) 12–5	NS	3.98 (0.8)	84	NS	Pathways in ASD cohort, Canada	No	CBCL Ext.	4	NS[c]	0.41
Piro-Gambetti, Rodriguez, Papp, Greenlee, and Hartley (2021)	188	7.9 (2.3)	11.7 months (3.7)	NS	86	34	Adverts, letters, clinics, USA	No	CBCL Total	4	↓	0.35
Rabbani et al. (2021)	150	NS	1 year	NS	80	NS	Clinics, Bangladesh	No	Study specific measure	52	Mixed due to COVID lockdowns	NA
^ Rattaz et al. (2018)	106	15 (1.3) 13–17	5 years	20.6 (1.5) 18–23	85	75	Epi-TED cohort, France	No	ABC Irritability, Agitation, & Crying	2	↔	NA
Reed et al. (2012) –Mainstream school group	NS	8.8 (3.5) 4–16)	10 months	NS	NS	NS[b]	Mainstream schools, UK	No	SDQ Total Difficulties	2	↔	NA
^ Reed et al. (2012) – Special education group	NS	8.2 (2.7) 4–16)	10 months	NS	NS	NS[b]	Special schools, UK	No	SDQ Total Difficulties	2	↓*	0.38
Rice et al. (2016)	98	8.14 (4.01)	18 years	24.03 (4.45)	86	45	ACAD study, Australia	No	DBC (set items)	5	Mixed	NA
Roselló Miranda, Berenguer Forner, Baixauli Fortea, & Miranda Casas (2022)	52	8.56 (1.38) 7–11	5 years	12.9 (0.9) 12–15	92.3	0	Schools, clinic, Spain	Medication	SDQ Total Difficulties	2	NS	NA
Sacrey et al. (2015)	62	0.51 (0.4)	1.5 years	2.5 (0.5)	77	NS[b]	Previous/ ongoing study, Canada	No	Non-S	6	NS	NA
Shattuck et al. (2007)	241	22.0 (9.7) 10–52	4.50 years	NS	766	69	AAA study, USA	No	SIB-R EXT, MBP	2	↓*	NA
^ Simonoff et al. (2013)	81	NS	4 years	15.5 (0.4) 14.7–16.8	93	NS[b]	SNAP cohort, UK	No	SDQ Total Difficulties	2	↓*	0.58
^ Siracusano et al. (2021) – Preschoolers		4.0	8 months	NS	76	12	Clinic, Italy	Therapy, parents support	CBCL Total, EXT	2	↔	0.13
^ Siracusano et al. (2021) – School attenders		9.0	10 months	NS	83	42	Clinic, Italy		CBCL Total, EXT	2	↔	0.00
^ Soke (2015)		6.0 (3.4)	1.04 years	NS	82	NS	ATN Registry	No	PCQ	2	NS	0.99
Taylor and Seltzer (2010)	242	16.3 (3.1) 10.1–23.5	10 years	NS	73	63	AAA study, USA	No	SIB-R Total, EXT	5	↓*before high school exit then ↔	NA
^ Totsika et al. (2013)	132	3	2 years	5.22 (NS) 5.18–5.26	82	22	Millennium cohort study, UK	No	SDQ Total Difficulties	2	NS[c]	−0.16
^ Vaillancourt et al. (2017)	367	3.2 (0.7)	3.24 years	NS	84	NS	Pathways in ASD cohort, Canada	No	CBCL EXT.	4	↓*	0.40
Woodman et al. (2015)	313	21.7 (9.45) 10–49	8.50 years	NS	75	69	AAA study, USA	No	SIB-R EXT, MBP	2	↓*	NA
^ Woodman et al. (2016)	364	21.85 (9.42) 10.15–52.08	8.22 years (3.07)	31.07 (8.85) 19.79–58.83	74	70	AAA study, USA	No	SIB-R Total, MBP	6	↓*	0.55
Zaidman-Zait et al. (2014)	184	3.4 (0.5)	3 years	NS	84	NS[b]	Pathways in ASD cohort, Canada	No	CBCL EXT	3	NS[c]	0.31

AAA = Adolescents and Adults with Autism Cohort, ATN-RCBA = Autism Treatment Network Registry Callback Assessment, BPS = Behaviour Problem scale, EXT = Externalising score, MBP = Maladaptive Behaviour Problems score, NCBRF = Nisonger Child Behaviour Rating Form, Non-S = Non-standardised measure, NS = Not stated, PCQ = Parental Concerns Questionnaire; TBP = Total Behaviour Problems.
↓ = Decline over time, ↑ = Increase over time, ↔ = no change over time, * = change is significant (No * may indicate no statistics reported by authors or non-significant change).

[a] Total sample, not specific to autism/longitudinal sample.

[b] Group IQ or adaptive behaviour provided.

[c] Mean CB score available at each time point but no statistical analysis of difference between the means presented.

[d] Unbiased estimate of Cohen's *d*. Studies included in the meta-analysis only.

3.2. Research team and study design

The research was conducted in 13 countries: the most frequent being United States ($n = 22$), United Kingdom ($n = 8$), Canada ($n = 5$) and Australia ($n = 5$). Due to the nature of the review, all studies were longitudinal, the number of data collection points ranging from two to 52. Time between baseline and final data collection ranged from 9 months to 16 years ($M = 3.55$ years).

A number of studies reported upon data collected as part of larger and/or ongoing cohorts, which is important to consider when interpreting the results of the systematic review due to the potential for overlapping study designs and participants. Nine reports (Baker, Seltzer, & Greenberg, 2011; Baker, Smith, Greenberg, Seltzer, & Taylor, 2011; Greenberg, Seltzer, Hong, & Orsmond, 2006; Hartley, Barker, Baker, Seltzer, & Greenberg, 2012; Kring, Greenberg, & Seltzer, 2009; Shattuck et al., 2007; Taylor & Seltzer, 2010; Woodman, Smith, Greenberg, & Mailick, 2015; Woodman, Smith, Greenberg, & Mailick, 2016) reported on data from the Adolescents and Adults with Autism (AAA) cohort; three (Peverill et al., 2019; Vaillancourt et al., 2017; Zaidman-Zait et al., 2014) reported on participants from the Pathways in ASD cohort; and two (Gray et al., 2012; Rice et al., 2016) reported on data from the Australian Child to Adult Development (ACAD) study. There was also overlap in participants and/or datasets used between Baghdadli et al. (2012) and Rattaz et al. (2018; Epi-TED cohort), Capal, Macklin, Lu, and Barnes (2020) and Mazurek et al. (2019; Autism Treatment Network Registry Call Back Assessment), and Berkovits, Eisenhower, and Blacher (2017) and Eisenhower, Blacher, and Bush (2015).

3.3. Quality of included studies

Ratings using the Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies (13 items; maximum 26 points) ranged between 13 and 22 ($M = 18.01$, $SD = 2.36$). None of the included studies met all the criteria on the Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies. The most commonly met criteria (met by 100% of studies) were those that aligned with the inclusion criteria of this review i.e. that articles specified and defined their study population and were aware of each participant's diagnosis of autism prior to measuring challenging behaviour. The least commonly met criteria, or ones that were least commonly reported upon, were sample size justification or power analysis (6% of studies) and whether the outcome assessors (i.e., those rating challenging behaviour) were blind to the participants' diagnosis of autism (0% of studies). It is notable, however, that fulfilling this latter criterion is not possible when exploring parent-reported challenging behaviour due to parents' knowledge of their child's diagnosis. Furthermore, loss to follow-up was <20% for only 46% of studies. Overall, in accordance with the GRADE approach, given that all studies were observational in nature, the initial certainty in an estimate of effect is low.

3.4. Participants

Sample sizes ranged from 13 to 1603. Larger samples tended to be drawn from large datasets or the ongoing cohorts noted above. Ten studies solely recruited through clinics or services, four solely through schools or education authorities. Three did not state their recruitment method.

The average percentage of males was 82.1%. Four studies did not report on the gender of the participants within the sample. The average mean cohort age was 10.16 years ($SD = 7.10$ years), with samples reporting age ranges of 6 months up to 66 years. Twenty-three studies did not report the proportion of participants with an ID. The average proportion of participants with an ID in the 34 studies that did report this was 42.8% ($SD = 34.88\%$). Eight studies did not include participants with an ID; in three studies all participants had ID. Medications and therapy (related or not related to challenging behaviour) accessed

between assessments was reported in 19 of the studies.

3.5. Measure of challenging behaviour used

Fifty-four (96%) of the studies used a standardised, published questionnaire. The majority of studies reported upon broad measures of behaviour, such as the Child Behaviour Checklist (CBCL; Achenbach & Rescorla, 2001; $n = 13$), Aberrant Behaviour Checklist (ABC; Aman & Singh, 1986; $n = 10$), the Scales of Independent Behaviour – Revised (SIB-R; Bruininks, Woodcock, Weatherman, & Hill, 1996; $n = 9$), the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997; $n = 8$), and the Developmental Behaviour Checklist (DBC; Einfeld & Tonge, 2002; $n = 4$). All of the studies reporting on the SIB-R were reporting on the same cohort of participants (i.e., AAA study).

Some studies reported total scores or total maladaptive behaviour scores, others reported subscale scores. For the CBCL, three studies reported the total score, 11 reported the Externalising subscale (three of which were reporting on the same cohort), and two the Aggressive Behaviour subscale. For the ABC, two studies reported the total score and five the Irritability, Agitation, and Crying subscale (which contains items relating to self-injury, aggression and other challenging behaviours). For the SIB-R, seven reported the Maladaptive Behaviour Problems total score (all from the same cohort) and five reported the Externalising subscale (three from the same cohort). Studies reporting on the SDQ predominantly used the total score (n = 9). Three of the four studies reporting on the DBC reported the total score, two also reported the Disruptive Behaviour subscale score, and one reported on specific items. Only three studies reported on data from informants other than parents/caregivers: two (Chandler et al., 2016; Hickey et al., 2020) used teacher data on a standardised questionnaire and one (Breeman et al., 2018) used peer nominations of disobedience.

3.6. Change in behaviour score over time

Due to the inclusion criteria for the study, all studies reported on data from the measure of behaviour across multiple time points of the study. However, the level or prevalence of behaviour was not the focus of all studies, so only 33 studies statistically explored the change in behaviour over time (see Table 1 for a summary). Of these, only two studies Breeman et al. (2018) reported a significant increase in challenging behaviour over time over time. Used peer nominations of disobedience, suggesting that over time, children rate a higher number of peers as being disobedient to the teacher. The other study to report a significant increase was in the context of mixed results; Berkovits et al. (2017) found mixedreport a significant increase in the CBCL score over time, but no significant difference in the CBCL total score. Five studies reported mixed outcomes across measures, subgroups, or across timepoints. For example, Taylor and Seltzer (2010) noted a significant decline on SIB-R Total score prior to leaving high school, then no significant change beyond high school.

Six studies reported no significant change on the measures of challenging behaviour over time: three of these reported on samples that were preschoolers at Time 1, and three reported on samples that were school aged at Time 1. Two studies used the ABC, two used the SDQ, one used the DBC, and one used the CBCL.

Twenty-one studies reported a significant decline in scores on the measures of challenging behaviour over time; however, it is important to note that six of these were from the same cohort (AAA). Fourteen of these studies reported on cohorts that were school aged at Time 1, five reported on adults from the AAA cohort, and two reported on a cohort of children who were preschoolers at Time 1. All seven of the studies that reported a significant decline on the SIB over time were using data from the AAA cohort. Four studies used the CBCL, two the DBC, two used the SDQ, two the ABC and one each used the CSI, TRF, NCBRF, and CBQ.

As there may be specific changes for specific subscales or behaviours, the findings for specific subscales and behaviours are synthesised below. Five studies reported statistical comparisons for two or more timepoints from the Irritability, agitation and crying subscale of the ABC (which contains items such as ""injures self", "does physical violence to self") and aggression (e.g. "aggressive to others". Of these, two reported a significant decline over time and three reported no change. Some statistical comparisons on behaviour-specific subscales were only reported within one study; these reported significant declines on the CSI Disruptive Behaviour subscale, DBC Disruptive behaviour subscale, and the CBQ self-injurious behaviour (present/absent) subscale.

3.7. Meta-analysis

Due to the significant heterogeneity reported in the literature surrounding challenging behaviour in autistic individuals (Mazurek et al., 2013; Summers et al., 2017), a random-effects model was used for all analyses. In addition, a standard test of heterogeneity (Cochrane's test; DerSimonian & Laird, 1986) was used to test heterogeneity within our sample. In the current sample, Cochran's $Q = 712.58$, $p < .001$. This statistically significant finding suggests that the variance across studies is due to more than just sampling error. Furthermore, I^2 was used as an indicator of the percentage of heterogeneity present between studies, with $I^2 = 25\%$ indicating a small amount of heterogeneity, $I^2 = 50\%$ indicating a medium amount of heterogeneity, and $I^2 = 75\%$ indicating a large amount of heterogeneity (Higgins & Thompson, 2002; Higgins, Thompson, Deeks, & Altman, 2003). A large amount of heterogeneity remained in the current meta-analysis, with the analysis explaining 7.89% of the total between-study variation ($I^2 = 92.11\%$). Such a finding is unsurprising given that the hypotheses of this meta-analysis relate to the use of different methods (e.g., different outcomes, different ages, etc.).

3.7.1. Levels of challenging behaviour over time

To statistically examine whether there is any change in the level of challenging behaviour reported with time, a meta-analysis was undertaken on 37 samples from 34 of the 56 articles identified in the systematic review. Cohen's d was calculated from 37 independent samples from 34 papers. Sample sizes of the studies within the meta-analysis ranged from 13 to 1603 participants ($M = 140.43$, $SD = 262.25$), with a total of 5196 participants represented in the meta-analysis overall. The calculated effect sizes ranged from −0.46 to 1.03 ($M = 0.28$, $SD = 0.32$), with 30 of the 37 effect sizes being positive (i.e., indicating improvement in challenging behaviour over time). Using a random-effects model, the weighted mean effect size was 0.28 (95% CI = 0.17, 0.39), and was found to be significantly greater than zero ($z = 4.87$, $p < .001$; see Fig. 2). Therefore, across studies, challenging behaviour improves to a significant, albeit small, extent over time.

A sensitivity analysis was conducted to ascertain any change in the estimated overall mean effect size in the context of a small (as opposed to large) level of heterogeneity. Assuming $I^2 = 25\%$, the estimated mean effect size was 0.34 (95% CI = 0.30, 0.38) and was found to be significantly greater than zero ($z = 17.19$, $p < .001$). Therefore, even with small levels of heterogeneity, challenging behaviour was still observed to improve to a significant and small extent over time. Of note is that when the model was run excluding the two dissertations ("gray literature"), the results remained unchanged: the mean effect size of the model with only peer-reviewed articles was 0.26 (95% CI = 0.15, 0.37) and remained significantly greater than zero ($p < .001$).

An influence analysis was conducted to determine the impact of individual studies on overall pooled effect size, allowing for identification of potential outliers. As the point estimate for each included study fell within the confidence interval of the combined analysis (i.e., between 0.17 and 0.39), no individual studies were removed. The point estimates for individual studies ranged from 0.26 to 0.30.

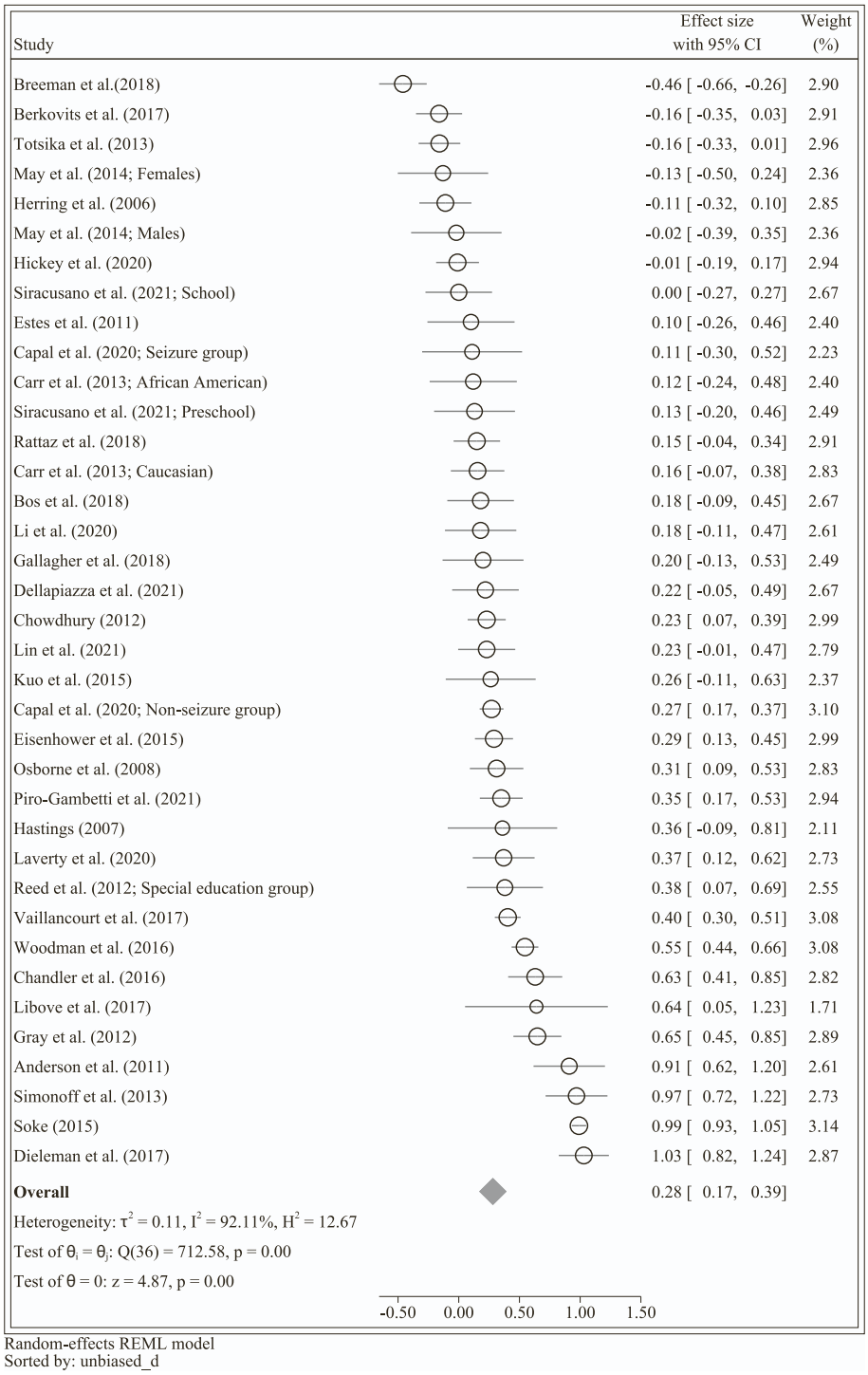

Fig. 2. Forest plot of effect sizes.

3.7.2. Impact of broad and narrow measures of behaviour on heterogeneity of results

Due to the presence of substantial heterogeneity in the current sample, a subgroup analysis was undertaken to investigate whether the rate of improvement in challenging behaviour varies based upon the outcome measure used. Of the 37 samples included, 20 compared changes in challenging behaviour over time as measured by a total challenging behaviour score (i.e., a broad measure), while 17 compared such change using an externalising behaviour score (i.e., a narrow measure) only (see upper half of Table 2). The weighted mean effect size was significantly greater than zero for samples that measured change in challenging behaviour over time using (1) a total score from a broad behaviour measure and (2) observed externalising behaviours specifically, suggesting observed challenging behaviour improves to a significant, albeit small extent over time in autistic individuals when measured by either a broad or narrow measure. While 92.1% of the total between-study variation in the overall meta-analysis was systematic, stratifying by outcome measure indicated 91.5% systematic variance within broad measures and 91.8% systematic variance within narrow measures, indicating similar levels of heterogeneity when externalising behaviours are measured specifically ($Q_b = 0.01$, $p = .904$).

Table 2
Cochran's Q and subtest test results for subgroup analyses.

	Heterogeneity				Pooled estimates	
	Q	*p*-value	I^2	*M*	*CI* (95%)	
Outcome Measure						
Total score	234.65	<0.001	91.5%	0.29	0.12	0.45
Externalising	406.31	<0.001	91.8%	0.27	0.12	0.43
Recruitment method						
Clinical	88.79	<0.001	84.7%	0.28	0.13	0.43
Non-clinical	47.27	<0.001	88.4%	0.10	−0.19	0.38
Mixed	398.07	<0.001	95.3%	0.35	0.16	0.55

Note. M refers to the pooled value of Cohen's *d*.

3.7.3. Impact of recruitment method on heterogeneity of results

A second subgroup analysis was undertaken to investigate whether the rate of improvement in challenging behaviour varies based upon the recruitment used (i.e. recruited from the community, clinic, or a mix). Of the 37 samples included, 15 recruited their participants through clinics or services, six through community or educational settings only, 15 through both clinics and schools and one did not state their recruitment method (see lower half of Table 2). The weighted mean effect size was significantly greater than zero for all studies that used all three recruitment methods, suggesting observed challenging behaviour improves to a significant, albeit small extent over time in autistic individuals regardless of whether the sample was recruited through a clinic, through non-clinical/community settings or a mix of both. While 92.4% of the total between-study variation in the overall meta-analysis was systematic, stratifying by recruitment method indicated 84.7% systematic variance within studies that recruited clinical samples, 88.48% systematic variance for the non-clinical samples and 95.4% systematic variance for samples recruited using both clinical and non-clinical settings. This indicates similar levels of heterogeneity when different recruitment methods are used ($Q_b = 2.17$, $p = .338$).

3.8. Moderator analyses

Meta-regression analyses were conducted to investigate whether age of participants at baseline, time between baseline and the final measure, the proportion of males and those with an ID within the sample, the outcome measure used (broad vs. specific), study quality, and dropout rate moderate the degree to which challenging behaviour changes over time in autistic individuals (see Table 3). The unbiased effect size of each study represents the outcome variable (such that positive effect sizes represent a reduction in challenging behaviour over time), and random-effect weights were used within the meta-regression model. Age at recruitment, age at final time point, time between baseline and the final measure, proportion of males within a sample, the proportion of participants with an ID within a sample, and study quality were continuous variables, and as such, dummy coding was not required. The type of outcome measure used had two levels (narrow vs. broad) and so was dummy coded such that the reference category was the category of studies with the smallest effect size (i.e., externalising behaviour). Similarly, dropout rate had two levels (low, <20% vs. high, >20%[2]) and so was dummy coded such that the reference category was the category of studies with the smallest effect size (i.e., low dropout). Recruitment method had three levels (clinical, non-clinical and mixed) so was dummy coded such that the reference category was the category of studies with the smallest effect size (i.e., non-clinical).

Table 3
Meta-regression test results for predicted moderator variables.

Comparison	β	SE_β	*p*-value
Time between baseline and final measure	0.04	0.02	0.017
Age at baseline	0.02	0.01	0.139
Age at final timepoint	0.24	0.01	0.010
Proportion of participants with an ID	0.00	0.00	0.860
Proportion of male participants	−0.00	0.00	0.254
Broad (i.e., Total score) or narrow (e.g., Externalising score) measure	0.02	0.12	0.896
Quality	−0.04	0.03	0.170
Dropout rate	0.22	0.11	0.055
Recruitment method: clinical	0.18	0.17	0.28
Recruitment method: mixed	0.26	0.17	0.12

Both age at final timepoint and time between the initial measure of challenging behaviour and final follow-up significantly predicted change in challenging behaviour over time. Specifically, the greater the amount of time between baseline and the final follow-up, and the older the participants were at the final data collection point, the larger the effect size and therefore the greater the reduction in challenging behaviour. However, age of participants at initial assessment (e.g., Time 1 of the longitudinal study) did not significantly predict change in challenging behaviour over time, suggesting it is the time between data collection points, rather than the age of the participants per se, that is associated with the change in level of challenging behaviour. The proportion of male participants, the proportion of participants with an ID, the use of a narrow or a broad measure of challenging behaviour (i.e., a total behaviour score or an externalising behaviour score) and recruitment method (clinical, non-clinical or mixed) did not significantly predict change in challenging behaviour over time. Finally, neither quality rating of the study nor dropout rate predicted the change in challenging behaviour over time. A large amount of heterogeneity remained for all moderator variables, with the time between the initial measure of challenging behaviour and the final follow-up explaining 9.23% (I^2 = 90.77%) of the between-study variation, age of participants explaining 9.36% (I^2 = 90.64%), age at following up 14.28% (I^2 = 85.72%), proportion of male participants 8.32% (I^2 = 91.68%), proportion of participants with an ID 12.35% (I^2 = 87.65%), use of a broad or narrow outcome measure 8.08% (I^2 = 91.92%), recruitment method 8.16% (91.84%), study quality 8.60% (I^2 = 91.40%), and dropout rate 8.69% (I^2 = 91.31%).

3.9. Publication bias and small study effects

The data were assessed for publication bias (i.e., over-representation of positive or negative results) and small study effects (i.e., the tendency for smaller studies to have larger effect sizes than larger studies) using multiple methodologies as recommended by Rücker et al. (2011). First, a contour-enhanced funnel plot of the effect size of each sample against its standard error was inspected (see Fig. 3). Visual inspection of the contour-enhanced funnel plot revealed a pattern of effects that were asymmetrically distributed, with the suggestion of missing studies. This finding suggests that the overall effect size may have been over- or under-estimated due to possible publication bias.

As a result of the visual asymmetry in the contour-enhanced funnel plot, the trim-and-fill method (Duval & Tweedie, 2000) was used in order to estimate the number of missing studies, and to impute the effect sizes of any identified missing studies. This method allows adjustment of the mean effect size as though there was no publication bias. The trim-and-fill method identified eight studies that needed to be filled in the current dataset, resulting in a corrected mean effect size of 0.40 (95% CI 0.28, 0.52) as compared to the original mean effect size of 0.28 (95% CI 0.17, 0.39; See Fig. 4).

In addition to the above, Egger's test was conducted on the overall and stratified analyses to further assess the presence of publication bias. Egger's test was non-significant ($p > .05$) for all analyses, providing evidence of the absence of publication bias or related small study effects

[2] Low dropout rate was defined as <20% dropout, while high dropout rate was defined as >20% dropout consistent with the National Institute for Health Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies.

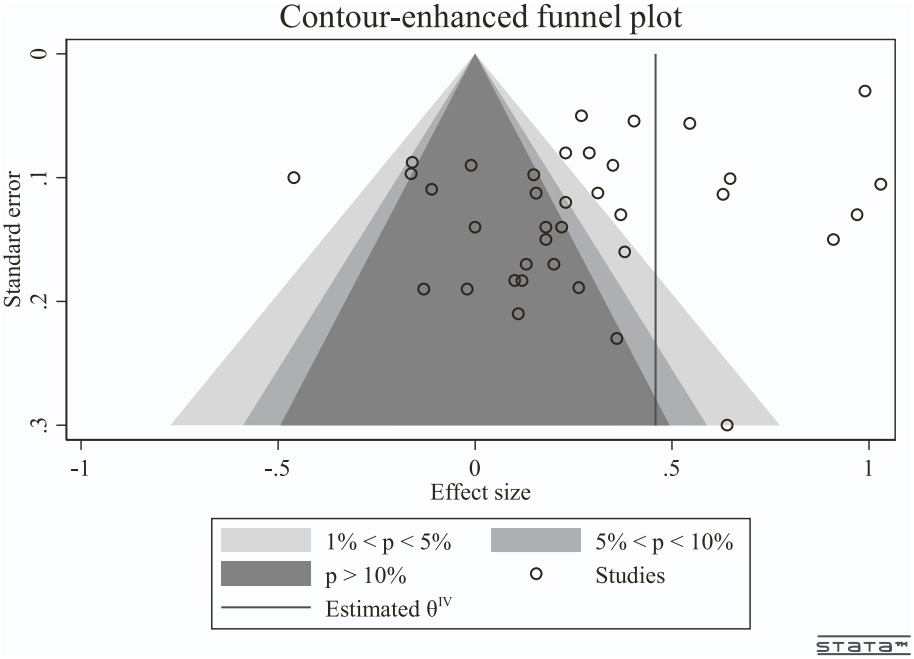

Fig. 3. Contour-enhanced funnel plot of the samples included in this meta-analysis.

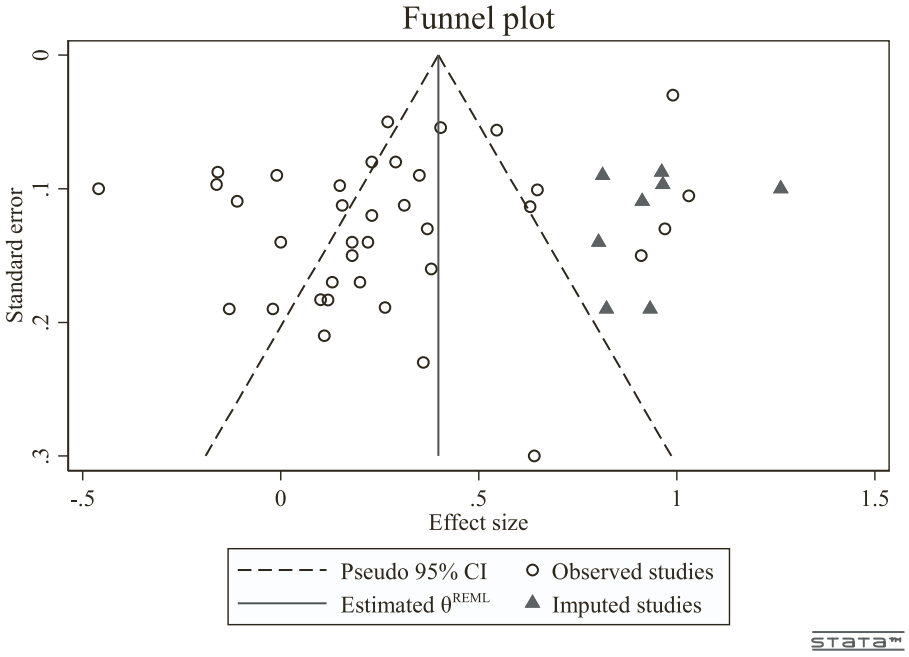

Fig. 4. Funnel plot of the samples included in this meta-analysis and imputed studies.

in this meta-analysis. A full list of Egger's test results is available here. Overall, the evidence for publication bias or related small study effects in the current results is inconsistent but may represent a limitation of our findings given the possibility that our effect sizes are underestimated in some cases.

4. Discussion

Despite challenging behaviours being associated with a broad range of negative outcomes for autistic individuals (e.g., Chiang & Wineman, 2014; Oliver et al., 2003) and those who care for them or support them (e.g., Estes et al., 2009; Yacoub et al., 2018), there have been only 56 observational articles reporting longitudinal data on challenging behaviour in autistic individuals, reporting on 43 different samples. For comparison, over two hundred studies were excluded during the systematic review process due to being longitudinal studies of challenging behaviour reporting on medical or psychological interventions. It is likely that the high prevalence and long-lasting negative impact of challenging behaviours has been a driving factor for this stronger focus on interventions research. While there is a clear need for intervention research, especially with longitudinal follow-up, such research may be better interpreted in relation to a large, established literature base documenting the course of challenging behaviours outside of intervention studies. The aim of this review was therefore to synthesise

longitudinal observational studies reporting data on challenging behaviour in autistic individuals and to use meta-analytic techniques to explore whether level of challenging behaviour over time is moderated by participant characteristics (level of ID, sex, age at initial assessment), study quality or study design factors (time between data collection points, measure used, subscale reported, dropout rate).

There were three main findings that emerged from this review. First, only 56 studies met the inclusion criteria and reported on longitudinal data on challenging behaviour in autistic individuals. These predominantly focussed on reporting on challenging behaviour across childhood, with the included studies reporting on only two independent samples focussed solely upon adulthood (i.e., study reports participants mean age of 18 or above at the study onset) and only eight independent samples which reported on data from childhood or adolescence into adulthood (i.e., study cohort had a mean age of 18 or above at the final data collection point). Second, almost all included studies reported either no significant change or (more commonly) a significant decline in scores on the behaviour measure over time. In addition, the meta-analysis on data from 5196 participants across 37 independent samples revealed a clear but small change in magnitude effect size which demonstrates that the frequency or severity of challenging behaviour reduces over time. It is important to note, however, that a reduction is not the same as remission. The longer the time period between baseline and the final follow-up, the greater the decline in challenging behaviour score. Finally, almost all of the observational cohort studies in this review measured challenging behaviour using a parent-based measure which has been developed to measure behaviour in neurotypical populations.

4.1. Age

The finding that most of the studies reported upon children is not unique to challenging behaviour. Despite most people being an adult for longer than they are a child, the volume of research on autism in adulthood is very limited compared to that on childhood (Howlin & Magiati, 2017; Lorenc et al., 2018). Within the present review, of the studies that provided an age range, the oldest participant was 66, and the oldest mean age of the sample was 37.8 years. This suggests that very little is known about the trajectory of challenging behaviours from adulthood into older adulthood. With increasing numbers of individuals receiving a diagnosis of autism, there will be an increasing ageing population of autistic individuals. There is therefore a clear need for more high-quality work focussing upon adulthood and older adulthood in autistic individuals (Happe & Frith, 2020), including research which explores age-related behavioural change in autistic older people, so as to anticipate any changing support and service requirements with age (Totsika et al., 2010).

4.2. Time

Almost all the studies in this review reported no significant change or, most commonly, a significant decline in the score for behavioural challenges over time. The results of the meta-analysis also indicated a small but significant reduction in the frequency or severity of challenging behaviour over time, with the moderator analysis highlighting that the longer the time between baseline and follow-up and the older the age at follow up (but not baseline), the larger the reduction in challenging behaviour score. While this is a very promising finding for parents and for service providers, it is important that this finding is further explored. First, reduction is not the same as remission: behaviours may still be present at a clinically significant level. As studies tended to report only mean scores, it is not known whether the reduction in scores on the measures used to describe challenging behaviour reflects a reduction in the number of behaviours, the severity of behaviours, or the frequency of behaviours. Even if the number or frequency of behaviour(s) decrease with age, the results of such behaviour from individuals who may be growing taller, heavier, or stronger with age may actually be more distressing or dangerous than the same behaviour in younger children (Levy & Perry, 2011; Shea & Mesibov, 2005). It is also possible that the parents or caregivers completing the questionnaires make informal changes or adjustments which impact upon their child's behaviour (e.g., Dreyfus & Dowse, 2018) or that they have accessed formal supports or therapy. Less than one third of the included studies reported on therapies or medications accessed by the participants, making it difficult to explore the potential role of therapy in the change in scores. Given that many of the measures ask for a relative rating (e.g., *never, sometimes, always*), if the behaviour remains present across the years, the subjective perspective of the frequency or severity of the behaviour may change, which may then result in a reduced rating. As none of the studies included objective or observational measures of behaviour, it is not possible to explore the extent of change over time on subjective measures compared to objective measures. Further, given that most studies reported on measures not designed for autism and that individuals on the spectrum typically show higher scores than other groups, some reversion to the mean is also a possible explanation.

4.3. Study methodology: recruitment methods and outcome measure

The results of the meta-analysis and meta-regression indicate that there was no significant difference in change in challenging behaviour over time based upon method of recruitment. This suggests the results presented are relevant to both clinical and community samples. In contrast, the meta-analysis did highlight the importance of careful measurement selection. The limited number of studies using each measure limited the ability to explore outcomes for specific measures or specific forms of challenging behaviour within the meta-analysis. This is not uncommon in autism research; Provenzani et al. (2020) systematic review identified 406 trials published between 1980 and 2016 which used 327 outcome measures, of which 69% had only been used in the literature once. This heterogeneous and often non-specific use of outcome measures may be significantly hampering comparisons between studies and, thus, progress of intervention research in this area. Despite large number of measures used, they were predominantly broad measures of behaviour resulting in a limited number of studies which reported on longitudinal data on behaviour-specific measures, such as those unique to self-injury or specific behaviours which would be classified as aggression. Combined with the mixed findings for some subscales (e.g. ABC Irritability, agitation and crying), this highlights an area of significant need for future research, as the aetiology and effective supports may differ based on the topography of the challenging behaviour (Oliver et al., 2022).

With the exception of the DBC and ABC, the majority measures used were designed for the neurotypical population. A systematic review of measures of challenging behaviour in young autistic children identified this lack of autism-specific work on measurement of challenging behaviour as a significant weakness in the field (Hanratty et al., 2015). Questions from within such measures, such as “can't get his/her mind off certain thoughts; obsessions” pose a risk of diagnostic overshadowing, as they may be measuring the presence of autism characteristics rather than challenging behaviour per se. This, in combination with a lack of autism-specific evidence on reliability and validity, means that it is difficult to know how much variability there may be in reporting over time, making it difficult to make judgements about the significant of change over time (Hanratty et al., 2015). Therefore, before these measures are used in intervention studies, further systematic work is required to explore and document the measurement properties of these tools for the child, adolescent, and adult populations, at home and school, with a particular focus on establishing their responsiveness to change.

Challenging behaviours can pose a particular challenge within school settings, potentially having a negative impact on the teachers, support staff, on other children, and on the learning environment

(Nicholls, Hastings, & Grindle, 2020). Higher levels of challenging behaviour are associated with a greater risk of a residential placement or of attending special school (Pilling, McGill, & Cooper, 2007). Given the challenge that schools face with such behaviours, it would be beneficial to be able to draw upon a large body of literature exploring the trajectory and moderators of challenging behaviour for autistic students at school. In contrast, this review identified only two studies which reported on longitudinal data from school. Therefore, although challenging behaviour decreases over time based on parent report (i.e., in the home and community settings), it is not known if a corresponding decrease is seen at school. The use of multiple informants (e.g., parents and teachers) is important to understand the role of the environment, something that is particularly important for autistic individuals (Ozonoff, Goodlin-Jones, & Solomon, 2005).

4.4. Sex, intellectual disability, and sociocultural factors

The existing literature has reported mixed findings in relation to the impact of sex and the presence of ID on the level or severity of challenging behaviour in autistic individuals. The meta-analysis in this study found that neither the proportion of male participants nor the proportion of participants with an ID significantly predicted change in challenging behaviour in autistic individuals over time. The included studies did not consistently report on sex and the presence of ID, however. Three studies in this review did not report on the proportion of males and females and 22 did not report on the proportion of individuals with a co-occurring ID. To better ascertain whether sex and the presence of ID moderate challenging behaviour in autistic individuals over time, it will be crucial that future longitudinal studies consistently report on the sex and cognitive ability of their participants. Furthermore, sociocultural factors, such as socioeconomic status or ethnicity, were seldom reported, prohibiting the inclusion of such factors as moderators. Numerous reviews of both descriptive and interventional studies (e.g., Pierce et al., 2014; Safer-Lichtenstein, Hamilton, & McIntyre, 2019;) have noted that less than one third of their included articles report on racial or ethnic backgrounds or indicators of socioeconomic status of their participants. Having a standard way of reporting factors associated with socioeconomic status (such as caregiver education level) would improve the quality of reporting as well as allowing for easier inclusion of such data into future meta-analyses.

4.5. Limitations

This study must be considered in relation to its limitations. While every effort was undertaken to include a broad range of search terms, it is possible that studies which do not use such terms, but describe a specific behaviour (e.g., smearing) may not have been identified. A priori decision to set the minimum time between assessments at 6 months will have led to the exclusion of additional studies with shorter follow-ups. The requirement for the study to report on the data at multiple timepoints led to the exclusion of some studies which had collected longitudinal data of challenging behaviour but did not report it at more than one time point (e.g. Richards, Moss, Nelson, & Oliver, 2016). Eleven impoverished articles were also identified (i.e., articles that met inclusion criteria, but were missing information essential to calculate an associated effect size for the meta-analysis). Although corresponding authors were emailed requesting the relevant data, no authors responded.

Each moderator explored explained up to 14.28% of the heterogeneity of the results, meaning that future research is needed to investigate other potential moderators of the level of challenging behaviour over time. It could also be suggested that more heterogeneity could be explained by investigating interactions between different combinations of the moderator variables; however, due to the limited number of studies, there was not sufficient power to be able to undertake this. As the date limit was set from 1994 covering both DSM-IV and DSM-V criteria for ASD (as well as iterations of the ICD-10) for the included articles, it is recognised that the participants may have received their autism diagnosis with differing diagnostic criteria, so earlier studies likely included those with PDD-NOS that may not have received an autism spectrum disorder diagnosis under DSM-5, for example. Therefore, results may not generalise from the older to newer studies or may not immediately generalise to studies collecting data at present.

5. Conclusion

This systematic review and meta-analysis reveals the limited extent of observational research reporting on challenging behaviour over time. The results highlight that based on parent (and in two cases, teacher) report, challenging behaviour decreases over time, with significantly larger decreases in studies with longer duration between baseline and final assessment. The predominant focus has been on children and adolescents, with only three independent samples reporting solely on the level of challenging behaviour over time within autistic adults. Few studies used measures of behaviour designed for individuals with neurodevelopmental disorders or for autistic individuals. Additional work is clearly needed to document the longitudinal trajectories of specific topographies of challenging behaviour over time, using sensitive and reliable measures for this cohort, so as to be able to provide a context to interpret intervention work in this area to ultimately support better outcomes for autistic individuals through evaluation of what constitutes a meaningful change over time in response to medical or psychological intervention.

Role of funding sources

No specific funding was received for this study. Dr. Nicole Dargue was on a post-doctoral fellowship during the data collection and analysis for this study. This post-doctoral fellowship was funded by Griffith University. Griffith University had no role in the study design, collection, analysis or interpretation of the data, writing the manuscript, or the decision to submit the paper for publication.

Contributors

Dawn Adams designed the study, wrote the protocol, extracted data, synthesised the results, and drafted the manuscript.

Nicole Dargue extracted data, competed the meta-analysis of the results, and contributed to the manuscript drafting.

Jessica Paynter designed the study and contributed to the drafting of the manuscript.

All authors contributed to and have approved the final manuscript.

Declaration of Competing Interest

All authors declare that they have no conflicts of interest.

Data availability

Data will be made available on request.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cpr.2023.102320.

References

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms and profiles*. University of Vermont, Research Center for Children, Youth and Families.

Adams, D., Clarke, S., Griffith, G., Howlin, P., Moss, J., Petty, J., Tunnicliffe, P., & Oliver, C. (2018). Mental health and well-being in mothers of children with rare genetic syndromes showing chronic challenging behavior: A cross-sectional and

longitudinal study. *American Journal on Intellectual and Developmental Disabilities, 123* (3), 241–253. https://doi.org/10.1352/1944-7558-123.3.241

Adams, D., Handley, L., Simkiss, D., Walls, E., Jones, A., Knapp, M., Romeo, R., & Oliver, C. (2016). Service use and access in young children with an intellectual disability or global developmental delay: Associations with challenging behaviour. *Journal of Intellectual & Developmental Disability, 43*(2), 232–241. https://doi.org/10.3109/13668250.2016.1238448

Aman, M., & Singh, N. N. (1986). *The aberrant behaviour checklist manual*. Slosson Educational Pubs.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.).

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Anderson, D. K., Maye, M. P., & Lord, C. (2011). Changes in maladaptive behaviors from mid-childhood to young adulthood in autism spectrum disorder. *American Journal on Intellectual and Developmental Disabilities, 116*(5), 381–397. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3185304/.

Baghdadli, A., Assouline, B., Sonie, S., Pernon, E., Darrou, C., Michelon, C., Picot, M.-C., Aussilloux, C., & Pry, R. (2012). Developmental trajectories of adaptive behaviors from early childhood to adolescence in a cohort of 152 children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 42*(7), 1314–1325. https://doi.org/10.1007/s10803-011-1357-z

Baker, J. K., Seltzer, M. M., & Greenberg, J. S. (2011). Longitudinal effects of adaptability on behavior problems and maternal depression in families of adolescents with autism. *Journal of Family Psychology, 25*(4), 601–609. https://doi.org/10.1037/a0024409

Baker, J. K., Smith, L. E., Greenberg, J. S., Seltzer, M. M., & Taylor, J. L. (2011). Change in maternal criticism and behavior problems in adolescents and adults with autism across a 7-year period. *Journal of Abnormal Psychology, 120*(2), 465–475. https://doi.org/10.1037/a0021900

Ballaban-Gil, K., Rapin, I., Tuchman, R., & Shinnar, S. (1996). Longitudinal examination of the behavioral, language, and social changes in a population of adolescents and young adults with autistic disorder. *Pediatric Neurology, 15*(3), 217–223. https://doi.org/10.1016/S0887-8994(96)00219-6

Belur, J., Tompson, L., Thornton, A., & Simon, M. (2021). Interrater reliability in systematic review methodology: Exploring variation in coder decision-making. *Sociological Methods & Research, 50*(2), 837–865. https://doi.org/10.1177/0049124118799372

Berkovits, L., Eisenhower, A., & Blacher, J. (2017). Emotion regulation in young children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 47*(1), 68–79. https://doi.org/10.1007/s10803-016-2922-2

Bos, M. G. N., Diamantopoulou, S., Stockmann, L., Begeer, S., & Rieffe, C. (2018). Emotion control predicts internalizing and externalizing behavior problems in boys with and without an autism spectrum disorder. *Journal of Autism and Developmental Disorders, 48*(8), 2727–2739. https://doi.org/10.1007/s10803-018-3519-8

Breeman, L. D., van Lier, P. A. C., Wubbels, T., Verhulst, F. C., van der Ende, J., Maras, A., … Tick, N. T. (2018). Developmental links between teacher-child closeness and disobedience for boys placed in special education. *Exceptionality, 26* (4), 230–244. https://doi.org/10.1080/09362835.2017.1283624

Brereton, A. V., Tonge, B. J., & Einfeld, S. L. (2006). Psychopathology in children and adolescents with autism compared to young people with intellectual disability. *Journal of Autism and Developmental Disorders, 36*(7), 863–870. https://doi.org/10.1007/s10803-006-0125-y

Bruininks, R. H., Woodcock, R. W., Weatherman, R. F., & Hill, B. K. (1996). *Scales of independent behavior-revised*. Itasca, IL: SIB-R. Riverside Publishing.

Capal, J. K., Macklin, E. A., Lu, F., & Barnes, G. (2020). Factors associated with seizure onset in children with autism spectrum disorder. *Pediatrics, 145*(Suppl. 1), S117–S125. https://doi.org/10.1542/peds.2019-1895O

Carr, T., & Lord, C. (2013). Longitudinal study of perceived negative impact in African American and Caucasian mothers of children with autism spectrum disorder. *Autism, 17*(4), 405–417. https://doi.org/10.1177/1362361311435155

Chandler, S., Howlin, P., Simonoff, E., O'Sullivan, T., Tseng, E., Kennedy, J., Charman, T., & Baird, G. (2016). Emotional and behavioural problems in young children with autism spectrum disorder. *Developmental Medicine and Child Neurology, 58*(2), 202–208. https://doi.org/10.1111/dmcn.12830

Chiang, H.-M., & Wineman, I. (2014). Factors associated with quality of life in individuals with autism spectrum disorders: A review of literature. *Research in Autism Spectrum Disorders, 8*(8), 974–986. https://doi.org/10.1016/j.rasd.2014.05.003

Chiarotti, F., & Venerosi, A. (2020). Epidemiology of autism spectrum disorders: A review of worldwide prevalence estimates since 2014. *Brain Sciences, 10*(5). https://doi.org/10.3390/brainsci10050274

Chowdhury, M. (2012). *Follow-up of maladaptive behaviors in youth with autism spectrum disorders: Changes and predictors over two to eight years. ProQuest Dissertations and Theses Global*.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). L. Erlbaum Associates https://go.exlibris.link/t28fZZHJ.

Cumming, G., Fidler, F., Kalinowski, P., & Lai, J. (2012). The statistical recommendations of the American Psychological Association publication manual: Effect sizes, confidence intervals, and meta-analysis. *Australian Journal of Psychology, 64*(3), 138–146. https://doi.org/10.1111/j.1742-9536.2011.00037.x

Davies, L., & Oliver, C. (2013). The age related prevalence of aggression and self-injury in persons with an intellectual disability: A review. *Research in Developmental Disabilities, 34*(2), 764–775. https://doi.org/10.1016/j.ridd.2012.10.004

Dellapiazza, F., Michelon, C., Picot, M.-C., & Baghdadli, A. (2021). A longitudinal exploratory study of changes in sensory processing in children with ASD from the ELENA cohort. *European Child & Adolescent Psychiatry*.. https://doi.org/10.1007/s00787-021-01746-1

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials, 7*(3), 177–188. https://doi.org/10.1016/j.cct.2015.09.002

Dieleman, L. M., De Pauw, S. S., Soenens, B., Beyers, W., & Prinzie, P. (2017). Examining bidirectional relationships between parenting and child maladjustment in youth with autism spectrum disorder: A 9-year longitudinal study. *Development and Psychopathology, 29*(4), 1199–1213. https://doi.org/10.1017/S0954579416001243

Dominick, K. C., Davis, N. O., Lainhart, J., Tager-Flusberg, H., & Folstein, S. (2007). Atypical behaviors in children with autism and children with a history of language impairment. *Research in Developmental Disabilities, 28*(2), 145–162. https://doi.org/10.1016/j.ridd.2006.02.003

Dreyfus, S. J., & Dowse, L. (2018). Experiences of parents who support a family member with intellectual disability and challenging behaviour: "This is what I deal with every single day". *Journal of Intellectual & Developmental Disability, 45*(1), 12–22. https://doi.org/10.3109/13668250.2018.1510117

Duerden, E. G., Oatley, H. K., Mak-Fan, K. M., McGrath, P. A., Taylor, M. J., Szatmari, P., & Roberts, S. W. (2012). Risk factors associated with self-injurious behaviors in children and adolescents with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 42*(11), 2460–2470. https://doi.org/10.1007/s10803-012-1497-9

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*(2), 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x

Einfeld, S., & Tonge, B. (2002). *Manual for the developmental behaviour checklist* (2nd ed.). Monash University Centre for Developmental Psychiatry and Psychology.

Eisenhower, A. S., Blacher, J., & Bush, H. H. (2015). Longitudinal associations between externalizing problems and student-teacher relationship quality for young children with ASD. *Research in Autism Spectrum Disorders, 9*, 163–173. https://doi.org/10.1016/j.rasd.2014.09.007

Emerson, E. (1995). *Challenging behaviour: Analysis and intervention in people with learning disabilities*. Cambridge University Press.

Esbensen, A. J., Seltzer, M. M., Lam, K. S., & Bodfish, J. W. (2009). Age-related differences in restricted repetitive behaviors in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39*(1), 57–66. https://doi.org/10.1007/s10803-008-0599-x

Estes, A., Munson, J., Dawson, G., Koehler, E., Zhou, X. H., & Abbott, R. (2009). Parenting stress and psychological functioning among mothers of preschool children with autism and developmental delay. *Autism, 13*(4), 375–387. https://doi.org/10.1177/1362361309105658

Estes, A., Rivera, V., Bryan, M., Cali, P., & Dawson, G. (2011). Discrepancies between academic achievement and intellectual ability in higher-functioning school-aged children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 41*(8), 1044–1052. https://doi.org/10.1007/s10803-010-1127-3

Esteves, J., Perry, A., Spiegel, R., & Weiss, J. A. (2021). Occurrence and predictors of challenging behavior in youth with intellectual disability with or without autism. *Journal of Mental Health Research in Intellectual Disabilities, 14*(2), 189–201. https://doi.org/10.1080/19315864.2021.1874577

Farmer, C., Butter, E., Mazurek, M. O., Cowan, C., Lainhart, J., Cook, E. H., … Aman, M. (2015). Aggression in children with autism spectrum disorders and a clinic-referred comparison group. *Autism, 19*(3), 281–291. https://doi.org/10.1177/1362361313518995

Ferri, S. L., Abel, T., & Brodkin, E. S. (2018). Sex differences in autism spectrum disorder: A review. *Current Psychiatry Reports, 20*(2), 9. https://doi.org/10.1007/s11920-018-0874-2

Fodstad, J. C., Rojahn, J., & Matson, J. L. (2012). The emergence of challenging behaviors in at-risk toddlers with and without autism spectrum disorder: A cross-sectional study. *Journal of Developmental and Physical Disabilities, 24*(3), 217–234. https://doi.org/10.1007/s10882-011-9266-9

Frazier, T. W., Georgiades, S., Bishop, S. L., & Hardan, A. Y. (2014). Behavioral and cognitive characteristics of females and males with autism in the Simons simplex collection. *Journal of the American Academy of Child & Adolescent Psychiatry, 53*(3), 329–340. https://doi.org/10.1016/j.jaac.2013.12.004

Fulton, A. M., Paynter, J. M., & Trembath, D. (2017). Gender comparisons in children with ASD entering early intervention. *Research in Developmental Disabilities, 68*, 27–34. https://doi.org/10.1016/j.ridd.2017.07.009

Gallagher, S., Pilch, M., & Hannigan, A. (2018). Prior depressive symptoms and persistent child problem behaviours predict future depression in parents of children with developmental disabilities: The growing up in Ireland cohort study [Psychological & Physical Disorders 3200]. *Research in Developmental Disabilities, 80*, 170–179. https://doi.org/10.1016/j.ridd.2018.07.001

Goodman, R. (1997). The strengths and difficulties questionnaire: A research note. *Journal of Child Psychology and Psychiatry, 38*(5), 581–586. https://doi.org/10.1111/j.1469-7610.1997.tb01545.x

Gray, K. M., Keating, C. M., Taffe, J. R., Brereton, A. V., Einfeld, S. L., & Tonge, B. J. (2012). Trajectory of behavior and emotional problems in autism. *American Journal on Intellectual and Developmental Disabilities, 117*(2), 121–133. https://doi.org/10.1352/1944-7588-117-2.121

Greenberg, J. S., Seltzer, M. M., Hong, J., & Orsmond, G. I. (2006). Bidirectional effects of expressed emotion and behavior problems and symptoms in adolescents and adults with autism. *American Journal of Mental Retardation, 111*(4), 229–249. https://doi.org/10.1352/0895-8017(2006)111[229:BEOEEA]2.0.CO;2

Hanratty, J., Livingstone, N., Robalino, S., Terwee, C. B., Glod, M., Oono, I. P., … McConachie, H. (2015). Systematic review of the measurement properties of tools used to measure behaviour problems in young children with autism. *PLoS One, 10* (12). https://doi.org/10.1371/journal.pone.0144649

Happe, F., & Frith, U. (2020). Annual research review: Looking back to look forward - changes in the concept of autism and implications for future research. *Journal of Child Psychology and Psychiatry, 61*(3), 218–232. https://doi.org/10.1111/jcpp.13176

Hartley, S. L., Barker, E. T., Baker, J. K., Seltzer, M. M., & Greenberg, J. S. (2012). Marital satisfaction and life circumstances of grown children with autism across 7 years. *Journal of Family Psychology, 26*(5), 688–697. https://doi.org/10.1037/a0029354

Hartley, S. L., & Sikora, D. M. (2009). Sex differences in autism spectrum disorder: An examination of developmental functioning, autistic symptoms, and coexisting behavior problems in toddlers. *Journal of Autism and Developmental Disorders, 39*(12), 1715–1722. https://doi.org/10.1007/s10803-009-0810-8

Hartley, S. L., Sikora, D. M., & McCoy, R. (2008). Prevalence and risk factors of maladaptive behaviour in young children with autistic disorder. *Journal of Intellectual Disability Research, 52*(10), 819–829. https://doi.org/10.1111/j.1365-2788.2008.01065.x

Hastings, R. P. (2007). Longitudinal relationships between sibling behavioral adjustment and behavior problems of children with developmental disabilities. *Journal of Autism and Developmental Disorders, 37*(8), 1485–1492. https://doi.org/10.1007/s10803-006-0230-y

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6*(2), 107–128. https://doi.org/10.3102/10769986006002107

Helverschou, S. B., Bakken, T. L., Berge, H., Bjorgen, T. G., Botheim, H., Hellerud, J. A., … Howlin, P. (2021). Preliminary findings from a nationwide, multicenter mental health service for adults and older adolescents with autism spectrum disorder and. *Journal of Policy and Practice in Intellectual Disabilities, 18*(2), 162–173. https://doi.org/10.1111/jppi.12366

Herring, S., Gray, K. M., Taffe, J., Tonge, B., Sweeney, D., & Einfeld, S. (2006). Behaviour and emotional problems in toddlers with pervasive developmental disorders and developmental delay: Associations with parental mental health and family functioning. *Journal of Intellectual Disability Research, 50*(12), 874–882. https://doi.org/10.1111/j.1365-2788.2006.00904.x

Hickey, E. J., Bolt, D., Rodriguez, G., & Hartley, S. L. (2020). Bidirectional relations between parent warmth and criticism and the symptoms and behavior problems of children with autism. *Journal of Abnormal Child Psychology, 48*(6), 865–879. https://doi.org/10.1007/s10802-020-00628-5

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21*(11), 1539–1558. https://doi.org/10.1002/sim.1186

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal, 327*(7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557

Hill, A. P., Zuckerman, K. E., Hagen, A. D., Kriz, D. J., Duvall, S. W., & Fombonne, E. (2014). Aggressive behavior problems in children with autism spectrum disorders: Prevalence and correlates in a large clinical sample. *Research in Autism Spectrum Disorders, 8*(9), 1121–1133. https://doi.org/10.1016/j.rasd.2014.05.006

Howlin, P., & Magiati, I. (2017). Autism spectrum disorder: Outcomes in adulthood. *Current Opinion in Psychiatry, 30*(2), 69–76. https://doi.org/10.1097/YCO.0000000000000308

Hsiao, M.-N., Tai, Y.-M., Wu, Y.-Y., Tsai, W.-C., Chiu, Y.-N., & Gau, S. S.-F. (2022). Psychopathologies mediate the link between autism spectrum disorder and bullying involvement: A follow-up study. *Journal of the Formosan Medical Association*. https://doi.org/10.1016/j.jfma.2021.12.030

Johnson, B. T. (2021). Toward a more transparent, rigorous, and generative psychology. *Psychological Bulletin, 147*(1), 1–15. https://doi.org/10.1037/bul0000317

Johnson, B. T., & Eagly, A. (2014). *Handbook of research methods in social and personality psychology: Meta-analysis of research in social and personality psychology*. Cambridge University Press. https://doi.org/10.1017/CBO9780511996481.031

Kanne, S. M., & Mazurek, M. O. (2011). Aggression in children and adolescents with ASD: Prevalence and risk factors. *Journal of Autism and Developmental Disorders, 41*(7), 926–937. https://doi.org/10.1007/s10803-010-1118-4

Kenny, L., Hattersley, C., Molins, B., Buckley, C., Povey, C., & Pellicano, E. (2016). Which terms should be used to describe autism? Perspectives from the UK autism community. *Autism, 20*(4), 442–462. https://doi.org/10.1177/1362361315588200

Kring, S. R., Greenberg, J. S., & Seltzer, M. M. (2009). The impact of health problems on behavior problems in adolescents and adults with autism spectrum disorders: Implications for maternal burden. *Social Work in Mental Health, 8*(1), 54–71. https://doi.org/10.1080/15332980902932441

Kuo, M. H., Magill-Evans, J., & Zwaigenbaum, L. (2015). Parental mediation of television viewing and videogaming of adolescents with autism spectrum disorder and their siblings. *Autism, 19*(6), 724–735. https://doi.org/10.1177/1362361314552199

Kurzius-Spencer, M., Pettygrove, S., Christensen, D., Pedersen, A. L., Cunniff, C., Meaney, F. J., … Rice, S. (2018). Behavioral problems in children with autism spectrum disorder with and without co-occurring intellectual disability. *Research in Autism Spectrum Disorders, 56*, 61–71. https://doi.org/10.1016/j.rasd.2018.09.002

Lai, J. K. Y., & Weiss, J. A. (2017). Priority service needs and receipt across the lifespan for individuals with autism spectrum disorder. *Autism Research, 10*(8), 1436–1447. https://doi.org/10.1002/aur.1786

Laverty, C., Oliver, C., Moss, J., Nelson, L., & Richards, C. (2020). Persistence and predictors of self-injurious behaviour in autism: A ten-year prospective cohort study. *Molecular Autism, 11*(1), 8. https://doi.org/10.1186/s13229-019-0307-z

Levy, A., & Perry, A. (2011). Outcomes in adolescents and adults with autism: A review of the literature. *Research in Autism Spectrum Disorders, 5*(4), 1271–1282. https://doi.org/10.1016/j.rasd.2011.01.023

Li, B., Bos, M. G. N., Stockmann, L., & Rieffe, C. (2020). Emotional functioning and the development of internalizing and externalizing problems in young boys with and without autism spectrum disorder [Developmental Disorders & Autism 3250]. *Autism, 24*(1), 200–210. https://doi.org/10.1177/1362361319874644

Libove, R. A., Frazier, T. W., O'Hara, R., Phillips, J. M., Jo, B., & Hardan, A. Y. (2017). A longitudinal pilot study of behavioral abnormalities in children with autism. *Journal of Psychiatry and Psychiatric Disorders, 1*(4), 215–223. https://doi.org/10.26502/jppd.2572-519X0022

Lin, Y.-N., Iao, L.-S., Lee, Y.-H., & Wu, C.-C. (2021). Parenting stress and child behavior problems in young children with autism spectrum disorder: Transactional relations across time. *Journal of Autism and Developmental Disorders.*. https://doi.org/10.1007/s10803-020-04720-z

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis* (Vol. 49). Sage publications.

Loomes, R., Hull, L., & Mandy, W. P. L. (2017). What is the male-to-female ratio in autism spectrum disorder? A systematic review and meta-analysis. *Journal of the American Academy of Child & Adolescent Psychiatry, 56*(6), 466–474. https://doi.org/10.1016/j.jaac.2017.03.013

Lorenc, T., Rodgers, M., Marshall, D., Melton, H., Rees, R., Wright, K., & Sowden, A. (2018). Support for adults with autism spectrum disorder without intellectual impairment: Systematic review. *Autism, 22*(6), 654–668. https://doi.org/10.1177/1362361317698939

Maenner, M. J., Shaw, K. A., Baio, J., Washington, A., Patrick, M., … Dietz, P. M. (2020). Prevalence of autism spectrum disorder among children aged 8 years - Autism and Developmental Disabilities Monitoring Network, 11 sites, United States, 2016. Morbidity and mortality weekly report. *Surveillance Summaries, 69*(4), 1–12. https://doi.org/10.15585/mmwr.ss6904a1

Mahendiran, T., Brian, J., Dupuis, A., Muhe, N., Wong, P. Y., Iaboni, A., & Anagnostou, E. (2019). Meta-analysis of sex differences in social and communication function in children with autism spectrum disorder and attention-deficit/hyperactivity disorder. *Frontiers in Psychiatry, 10*, 804. https://doi.org/10.3389/fpsyt.2019.00804

Margari, L., Palumbi, R., Peschechera, A., Craig, F., de Giambattista, C., Ventura, P., & Margari, F. (2019). Sex-gender comparisons in comorbidities of children and adolescents with high-functioning autism spectrum disorder. *Frontiers in Psychiatry, 10*, 159. https://doi.org/10.3389/fpsyt.2019.00159

Martin-Borreguero, P., Gómez-Fernández, A. R., De La Torre-Aguilar, M. J., Gil-Campos, M., Flores-Rojas, K., & Perez-Navero, J. L. (2021). Children with autism spectrum disorder and neurodevelopmental regression present a severe pattern after a follow-up at 24 months. *Frontiers in Psychiatry, 12*, 644324. https://doi.org/10.3389/fpsyt.2021.644324

Matson, J. L., Mahan, S., Hess, J. A., Fodstad, J. C., & Neal, D. (2010). Progression of challenging behaviors in children and adolescents with autism spectrum disorders as measured by the autism spectrum disorders-problem behaviors for children (ASD-PBC). *Research in Autism Spectrum Disorders, 4*(3), 400–404. https://doi.org/10.1016/j.rasd.2009.10.010

May, T., Cornish, K., & Rinehart, N. (2014). Does gender matter? A one year follow-up of autistic, attention and anxiety symptoms in high-functioning children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 44*(5), 1077–1086. https://doi.org/10.1007/s10803-013-1964-y

Mazurek, M. O., Dovgan, K., Neumeyer, A. M., & Malow, B. A. (2019). Course and predictors of sleep and co-occurring problems in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 49*(5), 2101–2115. https://doi.org/10.1007/s10803-019-03894-5

Mazurek, M. O., Kanne, S. M., & Wodka, E. L. (2013). Physical aggression in children and adolescents with autism spectrum disorders. *Research in Autism Spectrum Disorders, 7*(3), 455–465. https://doi.org/10.1016/j.rasd.2012.11.004

McClintock, K., Hall, S., & Oliver, C. (2003). Risk markers associated with challenging behaviours in people with intellectual disabilities: A meta-analytic study. *Journal of Intellectual Disability Research, 47*(6), 405–416. https://doi.org/10.1046/j.1365-2788.2003.00517.x

McTiernan, A., Leader, G., Healy, O., & Mannion, A. (2011). Analysis of risk factors and early predictors of challenging behavior for children with autism spectrum disorder. *Research in Autism Spectrum Disorders, 5*(3), 1215–1222. https://doi.org/10.1016/j.rasd.2011.01.009

Melville, C. A., Cooper, S.-A., Morrison, J., Smiley, E., Allan, L., Jackson, A., … Mantry, D. (2008). The prevalence and incidence of mental ill-health in adults with autism and intellectual disabilities. *Journal of Autism and Developmental Disorders, 38*(9), 1676–1688. https://doi.org/10.1007/s10803-008-0549-7

Nasca, B. C., Lopata, C., Donnelly, J. P., Rodgers, J. D., & Thomeer, M. L. (2020). Sex differences in externalizing and internalizing symptoms of children with ASD. *Journal of Autism and Developmental Disorders, 50*(9), 3245–3252. https://doi.org/10.1007/s10803-019-04132-8

Nevill, R. E., Istvan, E. M., & Havercamp, S. M. (2021). The effects of aggression subtypes on burnout and caregiver instability in direct support professionals. *Journal of Positive Behavior Interventions.*. https://doi.org/10.1177/10983007211013794

Nicholls, G., Hastings, R. P., & Grindle, C. (2020). Prevalence and correlates of challenging behaviour in children and young people in a special school setting. *European Journal of Special Needs Education, 35*(1), 40–54. https://doi.org/10.1080/08856257.2019.1607659

Oliver, C., Ellis, K., Agar, G., Bissell, S., Chung, J. C. Y., Crawford, H., … Woodcock, K. (2022). Distress and challenging behavior in people with profound or severe intellectual disability and complex needs: Assessment of causes and evaluation of intervention outcomes. In *International review of research in developmental disabilities, early-access*.

Oliver, C., McClintock, K., Hall, S., Smith, M., Dagnan, D., & Stenfert-Kroese, B. (2003). Assessing the severity of challenging behaviour: Psychometric properties of the challenging behaviour interview. *Journal of Applied Research in Intellectual Disabilities, 16*(1), 53–61. https://doi.org/10.1046/j.1468-3148.2003.00145.x

Orm, S., Øie, M. G., Fossum, I. N., Andersen, P. N., & Skogli, E. W. (2021). Declining trajectories of co-occurring psychopathology symptoms in attention-deficit/hyperactivity disorder and autism Spectrum disorder: A 10-year longitudinal study. *Frontiers in Psychiatry, 12*, 724759. https://doi.org/10.3389/fpsyt.2021.724759

Osborne, L. A., McHugh, L., Saunders, J., & Reed, P. (2008). The effect of parenting behaviors on subsequent child behavior problems in autistic spectrum conditions. *Research in Autism Spectrum Disorders, 2*(2), 249–263. https://doi.org/10.1016/j.rasd.2007.06.004

Ozonoff, S., Goodlin-Jones, B. L., & Solomon, M. (2005). Evidence-based assessment of autism spectrum disorders in children and adolescents. *Journal of Clinical Child and Adolescent Psychology, 34*(3), 523–540. https://doi.org/10.1207/s15374424jccp3403_8

Peters-Scheffer, N., Didden, R., & Korzilius, H. (2012). Maternal stress predicted by characteristics of children with autism spectrum disorder and intellectual disability. *Research in Autism Spectrum Disorders, 6*(2), 696–706. https://doi.org/10.1016/j.rasd.2011.10.003

Petticrew, M., & Roberts, H. (2008). *Systematic reviews in the social sciences*. John Wiley & Sons.

Peverill, S., Smith, I. M., Duku, E., Szatmari, P., Mirenda, P., Vaillancourt, T., … Ungar, W. J. (2019). Developmental trajectories of feeding problems in children with autism spectrum disorder. *Journal of Pediatric Psychology, 44*(8), 988–998. https://doi.org/10.1093/jpepsy/jsz033

Pierce, N. P., O'Reilly, M. F., Sorrells, A. M., Fragale, C. L., White, P. J., Aguilar, J. M., & Cole, H. A. (2014). Ethnicity reporting practices for empirical research in three autism-related journals. *Journal of Autism and Developmental Disorders, 44*(7), 1507–1519. https://doi.org/10.1007/s10803-014-2041-x

Pilling, N., McGill, P., & Cooper, V. (2007). Characteristics and experiences of children and young people with severe intellectual disabilities and challenging behaviour attending 52-week residential special schools. *Journal of Intellectual Disability Research, 51*(Pt 3), 184–196. https://doi.org/10.1111/j.1365-2788.2006.00852.x

Piro-Gambetti, B., Rodriguez, G., Papp, L. M., Greenlee, J. L., & Hartley, S. L. (2021). Parent couple conflict and emotional and behavioral problems in youth with autism: Longitudinal investigation of bidirectional effects. *Development and Psychopathology*, 1–11. https://doi.org/10.1017/S0954579421000596

Postorino, V., Fatta, L. M., De Peppo, L., & …& Mazzone, L.. (2015). Longitudinal comparison between male and female preschool children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 45*(7), 2046–2055. https://doi.org/10.1007/s10803-015-2366-0

Provenzani, U., Fusar-Poli, L., Brondino, N., Damiani, S., Vercesi, M., Meyer, N., Rocchetti, M., & Politi, P. (2020). What are we targeting when we treat autism spectrum disorder? A systematic review of 406 clinical trials. *Autism, 24*(2), 274–284. https://doi.org/10.1177/1362361319854641

Rabbani, M., Haque, M. M., Das Dipal, D., Zarif, M. I. I., Iqbal, A., Schwichtenberg, A., … Ahamed, S. I. (2021). An mCARE study on patterns of risk and resilience for children with ASD in Bangladesh. *Scientific Reports, 11*(1), 21342. https://doi.org/10.1038/s41598-021-00793-7

Rattaz, C., Michelon, C., Munir, K., & Baghdadli, A. (2018). Challenging behaviours at early adulthood in autism spectrum disorders: Topography, risk factors and evolution. *Journal of Intellectual Disability Research, 62*(7), 637–649. https://doi.org/10.1111/jir.12503

Reed, P., & Osborne, L. A. (2013). The role of parenting stress in discrepancies between parent and teacher ratings of behavior problems in young children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 43*(2), 471–477. https://doi.org/10.1007/s10803-012-1594-9

Reed, P., Osborne, L. A., & Waddington, E. M. (2012). A comparative study of the impact of mainstream and special school placement on the behaviour of children with autism spectrum disorders. *British Educational Research Journal, 38*(5), 749–763. https://doi.org/10.1080/01411926.2011.580048

Rice, L. J., Gray, K. M., Howlin, P., Taffe, J., Tonge, B. J., & Einfeld, S. L. (2016). The developmental trajectory of self-injurious behaviours in individuals with Prader Willi syndrome, autism spectrum disorder and intellectual disability. *Diseases, 4*(1). https://doi.org/10.3390/diseases4010009

Richards, C., Moss, J., Nelson, L., & Oliver, C. (2016). Persistence of self-injurious behaviour in autism spectrum disorder over 3 years: A prospective cohort study of risk markers. *Journal of Neurodevelopmental Disorders, 8*, 21. https://doi.org/10.1186/s11689-016-9153-x

Roselló Miranda, B., Berenguer Forner, C., Baixauli Fortea, I., & Miranda Casas, A. (2022). Personal and family childhood predictors of functional outcomes of adolescents with autism spectrum disorder. *Psicothema, 34*(4), 508–517.

Sacrey, L. A. R., Zwaigenbaum, L., Bryson, S., Brian, J., Smith, I. M., … Armstrong, V. (2015). Can parents' concerns predict autism spectrum disorder? A prospective study of high-risk siblings from 6 to 36 months of age. *Journal of the American Academy of Child and Adolescent Psychiatry, 54*(6), 470–478. https://doi.org/10.1016/j.jaac.2015.03.014

Safer-Lichtenstein, J., Hamilton, J. C., & McIntyre, L. L. (2019). Examining demographics in randomized controlled trials of group-based social skills interventions for individuals with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 49*(8), 3453–3461. https://doi.org/10.1007/s10803-019-04063-4

Shattuck, P. T., Seltzer, M. M., Greenberg, J. S., Orsmond, G. I., Bolt, D., Kring, S., … Lord, C. (2007). Change in autism symptoms and maladaptive behaviors in adolescents and adults with an autism spectrum disorder. *Journal of Autism and Developmental Disorders, 37*(9), 1735–1747.

Shea, V., & Mesibov, G. B. (2005). Adolescents and adults with autism. In F. R. Volkmar, R. Paul, A. Klin, & D. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders, Vol. 1: Diagnosis, development, neurobiology* (pp. 288–311). Wiley & Sons Inc.

Sikora, D. M., Hall, T. A., Hartley, S. L., Gerrard-Morris, A. E., & Cagle, S. (2008). Does parent report of behavior differ across ADOS-G classifications: Analysis of scores from the CBCL & GARS. *Journal of Autism and Developmental Disorders, 38*(3), 440–448. https://doi.org/10.1007/s10803-007-0407-z

Simonoff, E., Jones, C. R. G., Baird, G., Pickles, A., Happé, F., & Charman, T. (2013). The persistence and stability of psychiatric problems in adolescents with autism spectrum disorders. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 54*(2), 186–194. https://doi.org/10.1111/j.1469-7610.2012.02606.x

Simó-Pinatella, D., Mumbardó-Adam, C., Alomar-Kurz, E., Sugai, G., & Simonsen, B. (2019). Prevalence of challenging behaviors exhibited by children with disabilities: Mapping the literature. *Journal of Behavioral Education, 28*(3), 323–343. https://doi.org/10.1007/s10864-019-09326-9

Siracusano, M., Segatori, E., Riccioni, A., Emberti Gialloreti, L., Curatolo, P., & Mazzone, L. (2021). The impact of COVID-19 on the adaptive functioning, behavioral problems, and repetitive behaviors of Italian children with autism spectrum disorder: An observational study. *Children, 8*(2). https://doi.org/10.3390/children8020096

Soke, G. N. (2015). *Epidemiology of self-injurious behaviors in children with autism spectrum disorders (Doctoral dissertation). ProQuest Dissertations and Theses Global*.

Soke, G. N., Rosenberg, S. A., Hamman, R. F., … DiGuiseppi, C. (2016). Brief report: Prevalence of self-injurious behaviors among children with autism spectrum disorder—A population-based study. *Journal of Autism & Developmental Disorders, 46*(11), 3607–3614. https://doi.org/10.1007/s10803-016-2879-1

Soke, G. N., Rosenberg, S. A., Hamman, R. F., … DiGuiseppi, C. (2017). Factors associated with self-injurious behaviors in children with autism spectrum disorder: Findings from two large national samples. *Journal of Autism and Developmental Disorders, 47*(2), 285–296. https://doi.org/10.1007/s10803-016-2951-x

Sowislo, J. F., & Orth, U. (2013). Does low self-esteem predict depression and anxiety? A meta-analysis of longitudinal studies. *Psychological Bulletin, 139*(1), 213–240. https://doi.org/10.1037/a0028931

StataCorp. (2021). *Stata Statistical Software: Release 17*. StataCorp LLC.

Steenfeldt-Kristensen, C., Jones, C. A., & Richards, C. (2020). The prevalence of self-injurious behaviour in autism: A meta-analytic study. *Journal of Autism and Developmental Disorders, 50*, 3857–3873. https://doi.org/10.1007/s10803-020-04443-1

Summers, J., Shahrami, A., Cali, S., D'Mello, C., Kako, M., Palikucin-Reljin, A., Savage, M., Shaw, O., & Lunsky, Y. (2017). Self-injury in autism spectrum disorder and intellectual disability: Exploring the role of reactivity to pain and sensory input. *Brain Sciences, 7*(11). https://doi.org/10.3390/brainsci7110140

Taylor, J. L., & Seltzer, M. M. (2010). Changes in the autism behavioral phenotype during the transition to adulthood. *Journal of Autism and Developmental Disorders, 40*(12), 1431–1446. https://doi.org/10.1007/s10803-010-1005-z

Tonnsen, B. L., Boan, A. D., Bradley, C. C., Charles, J., Cohen, A., & Carpenter, L. A. (2016). Prevalence of autism spectrum disorders among children with intellectual disability. *American Journal on Intellectual and Developmental Disabilities, 121*(6), 487–500. https://doi.org/10.1352/1944-7558-121.6.487

Totsika, V., Felce, D., Kerr, M., & Hastings, R. P. (2010). Behavior problems, psychiatric symptoms, and quality of life for older adults with intellectual disability with and without autism. *Journal of Autism and Developmental Disorders, 40*(10), 1171–1178. https://doi.org/10.1007/s10803-010-0975-1

Totsika, V., Hastings, R. P., Emerson, E., Berridge, D. M., & Lancaster, G. A. (2011). Behavior problems at 5 years of age and maternal mental health in autism and intellectual disability. *Journal of Abnormal Child Psychology, 39*(8), 1137–1147. https://doi.org/10.1007/s10802-011-9534-2

Totsika, V., Hastings, R. P., Emerson, E., Lancaster, G. A., Berridge, D. M., & Vagenas, D. (2013). Is there a bidirectional relationship between maternal well-being and child behavior problems in autism spectrum disorders? Longitudinal analysis of a population-defined sample of young children. *Autism Research, 6*(3), 201–211. https://doi.org/10.1002/aur.1279

Tubio-Fungueirino, M., Cruz, S., Sampaio, A., Carracedo, A., & Fernandez-Prieto, M. (2021). Social camouflaging in females with autism spectrum disorder: A systematic review. *Journal of Autism and Developmental Disorders, 51*(7), 2190–2199. https://doi.org/10.1007/s10803-020-04695-x

Uljarevic, M., Hedley, D., Rose-Foley, K., Magiati, I., Cai, R. Y., Dissanayake, C., … Trollor, J. (2020). Anxiety and depression from adolescence to old age in autism spectrum disorder. *Journal of Autism and Developmental Disorders, 50*(9), 3155–3165. https://doi.org/10.1007/s10803-019-04084-z

Vaillancourt, T., Haltigan, J. D., Smith, I., Zwaigenbaum, L., Szatmari, P., … Bryson, S. (2017). Joint trajectories of internalizing and externalizing problems in preschool children with autism spectrum disorder. *Development and Psychopathology, 29*(1), 203–214. https://doi.org/10.1017/S0954579416000043

Vandewalle, K., & Melia, Y. (2021). Psychosocial and behavioural factors associated with self injurious behaviour (SIB) in individuals with autism spectrum disorders (ASD). *Research in Autism Spectrum Disorders, 81*. https://doi.org/10.1016/j.rasd.2020.101713

Varela, R. E., DuPont, R., Kamps, J. L., … Pucci, G. (2020). Age differences in expression of generalized and social anxiety among youth with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 50*(3), 730–740. https://doi.org/10.1007/s10803-019-04289-2

Walsh, C., Lydon, S., O'Dowd, E., & O'Connor, P. (2020). Barriers to healthcare for persons with autism: A systematic review of the literature and development of a taxonomy. *Developmental Neurorehabilitation, 23*(7), 413–430. https://doi.org/10.1080/17518423.2020.1716868

Wood-Downie, H., Wong, B., Kovshoff, H., Cortese, S., & Hadwin, J. A. (2021). Research review: A systematic review and meta-analysis of sex/gender differences in social interaction and communication in autistic and nonautistic children and adolescents.

Journal of Child Psychology and Psychiatry, 62(8), 922–936. https://doi.org/10.1111/jcpp.13337

Woodman, A. C., Smith, L. E., Greenberg, J. S., & Mailick, M. R. (2015). Change in autism symptoms and maladaptive behaviors in adolescence and adulthood: The role of positive family processes. *Journal of Autism and Developmental Disorders, 45*(1), 111–126. https://doi.org/10.1007/s10803-014-2199-2

Woodman, A. C., Smith, L. E., Greenberg, J. S., & Mailick, M. R. (2016). Contextual factors predict patterns of change in functioning over 10 years among adolescents and adults with autism spectrum disorders. *Journal of Autism & Developmental Disorders, 46*(1), 176–189. https://doi.org/10.1007/s10803-015-2561-z

Yacoub, E., Dowd, M., McCann, L., & Burke, L. (2018). Impact of challenging behaviour on siblings of people with autism. *Advances in Mental Health and Intellectual Disabilities, 12*(5/6), 145–152. https://doi.org/10.1108/amhid-01-2018-0001

Zaidman-Zait, A., Mirenda, P., Duku, E., Szatmari, P., Georgiades, S., … Thompson, A. (2014). Examination of bidirectional relationships between parent stress and two types of problem behavior in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 44*(8), 1908–1917. https://doi.org/10.1007/s10803-014-2064-3

Dawn Adams is the Director of the Autism Centre of Excellence at Griffith University and a registered Clinical Psychologist. Her research focus is upon anxiety and behaviour in individuals with autism, rare genetic syndromes and/or neurodevelopmental disorders.

Nicole Dargue is a Lecturer at the Autism Centre of Excellence at Griffith University and a Clinical Neuropsychologist. She has expertise in meta-analytic methods. Her current research is examining non-verbal behaviours and communication profiles in children on the autism spectrum.

Jessica Paynter is a Senior Lecturer in the School of Applied Psychology at Griffith University, Australia and a registered Clinical Psychologist. Her research focuses on evidence-based practices, implementation science and learning and literacy in students on the autism spectrum.