Behaviourally based interventions in children on the autism spectrum: A systematic review and dose-response meta-analysis

Version 2.0 – May 2023
Research and Evaluation Branch

ndis.gov.au

The contents of this document are OFFICIAL

Contents

Research and Evaluation Branch

The Research and Evaluation Branch is responsible for ensuring that National Disability Insurance Agency (NDIA) policies, practices and priorities are informed by trustworthy and robust evidence so that decisions can be based on an understanding of what works, what doesn't and the benefit to participants and the Agency.

This document

This report presents research findings from a systematic review and meta-analysis investigating the overall efficacy of behaviourally based interventions for children on the autism spectrum, as well as an investigation of contributing factors such as amount (dose) of intervention and other design (intervention type, person delivering, setting) and participant (age) factors.

Disclaimer

The NDIA accepts no responsibility for the accuracy or completeness of any material contained in this report. Further, the NDIA disclaims all liability to any person in respect of anything, and of the consequences of anything, done or not done by any such person in reliance, whether wholly or partly, upon any information presented in this report.

Material in this report is made available on the understanding that the NDIA is not providing professional advice. Before relying on any of the material in this report, users should obtain appropriate professional advice.

Views and recommendations of third parties, which may also be included in this report, do not necessarily reflect the views of the National Disability Insurance Agency, or indicate a commitment to a particular course of action.

Acknowledgements

The NDIA acknowledges the Traditional Owners and Custodians throughout Australia and their continuing connection to the many lands, seas, and communities. The NDIA pays respect to Elders past and present, and extends this acknowledgement and respect to any Aboriginal and Torres Strait Islander people who may be reading this Report.

The Research and Evaluation Branch would like to acknowledge Dr Megan Clark (Olga Tennison Autism Research Centre, LaTrobe University), Professor Valsamma Eapen (UNSW Sydney), Professor Cheryl Dissanayake (Olga Tennison Autism Research Centre, LaTrobe University), Professor Andrew Whitehouse (Telethon Kids Institute), and Professor Adam Guastella (Brain and Mind Centre, Children's Hospital Westmead Clinical School, University of Sydney) who assisted with the development of the project and provided expert advice throughout. We also acknowledge the broader NDIA team who assisted in the development of this research project.

Suggested citation: National Disability Insurance Agency 2023. 'Behaviourally based interventions in children on the autism spectrum: A systematic review and dose-response meta-analysis'. Prepared by E Ghezzi, A Lampit.

Enquiries about this publication should be addressed to:
National Disability Insurance Agency. Email: research@ndis.gov.au

Abbreviations

ABA	Applied behaviour analysis
ASELCCs	Autism Specific Early Learning and Care Centres
CABAS	Comprehensive Application of Behaviour Analysis to Schooling
CI	Confidence interval
CRC	Cooperative Research Centre
DIR	Developmental Individual-Difference Relationship-Based
DTT	Discrete Trial Teaching
EIBI	Early Intensive Behavioural Intervention
ESDM	Early Start Denver Model
GOLIAH	Gaming Open Library for Intervention in Autism at Home
IQ	Intelligence quotient
JASPER	Joint Attention, Symbolic Play, Engagement, and Regulation
LEAP	Learning Experiences and Alternative Program for Preschoolers and their Parents
NA	Not applicable
NDBI	Naturalistic developmental behavioural intervention
NDIA	National Disability Insurance Agency
NDIS	National Disability Insurance Scheme
PACT	Paediatric Autism and Communication Therapy
PCIT	Parent-Child Interaction Therapy
PECS	Picture exchange communication system
PRISMA	Preferred Reporting Items for Systematic reviews and Meta-Analyses
PRT	Pivotal response treatment
RoB 2.0	Cochrane Risk of bias 2
ROBINS-I	Risk Of Bias In Non-randomized Studies
TAU	Treatment as usual

Tau^2	Tau-squared
TEACCH	Treatment and Education of Autistic and Related Communication Handicapped Children
UK	United Kingdom
US	United States

Glossary

Term	Definition
Autism spectrum disorder	Autism spectrum disorder (also referred to as “autism”) is the collective term for a group of neurodevelopmental conditions affecting the brain’s growth and development. Autism is a life-long condition which can impact, to varying degrees, all areas of a person’s life, including social communication and social interaction. The behavioural features of autism are often present before a person is three years of age but in others they may not be recognised until their school years or later in life. The developmental challenges, signs and/or symptoms can vary widely in nature and degree between individuals, and in the same individual over time – that is why the term “spectrum” is used. We know that people prefer different terms to describe autism. We have used children on the autism spectrum (person-first language) to be consistent with how we refer to other target populations.
Behaviourally based intervention	We define behaviourally based interventions as interventions for children on the autism spectrum which are underpinned by behavioural principles. These interventions span several intervention categories defined within the Autism CRC umbrella review (external) of non-pharmacological interventions for children on the autism spectrum. These categories include behavioural interventions (for example, applied behaviour analysis [ABA]), naturalistic developmental behavioural interventions (NDBIs), technology-based interventions, developmental interventions, TEACCH, and other (uncategorised) intervention types. Behaviourally based interventions are typically delivered by trained clinicians but may also involve training parents or caregivers in behavioural principles to facilitate parent-delivered intervention.
Clinician	Throughout the report, we use the term ‘clinician’ to refer qualified or trained individuals who deliver interventions. These are typically the providers of the intervention.
Dose	To ensure results reflect current practice in the Australian context, dose was defined as clinician-delivered hours of intervention as this is the main delivery method funded through the NDIS. Dose was measured in two ways: 1. total clinician-delivered hours of the intervention, and 2. monthly clinician-delivered hours of intervention (i.e., dose intensity).

Term	Definition
Meta-analysis	A meta-analysis uses statistics to combine the results from these studies to find out how much of an effect the intervention has on selected outcomes (which we call the effect size) and what factors can predict the size of the reported effects.
Parent	For clarity of writing, throughout this report we use the term ‘parent’ to refer to any individual who has parenting responsibilities for a child.
Systematic review	A systematic review summarises the evidence from research studies focused on the same topic.

Summary

This report summarises findings from a systematic review and meta-analysis of research studies investigating the benefits of behaviourally based interventions in children (less than 7 years old) on the autism spectrum. A systematic review is a method for collecting evidence from studies of a particular topic. A meta-analysis involves synthesising this evidence statistically to arrive at quantifiable conclusions. Using this method, the benefits of behaviourally based interventions were investigated and how effects are related to the amount of intervention provided by clinicians (dose-response) was explored. An analysis of other intervention design factors and how these may relate to outcomes is also reported.

The project was conducted to assist the NDIA in developing evidence-based policy and practice guidance regarding the determination of reasonable and necessary supports for participants with autism under the age of 7. The Agency will consider this research evidence, alongside other important factors individual to each child and their circumstances, in determining what behaviourally based interventions are funded to help the child and family achieve their goals, aligned with the decision making criteria of the NDIS Act.

Key conclusions

Behaviourally based interventions are efficacious for key outcomes in children on the autism spectrum compared with children who undergo treatment as usual or non-behavioural interventions, but the pooled effect sizes are small (about 30% of a standard deviation) and vary considerably across studies.

Even for equivalent hours of clinician-delivered intervention, there is evidence for added benefit of behaviourally based intervention above that of treatment as usual (i.e., standard care or community-based intervention) or non-behaviourally based intervention.

With dose relationships varying based on the outcomes of the intervention, decisions regarding the amount of intervention provided should take into account the specific goals of the participant and the planned outcomes of the intervention.

- For goals related to the autism characteristics (i.e., socialisation, social affect, challenging behaviours, restricted repetitive behaviours, etc) or cognition and language, benefit of behaviourally based interventions can be seen at low total doses and dose intensities. There seems to be very little added benefit of increased hours and intensity of intervention for cognition and language, and no evidence for added benefit with increased dose in the case of autism characteristics, which means that many participants may benefit from less intervention hours.
- If the intervention is specifically intended to target adaptive functioning, there may be no benefit of behaviourally based interventions below approximately 800 total hours or an intensity of 65 hours per month. With such high dose requirements, intervention approaches other than behaviourally based interventions may be more feasible and cost-effective, unless supported by evidence for the specific intervention requested.

Importantly, just focusing on dose by itself is a weak predictor of outcomes of behaviourally based intervention in children on the autism spectrum and should only be considered as one factor in a treatment decision to ensure alignment to the holistic goals of the child.

No differences were found for benefit of behaviourally based interventions by type of intervention, type of comparison group, primary intervention setting, person delivering the intervention, or age of the participant. With no differences found in the intervention and design factors investigated, it is likely that all factors investigated could be useful in the right context. This emphasizes the importance of tailoring the intervention design to the unique context and goals of the child and their family.

With no evidence to suggest that interventions by a parent are inferior to those delivered by clinicians, this warrants further investigation of parent-delivered interventions. This report did not assess factors which may contribute to the success of these interventions, such as parent training (duration, content, etc), support for parents throughout the intervention (type, amount), and fidelity.

To further the work presented here, other meta-analysis methods would allow for examining differences in individual circumstances as well as other participant-level (e.g., autism severity) and intervention-related factors to be explored in more detail.

1. Background and NDIS context

Autism spectrum disorder, also referred to as “autism”, is the collective term for a group of neurodevelopmental conditions which affect brain growth and development. Characteristics of autism vary greatly in nature and degree, but may involve challenges with social interaction and communication, sensory issues, and restricted and repetitive behaviours, interests, or activities (American Psychiatric Association, 2013). Behavioural features of autism are often present before the age of 3 years old but may not be recognised until later in life.

Autism is the largest primary disability category in the NDIS, encompassing 34% of active participants. As of December 2022, 22,018 NDIS participants under the age of 6 had a primary disability of autism, representing approximately one-quarter of NDIS participants in this age group. Intervention during childhood represents an important opportunity to support early development and build on the child’s strengths. Through the NDIA’s early childhood approach, children under 7 years old can access NDIS funding for early intervention supports. The NDIA’s early childhood approach is based on the National Guidelines for Best Practice in Early Childhood Intervention (external), emphasising the central role of family as well as development in natural, everyday settings (Early Childhood Intervention Australia, 2016). These principles are promoted within the Autism CRC National Guideline (external) (Trembath et al., 2022) for supporting children on the autism spectrum and their families, where similar recommendations are made for approaches which are child and family-centred, individualised, and strengths-focused (Trembath et al., 2022).

A wide range of non-pharmacological interventions are available for children on the autism spectrum, all of which aim to assist early development and skill acquisition across domains (e.g., social affect, cognition, adaptive functioning). Behavioural principles underpin a considerable range of these interventions (see **Appendix 1** for list of interventions), which span several intervention categories defined within the Autism CRC umbrella review (external) of non-pharmacological interventions for children on the autism spectrum (Whitehouse & Eapen, 2020). These categories include behavioural interventions (for example, applied behaviour analysis [ABA]), naturalistic developmental behavioural interventions (NDBIs), technology-based interventions, developmental interventions, TEACCH, and other (uncategorised) intervention types.

Behaviourally based interventions are typically delivered by trained clinicians but may also involve training parents or caregivers in behavioural principles to facilitate parent-delivered intervention. It is estimated that at least 7,936 participants under the age of 7 with a primary disability of autism received some form of capacity building support (which may include behaviourally based interventions) funded through the NDIS early childhood services in 2021.

There is currently low to moderate evidence supporting the efficacy of behavioural interventions for core autism characteristics which includes communication, cognition, behaviour, school readiness and academic skills (Whitehouse & Eapen, 2020). While some evidence exists, there is limited and varied evidence which reports the effect of amount (“dose”) of behaviourally based interventions on outcomes. Additionally, results of dose response investigations vary by their focus and research methods. For example, one systematic review reported benefits of higher intensity (hours per week) of intervention for cognition and adaptive behaviour, but not language, and found no effect of total

intervention duration (Makrygianni & Reed, 2010). Another meta-analysis identified a potential linear association between hours of intervention and benefit to adaptive functioning and language outcomes (Virues-Ortega, 2010). Finally, a meta-analysis of individual participant data reported larger effect sizes for overall autism characteristics after 24 months of intervention compared to 12 months (Rodgers et al., 2020).

With limited understanding of the optimal dose of behaviourally based interventions for children on the autism spectrum, there are currently no evidence-based guides or best practice principles available, which has led to inconsistency in service provision and participants receiving varied hours of an intervention. Confusion associated with this can be distressing for parents, as it is unclear how many hours are necessary for their child to achieve the best outcomes. With such variable results and a general lack of investigation of contributing factors (i.e., setting, intervention characteristics, participant characteristics, etc) (Trembath et al., 2021), it is difficult to use existing evidence to guide NDIA policy and operations. Crucially, evidence regarding dose of interventions is only one variable which can inform an individualised decision about what is needed for a particular child within their unique environmental context and family circumstances. As such, it is important to have a body of evidence of what works, at what dose, for who, in what context and to what end, to reduce confusion, strengthen guidance, and ensure the best outcomes for participants are achieved. One of the national best practice principles for childhood intervention is that interventions and practice must be research-based, so creating a body of evidence to support NDIA policy and practice is an essential part of how the NDIA must discharge its decision making responsibilities regarding funding for reasonable and necessary supports under the NDIS Act.

2. What did we do?

A comprehensive systematic review and meta-analysis was undertaken to identify the overall efficacy of behaviourally based interventions for various outcomes in children on the autism spectrum and, importantly, how contributing factors impact these outcomes.

The objectives of the systematic review and meta-analysis were to:

- examine the evidence for the efficacy of behaviourally based interventions in children under 7 years on the autism spectrum on child (functional and developmental) and family outcomes; and
- investigate how effects are related to dose (amount) of intervention as well as other factors relating to study design, intervention, comparison group, and child characteristics.

2.1 Overview of methods used

Findings included in this report were identified through a systematic review and meta-analysis. A systematic review is a process to locate and summarise the results of all studies that ask a particular research question, usually by using different methods with a common underlying question (e.g., are behaviourally based interventions efficacious in improving adaptive functioning in children on the autism spectrum?). A meta-analysis is a statistical procedure that combines results from the studies identified in a systematic review to find a common estimate of effect between studies, as well as how effects might vary across settings and other factors (e.g., age, intervention type).

A full description of the study methods is available in **Appendix 1**.

2.1.1 Search and screening of articles

Five databases were searched to identify all published studies that examined the impact of a behaviourally based intervention on a range of outcomes in children under the age of 7 years on the autism spectrum. The search screening process is in **Appendix 1** and criterion for inclusion are outlined below by Population, Intervention, Comparison, and Outcome (i.e., PICO).

Population

Studies were eligible if they included children who:

- are 7-years old or younger at the beginning of the intervention,
- **AND** have a diagnosis of autism spectrum disorder (or have a high likelihood of autism spectrum disorder for children less than 3 years)

Intervention

Studies were eligible if they included behaviourally based interventions (typical interventions listed in **Appendix 1**).

Interventions may be delivered to children:

- face-to-face,
- **OR** via telehealth.

Interventions may be delivered to children by:

- qualified or trained individuals,
- parents,
- caregivers,
- teachers,
- **OR** a combination of these.

Interventions may be:

- one-to-one,
- **OR** in a small group format.

Comparison

Studies were eligible if they included a comparison group which comprised of children 7-years old or younger on the autism spectrum who:

- continued standard care or treatment as usual (i.e., community interventions),
- were on a waitlist,
- **OR** completed an alternate, non-behaviourally based intervention.

Studies without a comparison group (i.e., single arm studies) and case studies were excluded.

Outcomes

Studies were eligible if they reported outcomes that were measured both:

- before intervention has begun,
- **AND** following intervention.

Outcomes within the following five domains were eligible for inclusion:

1. **Autism characteristics**
    a. Global measures of autism characteristics and behaviours
    b. Emotional regulation
    c. Restricted repetitive behaviours/sensory
    d. Social affect (foundational social skills)
    e. Socialisation (application and competence in using social skills)
    f. Challenging behaviours
2. **Cognitive and language outcomes**
    a. Cognition (verbal and nonverbal cognitive abilities and motor skills)
    b. Language (receptive and expressive language and verbal communication)

3. **Functional outcomes**
    a. Adaptive behaviour (everyday functioning e.g., Vineland Adaptive Behaviour Scales)
    b. Education outcomes (e.g., education setting/level of support)
4. **Family outcomes**
    a. Caregiver or family wellbeing
    b. Quality of life (child, caregiver, overall family unit)
5. **Adverse effects**
    a. Child distress (e.g., anxiety/depression)
    b. Parent stress/burden (e.g., Parenting Stress Index)
    c. Reduced participation in mainstream settings (e.g., reduced participation in preschool)

For inclusion, each study must report **at least one** of these outcomes.

2.1.2 Combining effects from included studies

Intervention effects within each outcome domain (see **Section 4.1.1** for domain descriptions) were combined across studies using meta-analysis. The intervention effect was measured using standardised mean difference, calculated as Hedges' *g,* with 95% confidence interval (CI).

Hedges' *g* provides the difference (effect) between two groups in standard deviation units. This allows us to combine the intervention effects from the different outcome domains into a single analysis. A positive Hedges' *g* means that the intervention was beneficial over the comparison group.

The confidence interval estimated the precision of the estimate of effect. When the confidence interval includes the null (i.e., when the lower bound of the interval is below zero), the effect estimate is too imprecise to be considered statistically significant, meaning there is not enough information to determine whether the intervention is beneficial or not.

Each pooled estimate is provided along with a measure of statistical heterogeneity, denoted as tau-squared (Tau^2). This gives an estimation of the extent to which an effect estimate is inconsistent across studies.

2.1.3 Investigating a range of contributing factors to efficacy

To investigate variability in these combined efficacy estimates (i.e., heterogeneity), further analyses were conducted to explore the effect of dose as well as other study, intervention, and population-based factors. The factors examined within subgroup analyses were (1) person delivering intervention; (2) intervention type; (3) comparison group type; (4) age group; (5) primary intervention setting; and (6) study design.

These factors are critical as they are overarching characteristics of the study, intervention and participants which may impact the efficacy of the intervention. It is also useful to know what types of interventions work best, what settings are associated with best outcomes, whether there are better

outcomes when the intervention is delivered by certain individuals, and whether the behaviourally based interventions are more efficacious in certain age groups compared to usual alternatives.

The levels of each of these subgroups, in addition to descriptions and examples are described below.

Person delivering intervention

1. **Clinician** (i.e., clinician, facilitator, or provider)
2. **Clinician and parent**
3. **Parent** (i.e., parent or caregiver)
4. **Teacher** (i.e., early educator or teacher)

Intervention category

The six intervention categories are based on Autism CRC definitions (Whitehouse & Eapen, 2020).

1. **Behavioural**, for example:
    a. Early Intensive Behavioural Intervention (EIBI)
    b. Applied Behavioural Analysis (ABA)
    c. Discrete Trial Teaching (DTT)
    d. Picture exchange communication system (PECS)
2. **Naturalistic Developmental Behavioural Interventions (NDBIs)**, for example:
    a. Early Start Denver Model (ESDM)
    b. Pivotal response treatment (PRT)
    c. Joint Attention, Symbolic Play, Engagement, and Regulation (JASPER)
3. **Developmental**, for example:
    a. DIR Floortime
4. **Technology-based**, for example:
    a. GOLIAH
5. **TEACCH** (a discrete intervention)
6. **Other** interventions, for example:
    a. Autism 1-2-3
    b. Learning Experiences and Alternative Program for Preschoolers and their Parents (LEAP)
    c. Parent-Child Interaction Therapy (PCIT)

Comparison group

1. **Treatment as usual (TAU)**, for example:
    a. Waitlist controls
    b. Usual or routine care, often in the community (e.g., speech and language therapy).
    c. Regular or non-specific specialised school-based services.
    d. Public education or psychoeducation for parents.

2. **Eclectic** interventions (i.e., a specific early intervention program or intervention that is not behaviourally based and not part of routine or usual care)

Age group

1. **0-1 years** (up to but not including children 2 years old)
2. **2-4 years** (children from the age of 2, up to but not including children 5 years old)
3. **5-6 years** (children from the age of 5 years old)

Primary intervention setting

1. **Health** (i.e., interventions primarily delivered within clinical [e.g., psychology, university] specialist or private health settings)
2. **Community** (i.e., interventions primarily delivered in childcare centres, community or public agencies or service centres)
3. **Early education** (i.e., interventions primarily delivered in the child's preschool or school)
4. **Home** (i.e., interventions primarily delivered in child's home)

Study design

1. **Random** (i.e., randomised controlled trial)
2. **Non-random** (i.e., non-randomised controlled trial)
3. **Cohort study** (i.e., prospective comparison of intervention groups)

Investigating the effect of dose

To ensure results reflect current practice in the Australian context, dose was defined as **clinician-delivered hours** of intervention as this is the main delivery method funded through the NDIS. Specific parent-delivered interventions and interventions in early education settings (often delivered by teachers) were excluded from the dose analyses. It is important to note these studies were still included in the main efficacy analyses, as well as subgroup analyses investigating the effects of study, intervention and population characteristics.

Dose was measured in two ways:

3. **total** clinician-delivered hours of the intervention, and
4. **monthly** clinician-delivered hours of intervention (i.e., dose intensity).

It is worth noting that duration (weeks) of intervention was not accounted for or further explored within this report. Nevertheless, duration of intervention had a large association with both total hours of intervention ($r = 0.8$, $p < 0.001$) and monthly hours of intervention ($r = 0.8$, $p < 0.001$), and therefore its potential effect on the results is limited.

The effect of the dose of interventions was explored using three methods, listed here and further detailed below:

1. Relationship between dose and efficacy
2. Comparing efficacy for lower versus higher total and monthly dose
3. Relationship between dose and change from baseline to follow-up separately within the:
    a. Behaviourally based intervention group
    b. Comparison group

1. Relationship between dose and efficacy

Linear and non-linear models were used to explore the relationship between dose (**total** and **monthly** clinician-delivered hours) and efficacy (Hedges' *g*) of the intervention (see **Appendix 1** for description) across each of the outcome domains (where data permits).

As described previously, efficacy (Hedges' g) of the behaviourally based intervention (as compared to the comparative group) was calculated for each outcome domain reported in each study. For each outcome domain, the efficacy of each individual study was then plotted against the dose of intervention implemented within that study to visualise the dose relationship.

These relationships were then significance tested to assess the likelihood of a relationship between dose and efficacy. The model estimate (β) indicates the size of the relationship, with a positive model estimate indicating a positive relationship (as dose increases, so does efficacy) and a negative model estimate indicating a negative relationship (as dose increases, efficacy decreases). The model estimate (β) can be interpretated as the added benefit (in Hedges' *g*) associated with each additional hour of intervention. For example, if β=0.01, the effect size is estimated to increase by *g*=0.1 (i.e., 10% standard deviation difference) for every 10 additional hours.

2. Comparing efficacy for lower versus higher total and monthly dose

The dose analyses were corroborated by investigating differences in the efficacy of lower versus higher total dose as well as lower versus higher dose intensity (monthly hours), with lower and higher defined in both cases based on a median split. Lower and higher intervention doses were also directly compared within each outcome domain.

3. Relationship between dose and change from baseline to follow-up separately within the intervention group and the comparison group

This analysis involved the calculation of an effect size (Hedges' g) separately for the *behaviourally based intervention group* and the *comparison group* in each study. The resulting effect size is the change between two time-points: baseline (pre-intervention) and follow-up. A positive Hedges' *g* means that the group improved from baseline to follow-up on that outcome domain.

The effect size for the change in the *behaviourally based intervention group* represents the overall effect of the intervention, which includes the specific effect of intervention components as well as non-specific factors such as repeated measures and expectations ('placebo effect'). Conversely, the effect size for the change in the *comparison group* represents only the non-specific factors, such as those associated with treatment as usual in the community. Thus, if the effect size within the

intervention group is larger than that of the comparison group, the intervention offers a benefit beyond what would be expected from treatment as usual.

As was described for the dose response analyses (*1. Relationship between dose and efficacy*, above), linear and non-linear models were then used to explore the relationship between dose (total and monthly clinician-delivered hours) and change from baseline to follow-up (Hedges' *g*). These relationships were explored for each outcome domain (as outlined in **Section 4.1.1**), for both the intervention and comparison groups separately. Dose in the comparison groups was again recorded as clinician-delivered hours of intervention, and may include services such as occupational therapy, speech therapy, etc. This analysis allows an assessment of the difference in effect of behaviourally based intervention and treatment as usual, both with the same clinician-delivered hours.

3. What did we find?

The following section reports the key findings from the analysis. A more detailed description of results is available in **Appendix 2**.

3.1 Summary of studies

Overall, 98 studies were included, representing a total of 4,553 participants. These studies were conducted across 21 countries, predominantly the US (45 studies), followed by the UK (7 studies), Norway (6 studies), Australia (4 studies), Canada (4 studies), and Italy (4 studies), among others. Half of included studies were randomised controlled trials (50%), with the remainder non-randomly allocating participants to intervention or comparison groups (e.g., by caregiver preference or using existing groups), or using existing cohorts of participants. The age of study participants ranged from 9 months to 7.1 years with an overall mean of 3.8 years, and 84% of study participants were male. Further characteristics of the included studies are shown in **Appendix 2**.

Across the 98 studies, 1,560 outcome measures which met the criteria outlined in **Section 4.1.1** were reported, with an average of 16 outcome measures reported per study. The number of studies that reported outcomes within each domain varied: 81 studies reported autism characteristic outcomes; 47 reported adaptive functioning outcomes; 64 reported cognition and language outcomes; 20 reported family outcomes; and 27 reported adverse effect outcomes.

3.2 Characteristics of behavioural interventions

Interventions vary on several characteristics (subgroups) which include content, person delivering the intervention, and primary setting (see **Section 4.1.3**). The differences these characteristics have on outcomes following a behaviourally based intervention is explored in the following sections. To summarise the distribution of identified studies across these characteristics, **Table 1** shows the number of studies which fall under each category by outcome domain as well as the average dose of intervention across subgroup categories.

Table 1a. Number of studies and dose amount by person delivering intervention.

Notes: Number of studies is reported in total and across all five outcomes domains by subgroup level. Total and monthly clinician-delivered are across all studies which report that data within each subgroup level. NA = not applicable.

Subgroup level	Total	Autism characteristics	Adaptive functioning	Cognition and language	Family outcomes	Adverse effects	Total clinician-delivered hours: Median (Range)	Monthly clinician-delivered hours: Median (Range)
Clinician	**31**	25	14	20	2	3	96 (10 – 3616)	26 (3 – 152)
Clinician and parent	**25**	18	17	20	2	7	178 (7 – 5088)	32 (2 – 158)
Parent	**33**	31	13	18	16	16	NA	NA
Teacher	**10**	8	4	7	0	1	NA	NA

Table 1b. Number of studies and dose amount by intervention category.

Notes: Number of studies is reported in total and across all five outcomes domains by subgroup level. Total and monthly clinician-delivered are across all studies which report that data within each subgroup level. NA = not applicable.

Subgroup level	Total	Autism characteristics	Adaptive functioning	Cognition and language	Family outcomes	Adverse effects	Total clinician-delivered hours: Median (Range)	Monthly clinician-delivered hours: Median (Range)
Behavioural	**44**	30	27	28	11	12	1328 (10 – 5088)	103 (4 – 157)
Naturalistic Developmental Behavioural Interventions (NDBI)	**33**	31	11	23	6	9	156 (7 – 1912)	30 (2 – 80)
Developmental	**13**	13	5	7	2	2	20 (10 – 1040)	11 (3 – 43)
Technology-based	**3**	2	2	3	0	1	18 (18 – 18)	4 (4 – 4)
TEACCH	**4**	4	3	3	1	1	18 (18 – 18)	7 (7 – 7)
Other	**3**	3	1	2	0	2	NA	NA

Table 1c. Number of studies and dose amount by comparison group.

Notes: Number of studies is reported in total and across all five outcomes domains by subgroup level. Total and monthly clinician-delivered are across all studies which report that data within each subgroup level. NA = not applicable.

Subgroup level	Total	Autism characteristics	Adaptive functioning	Cognition and language	Family outcomes	Adverse effects	Total clinician-delivered hours: Median (Range)	Monthly clinician-delivered hours: Median (Range)
Treatment as usual	**78**	69	33	46	1	4	156 (7 – 5088)	31 (2 – 158)
Eclectic	**23**	15	13	16	18	22	172 (10 – 1820)	29 (4 – 158)

Table 1d. Number of studies and dose amount by age group.

Notes: Number of studies is reported in total and across all five outcomes domains by subgroup level. Total and monthly clinician-delivered are across all studies which report that data within each subgroup level. NA = not applicable.

Subgroup level	Total	Autism characteristics	Adaptive functioning	Cognition and language	Family outcomes	Adverse effects	Total clinician-delivered hours: Median (Range)	Monthly clinician-delivered hours: Median (Range)
0-1 years	**7**	7	3	6	2	2	876 (52 – 1912)	66 (4 – 80)
2-4 years	**73**	61	36	49	14	19	176 (10 – 5088)	39 (4 – 158)
5-6 years	**16**	13	7	7	4	6	12 (7 – 18)	4 (2 – 7)

Table 1e. Number of studies and dose amount by primary intervention setting.

Notes: Number of studies is reported in total and across all five outcomes domains by subgroup level. Total and monthly clinician-delivered are across all studies which report that data within each subgroup level. NA = not applicable.

Subgroup level	Total	Autism characteristics	Adaptive functioning	Cognition and language	Family outcomes	Adverse effects	Total clinician-delivered hours: Median (Range)	Monthly clinician-delivered hours: Median (Range)
Health	**38**	35	12	21	11	13	32 (10 – 1820)	17 (3 – 152)
Community	**10**	9	7	6	1	1	792 (10 – 5088)	66 (5 – 158)
Early education	**26**	17	16	20	0	3	NA	NA
Home	**28**	25	14	18	9	11	692 (7 – 2662)	48 (2 – 116)

Table 1f. Number of studies and dose amount by study design.

Notes: Number of studies is reported in total and across all five outcomes domains by subgroup level. Total and monthly clinician-delivered are across all studies which report that data within each subgroup level. NA = not applicable.

Subgroup level	Total	Autism characteristics	Adaptive functioning	Cognition and language	Family outcomes	Adverse effects	Total clinician-delivered hours: Median (Range)	Monthly clinician-delivered hours: Median (Range)
Randomised controlled trial	**49**	45	16	28	13	13	52 (10 – 1912)	13 (3 – 80)
Non-randomised controlled trial	**31**	23	16	20	6	12	156 (7 – 5088)	26 (2 – 158)
Prospective cohort study	**18**	13	15	16	1	2	1330 (80 – 3616)	94 (13 – 152)

3.2.1 Clinician-delivered dose of intervention

A total of 34 studies reported clinician-delivered dose. Due to limited evidence for family outcomes and adverse effects, dose-response analyses were not included for these outcome domains.

Total clinician-delivered dose (**Figure 1**) varied greatly across studies, ranging from 7 to 5088 hours (median = 137 hours). Monthly clinician-delivered hours (**Figure 2**) also had a wide range, from 2 to 158 hours per month (median = 27 hours).

Both total ($r = 0.8$, $p < 0.001$) and monthly ($r = 0.8$, $p < 0.001$) clinician-delivered hours were significantly associated with total duration (weeks) of the intervention, meaning that longer intervention durations typically meant more total and monthly clinician-delivered hours of intervention. Because of this association, the unique effect of duration is unlikely to be substantial and duration was not explored further within the analyses relating to dose.

The clinician-delivered dose of an intervention varied across subgroups. Average clinician-delivered dose and range by subgroup is shown in **Table 1**.

Figure 1. Total clinician-delivered hours of intervention within included studies

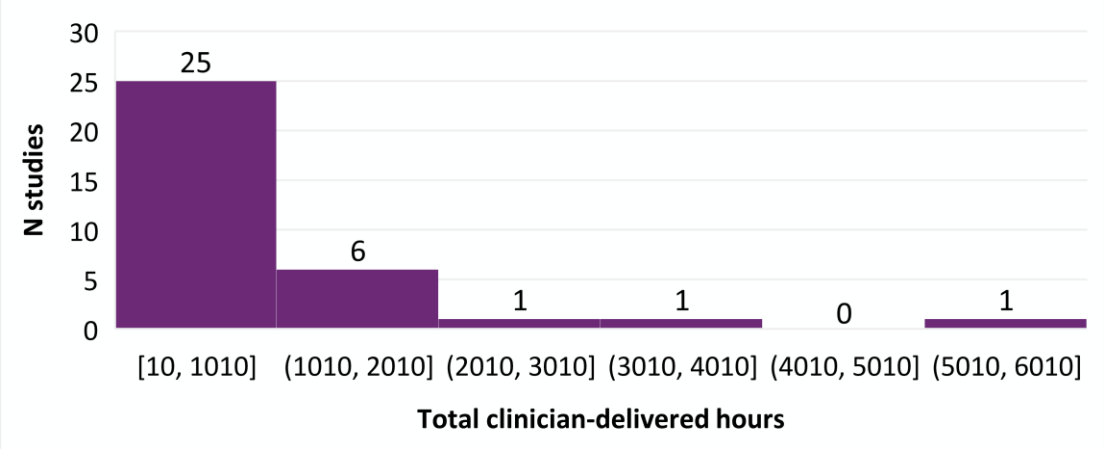
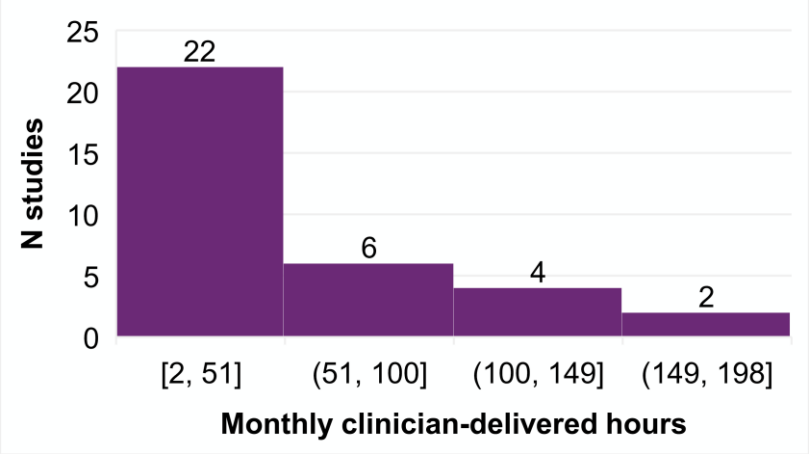


Figure 2. Monthly clinician-delivered hours of intervention within included studies



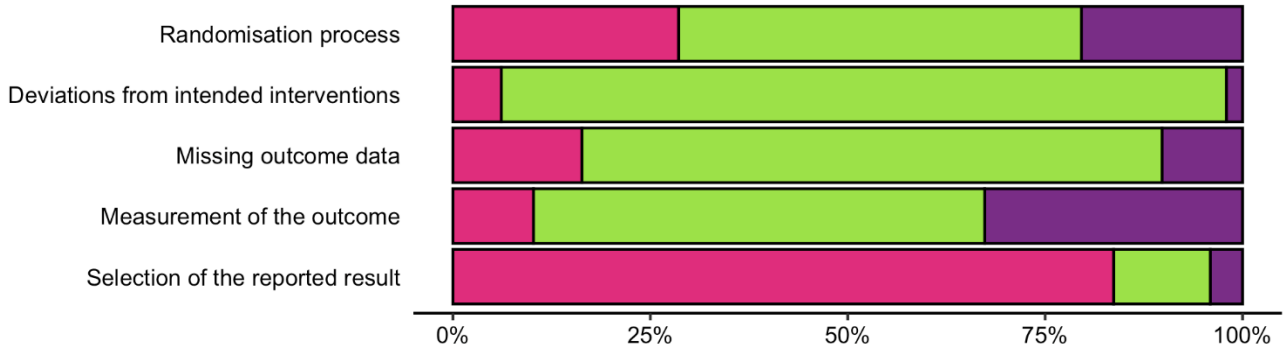
3.3 Quality of the evidence used within this report

3.3.1 Randomised controlled trials

Risk of bias was evaluated for the 49 randomised controlled trials included in this report. Of the overall assessments, 28 were deemed to have a high risk of bias, 20 had some concerns, and one was deemed low risk. Assessments of risk of bias within individual bias domains are summarised in **Figure 3** (see **Table B2** for individual domain assessments by study). The high proportion of unclear risk in the selection of the reported result was due to few studies providing a pre-specified analysis plan. A considerable number of studies (>25%) had high risk of bias in the measurement of outcomes due to the lack of blinding for outcome assessors. Approximately 25% of studies did not specify the use of a randomisation process which was concealed prior to enrollment and assignment to intervention, resulting in a high risk of bias for randomisation process. Relatively low risk of bias due to deviations from intended interventions and missing outcome data were identified.

Figure 3. Risk of bias across outcome domains in randomised controlled trials.

Note: Purple indicates high risk assessment. Green indicates low risk assessment. Pink indicates risk assessment of some concerns.

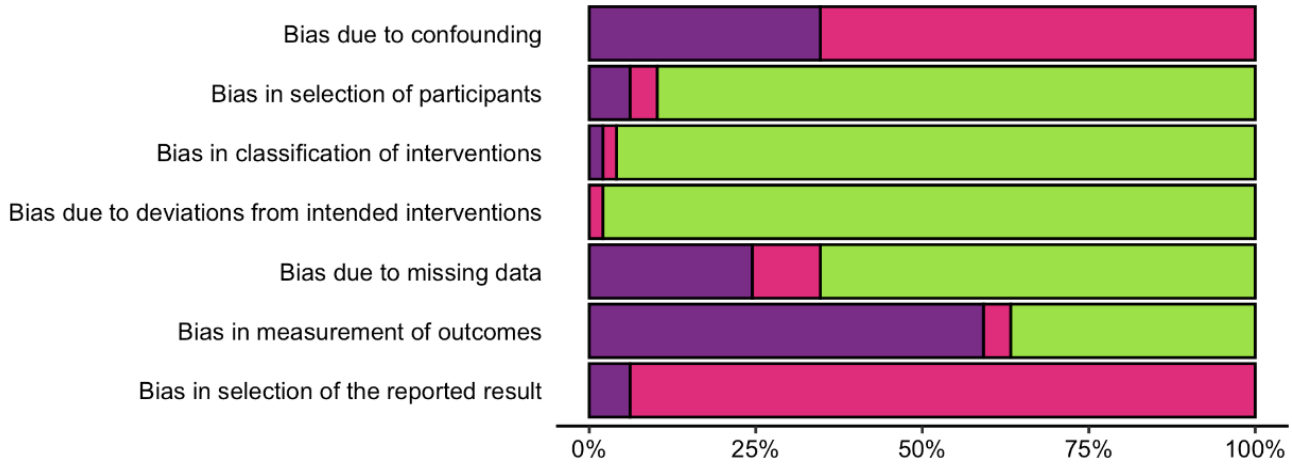


3.3.2 Non-randomised study designs

Risk of bias was assessed separately for the 49 studies which employed non-randomised study designs. Overall assessments deemed 38 to have serious risk of bias, and 11 to have a moderate risk of bias. Assessments of risk of bias within individual bias domains are summarised in **Figure 4** (see **Table B3** for individual domain assessments by study). Moderate or serious risk of bias due to confounding were seen across studies because of a lack of measurement or control of important cofounders (e.g., age, autism severity, IQ) within analyses. Moderate or serious risk of bias due to missing data were seen in >25% of studies due to the presence of substantial missing data (>10%) which was often unequal between groups. Greater than 50% of studies showed evidence of bias in the measurement of outcomes because assessors were unblinded. Predominantly moderate risk of bias in the selection of the reported result was shown due to the general absence of pre-specified analysis plans.

Figure 4. Risk of bias across outcome domains in non-randomised study designs.

Note: Purple indicates serious risk assessment. Green indicates low risk assessment. Pink indicates moderate risk assessment.



3.4 Overall results across outcome measures

Overall results across measures: a summary of findings

When considering the combined evidence across **all outcome measures** there was a **small benefit** of behaviourally based interventions when compared to treatment as usual, waitlist, or non-behaviourally based, "eclectic" interventions (the comparison groups). This small benefit was found, regardless of the dose intensity (**monthly** clinician-delivered hours) of the intervention.

A dose relationship was found, with increased **total** clinician-delivered hours associated with better outcomes following behaviourally based interventions than the comparative group.

- However, the associated real-world impact is arguably inconsequential. Every additional 100 hours of behaviourally based intervention a child receives translates to less than a 1% increase in standard deviation for the effect estimate.

3.4.1 Efficacy

Overall, there is evidence to support behaviourally based intervention in children less than 7 years old on the autism spectrum. Benefit of behaviourally based interventions when compared with treatment as usual, waitlist or a non-behaviourally based, "eclectic" intervention was shown across reported outcomes within the 98 studies but with a small effect size (Hedges' g = 0.32, 95%CI 0.26 – 0.38, Prediction Interval = -0.33 – 0.98, $tau^2 = 0.11$).

There was some evidence that effect sizes were larger in smaller studies, which can be indicative of over-estimation of treatment effect for this outcome. After estimating the bias, the effect estimate was reduced from 0.32 (95%CI 0.24 to 0.39) to 0.23 (95%CI 0.17 to 0.29). This estimated effect size which accounts for this bias is reported in **Appendix 2**.

3.4.2 Relationship between dose and efficacy

Across the 34 studies the linear models show a statistically significant relationship between increasing **total** clinician-delivered hours of a behaviourally based intervention and better outcomes (see **Figure 5**). However, while the linear dose response trend for increasing **total** hours of intervention was statistically significant, the associated real-world impact is arguably inconsequential, because every increase of 100 total hours of intervention translates to less than a 1% increase in standard deviation for the effect estimate.

There was no significant relationship between **monthly** clinician-delivered hours of behaviourally based intervention and effect estimate (Model estimates can be found in **Table B10**). The non-linear modelling demonstrated similar results (see **Figure 5**), with only slightly improved model fit.

Figure 5a. Linear model of monthly clinician-delivered dose by effect size (Hedges' g) for all outcomes.

Note: Hedges' g > 0 = better outcomes in the intervention group compared to the comparison group. Hedges' g < 0 = better outcomes in the comparison group compared to the intervention group.

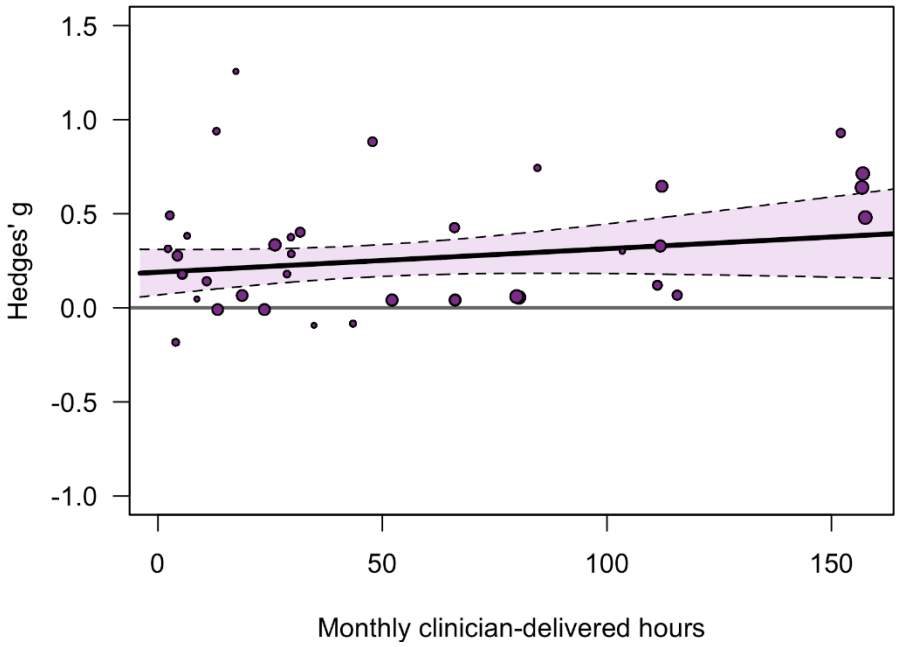


Figure 5b. Linear model of total clinician-delivered dose by effect size (Hedges' g) for all outcomes.

Note: Hedges' g > 0 = better outcomes in the intervention group compared to the comparison group. Hedges' g < 0 = better outcomes in the comparison group compared to the intervention group.

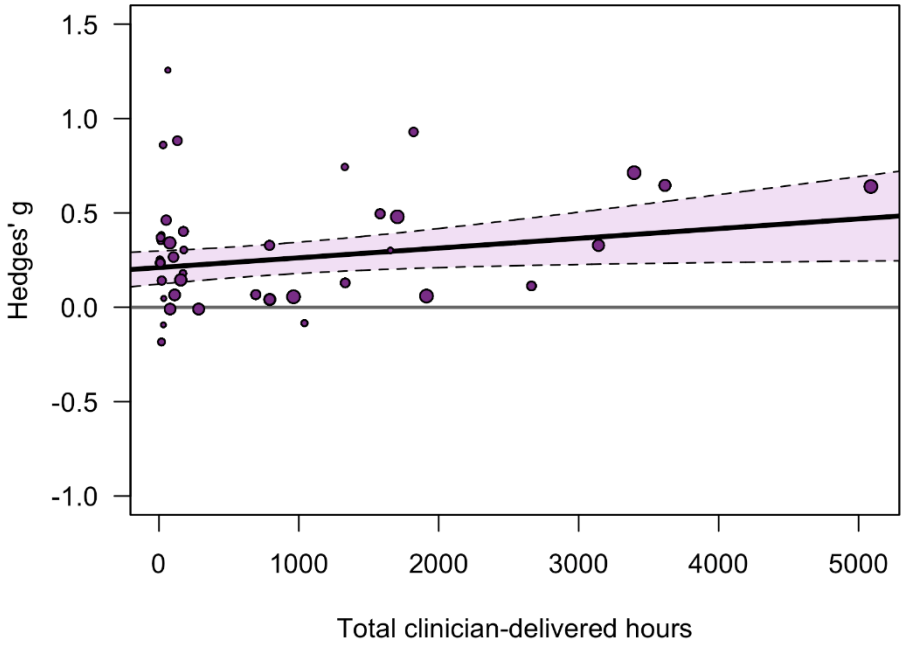


Figure 5c. Non-linear model of monthly clinician-delivered dose by effect size (Hedges' g) for all outcomes.

Note: Hedges' g > 0 = better outcomes in the intervention group compared to the comparison group. Hedges' g < 0 = better outcomes in the comparison group compared to the intervention group.

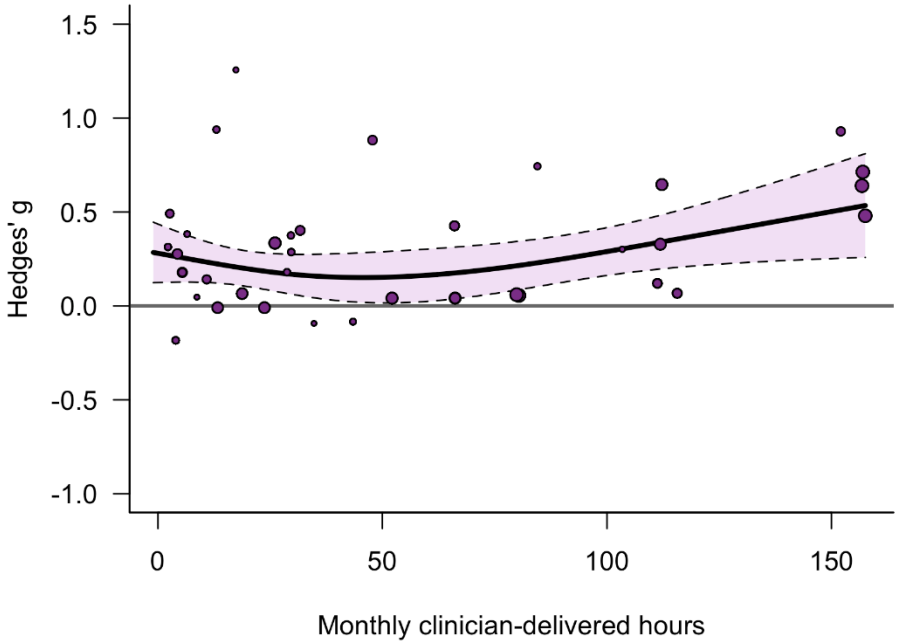
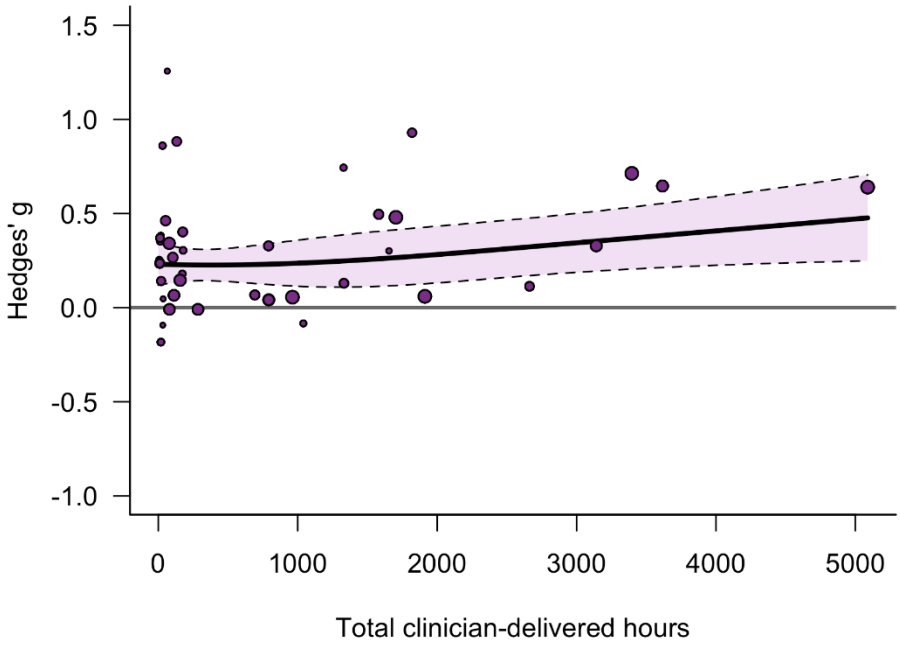


Figure 5d. Non-linear model of total clinician-delivered dose by effect size (Hedges' g) for all outcomes.

Note: Hedges' g > 0 = better outcomes in the intervention group compared to the comparison group. Hedges' g < 0 = better outcomes in the comparison group compared to the intervention group.



3.4.3 Comparing efficacy for lower versus higher total and monthly dose

The relationship between dose (both **total** and **monthly** hours of intervention) and efficacy was compared with lower and higher clinician-delivered hours (based on a median split of the data). No difference in efficacy was shown between lower and higher total and monthly doses of intervention, with small benefit shown at both dose levels (see **Figure 6** and **Table 2**).

Table 2. Results of analyses comparing lower and higher total and monthly doses for all outcome measures.

	Monthly clinician-delivered hours	Total clinician-delivered hours
Median hours	29.7 monthly hours	156 total hours
Lower dose: Less than median hours	**N studies:** 19 studies **Hedges' g (95% CI):** 0.27 (0.13-0.42)	**N studies:** 18 studies **Hedges' g (95% CI):** 0.30 (0.14-0.46)
Higher dose: Greater than median hours	**N studies:** 16 studies **Hedges' g (95% CI):** 0.29 (0.12-0.45)	**N studies:** 18 studies **Hedges' g (95% CI):** 0.27 (0.12-0.41)

Figure 6a. Linear dose relationship for lower versus higher monthly clinician hours (based on median) for all outcome measures.

Note: The green shaded area indicates less than the median number of monthly clinician-delivered hours, and the pink shaded area indicates higher than the median number of monthly clinician-delivered hours for this outcome.

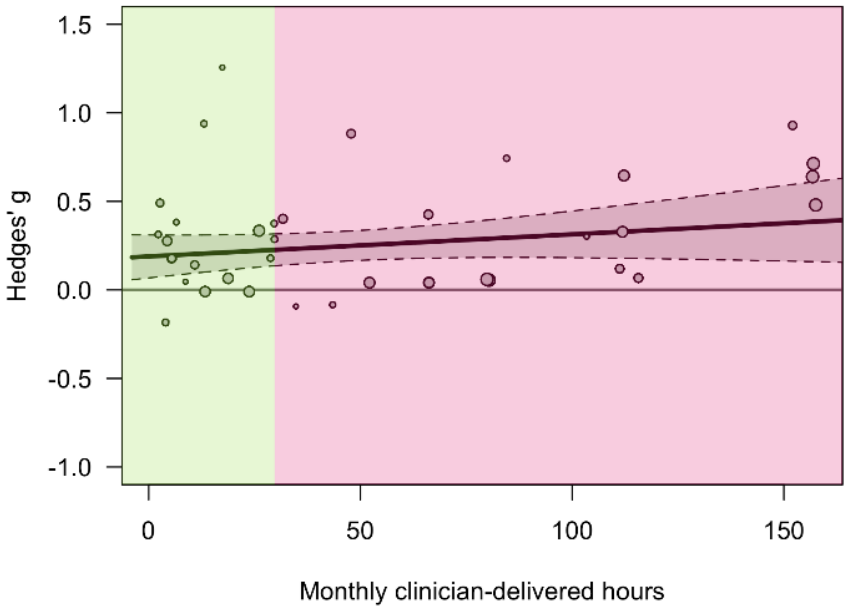
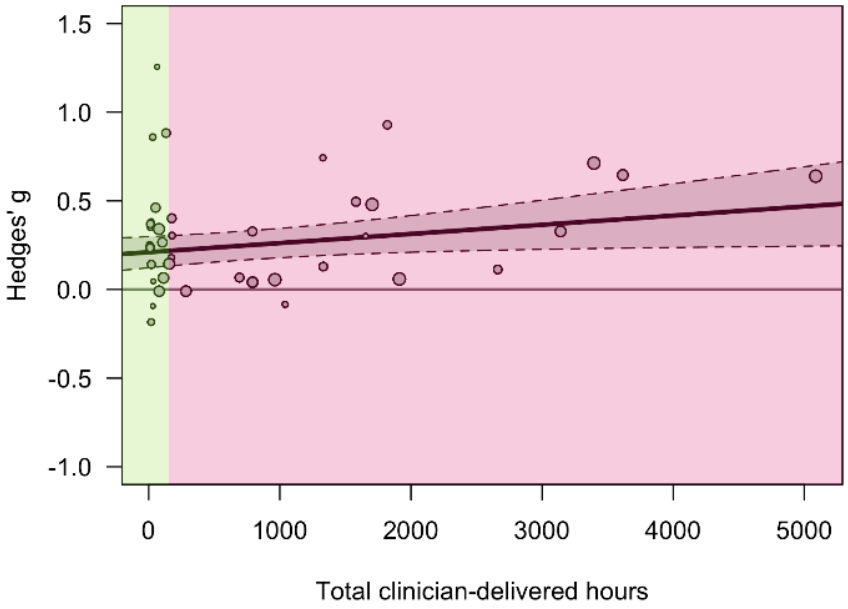


Figure 6b. Linear dose relationship for lower versus higher total clinician hours (based on median) for all outcome measures.

Note: The green shaded area indicates less than the median number of total clinician-delivered hours, and the pink shaded area indicates higher than the total number of monthly clinician-delivered hours for this outcome.



3.4.4 Relationship between dose and change from baseline to follow-up separately within the intervention group and the comparison group

When investigating change in outcomes over time from baseline to follow-up within the intervention group (regardless of the comparison group), linear models show a statistically significant relationship, indicating that increasing **total** and **monthly** clinician-delivered hours of an intervention translates to improved outcomes (see **Figure 7**). Model results and graphs of individual models with 95% confidence intervals can be found in **Table B11** and **Figure B14**. This means that, without controlling for the comparison group, increasing dose translates to better outcomes in the intervention group.

This relationship between dose (**total** and **monthly**) and change in outcomes over time from baseline to follow-up was not replicated in the comparison group (see **Table B12** for model results). Non-linear models can be seen in **Figure 7**, with models with 95% confidence intervals shown in **Figure B18**. This demonstrates that increasing clinician hours of treatment as usual intervention (most comparison groups involved treatment as usual or standard care in the community) does not translate to better outcomes.

Figure 7a. Linear model of monthly clinician-delivered dose by effect size (Hedges' g) for change in all outcomes from pre- to post-intervention within intervention and comparison groups.

Note: Hedges' g > 0 = improvement in outcomes from baseline to follow-up in the specified group. Hedges' g > 0 = decrease in outcomes from baseline to follow-up in the specified group.

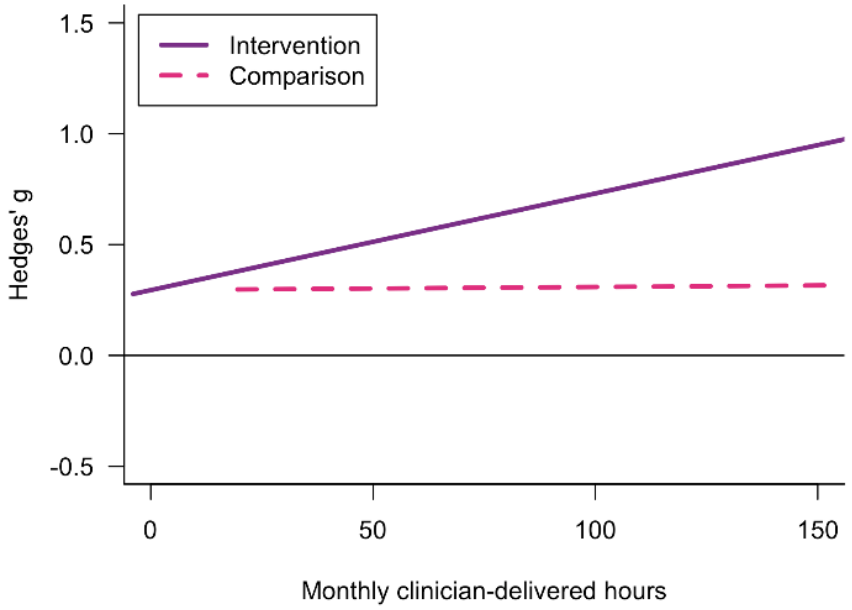


Figure 7b. Linear model of total clinician-delivered dose by effect size (Hedges' g) for change in all outcomes from pre- to post-intervention within intervention and comparison groups.

Note: Hedges' g > 0 = improvement in outcomes from baseline to follow-up in the specified group. Hedges' g > 0 = decrease in outcomes from baseline to follow-up in the specified group.

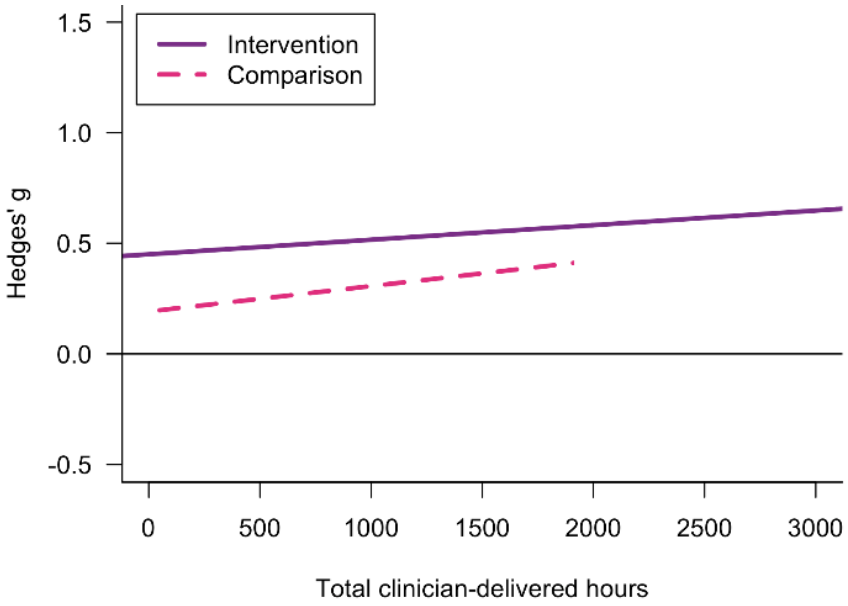


Figure 7c. Non-linear model of monthly clinician-delivered dose by effect size (Hedges' g) for change in all outcomes from pre- to post-intervention within intervention and comparison groups.

Note: Hedges' g > 0 = improvement in outcomes from baseline to follow-up in the specified group. Hedges' g > 0 = decrease in outcomes from baseline to follow-up in the specified group.

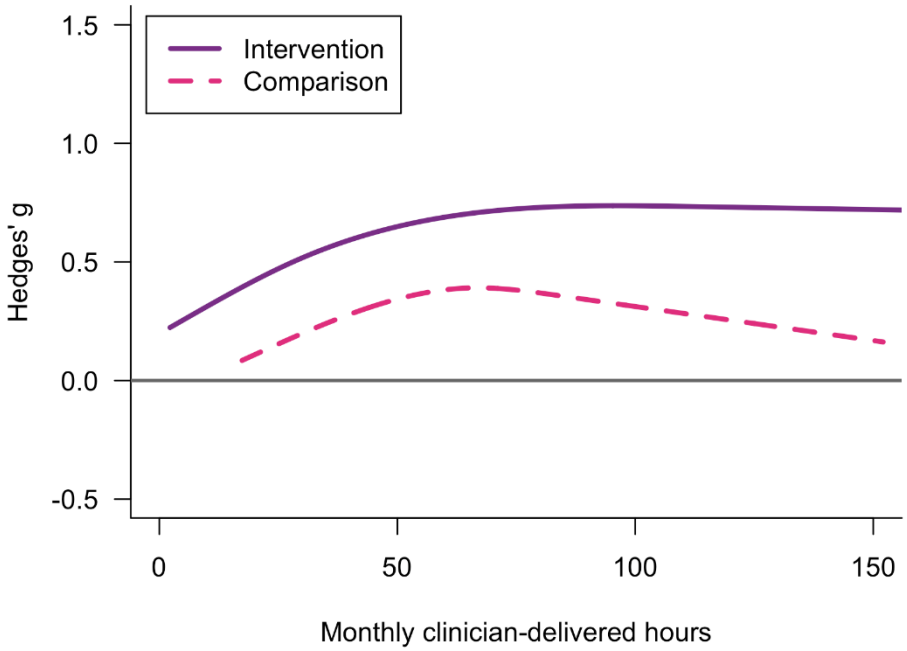
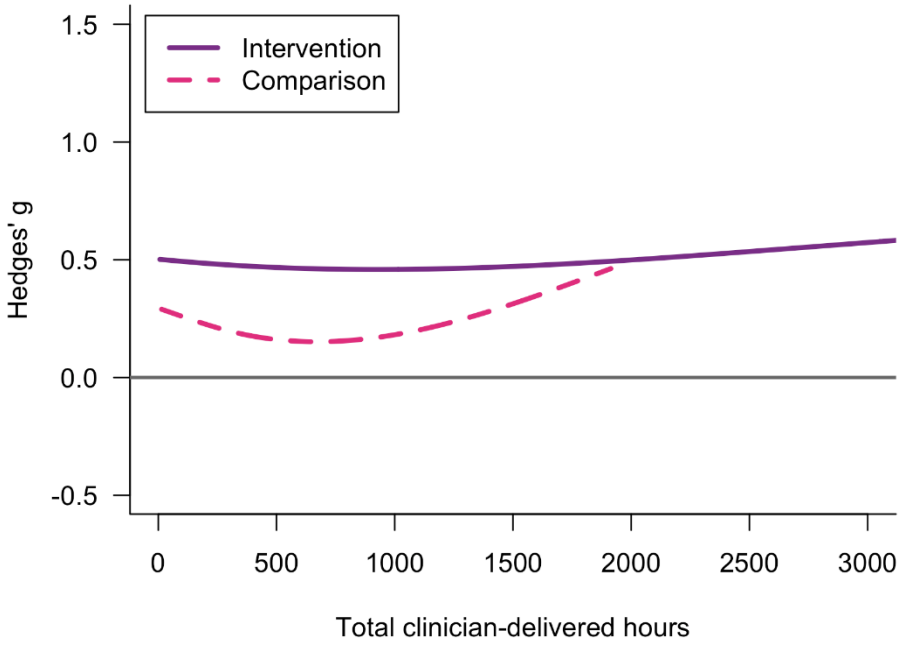


Figure 7d. Non-linear model of total clinician-delivered dose by effect size (Hedges' g) for change in all outcomes from pre- to post-intervention within intervention and comparison groups.

Note: Hedges' g > 0 = improvement in outcomes from baseline to follow-up in the specified group. Hedges' g > 0 = decrease in outcomes from baseline to follow-up in the specified group.



3.5 Efficacy within individual outcome domains

Efficacy within individual outcome domains: summary of findings

When the efficacy of the five outcome domains were investigated separately, evidence showed:

- There was a **small benefit** of behaviourally based interventions when compared to treatment as usual, waitlist, or non-behaviourally based, "eclectic" interventions on measures of autism characteristics, adaptive functioning, cognition and language, family outcomes and reductions in adverse effects (i.e., parent stress).
- There were large variances in effect estimates between studies (i.e., heterogeneity). This heterogeneity may be related to differences among studies, such as population, intervention, and study design factors. Analyses of some of these factors are provided in **Sections 6.6** and **6.7** below.

The efficacy of behaviourally based interventions was investigated across five outcome domains, which represent typical autism characteristics and behaviours, including autism characteristics, adaptive functioning, and cognition and language, as well as family outcomes and reductions in adverse effects (i.e., child and parent stress). These outcome domains are detailed in **Section 4.1.1**.

Small effect sizes for outcomes in all five outcome domains were identified (see **Figure 8** for individual estimates). However, they represent better performance in the behaviourally based intervention group when compared with the comparative group across all outcomes.

There was some evidence that effect sizes for the autism characteristics outcome domain were larger in smaller studies, which can be indicative of over-estimation of treatment effect for this outcome. After estimating the bias, the effect estimate was reduced from 0.32 (95%CI 0.24 to 0.39) to 0.22 (95%CI 0.15 to 0.29). This estimated effect size which accounts for this bias is reported in **Appendix 2**.

While pooled effect estimates for all five domains are statistically significant (confidence intervals do not include 0), there is large variability in the prediction intervals (i.e., the range of true effects across studies) (see **Figure 8**). This large variability indicates that some behaviourally based interventions may have no effect on investigated outcomes.

Figure 8. Forest plot of pooled effect sizes of overall efficacy for specific domains

Note: An accessible version of the data displayed in this figure is presented in Table 3 below. The prediction interval indicates the range of true effects across studies. Tau^2 is a measure of statistical heterogeneity, which gives an estimation of the extent to which an effect estimate is inconsistent across studies.

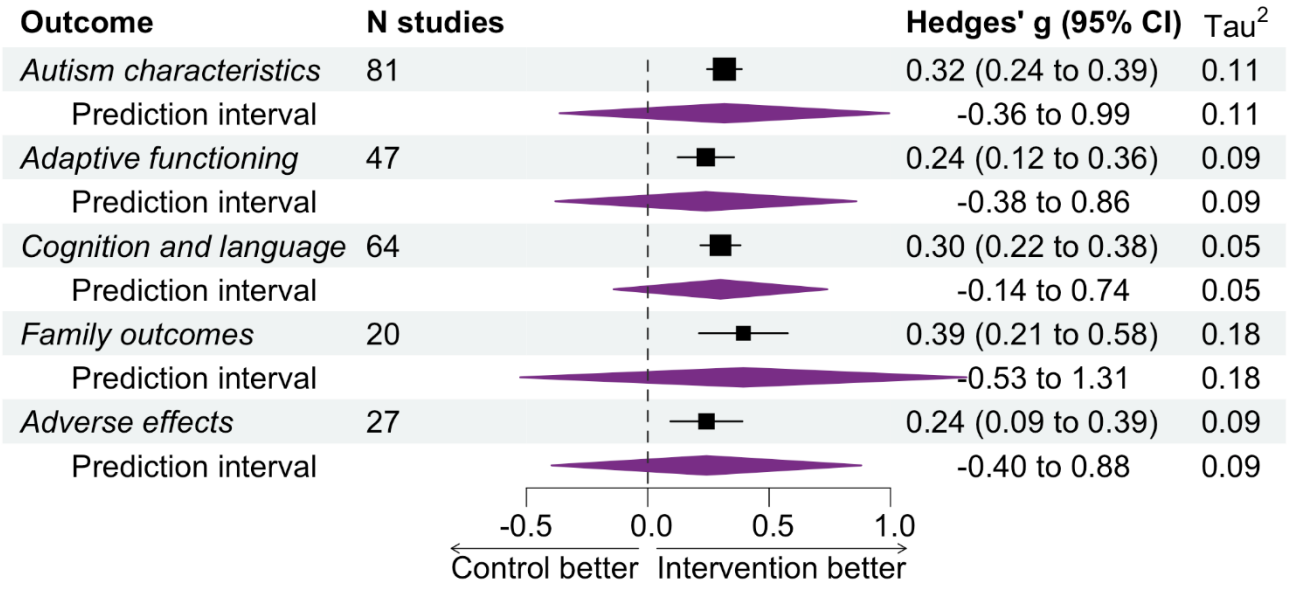


Table 3. Table version of forest plot of pooled effect sizes of overall efficacy for specific domains

Note: This table presents the information displayed in Figure 8 in an accessible format. The prediction interval indicates the range of true effects across studies. Tau^2 is a measure of statistical heterogeneity, which gives an estimation of the extent to which an effect estimate is inconsistent across studies.

Outcome	N studies	Hedges' g (96% CI)	Prediction interval	Tau^2
Autism characteristics	81	0.32 (0.24 to 0.39)	-0.36 to 0.99	0.11
Adaptive functioning	47	0.24 (0.12 to 0.36)	-0.38 to 0.86	0.09
Cognition and language	64	0.30 (0.22 to 0.38)	-0.14 to 0.74	0.05
Family outcomes	20	0.39 (0.21 to 0.58)	-0.53 to 1.31	0.18
Adverse effects	27	0.24 (0.09 to 0.39)	-0.40 to 0.88	0.09

It is important to explore this variability to find what factors are associated with best outcomes and what should be avoided. This was achieved by investigating changes in the effect across different factors (subgroups) and differences between these. In the following sections, the impact of dose, intervention (content, delivery, setting) and population (age) subgroups were investigated. These subgroups are detailed in **Section 4.1.3**.

3.6 Investigating the effect of dose

Investigating the effect of dose: summary of findings

Autism characteristics

No dose relationship was found between **total** or **monthly** clinician-delivered intervention hours and autism characteristic outcomes.

This means that there is no evidence for added benefit of increasing intervention hours or intensity for autism characteristic outcomes. Small benefits of behaviourally based interventions above that of the comparison group are consistently seen, regardless of the total or monthly clinician-delivered hours received.

Adaptive functioning

A dose relationship was found, with increased **total** and **monthly** clinician-delivered hours associated with better adaptive functioning outcomes following behaviourally based interventions.

The associated real-world impact of relatively smaller increases in intervention hours is minimal.

- Every increase in the total intervention dose by 100 hours translates to less than a 1% increase in standard deviation for the effect estimate.

- Every increase in the intervention dose intensity by 10 hours per month translates to an increase of only 3% of a standard deviation of the effect estimate.

Doses below approximately 65 hours per month or 800 total intervention hours may not produce the desired effect (inefficacious).

Increasing efficacy with increased total and monthly hours may be driven by decreasing outcomes with more clinician-delivered intervention hours in comparison group, rather than a specific benefit of behaviourally based interventions.

Cognition and language
A dose relationship was found, with increased **total** and **monthly** clinician-delivered hours associated with better cognition and language outcomes following behaviourally based interventions.

No evidence for a minimal required dose. Small, positive intervention effects are present even at lower total dose amounts and dose intensities, but slightly increasing with additional hours.

The associated real-world impact of relatively smaller increases in intervention hours is minimal.

- Every increase in the total intervention dose by 100 hours translates to less than a 1% increase in standard deviation for the effect estimate.
- Every increase in the intervention dose intensity by 10 hours per month translates to an increase of only 2% of a standard deviation of the effect estimate.

A relationship between increasing total clinician hours and improved outcomes was found for both the intervention and comparison groups, suggesting that the finding of better outcomes with increased hours is likely related to the amount of time spent with a clinician than the actual intervention taking place.

3.6.1 Autism characteristics

Relationship between dose and efficacy

The linear and non-linear models indicate no dose relationship between **total** or **monthly** clinician-delivered hours and efficacy of behaviourally based interventions for autism characteristics (see **Figure 9**). This shows that increasing the dose of an intervention (**total** or **monthly** clinician hours) does not change the efficacy of behaviourally based interventions for autism characteristics.

Linear model estimates are reported in **Table B10**.

Figure 9a. Linear model of monthly clinician-delivered dose by effect size (Hedges' g) for autism characteristics outcomes.

Note: Hedges' g > 0 = better outcomes in the intervention group compared to the comparison group. Hedges' g < 0 = better outcomes in the comparison group compared to the intervention group.

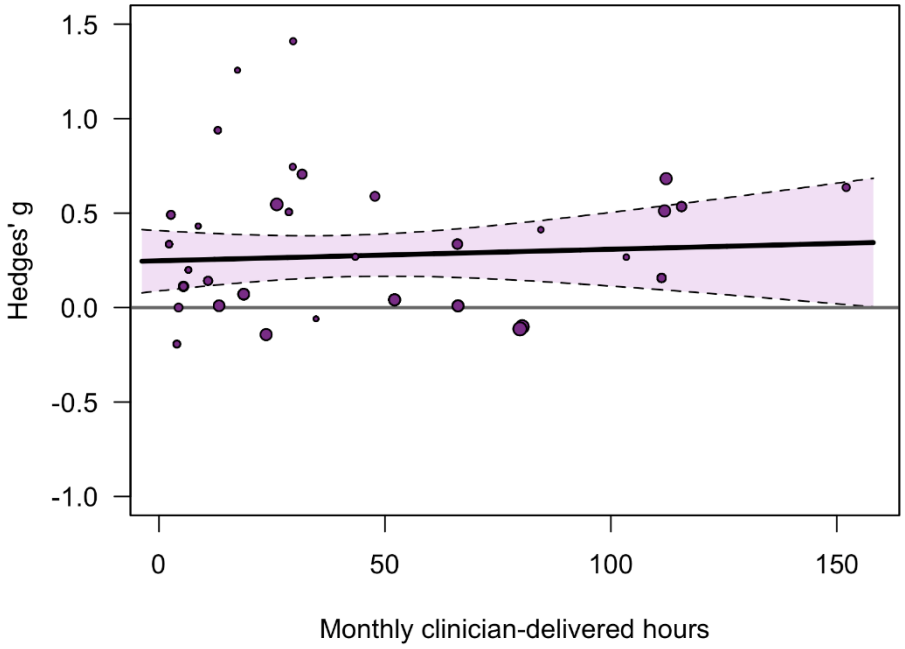


Figure 9b. Linear model of total clinician-delivered dose by effect size (Hedges' g) for autism characteristics outcomes.

Note: Hedges' g > 0 = better outcomes in the intervention group compared to the comparison group. Hedges' g < 0 = better outcomes in the comparison group compared to the intervention group.

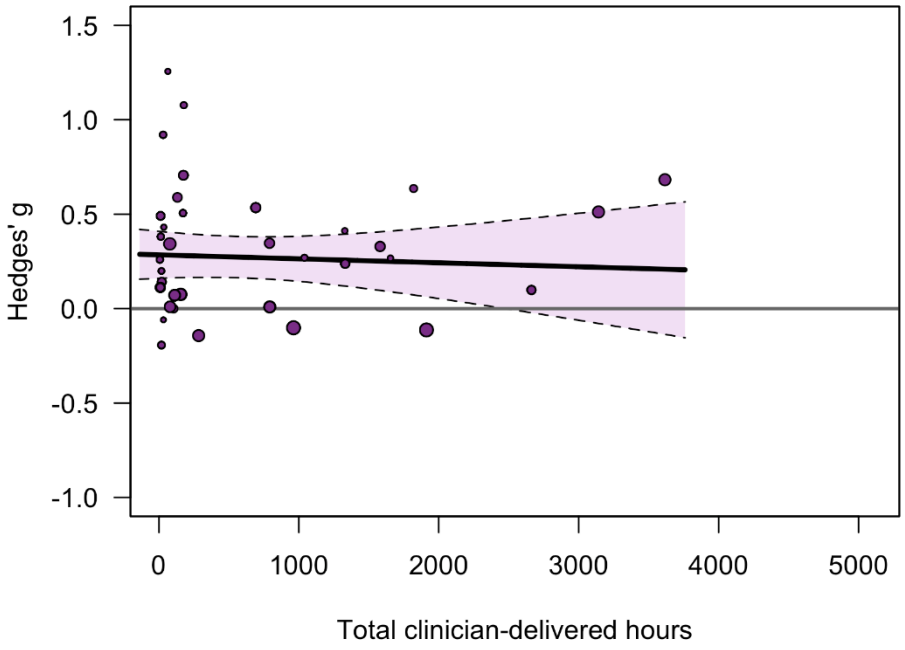


Figure 9c. Non-linear model of monthly clinician-delivered dose by effect size (Hedges' g) for autism characteristics outcomes.

Note: Hedges' g > 0 = better outcomes in the intervention group compared to the comparison group. Hedges' g < 0 = better outcomes in the comparison group compared to the intervention group.

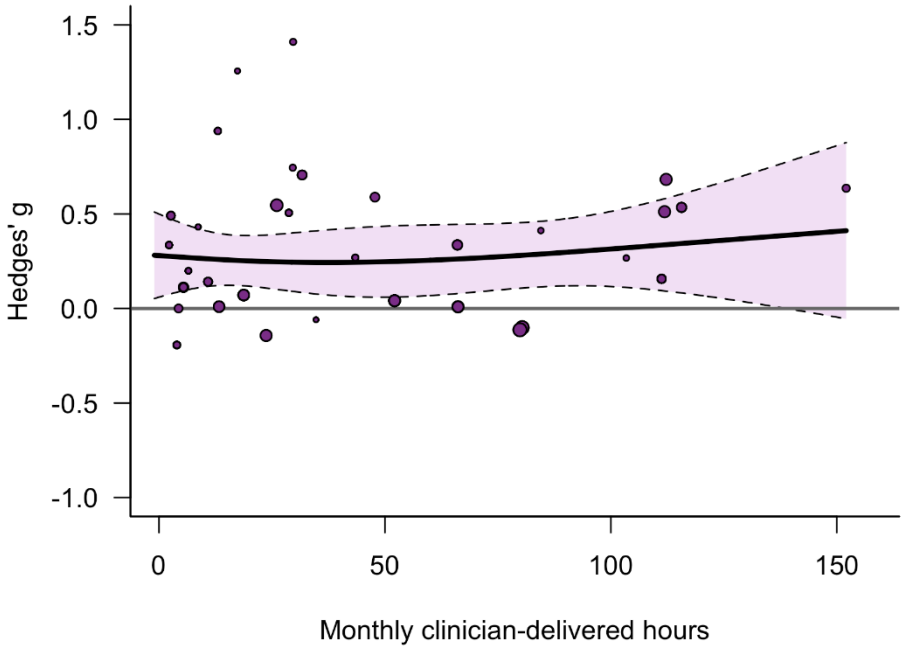
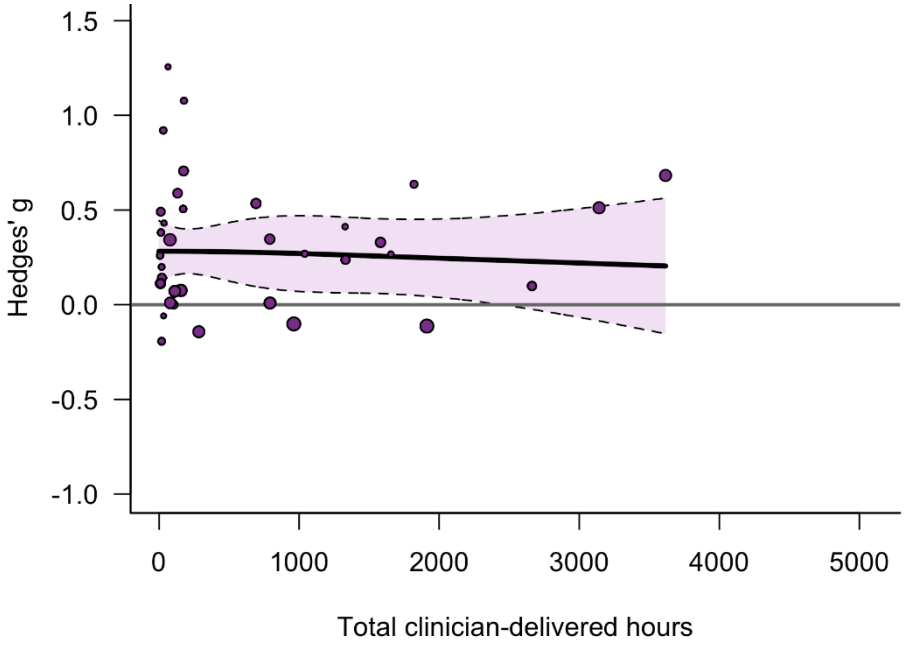


Figure 9d. Non-linear model of total clinician-delivered dose by effect size (Hedges' g) for autism characteristics outcomes.

Note: Hedges' g > 0 = better outcomes in the intervention group compared to the comparison group. Hedges' g < 0 = better outcomes in the comparison group compared to the intervention group.



Comparing efficacy for lower versus higher total and monthly dose

The results did not show a difference in efficacy between behaviourally based interventions delivered for less than the median dose versus more than the median dose, for both **total** and **monthly** clinician-delivered hours on autism characteristics (see **Figure 10** and **Table 4**). This confirms that efficacy of behaviourally based interventions on autism characteristics does not differ between lower and higher hours of intervention.

Table 4. Results of analyses comparing lower and higher total and monthly doses for autism characteristics.

	Monthly clinician-delivered hours	Total clinician-delivered hours
Median hours	29.66 monthly hours	156 total hours
Lower dose: Less than median hours	**N studies:** 156 total hours **Hedges' g (95% CI):** 0.28 (0.09-0.46)	**N studies:** 16 studies **Hedges' g (95% CI):** 0.27 (0.09-0.45)
Higher dose: Greater than median hours	**N studies:** 15 studies **Hedges' g (95% CI):** 0.30 (0.12-0.48)	**N studies:** 17 studies **Hedges' g (95% CI):** 0.30 (0.12-0.47)

Figure 10a. Linear dose relationship for lower versus higher monthly clinician hours (based on median) for autism characteristic outcomes.

Note: The green shaded area indicates less than the median number of monthly clinician-delivered hours, and the pink shaded area indicates higher than the median number of monthly clinician-delivered hours for this outcome.

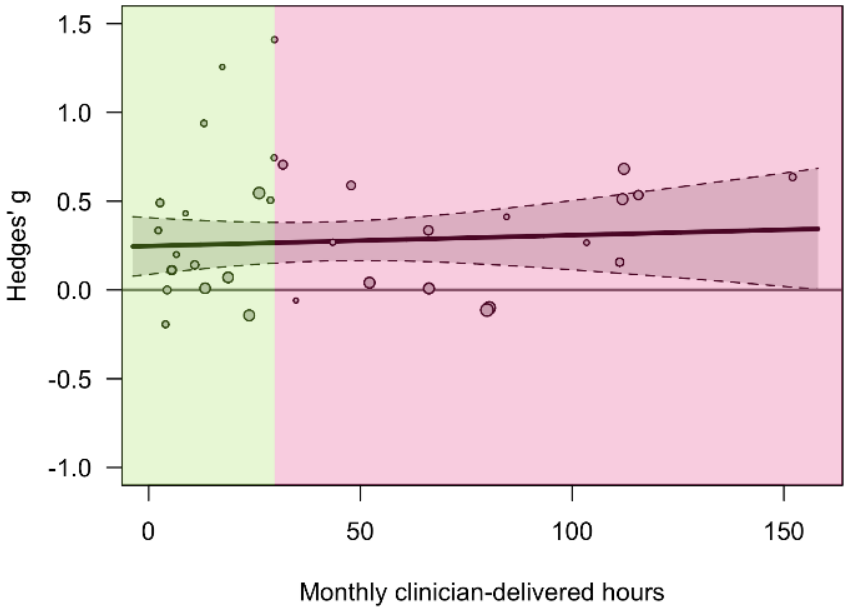
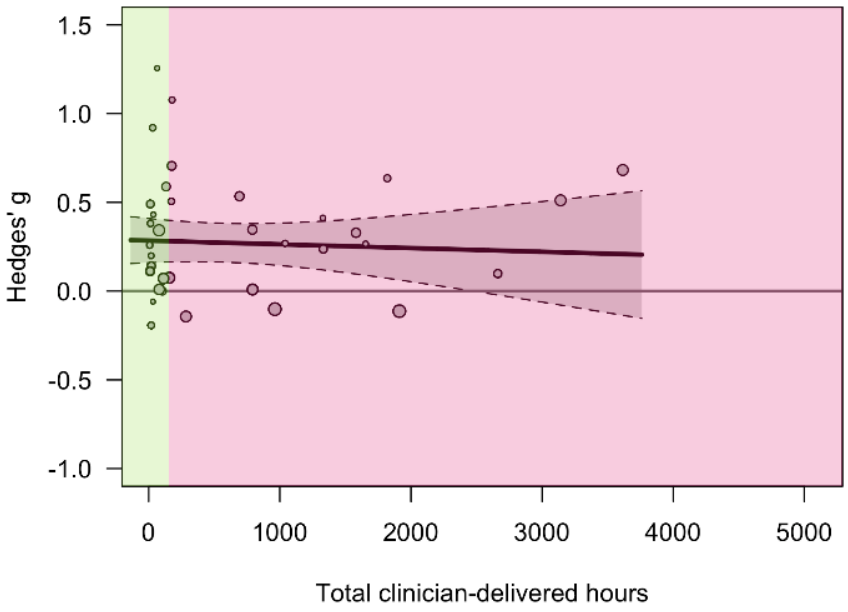


Figure 10b. Linear dose relationship for lower versus higher total clinician hours (based on median) for autism characteristic outcomes.

Note: The green shaded area indicates less than the median number of monthly clinician-delivered hours, and the pink shaded area indicates higher than the median number of monthly clinician-delivered hours for this outcome.



Relationship between dose and change from baseline to follow-up separately within the intervention group and the comparison group

The lack of relationship between dose and effect size was corroborated by analyses exploring the relationship between dose and change from baseline to follow-up separately in both the intervention and comparison groups (see **Figure 11**). Neither those who received behaviourally based interventions nor those who received some level of clinician contact hours in the comparative group (through treatment as usual, e.g., speech pathology) benefited from more intensive doses or higher total contact hours with a clinician (see **Tables B11 & B12** for model results).

The non-linear model suggested a gradual increase in effect with increasing monthly hours for the behaviourally based intervention group (**Figure 11**). However, confidence in this result is low due to a limited number of studies providing high intensity interventions. Individual linear and non-linear models with 95% confidence intervals are shown in **Figures B15 & B19**.

Figure 11a. Linear model of monthly clinician-delivered dose by effect size (Hedges' g) for change in autism characteristic outcomes from pre- to post-intervention within intervention and comparison groups.

Note: Hedges' g > 0 = improvement in outcomes from baseline to follow-up in the specified group. Hedges' g > 0 = decrease in outcomes from baseline to follow-up in the specified group.

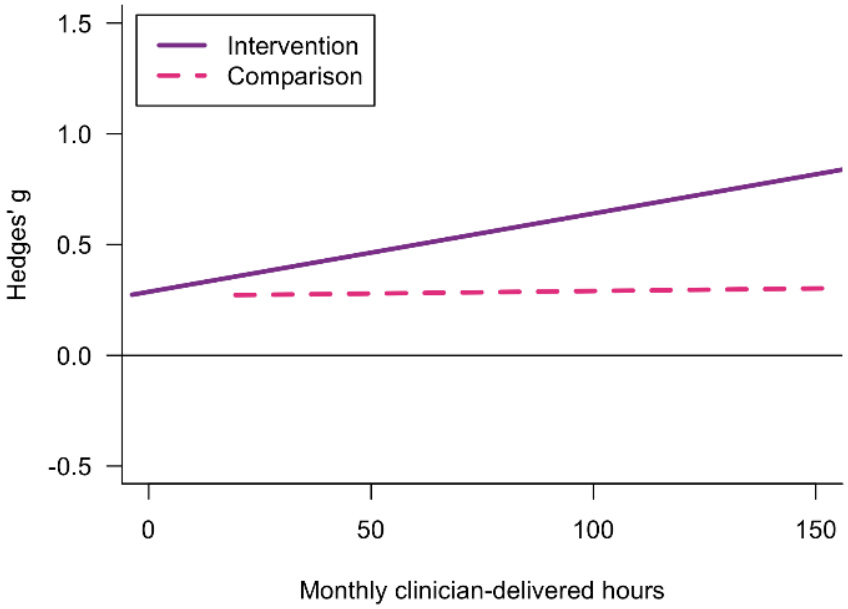


Figure 11b. Linear model of total clinician-delivered dose by effect size (Hedges' g) for change in autism characteristic outcomes from pre- to post-intervention within intervention and comparison groups.

Note: Hedges' g > 0 = improvement in outcomes from baseline to follow-up in the specified group. Hedges' g > 0 = decrease in outcomes from baseline to follow-up in the specified group.

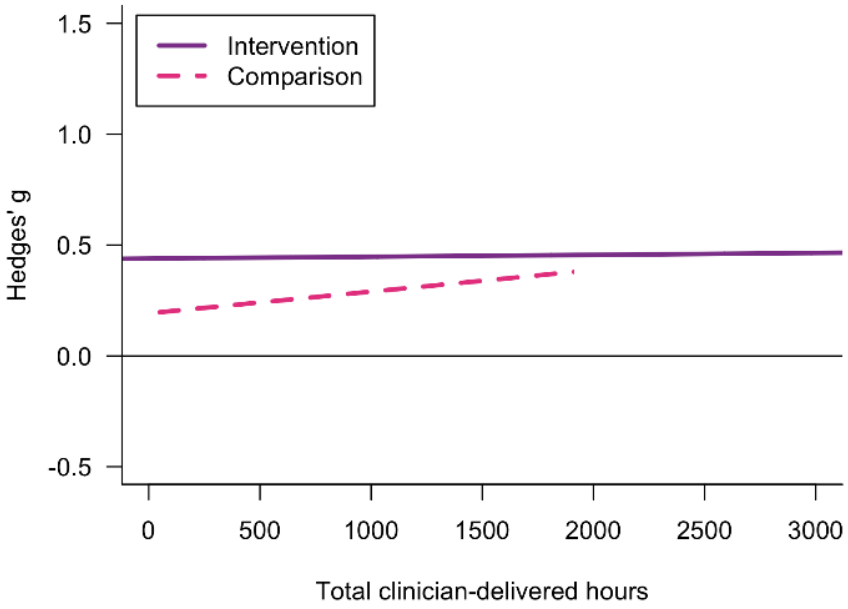


Figure 11c. Non-linear model of monthly clinician-delivered dose by effect size (Hedges' g) for change in autism characteristic outcomes from pre- to post-intervention within intervention and comparison groups.

Note: Hedges' g > 0 = improvement in outcomes from baseline to follow-up in the specified group. Hedges' g > 0 = decrease in outcomes from baseline to follow-up in the specified group.

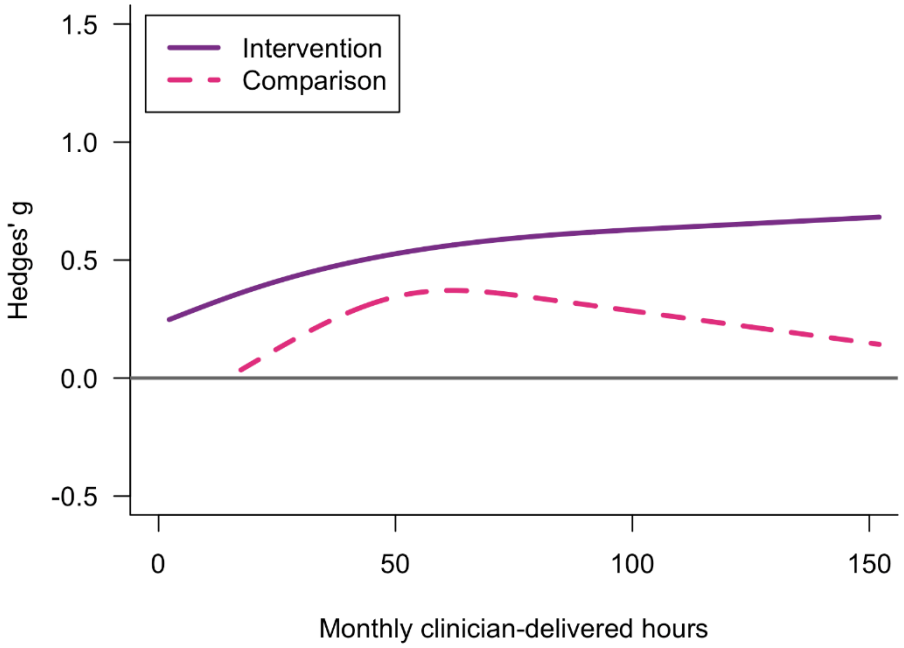
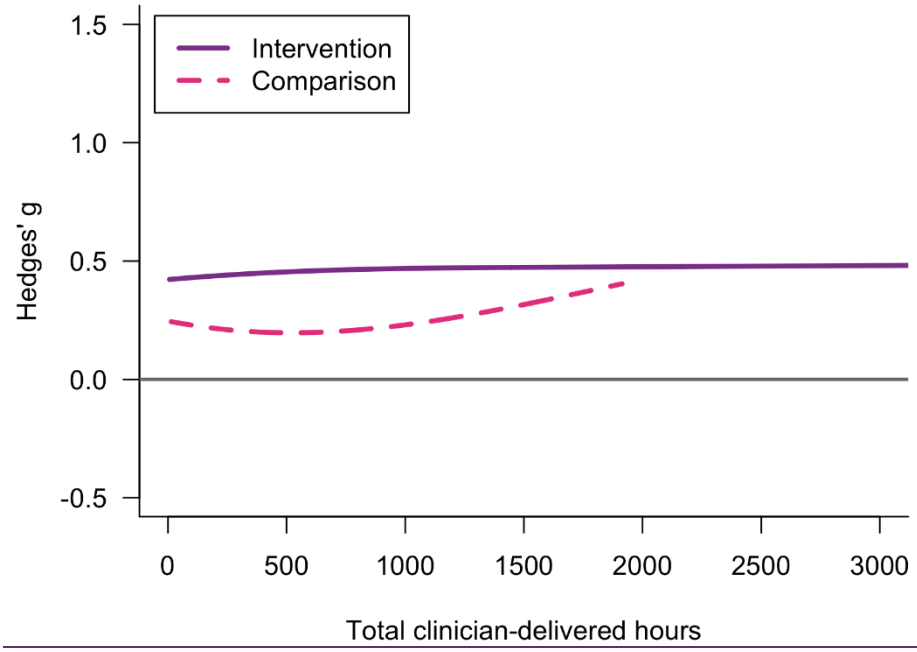


Figure 11d. Non-linear model of total clinician-delivered dose by effect size (Hedges' g) for change in autism characteristic outcomes from pre- to post-intervention within intervention and comparison groups.

Note: Hedges' g > 0 = improvement in outcomes from baseline to follow-up in the specified group. Hedges' g > 0 = decrease in outcomes from baseline to follow-up in the specified group.



3.6.2 Adaptive functioning

Relationship between dose and efficacy

The linear models show a statistically significant relationship between increasing **total** and **monthly** clinician-delivered hours of intervention and better adaptive functioning outcomes following behaviourally based intervention (see **Figure 12** and **Table B10** for model estimates). Non-linear modelling demonstrated similar results (see **Figure 12**), with only slightly improved model fit. A visual investigation of the non-linear model suggested that adaptive function may require large doses to achieve a clinically meaningful effect size, with negligible effects shown at lower doses.

Although the linear dose response trends were statistically significant, the associated real-world impact of relatively smaller increases in intervention hours is minimal. Every increase of the intervention dose by 10 hours per month translates to an increase of only 3% of a standard deviation of the effect estimate. Similarly, every increase of 100 total hours of intervention overall translates to less than a 1% increase in standard deviation for the effect estimate.

Figure 12a. Linear model of monthly clinician-delivered dose by effect size (Hedges' g) for adaptive functioning outcomes.

Note: Hedges' g > 0 = better outcomes in the intervention group compared to the comparison group. Hedges' g < 0 = better outcomes in the comparison group compared to the intervention group.

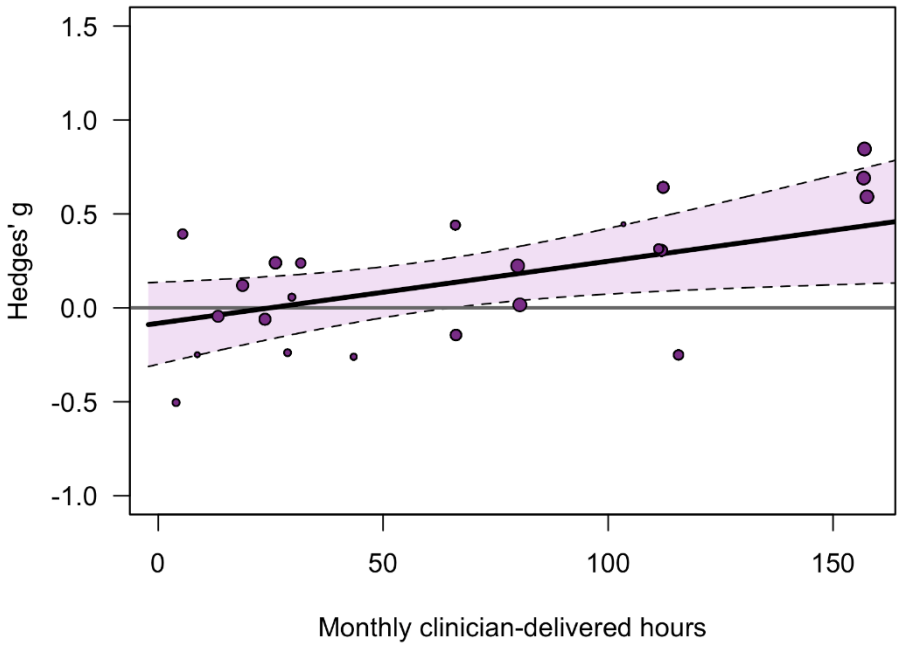


Figure 12b. Linear model of total clinician-delivered dose by effect size (Hedges' g) for adaptive functioning outcomes.

Note: Hedges' g > 0 = better outcomes in the intervention group compared to the comparison group. Hedges' g < 0 = better outcomes in the comparison group compared to the intervention group.

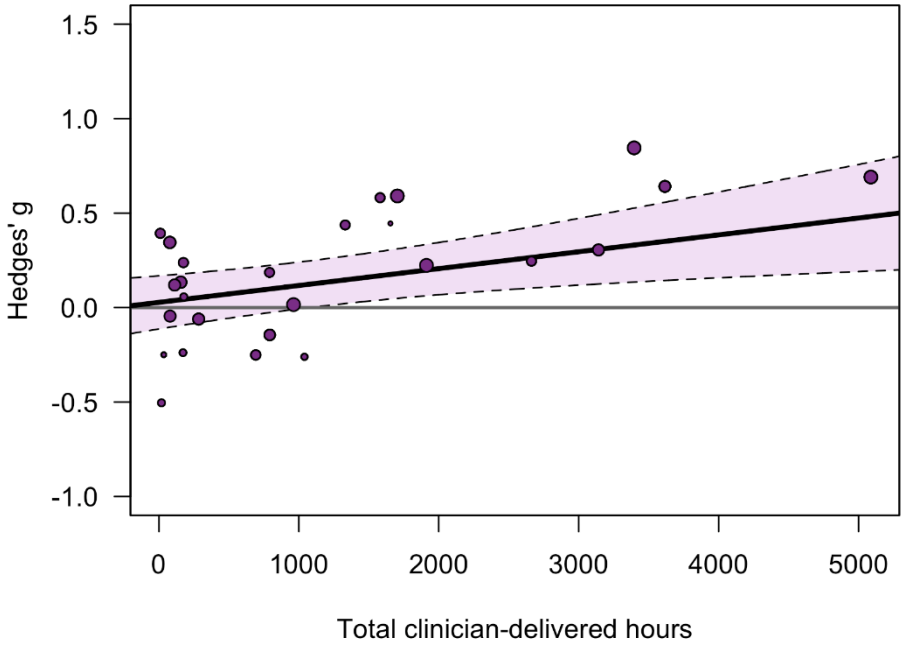


Figure 12c. Non-linear model of monthly clinician-delivered dose by effect size (Hedges' g) for adaptive functioning outcomes.

Note: Hedges' g > 0 = better outcomes in the intervention group compared to the comparison group. Hedges' g < 0 = better outcomes in the comparison group compared to the intervention group.

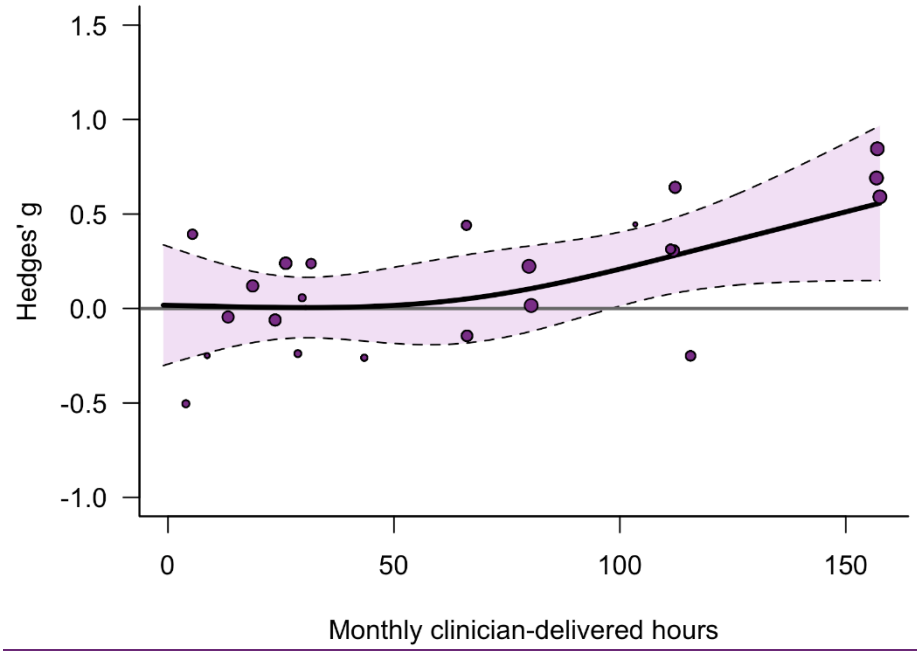
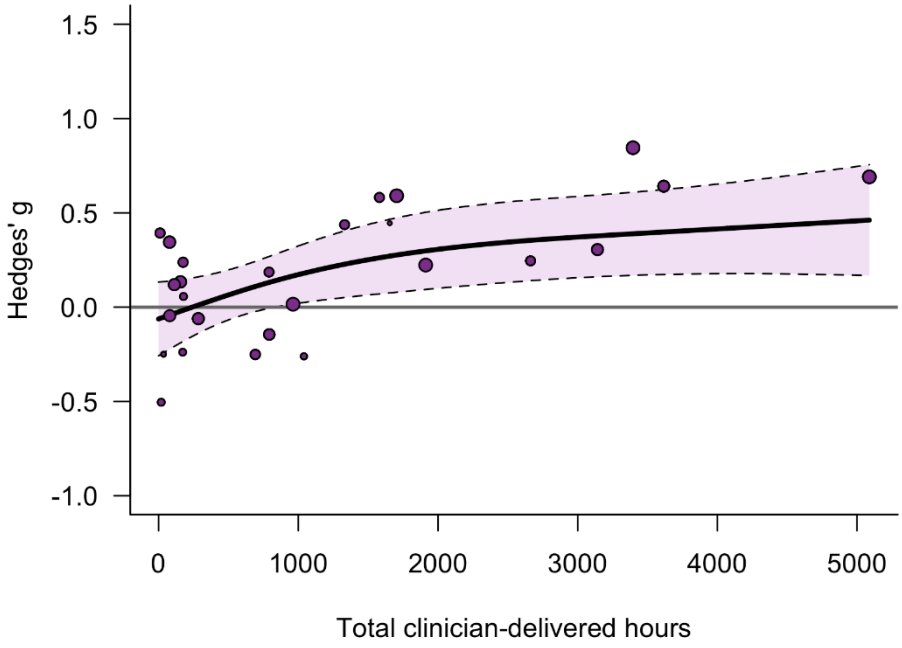


Figure 12d. Non-linear model of total clinician-delivered dose by effect size (Hedges' g) for adaptive functioning outcomes.

Note: Hedges' g > 0 = better outcomes in the intervention group compared to the comparison group. Hedges' g < 0 = better outcomes in the comparison group compared to the intervention group.



Comparing efficacy for lower versus higher total and monthly dose

A difference in effect sizes was found between studies that delivered less than the median dose and those that delivered doses higher than the median, both for total clinician-delivered hours as well as hours per month (see **Figure 13** for visual representation). Effect sizes were larger in studies that delivered doses at a higher versus lower dose, where a negligible effect size was found for the lower dose, and a small effect size for the higher dose. Practically, this suggests there are negligible effects below approximately 65 hours per month or 800 clinician hours overall.

However, there was not enough information (i.e., statistical power and precision) to detect a statistically significant difference between groups as only 10 and 8 studies, and 11 and 8 studies were available for analyses of both dose types. More detailed results can be found in **Table 5**.

Table 5. Results of analyses comparing lower and higher total and monthly doses for adaptive functioning.

	Monthly clinician-delivered hours	Total clinician-delivered hours
Median hours	66.04 monthly hours	791.20 total hours
Lower dose: Less than median hours	**N studies:** 10 studies **Hedges' g (95% CI):** 0.03 (-0.16-0.22)	**N studies:** 11 studies **Hedges' g (95% CI):** 0.03 (-0.15-0.21)
Higher dose: Greater than median hours	**N studies:** 8 studies **Hedges' g (95% CI):** 0.26 (-0.05-0.56)	**N studies:** 8 studies **Hedges' g (95% CI):** 0.30 (-0.03-0.62)

Figure 13a. Linear dose relationship for lower versus higher total and monthly clinician hours (based on median) for adaptive functioning outcomes.

Note: The green shaded area indicates less than the median number of monthly clinician-delivered hours, and the pink shaded area indicates higher than the median number of monthly clinician-delivered hours for this outcome.

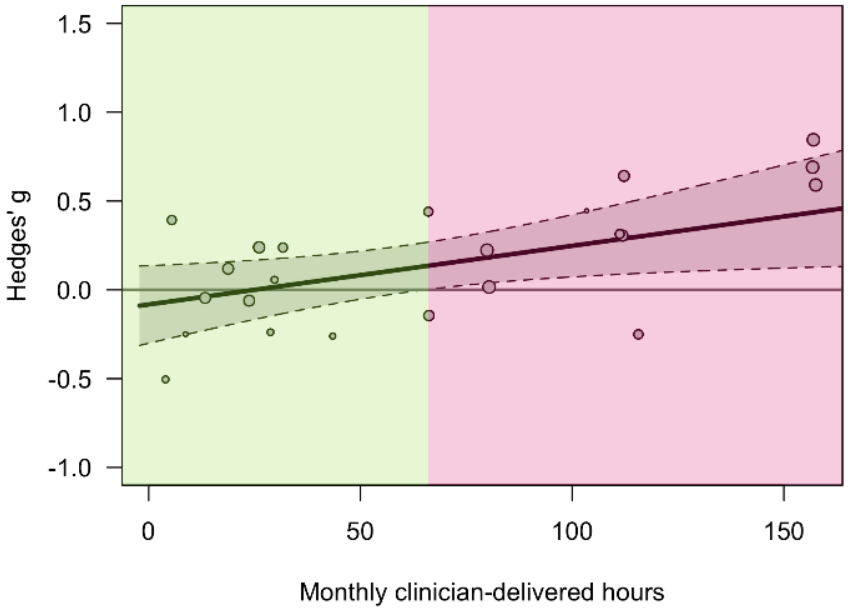
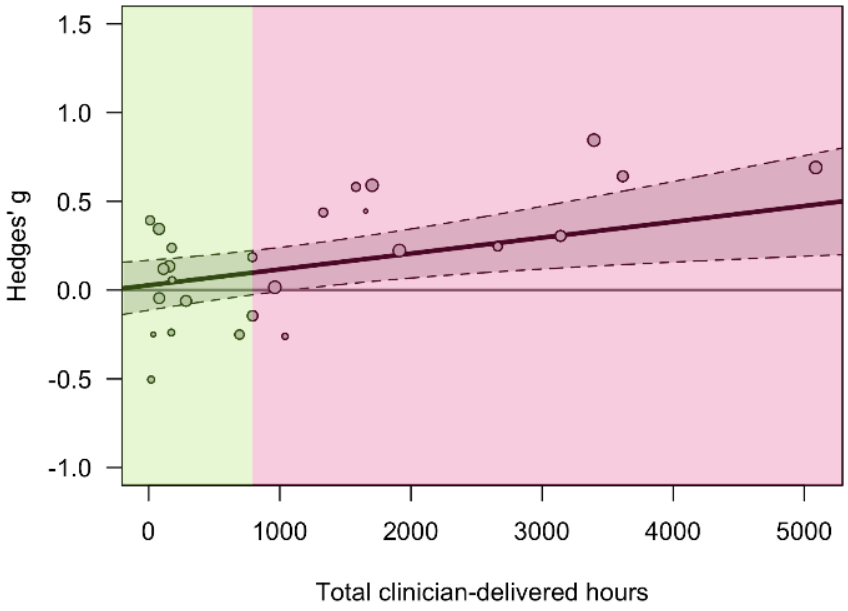


Figure 13b. Linear dose relationship for lower versus higher total and monthly clinician hours (based on median) for adaptive functioning outcomes.

Note: The green shaded area indicates less than the median number of monthly clinician-delivered hours, and the pink shaded area indicates higher than the median number of monthly clinician-delivered hours for this outcome.



Relationship between dose and change from baseline to follow-up separately within the intervention group and the comparison group

Analyses of the change from baseline to follow-up separately within the intervention and comparison groups provide some indication that the relationship between clinician time and effect size are specific to the intervention group. Linear models (see **Figure 14** and **Table B11**) revealed a statistically significant relationship (albeit small) between increasing **total** and **monthly** clinician-delivered hours of intervention and better adaptive functioning outcomes at follow-up for the behaviourally based intervention group, while no such relationship was observed between clinician hours and change in outcomes from baseline to follow-up within the comparison groups (see **Figure 14** and **Table B12**). Individual models with 95% confidence intervals are displayed in **Figures B16 & B20**.

Additionally, the non-linear models (**Figure 14**) indicate that improvements from baseline to follow-up within the behaviourally based intervention groups decrease after approximately 100 monthly hours. This may imply that the apparent increase in efficacy (comparing the intervention to the comparison group) with increasing dose (**Figure 14**) is driven by a decrease in effect within comparison groups rather than a specific benefit of more intervention hours for the behaviourally based intervention group. However, these results are based on a small number of studies and not all comparison group data are available, and therefore should be interpretated with caution.

Figure 14a. Linear model of monthly clinician-delivered dose by effect size (Hedges' g) for change in adaptive functioning outcomes from pre- to post-intervention in the intervention and comparison group.

Note: Hedges' g > 0 = improvement in outcomes from baseline to follow-up in the specified group. Hedges' g > 0 = decrease in outcomes from baseline to follow-up in the specified group.

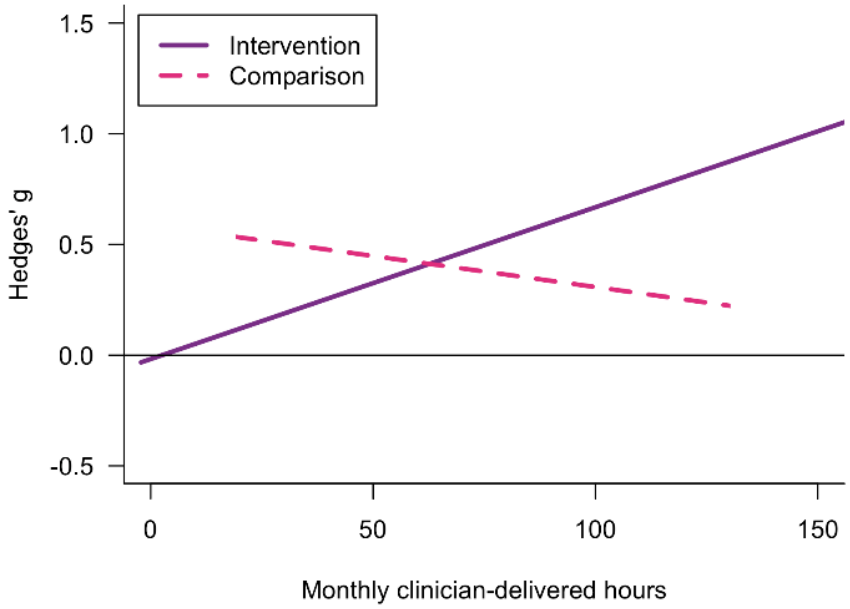


Figure 14b. Linear model of total clinician-delivered dose by effect size (Hedges' g) for change in adaptive functioning outcomes from pre- to post-intervention in the intervention and comparison group.

Note: Hedges' g > 0 = improvement in outcomes from baseline to follow-up in the specified group. Hedges' g > 0 = decrease in outcomes from baseline to follow-up in the specified group.

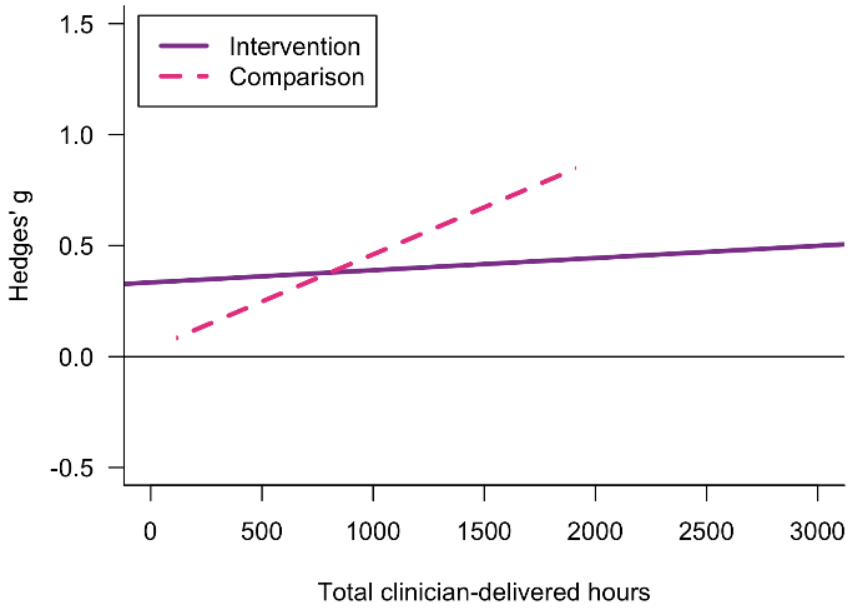


Figure 14c. Non-linear model of monthly clinician-delivered dose by effect size (Hedges' g) for change in adaptive functioning outcomes from pre- to post-intervention in the intervention and comparison group.

Note: Hedges' g > 0 = improvement in outcomes from baseline to follow-up in the specified group. Hedges' g > 0 = decrease in outcomes from baseline to follow-up in the specified group.

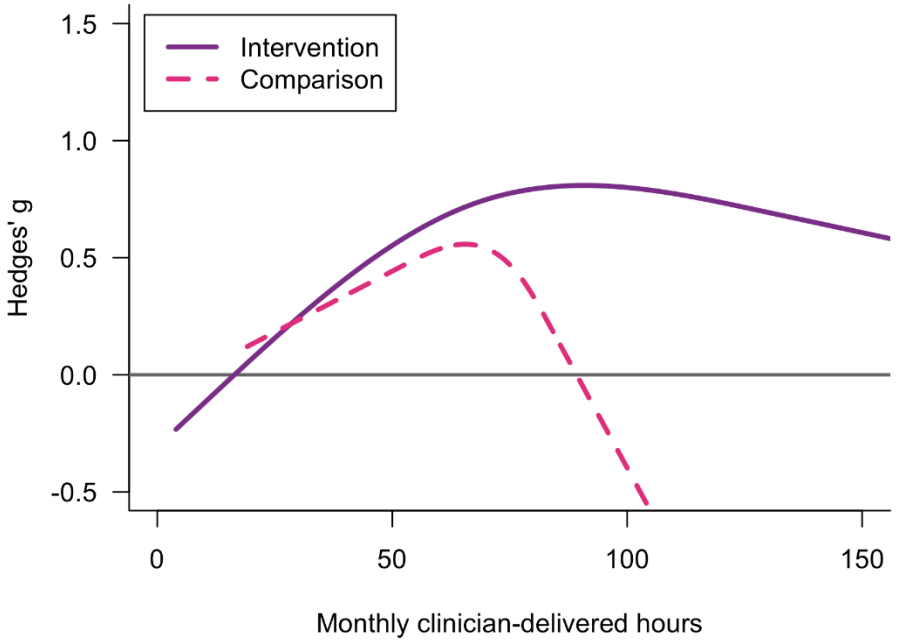
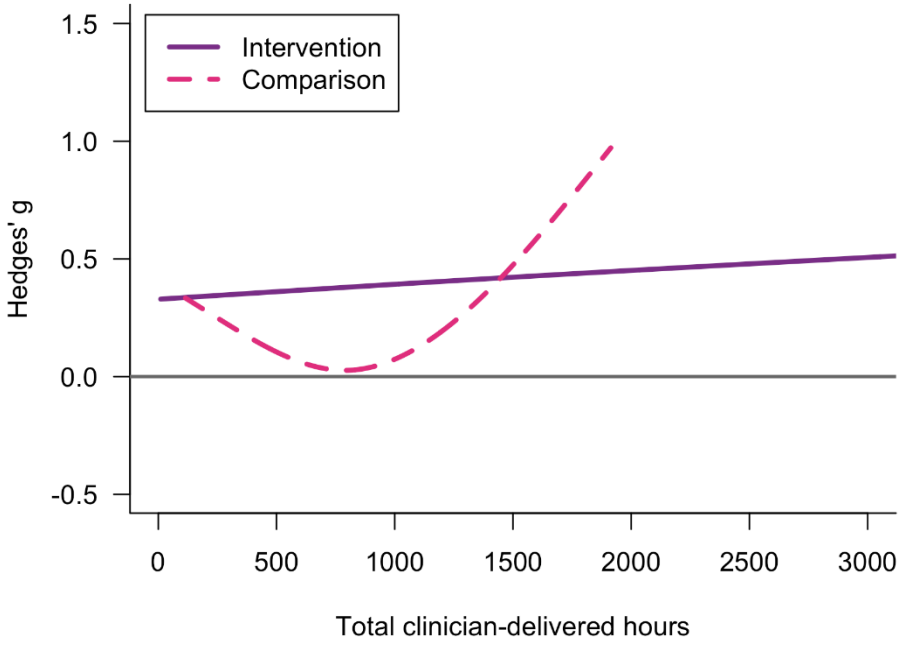


Figure 14d. Non-linear model of total clinician-delivered dose by effect size (Hedges' g) for change in adaptive functioning outcomes from pre- to post-intervention in the intervention and comparison group.

Note: Hedges' g > 0 = improvement in outcomes from baseline to follow-up in the specified group. Hedges' g > 0 = decrease in outcomes from baseline to follow-up in the specified group.



3.6.3 Cognition and language

Relationship between dose and efficacy

The linear models show a statistically significant relationship between increasing **total** and **monthly** clinician-delivered hours of intervention and better cognition and language outcomes following behaviourally based intervention as compared to a comparison group (see **Figure 15**). Linear model estimates are reported in **Table B10**. Non-linear modelling demonstrated similar results (see **Figure 15**), with only slightly improved model fit.

While the dose response trends for the linear models were statistically significant, similar to adaptive functioning outcomes, the associated real-world impact is of unclear clinical value. Every increase in dose of an additional 10 hours per month translates to an increase of 2% of a standard deviation of the effect estimate, and an increase in total dose of 100 additional intervention hours translates to less than a 1% increase. There is also less available evidence and thus less confidence (wider confidence intervals) in estimates for higher dose hours.

Figure 15a. Linear model of monthly clinician-delivered dose by effect size (Hedges' g) for cognition and language outcomes.

Note: Hedges' g > 0 = better outcomes in the intervention group compared to the comparison group. Hedges' g < 0 = better outcomes in the comparison group compared to the intervention group.

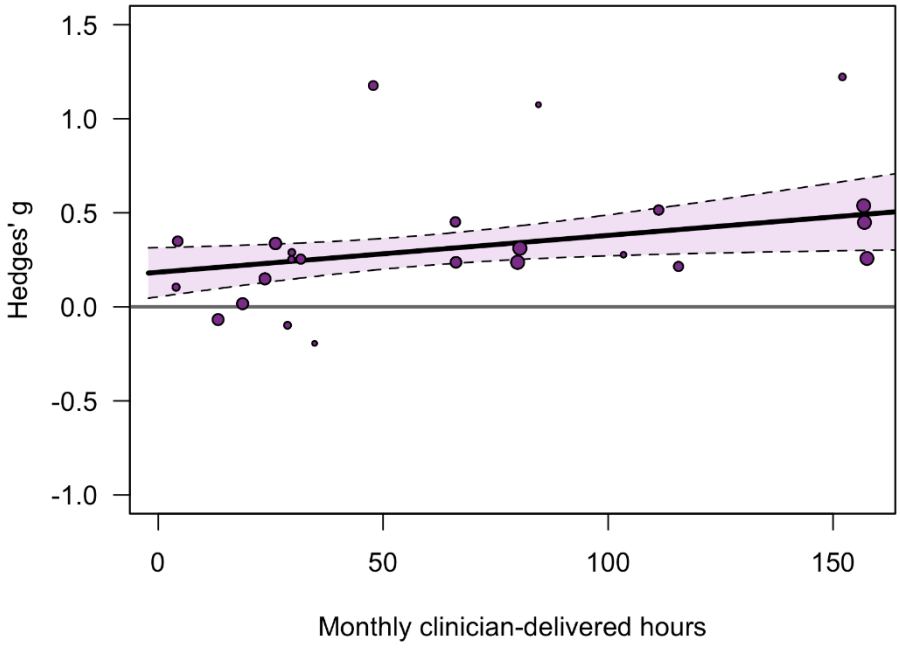


Figure 15b. Linear model of total clinician-delivered dose by effect size (Hedges' g) for cognition and language outcomes.

Note: Hedges' g > 0 = better outcomes in the intervention group compared to the comparison group. Hedges' g < 0 = better outcomes in the comparison group compared to the intervention group.

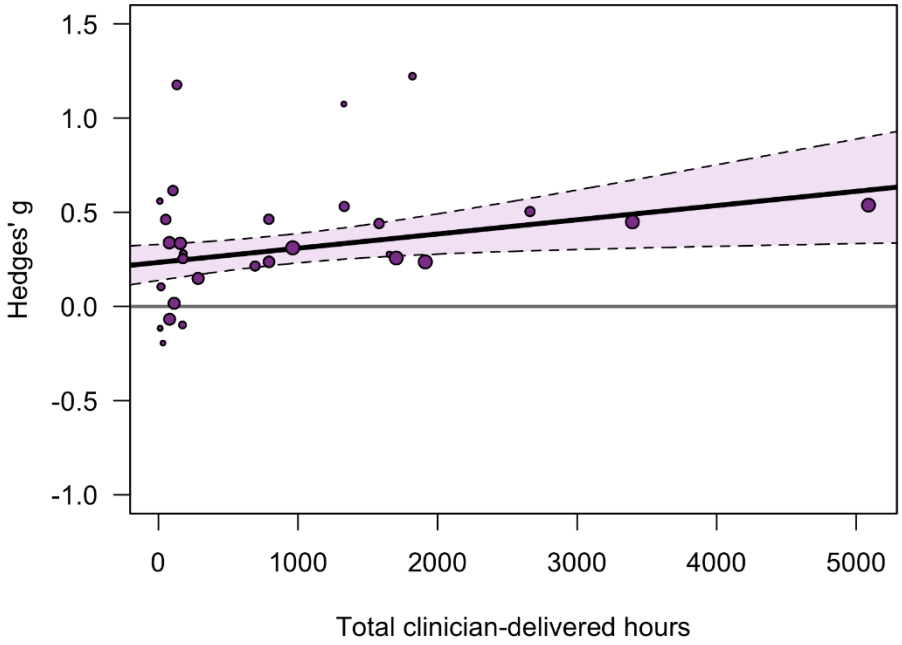


Figure 15c. Non-linear model of monthly clinician-delivered dose by effect size (Hedges' g) for cognition and language outcomes.

Note: Hedges' g > 0 = better outcomes in the intervention group compared to the comparison group. Hedges' g < 0 = better outcomes in the comparison group compared to the intervention group.

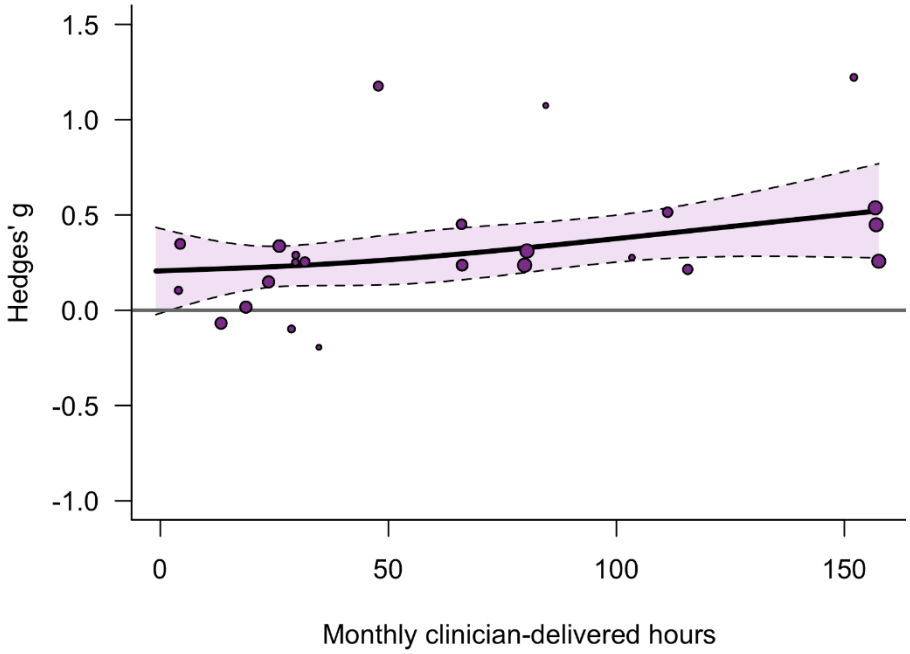
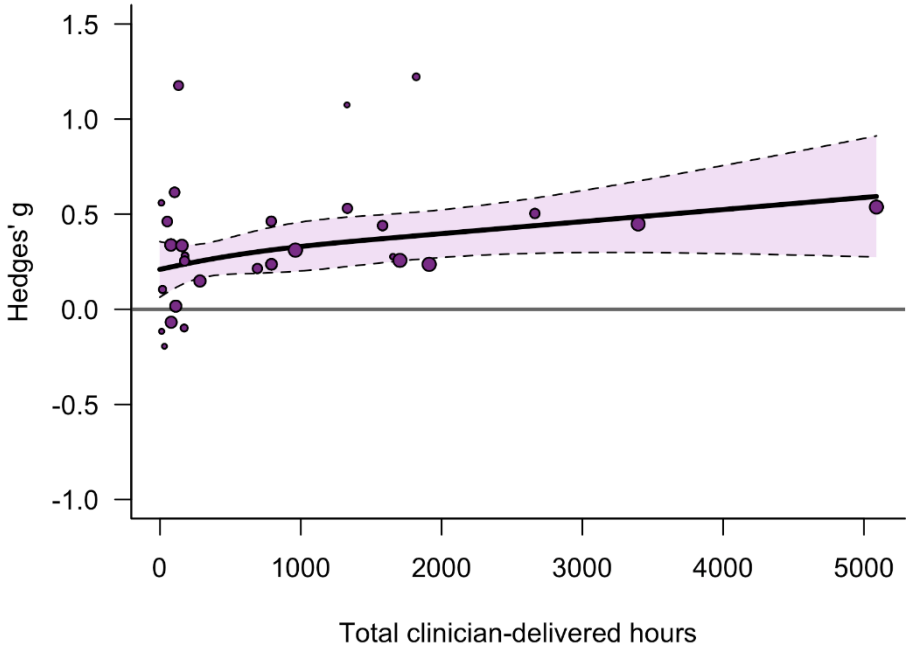


Figure 15d. Non-linear model of total clinician-delivered dose by effect size (Hedges' g) for cognition and language outcomes.

Note: Hedges' g > 0 = better outcomes in the intervention group compared to the comparison group. Hedges' g < 0 = better outcomes in the comparison group compared to the intervention group.



Comparing efficacy for lower versus higher total and monthly dose

The results did not show a difference in efficacy between the behaviourally based interventions delivered for less than the median dose of intervention versus more than the median dose of intervention, for both **total** and **monthly** clinician-delivered hours, on cognition and language outcomes (see **Figure 16** and **Table 6**). This confirms that efficacy of behaviourally based interventions on cognition and language outcomes does not differ between lower and higher **total** and **monthly** clinician-delivered intervention hours.

Table 6. Results of analyses comparing lower and higher total and monthly doses for cognition and language.

	Monthly clinician-delivered hours	Total clinician-delivered hours
Median hours	47.79 monthly hours	487.96 total hours
Lower dose: Less than median hours	**N studies:** 11 studies **Hedges' g (95% CI):** 0.20 (0.04-0.36)	**N studies:** 11 studies **Hedges' g (95% CI):** 0.21 (0.05-0.38)
Higher dose: Greater than median hours	**N studies:** 11 studies **Hedges' g (95% CI):** 0.40 (0.14-0.66)	**N studies:** 10 studies **Hedges' g (95% CI):** 0.39 (0.13-0.65)

Figure 16a. Linear dose relationship for lower versus higher monthly clinician hours (based on median) for cognition and language outcomes.

Note: The green shaded area indicates less than the median number of monthly clinician-delivered hours, and the pink shaded area indicates higher than the median number of monthly clinician-delivered hours for this outcome.

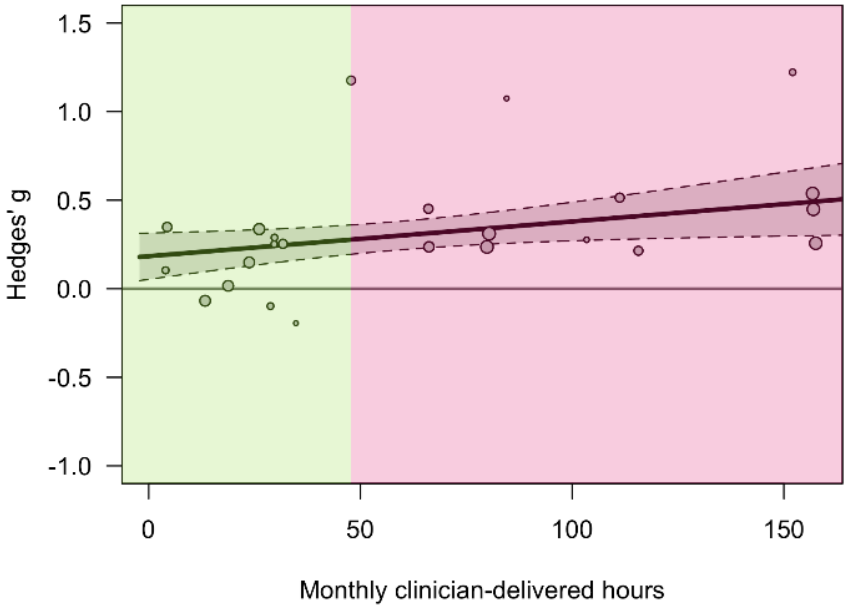
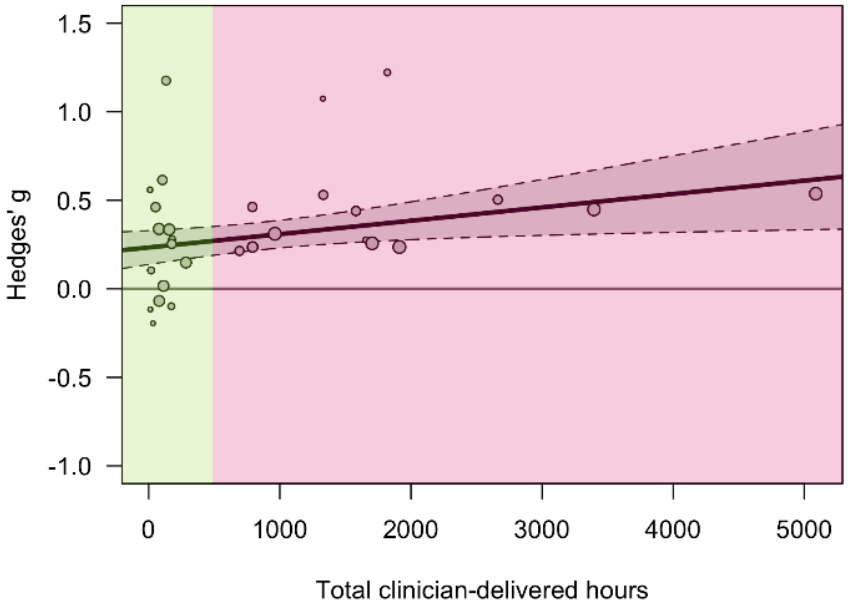


Figure 16b. Linear dose relationship for lower versus higher total clinician hours (based on median) for cognition and language outcomes.

Note: The green shaded area indicates less than the median number of monthly clinician-delivered hours, and the pink shaded area indicates higher than the median number of monthly clinician-delivered hours for this outcome.



Relationship between dose and change from baseline to follow-up separately within the intervention group and the comparison group

Analyses of change from baseline to follow-up separately within the intervention and comparison groups suggest that the relationship dose and effect size is not specific to the intervention. Linear models (see **Figure 17** and **Tables B11 & B12**) revealed statistically significant relationships between increasing **total** clinician hours and an improved outcomes from baseline to follow-up for both intervention and comparative groups, again with small coefficients (β = 0.0001 and 0.0002, respectively). Such relationships were not found for **monthly** clinician-delivered hours. This suggests that the slightly better outcomes with more hours are more likely to be related to amount of time spent with a clinician than the actual intervention taking place. The non-linear models (**Figure 17**) did not provide a clear indication of dose-response. Individual models with 95% confidence intervals can be seen in **Figures B17 & B21**.

Figure 17a. Linear model of monthly clinician-delivered dose by effect size (Hedges' g) for change in cognition and language outcomes from pre- to post-intervention in the intervention and comparison group.

Note: Hedges' g > 0 = improvement in outcomes from baseline to follow-up in the specified group. Hedges' g > 0 = decrease in outcomes from baseline to follow-up in the specified group.

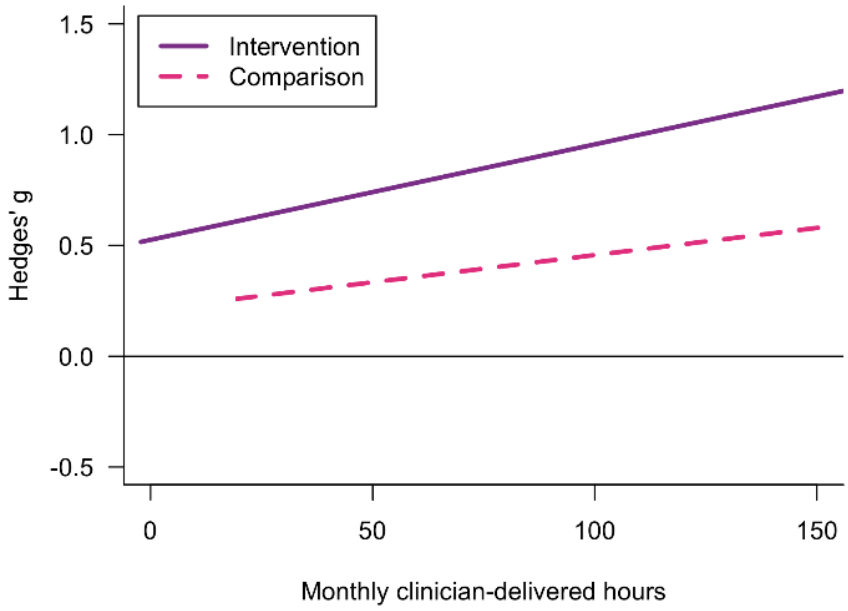


Figure 17b. Linear model of total clinician-delivered dose by effect size (Hedges' g) for change in cognition and language outcomes from pre- to post-intervention in the intervention and comparison group.

Note: Hedges' g > 0 = improvement in outcomes from baseline to follow-up in the specified group. Hedges' g > 0 = decrease in outcomes from baseline to follow-up in the specified group.

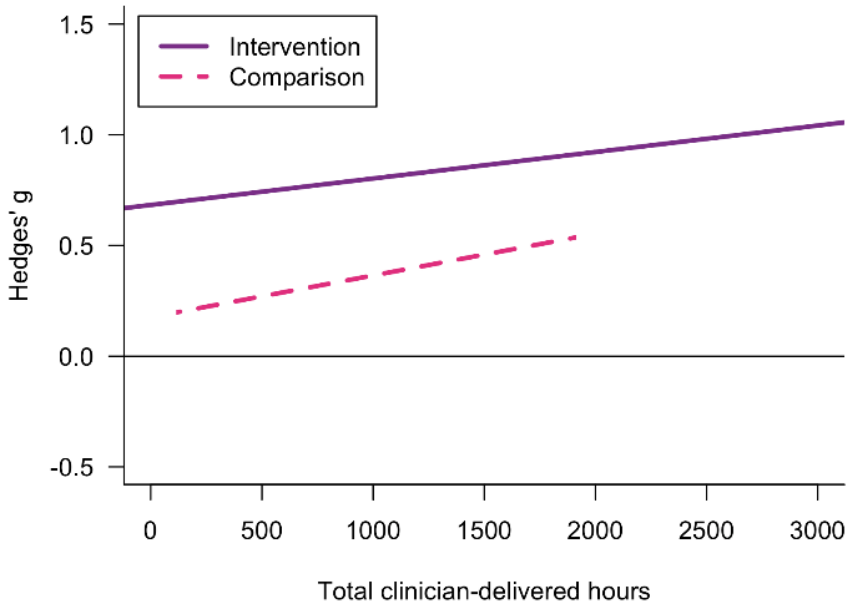


Figure 17c. Non-linear model of monthly clinician-delivered dose by effect size (Hedges' g) for change in cognition and language outcomes from pre- to post-intervention in the intervention and comparison group.

Note: Hedges' g > 0 = improvement in outcomes from baseline to follow-up in the specified group. Hedges' g > 0 = decrease in outcomes from baseline to follow-up in the specified group.

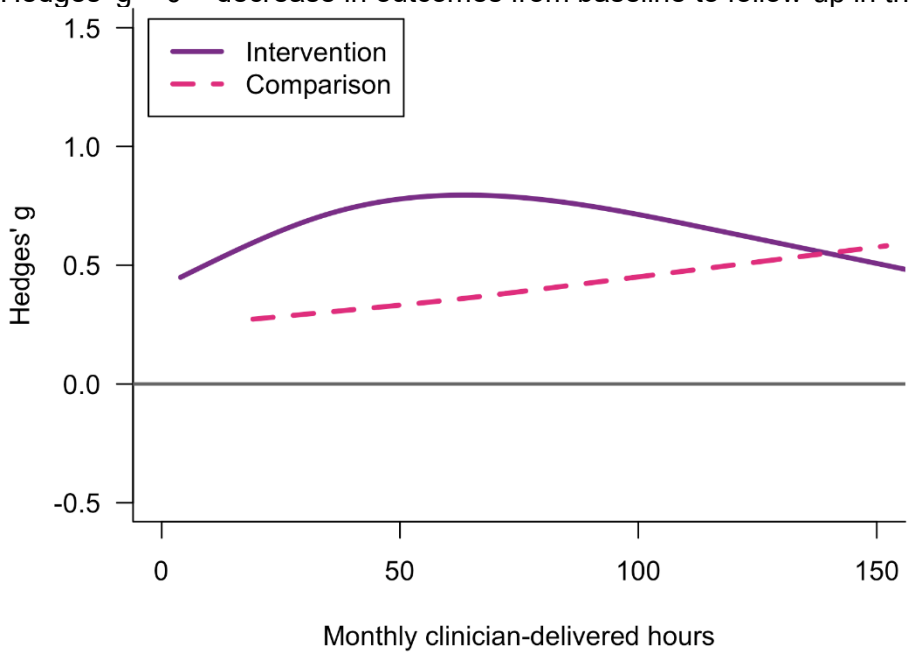
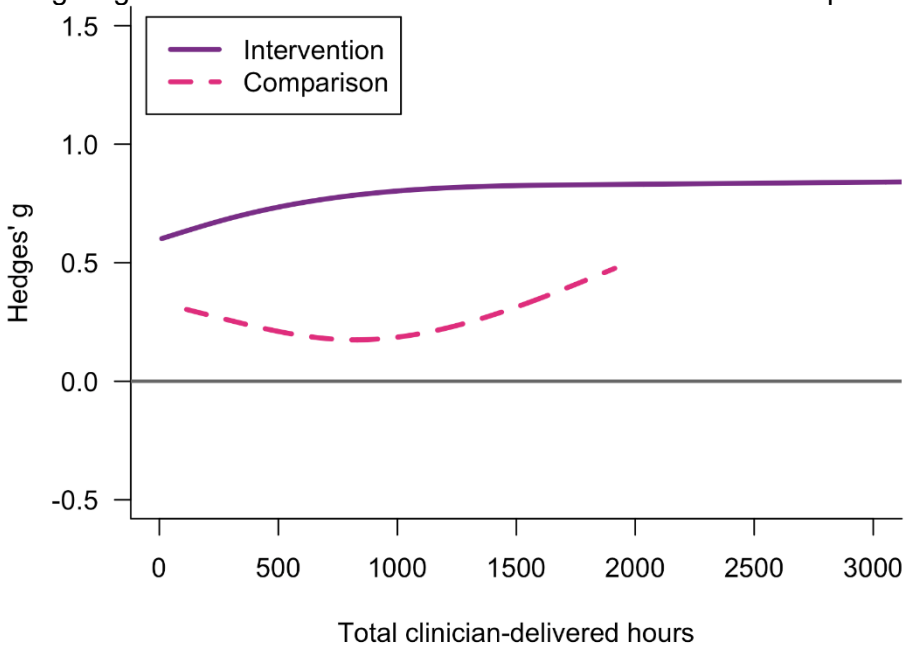


Figure 17d. Non-linear model of total clinician-delivered dose by effect size (Hedges' g) for change in cognition and language outcomes from pre- to post-intervention in the intervention and comparison group.

Note: Hedges' g > 0 = improvement in outcomes from baseline to follow-up in the specified group. Hedges' g > 0 = decrease in outcomes from baseline to follow-up in the specified group.



3.7 Investigating the effect of population, intervention, and study design factors on efficacy

Investigating the effect of population, intervention, and study design factors: summary of findings

No notable differences in the efficacy of behaviourally based interventions were identified based on any investigated factors which includes the person delivering the intervention, intervention category, comparison group, age of children, primary intervention setting, and study design.

There was no evidence to suggest that interventions delivered by a parent are inferior to those delivered by clinicians.

There was no evidence to suggest that the type of comparison group (i.e., treatment as usual or non-behaviourally based, "eclectic" intervention) biases the results.

There was no evidence to suggest that one choice of intervention characteristic is better than another. Interventions should be tailored to what is best for the child and their family's unique circumstances and needs.

3.7.1 Differences in effects within clinician, parent, or teacher-delivered interventions

The differences in effects when interventions were delivered either by a clinician, parent, or teacher was investigated using the evidence from the 98 studies (subgroup level data shown in **Table 1**). However, there was no evidence available for how teacher-delivered interventions effected family outcomes and adverse effects. Negligible effects were found in:

- teacher-delivered interventions for adaptive functioning outcomes;
- clinician-delivered interventions for family outcomes and adverse effects; and
- interventions which involved both clinician and parent-delivery on adverse effects.

Otherwise, consistent but small effects were identified for outcomes, regardless of the person delivering the behaviourally based intervention (see **Table 7** and **Figures B22-B26**).

Importantly, there were no significant differences in efficacy of intervention for any outcome depending on who delivered the intervention. While not significant, there was a difference in effect size in favour of parent-delivered interventions compared to clinician-delivered interventions for reductions in adverse effects (i.e., parent stress).

Table 7. Subgroup analysis of person delivering intervention for each outcome domain

Note: This table contains pooled effect estimates (Hedges' g) within each subgroup level for each of the five outcome domains. The colour of the text box indicates the direction of effect (worse outcomes for intervention versus comparison in pink; better outcomes for intervention versus comparison in green), with the opacity indicating the size of the effect (darker shade = stronger effect). Worse outcomes for the intervention versus comparison are also indicated by a negative effect size. Bold text and the presence of an asterisk (*) represents a significant effect estimate ($p < 0.05$). NA indicates insufficient data for analysis.

Subgroup level	Autism characteristics	Adaptive functioning	Cognition and language	Family outcomes	Adverse effects
Clinician	**0.33***	**0.32***	**0.41***	-0.07	-0.06
Clinician and parent	**0.30***	**0.26***	**0.32***	0.28	0.00
Parent delivered	**0.33***	**0.27***	**0.22***	**0.44***	**0.34***
Teacher delivered	**0.26***	-0.11	0.22	NA	NA

3.7.2 Differences in effects from each intervention category

Differences in effects relating to intervention categories (including behavioural, developmental, NDBI, technology-based, and TEACCH, as described in **Section 4.1.3**) were investigated. The number of studies which report data for each intervention category for each outcome domain are shown in **Table 1**. Results of this subgroup analysis are in **Table 8** and detailed in **Figures B22-B26**.

The evidence indicates no difference between intervention categories in efficacy across outcomes. Mostly small effect sizes in favour of behaviourally based interventions were seen, except where:

- a medium, positive effect size was observed for NDBI interventions on family outcomes;
- a negligible effect size was found for TEACCH and developmental interventions on cognition and language outcomes;
- a negligible effect size for technology-based and TEACCH interventions on autism characteristics; and
- a small, but negative effect of TEACCH on adaptive functioning outcomes was identified, although this was not significant.

Overall, the findings show that intervention content or the theoretical underpinnings of behaviourally based interventions do not effect the efficacy of the intervention on outcomes.

Table 8. Subgroup analysis of intervention category for each outcome domain

Note: This table contains pooled effect estimates (Hedges' g) within each subgroup level for each of the five outcome domains. The colour of the text box indicates the direction of effect (worse outcomes for intervention versus comparison in pink; better outcomes for intervention versus comparison in green), with the opacity indicating the size of the effect (darker shade = stronger effect). Worse outcomes for the intervention versus comparison are also indicated by a negative effect size. Bold text and the presence of an asterisk (*) represents a significant effect estimate ($p < 0.05$). NA indicates insufficient data for analysis.

Subgroup level	Autism characteristics	Adaptive functioning	Cognition and language	Family outcomes	Adverse effects
Behavioural	**0.37***	**0.32***	**0.41***	**0.35***	0.20
Developmental	**0.23***	0.36	0.06	0.21	0.33
NDBI	**0.35***	0.17	**0.29***	0.50	0.24
Other	0.27	NA	0.21	NA	0.44
TEACCH	0.05	-0.23	0.11	NA	NA
Technology-based	0.02	-0.01	0.25	NA	NA

3.7.3 Differences in effects by comparison group

Differences in effects relating to comparison group type (categories described in **Section 4.1.3**) were investigated. The number of studies which report data for each comparison group for each outcome domain are shown in **Table 1**. Results of this subgroup analysis are in **Table 9** and detailed in **Figures B22-B26**.

The evidence indicates there is no difference between the type of comparison group including treatment as usual and eclectic in efficacy across outcomes. Mostly small effect sizes are seen, except for a negligible, negative effect size for behaviourally based interventions when compared to eclectic interventions on the reduction of adverse effects. Overall, there is no evidence that the type of comparison group affects the efficacy of the intervention on outcomes.

Table 9. Subgroup analysis of comparison group for each outcome domain

Note: This table contains pooled effect estimates (Hedges' g) within each subgroup level for each of the five outcome domains. The colour of the text box indicates the direction of effect (worse outcomes for intervention versus comparison in pink; better outcomes for intervention versus comparison in green), with the opacity indicating the size of the effect (darker shade = stronger effect). Worse outcomes for the intervention versus comparison are also indicated by a negative effect size. Bold text and the presence of an asterisk (*) represents a significant effect estimate ($p < 0.05$). NA indicates insufficient data for analysis.

Subgroup level	Autism characteristics	Adaptive functioning	Cognition and language	Family outcomes	Adverse effects
Eclectic	**0.38***	0.17	**0.23***	0.69	-0.02
TAU	**0.31***	**0.27***	**0.34***	**0.35***	**0.32***

3.7.4 Differences in effects by age group

Differences in effects relating to age group (0-1 years, 2-4 years, 5-7 years) were investigated. The number of studies which report data within each age group for each outcome domain are shown in **Table 1**. Results of this subgroup analysis are in **Table 10** and detailed in **Figures B22-B26**.

The evidence shows no difference between age groups in efficacy across outcomes. Mostly small effect sizes are seen across age groups, except for:

- a medium effect size for 5-7 year-olds on cognitive outcomes;
- negligible effect sizes for 0-1 year-olds on autism characteristics and adaptive functioning outcomes; and
- a negligible effect size for 5-7 year-olds on adverse effects outcomes.

Overall, evidence indicates that age does not affect the efficacy of the intervention, although less studies currently report evidence for behavioural interventions for children aged less than 2 years old, and there is some indication of less benefit of intervention for autism characteristic and adaptive functioning outcomes in this age group.

Table 10. Subgroup analysis of age group for each outcome domain

Note: This table contains pooled effect estimates (Hedges' g) within each subgroup level for each of the five outcome domains. The colour of the text box indicates the direction of effect (worse outcomes for intervention versus comparison in pink; better outcomes for intervention versus comparison in green), with the opacity indicating the size of the effect (darker shade = stronger effect). Worse outcomes for the intervention versus comparison are also indicated by a negative effect size. Bold text and the presence of an asterisk (*) represents a significant effect estimate ($p < 0.05$). NA indicates insufficient data for analysis.

Subgroup level	Autism characteristics	Adaptive functioning	Cognition and language	Family outcomes	Adverse effects
0-1 years	0.15	0.12	**0.32***	0.4	0.24
2-4 years	**0.36***	**0.21***	**0.28***	**0.39***	**0.30***
5-6 years	**0.21***	0.42	**0.47***	**0.42***	0.06

3.7.5 Differences in effect by primary intervention setting

Differences in effects relating to primary setting of intervention (settings described in **Section 4.1.3**) were investigated. The number of studies that report data for each primary intervention setting and for each outcome domain are shown in **Table 1**. Results of this subgroup analysis are in **Table 11** and detailed in **Figures B22-B26**.

Available evidence shows no significant difference in the efficacy of behavioural interventions across settings for most outcomes. However, there was a significant difference between behaviourally based interventions primarily delivered in health and home-based settings in their efficacy for family outcomes. It is likely that health settings are more efficacious (medium effect size) compared to home settings (small effect size) for this outcome domain. Importantly, there is still evidence that interventions primarily delivered in the home are efficacious for family outcomes, the evidence just indicates that they are efficacious to a lesser degree than interventions delivered in health settings.

Table 11. Subgroup analysis of primary intervention setting for each outcome domain

Note: This table contains pooled effect estimates (Hedges' g) within each subgroup level for each of the five outcome domains. The colour of the text box indicates the direction of effect (worse outcomes for intervention versus comparison in pink; better outcomes for intervention versus comparison in green), with the opacity indicating the size of the effect (darker shade = stronger effect). Worse outcomes for the intervention versus comparison are also indicated by a negative effect size. Bold text and the presence of an asterisk (*) represents a significant effect estimate ($p < 0.05$). NA indicates insufficient data for analysis.

Subgroup level	Autism characteristics	Adaptive functioning	Cognition and language	Family outcomes	Adverse effects
Community	**0.28***	0.27	0.17	NA	NA
Early education	**0.37***	**0.31***	**0.37***	NA	0.41
Health	**0.31***	0.12	**0.29***	**0.53***	**0.27***
Home	**0.31***	**0.26***	**0.30***	**0.20***	0.16

3.7.6 Differences in effects by study design

Differences in effects relating to study design (study designs described in **Section 4.1.3**) were investigated. The number of studies which report data for each study design for each outcome domain are shown in **Table 1**. Results of this subgroup analysis are in **Table 12** and detailed in **Figures B22-B26**.

Available evidence shows no significant difference in the efficacy of behavioural interventions across study designs for most outcomes. However, there was a significant difference between cohort studies and randomised/non-randomised controlled trials on the reduction of adverse effects. It is likely that randomised and non-randomised controlled trials are more efficacious (small effect size) compared to cohort studies (small effect size), which demonstrate a significant increase in adverse effects (although with a small effect size).

Table 12. Subgroup analysis of study design for each outcome domain

Note: This table contains pooled effect estimates (Hedges' g) within each subgroup level for each of the five outcome domains. The colour of the text box indicates the direction of effect (worse outcomes for intervention versus comparison in pink; better outcomes for intervention versus comparison in green), with the opacity indicating the size of the effect (darker shade = stronger effect). Worse outcomes for the intervention versus comparison are also indicated by a negative effect size. Bold text and the presence of an asterisk (*) represents a significant effect estimate ($p < 0.05$). NA indicates insufficient data for analysis.

Subgroup level	Autism characteristics	Adaptive functioning	Cognition and language	Family outcomes	Adverse effects
Cohort	**0.26***	**0.27***	**0.33***	NA	**-0.27***
Non-random	**0.35***	0.20	**0.39***	0.50	**0.32***
Random	**0.32***	0.25	**0.22***	**0.39***	**0.25***

4. Limitations

To the best of our knowledge, this report is the largest systematic review and meta-analysis of behavioural interventions in children on the autism spectrum conducted (Whitehouse & Eapen, 2020). However, although 98 studies met eligibility criteria, approximately two thirds of these did not report quantifiable clinician-delivered hours of intervention. These studies were either parent- or teacher-delivered or did not report dose information at all. Therefore, the dose-response analyses in this report are based on a more limited evidence pool of 34 studies, which limits the precision (i.e., statistical power) of the analyses.

This report includes a statistical summary of best available evidence across the literature. In doing this, a high-level summary of over 4,500 children was made. Individual experiences and responses of children to these behaviourally based interventions will vary. It is beyond the scope of this investigation to evaluate philosophical or qualitative information around responses to behaviourally based interventions.

The quality and accuracy of reported dose information varied across studies. Often only the planned dose was reported rather than actual dose delivered. As actual dose delivered is commonly less than what was planned, the dose estimates reported here are likely an overestimation of what was delivered. Additionally, assessments of the potential risk of bias within included studies suggested a high, or serious, overall risk.

All children, even those in the comparison group, likely received some level of intervention. Studies varied in how comprehensively they reported details of alternative or standard-care intervention in the comparison group. Because of this, it was difficult to categorise and investigate the effects compared to different comparison groups. We acknowledge that the comparison group definitions used here (TAU and eclectic) are arbitrary, and there is potentially overlap across these groups.

Any behaviourally based interventions for children on the autism spectrum were included. While all-inclusive and comprehensive, this means that included interventions vary in their evidence for efficacy, with some based on stronger evidence. Additionally, fidelity was not adjusted for within analyses. The inclusion of all interventions, including those with less evidence and lower fidelity, may have resulted in an underestimation of the effect size.

Finally, the impact of important participant characteristics, such as autism severity, on the efficacy of behaviourally based interventions were unable to be assessed in this report. This was due to limited evidence and the lack of uniformity in the measures used to quantify autism severity within included studies.

5. What did we learn?

5.1 What benefits are likely?

Overall, there is evidence for benefit, of a small effect size (extent of benefit), of behaviourally based interventions for children less than 7 years old on the autism spectrum for key clinical outcome domains. Better outcomes following behaviourally based interventions than those in comparative groups were identified for all five investigated outcome domains: autism characteristics (e.g., socialisation, challenging behaviours), adaptive functioning, cognition and language, family outcomes, and in the reduction of adverse effects (i.e., child and parent stress/burden). Although, of note, the extent of the benefit (effect size) was smaller and more varied than reported in previous systematic reviews (Whitehouse & Eapen, 2020).

Importantly, while small in effect size, the benefit of behaviourally based interventions was found when compared against children undergoing both usual care and other “eclectic” intervention types. Importantly, larger improvements were consistently seen for children who underwent a behaviourally based intervention compared to children who experienced equivalent clinician-delivered hours (total and monthly) of treatment as usual or alternative, “eclectic” interventions. Due to limited data, this comparison of change following intervention by dose of intervention was only possible for autism characteristics, adaptive functioning and cognition and language outcomes.

The small effect size and variability identified across outcomes means that benefit of behaviourally based interventions cannot be guaranteed across all interventions, settings, and participants. This indicates that multiple factors must be considered when making treatment decisions for a child on the autism spectrum and decisions pertaining to a child’s goals and family values should reflect their individual needs. This report has explored factors including dose, primary setting, the person delivering the intervention and child characteristics (e.g., age). The impact of these on the efficacy of behaviourally based interventions is discussed in the sections which follow. Of note, some important factors to consider (e.g., autism severity) were unable to be investigated here.

5.2 Is more intervention better?

To answer the question of whether more clinician-delivered hours of intervention lead to better outcomes depends on the child’s outcome/s of interest (i.e., autism characteristics, adaptive functioning, cognition and language).

Evidence shows that there is no benefit in increasing clinician hours (total or intensity) for autism characteristic outcomes (e.g., global autism measures, social affect, socialisation, challenging behaviours, etc). For this outcome, small benefits of behaviourally based interventions are consistently seen, regardless of dose. This means that lower dose intensities and total doses may be sufficient to see maximum benefit for autism characteristics, and increased benefit is unlikely to occur with alterations of dose of the intervention.

Contrastingly, evidence indicates that increasing both total clinician hours as well as the intervention intensity is associated with improved adaptive functioning and cognition and language outcomes. However, incremental increases (e.g., from 10 to 20 monthly hours) show little added value, so decisions about the amount of intervention received should be made upon a child's progress towards their overall goals and what is most beneficial to the child, rather than dose alone. Improvements in cognition and language outcomes were shown at all dose levels, even low total clinician hours and dose intensities. This was not the case for adaptive functioning, where there is little evidence for benefit of behaviourally based interventions when delivered for less than approximately 800 hours, or 65 hours per month.

The results of the dose analyses warrant further validation due to the relatively small number of clinician-led studies reporting outcomes, especially at higher doses. The small number of studies meant there was large variability in estimates and lower confidence in the results, particularly for higher doses. Even if a dose relationship is found with more evidence, the potential benefit of increasing hours identified here was found to be minimal, translating to negligible real-world impact.

5.3 Who is best placed to deliver interventions?

Parent-delivered behaviourally based interventions (typically following parent training and ongoing support from providers) may be as useful as clinician-delivered designs, with no difference found between these designs on benefit of the intervention. Support for the benefit of parent-delivered interventions echoes the recent Autism CRC report (external), "Interventions for children on the autism spectrum", where the important and beneficial role of parents or caregivers in delivering early interventions is highlighted (Whitehouse & Eapen, 2020), as well the National Guidelines for Best Practice in Early Childhood Intervention (external) (Early Childhood Intervention Australia, 2016), which emphasize the importance of family-centred supports and the involvement of family in the intervention process.

5.4 What other intervention design, implementation or participant factors impact outcomes?

The current evidence does not indicate that the benefits of behaviourally based interventions differ based on the age group of children receiving the intervention, primary intervention setting (i.e., home or health setting), intervention category, type of comparison group (i.e., TAU or 'eclectic'), or study design.

5.5 Considerations for practice

Behaviourally based interventions can be efficacious for children on the autism spectrum under 7 years. However, effects will vary depending on individual and intervention-specific factors. Importantly, no particular benefit of specific intervention characteristics was found (e.g., intervention category, primary intervention setting, person delivering the intervention). This implies that all factors may be useful in the right context and behaviourally based interventions can, and should be,

individualised and take into consideration the needs, preferences, and individual circumstances of the child and their family.

In clinician-led behaviourally based interventions, the number and intensity of clinician contact hours cannot, on its own, account for the variability in the effects found. This means that intervention planning decisions should consider dose, but not in the absence of considerations relating to the child's goals, context, and family circumstances.

The specific goals of the participant and the planned outcomes of the intervention are of particular importance. Evidence presented here shows a benefit of more total hours as well as more intense intervention (more monthly hours of clinician-delivered intervention) for adaptive functioning and cognition and language outcomes. No impact of dose was found for autism characteristic outcomes. This difference in the effect of dose based on outcomes measured indicates that the justification for increased dose intensity should be based upon the needs and the goals of the participant.

Evidence suggests that higher doses of behaviourally based intervention may be required to see benefit for adaptive functioning outcomes. It may be the case that if the goals which have prompted the child to seek intervention relate to adaptive functioning, at least 65 hours per month of intervention would achieve greatest benefit, and lower doses may be futile. It is important to note that this is an estimated amount and results will vary, based upon a child's overall goals and other intervention-related factors. With such high doses required to see benefit for adaptive functioning, this may suggest that behaviourally based interventions may be less efficient (in terms of contact hours) for some outcomes. Thus, if treatment goal is adaptive functioning, participants may want to consider alternative approaches.

Additionally, the potential added benefits of incrementally increasing total and monthly dose across outcomes were shown to be minimal and unlikely to be clinically meaningful. Decisions to increase intervention intensity must be considered within the child's context and dose should only be considered as one factor within treatment decisions as it is not always related to better outcomes. Importantly, decisions regarding the amount and duration of intervention should be made in consideration of concerns around the impact of intensive therapies on a child's development (as highlighted within Recommendation 56 of the Autism CRC National Guideline (external) (Trembath et al., 2022)). For example, time spent in health and clinical settings may come at the cost of time for learning and development in more naturalistic settings that are family centred, which is beneficial to child development.

5.6 Potential areas and considerations for follow-up work

All analyses reported here are at group-level (grouped by characteristics within and between studies). To further this work, it is important to adjust for differences in individual circumstances. An analysis of individual participant data from the literature, clinical partners (e.g., from the Autism Specific Early Learning and Care Centres, ASELCCs (Masi et al., 2021)) or other NDIS providers will allow for factors relating to the interaction between dose and individual or intervention design factors to be explored in more detail.

The robust evidence for efficacy of parent-led interventions warrants further investigation into the factors underlying efficacious interventions. Potential questions may include, among others:

(1) what interventions or intervention components can be delivered effectively by parents,
(2) how to balance clinician- and parent-time, and
(3) how parents could be better supported to deliver interventions.

This can be achieved using network meta-analysis, which can investigate the components of interventions and synthetically compare intervention approaches head-to-head.

There is limited available information on enduring change over time following the conclusion of these interventions. Future research must include longer-term follow-ups in the children who receive these interventions in order to address this gap.

Importantly, research in this area is of poor quality. High risk of bias was identified in two thirds of studies included in this report. Improving the quality of studies in this field is vital. Common points for consideration in future study designs to improve study quality include ensuring the blinding of outcome assessors, concealing randomisation prior to assignment to intervention (for randomised designs), measuring or controlling for important confounders (e.g., age, autism severity, IQ), and using appropriate statistical methods for missing data.

6. References

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders - 5th Edition*. American Psychiatric Association.

Early Childhood Intervention Australia. (2016). *National guidelines: Best practice in early childhood intervention*. https://www.ecia.org.au/documents/item/1419

Makrygianni, M. K., & Reed, P. (2010). A meta-analytic review of the effectiveness of behavioural early intervention programs for children with Autistic Spectrum Disorders. *Research in Autism Spectrum Disorders*, *4*(4), 577-593. https://doi.org/https://doi.org/10.1016/j.rasd.2010.01.014

Masi, A., Dissanayake, C., Alach, T., Cameron, K., Fordyce, K., Frost, G., . . . Eapen, V. (2021). Clinical outcomes and associated predictors of early intervention in autism spectrum disorder: a study protocol. *BMJ Open*, *11*(8), e047290. https://doi.org/10.1136/bmjopen-2020-047290

Rodgers, M., Marshall, D., Simmonds, M., Le Couteur, A., Biswas, M., Wright, K., . . . Hodgson, R. (2020). Interventions based on early intensive applied behaviour analysis for autistic children: a systematic review and cost-effectiveness analysis. *Health Technology Assessment*, *24*(35), 1-306. https://doi.org/10.3310/hta24350

Trembath, D., Varcin, K., Waddington, H., Sulek, R., Pillar, S., Allen, G., . . . Whitehouse, A. (2022). *National guideline for supporting the learning, participation, and wellbeing of autistic children and their families in Australia*. Autism CRC. https://www.autismcrc.com.au/access/supporting-children

Trembath, D., Waddington, H., Sulek, R., Varcin, K., Bent, C., Ashburner, J., . . . Whitehouse, A. (2021). An evidence-based framework for determining the optimal amount of intervention for autistic children. *Lancet Child Adolesc Health*, *5*(12), 896-904. https://doi.org/10.1016/s2352-4642(21)00285-6

Virues-Ortega, J. (2010). Applied behavior analytic intervention for autism in early childhood: meta-analysis, meta-regression and dose-response meta-analysis of multiple outcomes. *Clinical Psychology Review*, *30*(4), 387-399. https://doi.org/10.1016/j.cpr.2010.01.008

Whitehouse, A., Varcin, K., Waddington, H., Sulek, R., Bent, C., Ashburner, J.,, & Eapen, V., Goodall, E., Hudry, K., Roberts, J., Silove, N., Trembath, D. (2020). *Interventions for children on the autism spectrum: A synthesis of research evidence*. A. CRC. https://www.autismcrc.com.au/interventions-evidence

Appendix A: Detailed study methodology

This systematic review adheres to guidelines from the 2020 update of the Preferred Reporting Items for Systematic reviews and Meta-Analyses statement (PRISMA 2020 (Page et al., 2021)).

A1. Study objectives

The review sought to synthesise the available evidence for the efficacy of behavioural interventions in children, aged 7-years or younger on the autism spectrum.

The following objectives were examined in the current systematic review and meta-analysis:

1. What is the evidence for the general efficacy and effectiveness of behavioural interventions?
    a. What effect sizes should be expected on common composite and domain-specific assessments of autism characteristics, functional or community outcomes?
    b. What adverse effect are reported and how common are they?
    c. How do effect sizes vary across outcome measures and domains?
    d. To what extent are any observed effects confounded by common sources of bias within and between studies?
2. How do effect sizes vary across settings?
    a. To what extent do effect sizes vary across studies (i.e., heterogeneity in true effects)?
    b. What common population, intervention and study design factors are possible moderators of heterogeneity?
3. How are effects associated with behavioural interventions related to intervention dose?
    a. What effects should be expected across different intensities (i.e., hours per week) and durations of interventions?
    b. How do intensity and duration interact across different delivery formats?
    c. What are the shapes of the dose-response curves for different outcomes?
    d. Are such dose-response relationships moderated or confounded by other design factors?
    e. How do the outcomes of each intervention compare to other behavioural interventions at different levels of intensity?

A2. Electronic search strategy

A single search of MEDLINE, EMBASE, CENTRAL and PsycINFO via OVID was conducted on 15 November 2021 for studies examining the effects of behavioural interventions (based on ABA principles) in children aged 7-years or less on the autism spectrum on at least one outcome involving autism characteristics, adaptive functioning, cognition and language, family outcomes, or adverse effects. The Medline search strategy is shown below.

The Ovid MEDLINE search strategy (including ALL from 1946 to November 15, 2021) was:

1. exp Autism/ or exp Autistic Disorder/
2. exp Autism Spectrum Disorder/ or exp Asperger Syndrome/
3. (autis$ or Asperger$ or Kanner$ or ASD or ASC or AAC).ti,ab,kw.
4. exp child development disorders, pervasive/
5. exp Developmental Disabilities/
6. Pervasive development$ disorder$.ti,ab,kw.
7. 1 or 2 or 3 or 4 or 5 or 6
8. exp Applied Behavior Analysis/
9. exp Behavior Therapy/
10. early intervention therap$.ti,ab.
11. (high intensity adj2 (analys$ or behavior$ or behaviour$ or intervention$ or model$ or program$ or therap$ or treat$)).ti,ab.
12. (low intensity adj2 (analys$ or behavior$ or behaviour$ or intervention$ or model$ or program$ or therap$ or treat$)).ti,ab.
13. (intensive behavior$ adj2 (analys$ or intervention$ or model$ or program$ or therap$ or treat$)).ti,ab.
14. (intensive behaviour$ adj2 (analys$ or intervention$ or model$ or program$ or therap$ or treat $)).ti,ab.
15. (early behavior$ adj2 (analys$ or intervention$ or model$ or program$ or therap$ or treat$)).ti,ab.
16. (early behaviour$ adj2 (analys$ or intervention$ or model$ or program$ or therap$ or treat$)).ti,ab.
17. (comprehensive behavior$ adj2 (analys$ or intervention$ or model$ or program$ or therap$ or treat$)).ti,ab.
18. (comprehensive behaviour$ adj2 (analys$ or intervention$ or model$ or program$ or therap$ or treat$)).ti,ab.
19. (applied behavior$ adj2 (analy$ or intervention$ or model$ or program$ or therap$ or treat$)).ti,ab.
20. (applied behaviour$ adj2 (analy$ or intervention$ or model$ or program$ or therap$ or treat$)).ti,ab.
21. (ABA$ adj2 (analys$ or intervention$ or model$ or program$ or therap$ or treat$)).ti,ab.
22. (IBI or EIBI or ABA).ti,ab.
23. Lovaas$.mp.
24. discrete trial train$.ti,ab.
25. Picture exchange communication system$.ti,ab.
26. functional communication training$.ti,ab.
27. (intens$ adj2 (analys$ or behav$ or intervention$ or program$ or therap$ or treat$)).ti,ab,kw.
28. (behavio?r$ adj2 (analy$ or intervention$ or model$ or program$ or therap$ or treat$)).ti,ab,kw.
29. (behav$ adj2 (analy$ or intervention$ or model$ or program$ or therap$ or treat$)).ti,ab,kw.
30. Comprehensive application of behav* analysis to school*.mp.
31. (Comprehensive application of behav* analysis to school* or CABAS).ti,ab,kw.

32. PECS.ti,ab,kw.
33. Direct instruction$.ti,ab,kw.
34. "treatment and education of autistic and communication related handicapped children".mp.
35. TEACCH.mp.
36. (Early Start Denver Model or ESDM).mp.
37. (Naturalistic Developmental behav* or NDBI*).mp.
38. (joint attention adj (training or skills or learning or intervention or program or therap$)).mp.
39. (Joint Attention Symbolic Play or JASPER).mp.
40. (Pivotal response adj1 (training or skills or learning or intervention or program or therap*)).mp.
41. reciprocal imitation.mp.
42. positive behav$ support.mp.
43. (developmental individual difference relationship based or DIR or floortime or floor time or interactive play).mp.
44. (developmental individual difference relationship or floortime or floor time or interactive play).mp.
45. (autism adj communication therapy).mp.
46. language training.mp.
47. Functional Communication Training.mp.
48. 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47
49. 7 and 48

The search was not limited by time, location, or language. Articles written in a language other than English were translated. Additional articles were identified by scanning the reference lists of existing reviews. One reviewer (Nicole Hill) conducted the initial search. Screening of title and abstracts and the review of full texts was conducted by five reviewers (Nicole Hill, Ivana Randjelovic, Amit Lampit, Erica Ghezzi, Matthew McQueen). Each article was screened by two of the five reviewers. Discrepancies were resolved by Amit Lampit who also contacted corresponding authors for additional information when required.

A3. Study selection and eligibility criteria

A3.1 Types of studies

Eligible studies included in the meta-analysis were randomised or non-randomised. Eligible studies must have been published in peer-reviewed journals or included in previous systematic reviews, but data extracted from those studies may have been unpublished (e.g., obtained from study authors).

A3.2 Types of participants

Studies were eligible if they included children (mean age ≤7 years at baseline) with a diagnosis of autism spectrum disorder or reported as at high likelihood for autism spectrum disorder if too young for formal diagnosis (less than 3 years old). Autism spectrum disorder comorbid with other conditions (including established or evident intellectual disability) will be eligible.

A3.3 Types of interventions

Behavioural interventions included those which:

- Used behaviourally based teaching strategies as the core components
- Used a comprehensive approach, to increase social engagement and learning while targeting a range of behaviours, skills (i.e., social, interpersonal, and daily living skills) and developmental domains (i.e., language, social communication, cognition, adaptive functioning, play development)
- Delivered face to face or using telehealth by qualified or trained individuals, on a one-to-one or small group basis to children directly, or via parents, caregivers, teachers, or combinations thereof
- Delivered at centre, home, school, or the community, or across multiple settings

Above criteria were based on Rodgers et al. (2020) and the Autism CRC (Whitehouse & Eapen, 2020). Studies of eligible interventions combined with other approaches were included if ≥50% of intervention time met above criteria.

Examples of typical interventions which meet the above criteria include:

- Early Intensive Behavioral Treatment (University of California/Lovaas Model)
- Intensive ABA
- Non-intensive ABA
- Comprehensive Application of Behaviour Analysis to Schooling (CABAS)
- Verbal behavior
- Discrete trial training
- Direct Instruction
- Picture Exchange Communication System (PECS)
- Treatment and Education of Autistic and Communication related handicapped CHildren (TEACCH)
- Early Start Denver Model (ESDM)
- Comprehensive intensive early intervention
- Naturalistic Developmental Behavioural Interventions (NDBIs)
- Early Social Interaction Project
- Joint attention and imitation skill-building
- Joint Attention, Symbolic Play, Engagement, and Regulation (JASPER)
- Learning Experiences Alternative Program (LEAP)
- Pivotal Response Training (PRT, also called Pivotal Response Treatment)
- Reciprocal Imitation Training
- Positive Behaviour Support
- Developmental Individual-Difference Relationship-Based (DIR) / Floortime
- Paediatric Autism and Communication Therapy (PACT)
- Language training
- Functional Communication Training

Eligibility of behavioural interventions were determined in consultation with Megan Clark, Postdoctoral Research Fellow and Provisional Psychologist at the Olga Tennison Autism Research Centre of La Trobe University.

There will be no limitation on intervention dose or intensity (hours per week, total number of hours, overall duration). All eligible intervention arms in multi-arm studies will be included.

A3.4 Types of comparators

Studies had to report data for at least one comparison group which was also comprised of children less than 7 years on the autism spectrum. Eligible comparisons include passive/waitlist control, treatment as usual (TAU), alternative community-based interventions (e.g., eclectic treatments) or non-evidence supported treatments.

A3.5 Types of outcome measures

Outcomes assessed at two time points (before and after the intervention) were eligible. Eligible outcomes included any measure which came under the following five categories:

Autism characteristics: Describes specific characteristics of autism, as well as global autism characteristic measures. Includes characteristics such as emotion regulation, restricted repetitive behaviours, sensory problems, social affect, socialisation, challenging behaviours.

Cognition and language: Describes the child's cognitive and language abilities. Includes measures of IQ, developmental age, motor skills, as well as receptive and expressive language.

Adaptive functioning: Describes measures of the child's everyday functioning. Includes functional behaviours such as toileting, helping with chores, answering the phone.

Family outcomes: Describes wellbeing or quality of life of the child, caregiver, or overall family unit, as well as parent sense of competency.

Adverse effects: Describes adverse effects of the intervention. Includes child distress (e.g., anxiety/depression) as well as parent stress or burden. These effects were coded so that higher scores indicated better outcomes (i.e., reduction in adverse effects).

A4. Data collection and coding

Coding of outcome measures was conducted by Erica Ghezzi who double-checked all data for accuracy. Data was coded into an excel spreadsheet for analysis in R. Data from studies were usually entered as means and standard deviations for pre-post measures for the intervention and comparison group, but if this was unavailable, any data from which an effect size for the difference between intervention and comparison groups in change from pre- to post-measures could be calculated was entered.

If a study included multiple follow-up timepoints **during** the intervention (e.g., after 1 year of intervention AND after 2 years of intervention), both were collected as they represent different dose amounts for the dose analysis. If a study had multiple follow-up points (with no further intervention delivered), data from the first time-point, immediately after completion of the intervention (maximum dose) was collected.

In addition to the primary outcome measures, information on the study design and characteristics were extracted for each eligible article which included, author, publication year, country, study design, intervention description, comparison group description, participant characteristics (e.g., age, gender), intervention settings, intervention dose (duration and frequency), mode of delivery (e.g., parent or clinical supervised).

A5. Assessing the quality of the evidence

Risk of bias was assessed using the Cochrane RoB 2.0 (Sterne et al., 2019) for included studies which were randomised controlled trials, and the ROBINS-I (Sterne et al., 2016) tool for included studies which employed non-randomised designs (e.g., non-randomised controlled trials, cohort studies).

Risk of bias was assessed for each reported outcome domain (e.g., autism characteristics, adaptive functioning, etc) within each included manuscript. Risk assessments were then summarised at the study level by taking the highest risk assessment for each risk of bias domain. Overall assessments were made as per the respective risk of bias tool's guidelines.

A6. Data analysis

All analyses were conducted using the R packages `metafor` and `robumeta`.

A6.1 Combining effects from included studies

The primary outcome was standardised mean difference (calculated as Hedges' g) of difference between intervention and comparison groups in change from pre- to post-intervention. Precision of the Hedges' g was calculated for each outcome measure by the 95% confidence interval (CI). A positive Hedges' g implies better therapeutic effects over time in the intervention group compared to the comparison group. By convention, Hedges' g values of 0.2, 0.5 and 0.8 are considered small, moderate or large effect sizes, respectively.

When studies provided multiple effect sizes or subgroups, all eligible effect sizes and subgroups were pooled using robust variance estimation models. Heterogeneity across studies was quantified using the tau^2 statistic.

Small-study effect ('publication bias') was assessed by visually inspecting funnel plots of effect sizes versus standard error. Where there were at least 10 studies in analyses, the small-study effect was formally tested using Egger's test. If evidence was found for this effect (if $p < 0.1$), the trim and fill method was used to create an adjusted effect estimate.

A6.2 Dose response analyses

Relationship between dose and efficacy

Linear models

The relationship between dose and effect size was modelled using multivariate linear meta-regression. When studies provided multiple effect sizes for the same dose of intervention, all eligible effect sizes and subgroups were pooled using robust variance estimation models, and then a linear regression was run.

These models were run separately for the two measures of dose (**total** clinician-delivered hours, and **monthly** clinician-delivered hours) for each of the three outcome domains which reported sufficient data for dose analyses (autism characteristics, adaptive functioning, cognition and language).

The model statistics were recorded, and the model significance was tested using the p-value ($p < 0.05$ represents a statistically significant lineal model). The model was then plotted, including the 95% confidence interval, which represents the precision of the model.

Non-linear models

The same relationships (effect size and dose for relevant outcome domains) were then explored using non-linear meta-regression models. This involved the same process, except now the relationship was not assumed to be linear.

For each analysis of dose and effect size by outcome domain, three types of non-linear models were investigated: cubic polynomial, restricted cubic spline, and thin plate spline. Across outcomes, the restricted cubic spline model was shown to have the best fit, and so this method was used for all non-linear models within this report. These models were fitted with three knots at the 10^{th}, 50^{th} and 90^{th} percentiles of dose.

Once again, model statistics were recorded, and the model significance was tested using the p-value ($p < 0.05$ represents a statistically significant lineal model). Plots of non-linear models included the 95% confidence interval to represent model precision.

Comparing efficacy for lower versus higher total and monthly dose

This analysis involved pooling the effect sizes (Hedges' g), as was done the main analysis described above, but now within subgroups. In this case, the subgroups were lower and higher dose, split by the median dose across studies with available data within each outcome domain (autism characteristics, adaptive functioning, and cognition and language). Separate analyses were conducted for the two definitions of dose: total clinician-delivered hours and monthly clinician-delivered hours (dose intensity).

Differences in effect size between lower versus higher dose levels were assessed for each dose type and outcome domain using a Wald-type test. If statistically significant ($p < 0.05$), this indicates that there is a difference in efficacy of the intervention between levels of the subgroup. In this case, it would mean there is a difference in efficacy for that outcome domain between lower and higher dose levels (per the definition of the specified analysis).

These analyses supplement and corroborate the previous dose analysis (Relationship between dose and efficacy). While the previous analyses investigated the relationship of dose and efficacy by investigating dose as a continuous variable, these analyses treat dose as a dichotomous variable (lower versus higher, defined based on a median split).

Relationship between dose and change from baseline to follow-up separately within the intervention group and the comparison group

The linear and non-linear model analysis is identical to the *Relationship between dose and efficacy* analysis described above. However, different data is input for the effect size (Hedges' g) and the interpretation differs.

In this case, Hedges' g represents the change between two time-points: baseline (pre-intervention) and follow-up. This effect size was calculated separately for the behaviourally based intervention group and the comparison group. As such, linear and non-linear models of dose by effect size were conducted separately for behavioural intervention and comparison groups.

Dose once again was defined as clinician-delivered hours (total and monthly). Studies were only included in the analyses if dose was reported. For the comparison group, the study had to report clinician-delivered hours of alternative intervention (including those in the community, occupational therapy, speech pathology, etc).

To compare the difference from baseline to follow-up across dose (total and monthly clinician-delivered hours) between behaviourally based intervention and comparison groups, plots were created. These plots included the linear and non-linear models (without 95% confidence intervals) for behaviourally based intervention and comparison groups on the same plot. This was to allow for comparison of effect sizes between the two groups for the same dose amount.

It is important to note that, while plotted in this way, different studies contributed to each analysis (less studies in the comparison group analyses). Additionally, even if studies contributed to both, the dose amount between the behaviourally based intervention group and the comparison group is not necessarily equal.

A6.3 Subgroup analyses

Heterogeneity (variance between studies) was further investigated through subgroup analyses. These involved pooling the effect sizes (Hedges' g), as was done the main analysis described above (beneath the *Data Analysis* header), but now within subgroups. Subgroup analyses assess whether there were any differences in efficacy of behaviourally based interventions based on differences in study design, intervention characteristics, and population characteristics.

All seven subgroups (low versus high dose [outlined in the previous section], person delivering, intervention category, comparison group, age group, primary intervention setting, and study design) were assessed for each of the five outcome domains. Each subgroup level that two or more studies reported data for a particular outcome domain was included in analyses.

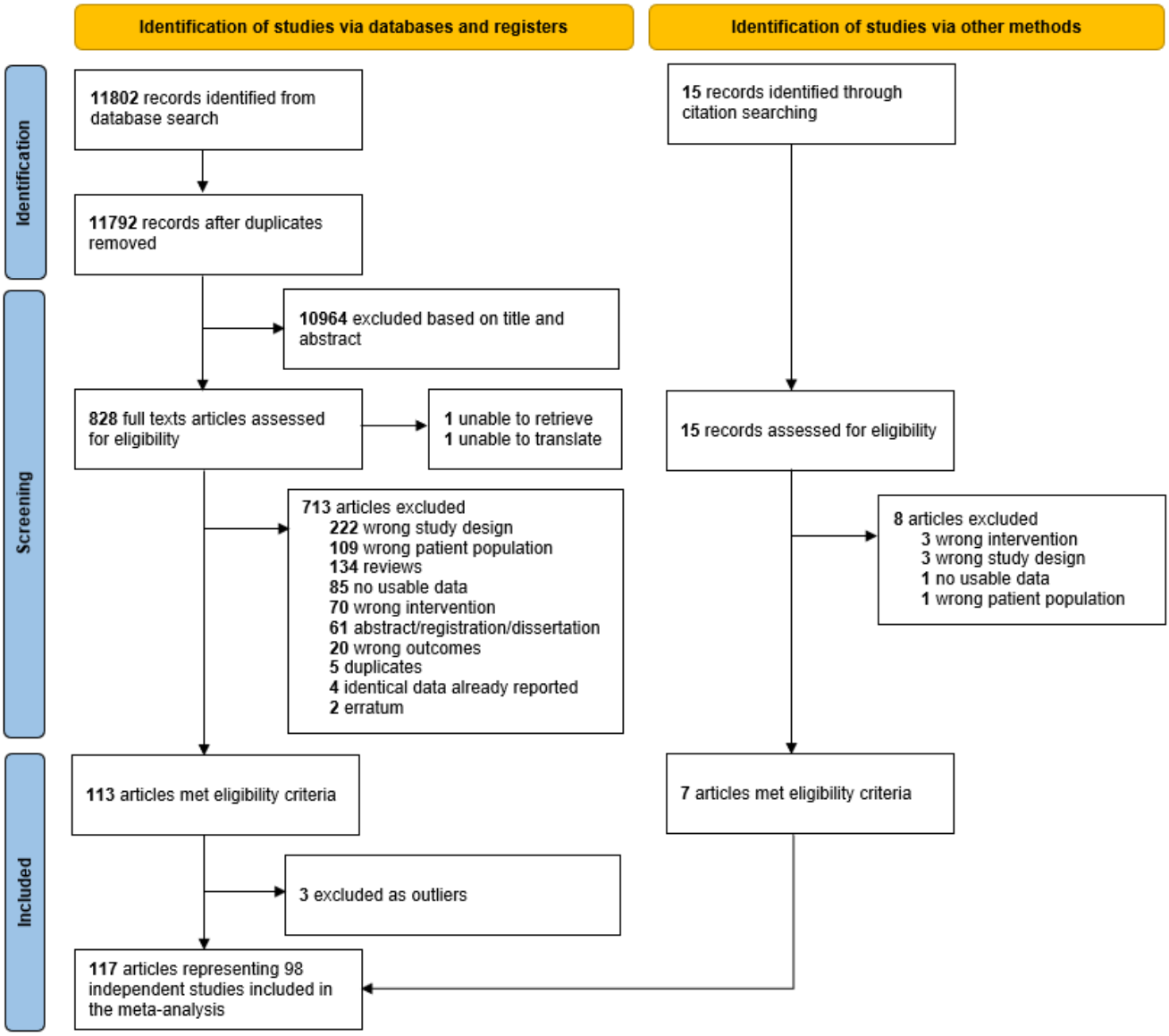
Pooled effect sizes (Hedges' g) and confidence intervals were estimated for each subgroup level individually. Differences in effect size between subgroup levels was then assessed using a Wald-type test. If statistically significant ($p < 0.05$), this indicates that there is a difference in efficacy of the intervention between levels of the subgroup.

Appendix B: Results

B1. Study selection

The initial search identified 11802 records, of which 10 were duplicates. A total of 11792 records were screened based on title an abstract (**Figure B1**). The full-text of 816 records were assessed, of which 113 records met the eligibility criteria. Of 15 articles identified through citation searching, 7 met eligibility criteria. After removing three articles as outliers, 117 records were included in analyses. A total of 35 records reported data from the same overarching study as at least one other record. These 35 records were combined within studies, to form 14 independent studies. The final dataset included 98 independent studies (14 of which included multiple records).

Figure B1. Summary of study selection



B2. Characteristics of included studies

A total of 117 records representing 98 studies were included in the meta-analysis. Details of all included records are shown in **Table B1**. All studies included interventions based on behaviourally based principles, but the characteristics of the interventions varied largely across studies (see **Section 4.1.2** of the main report). The intensity and duration of behaviourally based behavioural interventions ranged between 2.2 to 157.53 clinician-delivered hours per month, delivered over 4 to 141 weeks.

Table B1: Description of included studies

Note: NDBI = Naturalistic Developmental Behavioural Intervention; TAU = treatment as usual; NR = not reported.

Study	Article author (year)	Country	Study design	N	Boys (%)	Age (mths)	Outcome domains reported	Intervention category	Control group	Primary setting	Person delivering
1	Argumedes et al. (2021)	Canada	Random	23	78	51.6	Autism characteristics	Behavioural	TAU	Home	Parent delivered
2	Azarbehi (2012)	Canada	Cohort	20	85	41	Autism characteristics, adaptive functioning, cognition	Behavioural	TAU	Home	Clinician
3	Iadarola, Levato, et al. (2018)	USA	Random	180	87.78	50.4	Autism characteristics, family outcomes, adverse effects	Behavioural	Eclectic	Home	Parent delivered
3	Bearss et al. (2015)	USA	Random	180	87.78	56.4	Autism characteristics	Behavioural	Eclectic	Health	Parent delivered
3	Scahill et al. (2016)	USA	Random	180	87.78	56.4	Adaptive functioning	Behavioural	Eclectic	Health	Parent delivered
4	Bentenuto et al. (2020)	Italy	Non-random	37	NR	41.6	Autism characteristics, cognition	NDBI	TAU	Health	Clinician
5	Bernard-Opitz et al. (2004)	Singapore	Non-random	8	NR	38.75	Autism characteristics	Behavioural	Eclectic	Health	Clinician and parent
6	Blackman et al. (2020)	USA	Non-random	12	66.67	48.48	Autism characteristics, family outcomes, adverse effects	Behavioural	TAU	Health	Parent delivered
7	Bordini et al. (2020)	Brazil	Random	66	80.3	57.6	Autism characteristics, adaptive functioning, cognition	Behavioural	TAU	Health	Parent delivered

Study	Article author (year)	Country	Study design	N	Boys (%)	Age (mths)	Outcome domains reported	Intervention category	Control group	Primary setting	Person delivering
8	Boyd et al. (2014)	USA	Cohort	113	83.19	48.24	Autism characteristics, adaptive functioning, cognition	Other	TAU	Early education	Teacher delivered
8	Coman (2014)	USA	Cohort	144	85.42	48.36	Autism characteristics, adaptive functioning, cognition	TEACCH	TAU	Early education	Teacher delivered
9	Cariveau et al. (2019)	USA	Random	10	NR	34.84	Autism characteristics, cognition	Behavioural	TAU	Health	Clinician
10	Carr and Felce (2007)	UK	Non-random	10	NR	69.6	Autism characteristics, cognition	Behavioural	TAU	Early education	Clinician
11	Chang et al. (2016)	USA	Random	66	89	50.26	Autism characteristics, cognition	NDBI	TAU	Early education	Teacher delivered
12	Charman et al. (2021)	England	Random	62	80.6	80.04	Autism characteristics, family outcomes, adverse effects	Behavioural	TAU	Home	Parent delivered
13	Chiang et al. (2016)	Taiwan	Non-random	34	NR	37.6	Autism characteristics	Developmental	TAU	Health	Clinician
14	Cohen et al. (2006)	USA	Non-random	42	83.33	NR	Adaptive functioning, cognition	Behavioural	TAU	Home	Clinician
15	Coleman (2017)	USA	Random	19	81.5	45.4	Autism characteristics	Behavioural	Eclectic	Home	Parent delivered
16	Colombi et al. (2018)	Italy	Non-random	92	NR	34.22	Autism characteristics, adaptive functioning, cognition	NDBI	TAU	Health	Clinician
17	D'Elia et al. (2014)	Italy	Non-random	30	80	49.2	Autism characteristics, adaptive functioning, cognition, adverse effects	TEACCH	TAU	Early education	Clinician and parent
18	Dai et al. (2018)	Albania	Non-random	29	89.66	45.17	Family outcomes	Behavioural	TAU	Home	Parent delivered

Study	Article author (year)	Country	Study design	N	Boys (%)	Age (mths)	Outcome domains reported	Intervention category	Control group	Primary setting	Person delivering
19	Dawson et al. (2010)	USA	Random	48	NR	23.5	Autism characteristics, adaptive functioning, cognition	NDBI	TAU	Health	Clinician and parent
19	Estes et al. (2015)	USA	Random	48	76.92	23.5	Autism characteristics, adaptive functioning	NDBI	NA	Health	Clinician and parent
20	Dixon et al. (2019)	USA	Non-random	20	84.21	65.4	Cognition	Behavioural	TAU	Health	Clinician
21	Drew et al. (2002)	England	Random	24	79.17	22.5	Autism characteristics, cognition, adverse effects	Developmental	TAU	Home	Parent delivered
22	Duifhuis et al. (2017)	Netherlands	Non-random	47	83.33	69.48	Autism characteristics, adverse effects	NDBI	TAU	Home	Clinician and parent
23	Eikeseth et al. (2002)	Norway	Non-random	25	76	65.68	Adaptive functioning, cognition	Behavioural	Eclectic	Early education	Clinician and parent
24	Eikeseth et al. (2012)	Norway	Non-random	59	83.05	49.2	Adaptive functioning	Behavioural	Eclectic	Early education	Clinician
25	Eldevik et al. (2006)	Norway	Cohort	28	85.71	50.86	Autism characteristics, adaptive functioning, cognition	Behavioural	Eclectic	Early education	Clinician and parent
26	Eldevik et al. (2010)	Norway	Cohort	25	76	49.64	Adaptive functioning, cognition	Behavioural	Eclectic	Early education	Clinician
27	Eldevik et al. (2012)	Norway	Cohort	43	76.74	43.32	Adaptive functioning, cognition	Behavioural	Eclectic	Early education	Clinician
28	Elder (2012)	USA	Random	97	77	21	Autism characteristics, family outcomes	NDBI	TAU	Health	Parent delivered
28	Estes et al. (2014)	USA	Random	82	75.61	21.01	Adverse effects	NDBI	TAU	Health	Parent delivered

Study	Article author (year)	Country	Study design	N	Boys (%)	Age (mths)	Outcome domains reported	Intervention category	Control group	Primary setting	Person delivering
29	Fava et al. (2011)	Italy	Non-random	22	86.36	48.23	Autism characteristics, adaptive functioning, cognition, adverse effects	Behavioural	Eclectic	Home	Clinician and parent
30	Felzer-Kim and Hauck (2020)	USA	Random	14	71.43	53.86	Cognition	Behavioural	Eclectic	Health	Clinician
31	Feng et al. (2019)	China	Non-random	67	78.46	28.46	Autism characteristics	NDBI	TAU	Health	Clinician
32	Flanagan (2011)	Canada	Cohort	134	NR	35.2	Autism characteristics, adaptive functioning	Behavioural	TAU	Community	Clinician
33	Fox (2018)	USA	Random	10	70	32.8	Autism characteristics, cognition, family outcomes, adverse effects	NDBI	TAU	Health	Parent delivered
34	Frey et al. (2015)	USA	Random	34	NR	49.2	Autism characteristics, adaptive functioning	Behavioural	TAU	Early education	Clinician and parent
35	Furukawa et al. (2018)	Japan	Non-random	21	81	62.88	Autism characteristics, adverse effects	Other	TAU	Health	Parent delivered
36	Gengoux et al. (2019)	USA	Random	43	88.37	48.43	Autism characteristics, adaptive functioning, cognition	NDBI	TAU	Home	Clinician and parent
37	Gengoux et al. (2021)	USA	Random	44	95.45	60	Autism characteristics, adaptive functioning	Developmental	TAU	Health	Clinician
38	Ginn et al. (2017)	USA	Random	30	80	56.4	Autism characteristics, cognition, adverse effects	Other	TAU	Health	Parent delivered
39	Gomes et al. (2019)	Brazil	Cohort	33	87.5	59.3	Adaptive functioning, cognition	Behavioural	TAU	Home	Parent delivered

Study	Article author (year)	Country	Study design	N	Boys (%)	Age (mths)	Outcome domains reported	Intervention category	Control group	Primary setting	Person delivering
40	Goods et al. (2013)	USA	Random	15	NR	51.9	Autism characteristics, cognition	NDBI	TAU	Early education	Clinician
41	Grahame et al. (2015)	UK	Random	45	86.67	61.47	Autism characteristics, cognition, family outcomes	Behavioural	TAU	Health	Parent delivered
42	Grindle et al. (2012)	UK	Cohort	29	86.21	61.7	Adaptive functioning, cognition	Behavioural	TAU	Early education	Clinician and parent
43	Gulsrud et al. (2019)	USA	Random	20	65	49.36	Autism characteristics, cognition	NDBI	TAU	Early education	Clinician
44	Haglund et al. (2021)	Sweden	Non-random	94	81.91	51.6	Autism characteristics	NDBI	TAU	Early education	Clinician and parent
45	Hampton et al. (2020)	USA	Random	68	77.94	43	Autism characteristics, cognition	NDBI	TAU	Health	Clinician and parent
46	Haraguchi et al. (2020)	Japan	Non-random	61	85.25	46.75	Autism characteristics, adaptive functioning, cognition, adverse effects	Behavioural	TAU	Health	Clinician and parent
47	Hardan et al. (2015)	USA	Random	48	75	49.2	Autism characteristics, adaptive functioning, cognition	NDBI	TAU	Health	Parent delivered
48	Ho and Lin (2020)	Taiwan	Random	24	100	48.5	Autism characteristics, adaptive functioning, cognition	Developmental	TAU	Home	Parent delivered
49	Holzinger et al. (2019)	Austria	Non-random	13	100	43.32	Autism characteristics, adaptive functioning, cognition, family outcomes, adverse effects	NDBI	TAU	Home	Clinician

Study	Article author (year)	Country	Study design	N	Boys (%)	Age (mths)	Outcome domains reported	Intervention category	Control group	Primary setting	Person delivering
50	Howard et al. (2005)	USA	Non-random	45	84.44	32.17	Adaptive functioning, cognition	Behavioural	TAU	Early education	Clinician and parent
50	Howard et al. (2014)	USA	Non-random	45	NR	32.24	Adaptive functioning, cognition	Behavioural	TAU	Early education	Clinician and parent
51	Iadarola, Shih, et al. (2018)	USA	Random	150	87.33	85.14	Autism characteristics, adaptive functioning	Behavioural	TAU	Early education	Teacher delivered
52	Ingersoll (2010)	USA	Random	21	85.71	39.38	Autism characteristics	NDBI	TAU	Health	Clinician
52	Ingersoll (2012)	USA	Random	27	88.89	37.95	Autism characteristics	NDBI	TAU	Health	Clinician
53	Johnson et al. (2019)	USA	Random	42	95.24	61.2	Autism characteristics, adaptive functioning, family outcomes, adverse effects	Behavioural	TAU	Health	Parent delivered
54	Jouen et al. (2017)	France	Non-random	24	100	83.76	Autism characteristics, adaptive functioning, cognition, adverse effects	Technology-based	TAU	Health	Clinician and parent
55	Kaale et al. (2012)	Norway	Random	61	78.69	48.8	Autism characteristics	Developmental	TAU	Early education	Teacher delivered
55	Kaale et al. (2014)	Norway	Random	61	78.69	48.8	Autism characteristics, cognition	Developmental	TAU	Early education	Teacher delivered
56	Arora (2008)	USA	Random	36	NR	42.62	Autism characteristics	Developmental	TAU	Early education	Clinician
56	Kasari et al. (2006)	USA	Random	38	81.08	42.34	Autism characteristics	Developmental	TAU	Early education	Clinician

Study	Article author (year)	Country	Study design	N	Boys (%)	Age (mths)	Outcome domains reported	Intervention category	Control group	Primary setting	Person delivering
56	Kasari et al. (2008)	USA	Random	38	81.08	42.34	Autism characteristics, cognition	Developmental	TAU	Early education	Clinician
56	Lawton and Kasari (2012a)	USA	Random	36	78.12	41.36	Autism characteristics	Developmental	Eclectic	Early education	Clinician
57	Kasari et al. (2010)	USA	Random	38	76.32	30.83	Autism characteristics	Developmental	TAU	Health	Parent delivered
58	Kasari et al. (2015)	USA	Random	86	81	31.5	Autism characteristics, cognition, adverse effects	NDBI	Eclectic	Health	Parent delivered
58	Schlink et al. (2022)	USA	Random	86	81	31.5	Adverse effects	NDBI	Eclectic	Health	Parent delivered
58	Gulsrud et al. (2016)	USA	Random	86	81	31.5	Family outcomes	NDBI	Eclectic	Health	Parent delivered
58	Dimachkie (2021)	USA	Random	75	82.7	31.5	Autism characteristics	NDBI	Eclectic	Health	Parent delivered
59	Lawton and Kasari (2012b)	USA	Random	16	NR	44.69	Autism characteristics	NDBI	TAU	Early education	Teacher delivered
60	Leaf et al. (2017)	USA	Random	15	NR	56.4	Autism characteristics	Behavioural	TAU	Health	Clinician
61	Magiati et al. (2007)	UK	Cohort	44	NR	39.64	Autism characteristics, adaptive functioning, cognition	Behavioural	Eclectic	Home	Parent delivered
62	Manohar et al. (2019)	India	Random	50	88	41.4	Autism characteristics, family outcomes, adverse effects	NDBI	TAU	Home	Parent delivered
63	Matthews et al. (2018)	USA	Non-random	18	97.22	40.78	Autism characteristics, family outcomes, adverse effects	Behavioural	TAU	Health	Parent delivered

Study	Article author (year)	Country	Study design	N	Boys (%)	Age (mths)	Outcome domains reported	Intervention category	Control group	Primary setting	Person delivering
64	Nojiri and Yanagawa (2019)	Japan	Random	36	86.11	43.4	Autism characteristics, family outcomes	Behavioural	TAU	Health	Parent delivered
65	Nowell et al. (2019)	USA	Random	17	76.47	81.84	Autism characteristics, family outcomes	TEACCH	TAU	Health	Clinician and parent
66	Oosterling et al. (2010)	The Netherlands	Non-random	67	77.61	34.32	Autism characteristics, cognition, family outcomes	Developmental	TAU	Home	Parent delivered
67	Pajareya and Nopmaneejumruslers (2011)	Thailand	Random	32	87.5	54.05	Autism characteristics, adaptive functioning	Developmental	TAU	Home	Parent delivered
68	Peters-Scheffer et al. (2013)	The Netherlands	Cohort	40	90	62.52	Autism characteristics, adaptive functioning, cognition, adverse effects	Behavioural	TAU	Early education	Clinician
69	Reitzel et al. (2013)	Canada	Random	15	NR	58.5	Autism characteristics, adaptive functioning, family outcomes, adverse effects	Behavioural	TAU	Health	Clinician
70	Remington et al. (2007)	England	Cohort	44	NR	36.99	Autism characteristics, adaptive functioning, cognition, family outcomes, adverse effects	Behavioural	TAU	Home	Clinician and parent
70	Kovshoff et al. (2011)	England	Cohort	44	NR	36.99	Autism characteristics, adaptive functioning, cognition	Behavioural	TAU	Home	Clinician and parent
71	Rogers et al. (2006)	USA	Random	10	100	38.4	Cognition	NDBI	Eclectic	Health	Clinician and parent

Study	Article author (year)	Country	Study design	N	Boys (%)	Age (mths)	Outcome domains reported	Intervention category	Control group	Primary setting	Person delivering
72	Rogers et al. (2019)	USA	Random	118	77.97	21.02	Autism characteristics, adaptive functioning, cognition	NDBI	TAU	Home	Parent delivered
72	Rogers et al. (2012)	USA	Random	98	77.55	20.98	Autism characteristics, adaptive functioning, cognition	NDBI	TAU	Health	Parent delivered
73	Rogers et al. (2014)	USA	Non-random	14	63.64	9	Autism characteristics, cognition	NDBI	TAU	Health	Parent delivered
74	Ruiz (2020)	USA	Random	40	97.5	60.38	Autism characteristics	NDBI	TAU	Community	Clinician
75	Shawler (2017)	USA	Random	51	86.27	27.69	Cognition	Behavioural	TAU	NA	Clinician and parent
76	Sheinkopf and Siegel (1998)	USA	Cohort	22	NR	34.55	Autism characteristics, cognition	Behavioural	TAU	Home	Clinician
77	Shire et al. (2017)	USA	Random	113	77.88	31.63	Autism characteristics	NDBI	TAU	Community	Teacher delivered
78	Sinai-Gavrilov et al. (2020)	Israel	Non-random	51	82.35	44.37	Adaptive functioning, cognition	NDBI	Eclectic	Early education	Teacher delivered
79	Solomon et al. (2014)	USA	Random	128	82.03	50.19	Autism characteristics, cognition, family outcomes, adverse effects	Developmental	TAU	Home	Parent delivered
80	Spjut Jansson et al. (2016)	Sweden	Cohort	52	72.5	36.09	Autism characteristics, adaptive functioning	Developmental	Eclectic	Home	Clinician
81	Stadnick et al. (2015)	USA	Non-random	30	80	54.83	Autism characteristics, adaptive functioning, adverse effects	NDBI	TAU	Community	Parent delivered

Study	Article author (year)	Country	Study design	N	Boys (%)	Age (mths)	Outcome domains reported	Intervention category	Control group	Primary setting	Person delivering
82	Stahmer et al. (2020)	USA	Non-random	25	68	22.76	Autism characteristics, adaptive functioning, cognition, family outcomes	NDBI	TAU	Community	Parent delivered
83	Strauss et al. (2012)	NA	Non-random	44	93.18	49.43	Autism characteristics, adaptive functioning, cognition, adverse effects	Behavioural	Eclectic	Home	Clinician and parent
84	Sullivan (2014)	USA	Random	48	77.08	23.5	Autism characteristics, cognition	NDBI	TAU	Home	Clinician and parent
85	Tonge et al. (2006)	Australia	Random	70	82.86	46.41	Family outcomes, adverse effects	Behavioural	TAU	Health	Parent delivered
85	Tonge et al. (2014)	Australia	Random	70	82.86	46.67	Autism characteristics, adaptive functioning, cognition	Behavioural	TAU	Health	Parent delivered
86	Tsang et al. (2007)	Hong Kong	Non-random	34	85.29	48.68	Autism characteristics, adaptive functioning, cognition	TEACCH	TAU	Early education	Teacher delivered
87	Van der Paelt et al. (2016)	Belgium	Cohort	55	80	47.41	Autism characteristics, adaptive functioning, cognition	Behavioural	Eclectic	Community	Clinician
87	Van der Paelt et al. (2016)	Belgium	Cohort	65	81.54	50.34	Autism characteristics, adaptive functioning, cognition	Developmental	Eclectic	Community	Clinician
88	Vernon et al. (2019)	USA	Random	23	86.96	35.13	Autism characteristics, adaptive functioning, cognition	NDBI	TAU	Community	Clinician and parent
88	Barrett et al. (2020)	USA	Random	21	NR	36.8	Autism characteristics, cognition	NDBI	TAU	Home	Clinician and parent

Study	Article author (year)	Country	Study design	N	Boys (%)	Age (mths)	Outcome domains reported	Intervention category	Control group	Primary setting	Person delivering
89	Vinen et al. (2018)	Australia	Cohort	59	88.14	37.4	Autism characteristics, cognition	NDBI	Eclectic	Community	Clinician
90	Vivanti et al. (2014)	Australia	Cohort	57	87.72	41.18	Autism characteristics, adaptive functioning, cognition	NDBI	Eclectic	Community	Clinician
91	Warreyn and Roeyers (2014)	Belgium	Random	36	75	77.9	Autism characteristics	Developmental	TAU	Health	Clinician
92	Waters et al. (2020)	USA	Non-random	94	95.74	40.1	Adaptive functioning, cognition	Behavioural	TAU	Community	Clinician and parent
93	Whalen et al. (2010)	USA	Random	24	NR	NR	Cognition	Technology-based	TAU	Early education	Teacher delivered
94	Whitehouse et al. (2017)	Australia	Random	75	78.75	39.78	Autism characteristics, adaptive functioning, cognition	Technology-based	TAU	Home	Parent delivered
95	Xu et al. (2018)	China	Random	36	88.89	44.94	Autism characteristics	NDBI	Eclectic	Early education	Teacher delivered
95	Xu et al. (2017)	China	Random	36	94.44	44.94	Autism characteristics, cognition, adverse effects	NDBI	TAU	Early education	Teacher delivered
96	Zachor et al. (2007)	NA	Cohort	39	94.87	28.24	Autism characteristics, cognition	Behavioural	Eclectic	Health	Clinician
97	Zachor and Ben Itzchak (2010)	Israel	Cohort	78	91.03	25.4	Adaptive functioning, cognition	Behavioural	Eclectic	Early education	Clinician and parent
98	Zhou et al. (2018)	China	Non-random	43	88.37	26.55	Autism characteristics, cognition, adverse effects	NDBI	TAU	Health	Parent delivered

B3. Quality of the evidence used within this report

B3.1 Randomised controlled trials

The assessed risk of bias level (low risk, some concerns, high risk) for each of the five domains within the Cochrane RoB 2.0 tool (Sterne et al., 2019) for each of the randomised controlled trial studies included within this report are displayed in **Table B2**.

Table B2: Domain and overall risk of bias assessments for included randomised controlled trials using Cochrane RoB 2.0

Notes: Study refers to the IDs in Table B1.

Low risk assessments are highlighted in green.

Risk assessments of some concern are highlighted in yellow.

High risk assessments are highlighted in red.

Study	Randomisation process	Deviations from intended interventions	Missing outcome data	Measurement of the outcome	Selection of the reported result	Overall
1	Some concerns	Low risk	Low risk	Low risk	Low risk	Some concerns
3	Low risk	Low risk	Low risk	Low risk	Some concerns	Some concerns
7	High risk	Low risk	Low risk	Low risk	Low risk	High risk
9	Low risk	Low risk	Low risk	Some concerns	Some concerns	Some concerns
11	Some concerns	Some concerns	Low risk	Low risk	Some concerns	Some concerns
12	Some concerns	Low risk	Low risk	High risk	Low risk	High risk
15	Some concerns	Low risk	Low risk	High risk	Some concerns	High risk
19	Some concerns	Low risk	Low risk	Low risk	Some concerns	Some concerns
21	High risk	Low risk	Some concerns	Some concerns	Some concerns	High risk
28	Low risk	Low risk	Low risk	Low risk	Some concerns	Some concerns
30	Some concerns	Low risk	Low risk	Low risk	Some concerns	Some concerns
33	Low risk	Low risk	Some concerns	Some concerns	Some concerns	Some concerns
34	Some concerns	Some concerns	Low risk	High risk	Some concerns	High risk
36	Low risk	Low risk	Low risk	Low risk	High risk	High risk

Study	Randomisation process	Deviations from intended interventions	Missing outcome data	Measurement of the outcome	Selection of the reported result	Overall
37	Low risk	Low risk	Low risk	Low risk	Some concerns	Some concerns
38	Some concerns	Low risk	Some concerns	High risk	Some concerns	High risk
40	Low risk	Low risk	Some concerns	Low risk	Some concerns	Some concerns
41	Low risk	Low risk	Low risk	High risk	Some concerns	High risk
43	Low risk	Low risk	Low risk	Low risk	Some concerns	Some concerns
45	Low risk	Low risk	Some concerns	Low risk	Some concerns	Some concerns
47	Low risk	Low risk	Low risk	High risk	Low risk	High risk
48	Some concerns	Low risk	Low risk	High risk	Some concerns	High risk
51	Low risk	Low risk	Low risk	High risk	Some concerns	High risk
52	High risk	Low risk	Low risk	Low risk	High risk	High risk
53	Low risk	Some concerns	High risk	High risk	Some concerns	High risk
55	Low risk	Low risk	Low risk	Low risk	Some concerns	Some concerns
56	Some concerns	Low risk	Low risk	Low risk	Some concerns	Some concerns
57	High risk	Low risk	Low risk	Low risk	Some concerns	High risk
58	Low risk	Low risk	High risk	High risk	Some concerns	High risk
59	Low risk	Low risk	Low risk	Low risk	Some concerns	Some concerns
60	Some concerns	Low risk	Low risk	Low risk	Some concerns	Some concerns
62	Low risk	Low risk	Low risk	High risk	Some concerns	High risk
64	Low risk	Low risk	Some concerns	High risk	Some concerns	High risk
65	Some concerns	Low risk	Low risk	Low risk	Some concerns	Some concerns
67	High risk	Low risk	Low risk	Low risk	Some concerns	High risk
69	High risk	Low risk	Some concerns	Some concerns	Some concerns	High risk
71	High risk	Low risk	Low risk	Low risk	Some concerns	High risk

Study	Randomisation process	Deviations from intended interventions	Missing outcome data	Measurement of the outcome	Selection of the reported result	Overall
72	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk
74	Low risk	Low risk	High risk	Low risk	Some concerns	High risk
75	Some concerns	Low risk	Low risk	High risk	Some concerns	High risk
77	Low risk	Low risk	Low risk	Low risk	Some concerns	Some concerns
79	Low risk	Low risk	Some concerns	High risk	Low risk	High risk
84	High risk	Low risk	High risk	Low risk	Some concerns	High risk
85	Low risk	Low risk	Low risk	Low risk	Some concerns	Some concerns
88	Low risk	Low risk	Low risk	Low risk	Some concerns	Some concerns
91	High risk	High risk	High risk	Low risk	Some concerns	High risk
93	High risk	Low risk	Low risk	Some concerns	Some concerns	High risk
94	Some concerns	Low risk	Low risk	High risk	Some concerns	High risk
95	Low risk	Low risk	Low risk	High risk	Some concerns	High risk

B3.2 Non-randomised study designs

The assessed risk of bias level (low risk, moderate risk, serious risk) for each of the five domains within the ROBINS-I tool (Sterne et al., 2016) For each of the non-randomised studies included within this report are displayed in **Table B4**.

Table B3: Domain and overall risk of bias assessments for included non-randomised studies using ROBINS-I

Notes: Study refers to the IDs in Table B1.

Low risk assessments are highlighted in green.

Moderate risk assessments are highlighted in yellow.

Serious risk assessments are highlighted in red.

Study	Bias due to confounding	Bias in selection of participants	Bias in classification of interventions	Bias due to deviations from intended interventions	Bias due to missing data	Bias in measurement of outcomes	Bias in selection of the reported result	Overall
2	Moderate risk	Low risk	Serious risk	Low risk	Serious risk	Serious risk	Serious risk	Serious risk
4	Serious risk	Low risk	Low risk	Low risk	Low risk	Low risk	Moderate risk	Serious risk
5	Moderate risk	Moderate risk	Low risk	Low risk	Low risk	Low risk	Serious risk	Serious risk
6	Serious risk	Low risk	Low risk	Low risk	Serious risk	Serious risk	Moderate risk	Serious risk
8	Moderate risk	Low risk	Low risk	Low risk	Low risk	Serious risk	Moderate risk	Serious risk
10	Serious risk	Low risk	Low risk	Low risk	Serious risk	Serious risk	Serious risk	Serious risk
13	Moderate risk	Low risk	Low risk	Low risk	Serious risk	Low risk	Moderate risk	Serious risk
14	Moderate risk	Low risk	Low risk	Low risk	Moderate risk	Low risk	Moderate risk	Moderate risk
16	Serious risk	Low risk	Low risk	Low risk	Low risk	Serious risk	Moderate risk	Serious risk

Study	Bias due to confounding	Bias in selection of participants	Bias in classification of interventions	Bias due to deviations from intended interventions	Bias due to missing data	Bias in measurement of outcomes	Bias in selection of the reported result	Overall
17	Serious risk	Serious risk	Low risk	Low risk	Low risk	Low risk	Moderate risk	Serious risk
18	Serious risk	Low risk	Low risk	Low risk	Serious risk	Serious risk	Moderate risk	Serious risk
20	Serious risk	Low risk	Low risk	Low risk	Low risk	Serious risk	Moderate risk	Serious risk
22	Serious risk	Low risk	Low risk	Low risk	Low risk	Serious risk	Moderate risk	Serious risk
23	Serious risk	Serious risk	Low risk	Low risk	Low risk	Low risk	Moderate risk	Serious risk
24	Serious risk	Serious risk	Low risk	Low risk	Serious risk	Serious risk	Moderate risk	Serious risk
25	Serious risk	Low risk	Moderate risk	Low risk	Low risk	Low risk	Moderate risk	Serious risk
26	Serious risk	Low risk	Low risk	Low risk	Low risk	Serious risk	Moderate risk	Serious risk
27	Moderate risk	Low risk	Low risk	Low risk	Low risk	Low risk	Moderate risk	Moderate risk
29	Moderate risk	Low risk	Low risk	Low risk	Low risk	Serious risk	Moderate risk	Serious risk
31	Moderate risk	Low risk	Low risk	Low risk	Low risk	Serious risk	Moderate risk	Serious risk
32	Moderate risk	Low risk	Low risk	Low risk	Moderate risk	Serious risk	Moderate risk	Serious risk
35	Moderate risk	Low risk	Low risk	Low risk	Moderate risk	Serious risk	Moderate risk	Serious risk
39	Serious risk	Low risk	Low risk	Low risk	Low risk	Serious risk	Moderate risk	Serious risk

Study	Bias due to confounding	Bias in selection of participants	Bias in classification of interventions	Bias due to deviations from intended interventions	Bias due to missing data	Bias in measurement of outcomes	Bias in selection of the reported result	Overall
42	Moderate risk	Low risk	Low risk	Low risk	Serious risk	Low risk	Moderate risk	Serious risk
44	Serious risk	Low risk	Low risk	Low risk	Low risk	Low risk	Moderate risk	Serious risk
46	Moderate risk	Low risk	Low risk	Low risk	Low risk	Low risk	Moderate risk	Moderate risk
49	Moderate risk	Low risk	Low risk	Low risk	Low risk	Moderate risk	Moderate risk	Moderate risk
50	Serious risk	Low risk	Low risk	Low risk	Low risk	Serious risk	Moderate risk	Serious risk
54	Moderate risk	Low risk	Low risk	Low risk	Low risk	Serious risk	Moderate risk	Serious risk
61	Moderate risk	Low risk	Low risk	Low risk	Low risk	Serious risk	Moderate risk	Serious risk
63	Serious risk	Low risk	Low risk	Low risk	Serious risk	Serious risk	Moderate risk	Serious risk
66	Moderate risk	Low risk	Low risk	Low risk	Low risk	Serious risk	Moderate risk	Serious risk
68	Moderate risk	Low risk	Low risk	Low risk	Moderate risk	Low risk	Moderate risk	Moderate risk
70	Moderate risk	Low risk	Low risk	Low risk	Low risk	Serious risk	Moderate risk	Serious risk
73	Moderate risk	Moderate risk	Low risk	Low risk	Low risk	Low risk	Moderate risk	Moderate risk
76	Moderate risk	Low risk	Low risk	Low risk	Low risk	Low risk	Moderate risk	Moderate risk
78	Moderate risk	Low risk	Low risk	Low risk	Low risk	Low risk	Moderate risk	Moderate risk

Study	Bias due to confounding	Bias in selection of participants	Bias in classification of interventions	Bias due to deviations from intended interventions	Bias due to missing data	Bias in measurement of outcomes	Bias in selection of the reported result	Overall
80	Moderate risk	Low risk	Low risk	Low risk	Moderate risk	Low risk	Moderate risk	Moderate risk
81	Moderate risk	Low risk	Low risk	Low risk	Low risk	Serious risk	Moderate risk	Serious risk
82	Serious risk	Low risk	Low risk	Low risk	Serious risk	Moderate risk	Moderate risk	Serious risk
83	Moderate risk	Low risk	Low risk	Low risk	Low risk	Serious risk	Moderate risk	Serious risk
86	Moderate risk	Low risk	Low risk	Low risk	Serious risk	Serious risk	Moderate risk	Serious risk
87	Moderate risk	Low risk	Low risk	Low risk	Serious risk	Serious risk	Moderate risk	Serious risk
89	Moderate risk	Low risk	Low risk	Low risk	Low risk	Low risk	Moderate risk	Moderate risk
90	Moderate risk	Low risk	Low risk	Low risk	Low risk	Serious risk	Moderate risk	Serious risk
92	Moderate risk	Low risk	Low risk	Moderate risk	Low risk	Low risk	Moderate risk	Moderate risk
96	Moderate risk	Low risk	Low risk	Low risk	Low risk	Serious risk	Moderate risk	Serious risk
97	Moderate risk	Low risk	Low risk	Low risk	Low risk	Serious risk	Moderate risk	Serious risk
98	Moderate risk	Low risk	Low risk	Low risk	Serious risk	Serious risk	Moderate risk	Serious risk

B4. Efficacy across all and within individual outcome domains

B4.1 All outcome measures

The analysis of all outcome measures included 98 studies. The combined effect size was small and significant (g = 0.32, 95% CI 0.26 to 0.38, $tau^2 = 0.11$; **Figure B2**). The funnel plot did indicate evidence of small study effect (**Figure B3**), which was confirmed through formal testing (Egger's intercept = 1.22, p = 0.002). Adjusting for this effect (imputing 19 studies) resulted in a reduction in effect size (Hedges' g = 0.23, 95%CI 0.17 – 0.29, p < 0.001), although still small and statistically significant.

Figure B2.1. Forest plot of all outcome measures

Note: Figures B2.1-B2.3 comprise one figure, displayed across multiple pages to ensure readability. An accessible version of the data displayed in this figure is presented in Table B4 below.

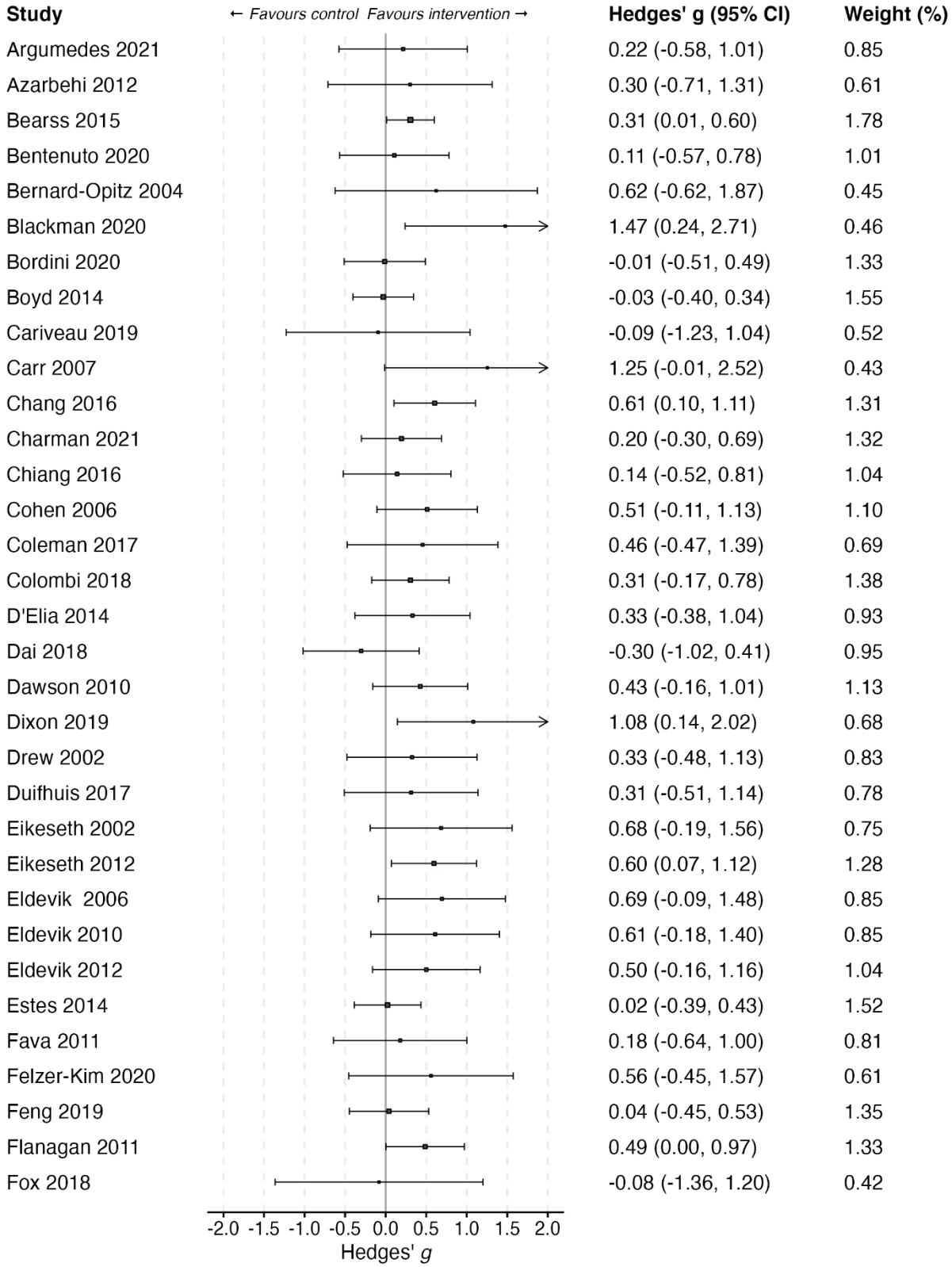


Figure B2.2. Forest plot of all outcomes

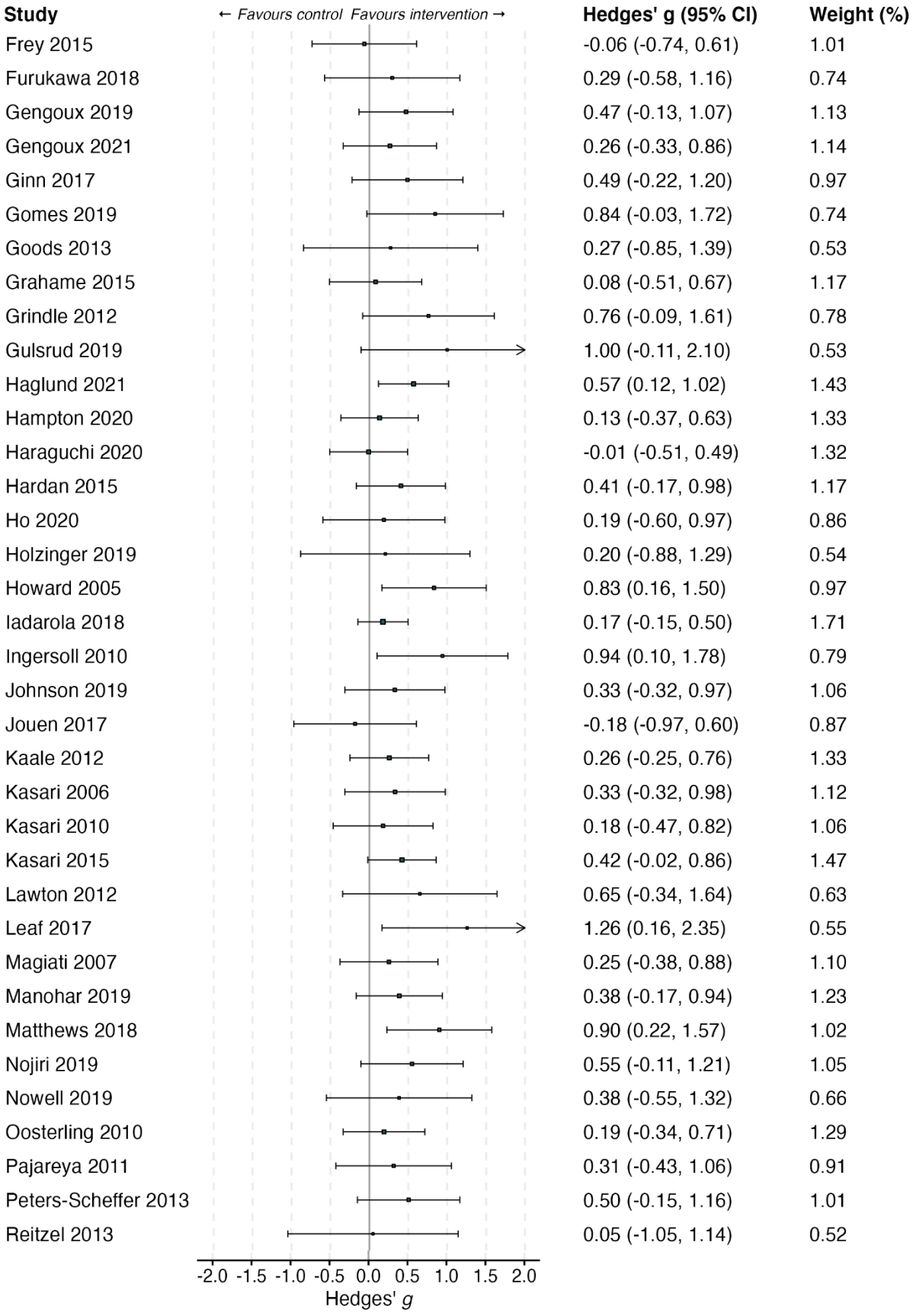


Figure B2.3. Forest plot of all outcomes

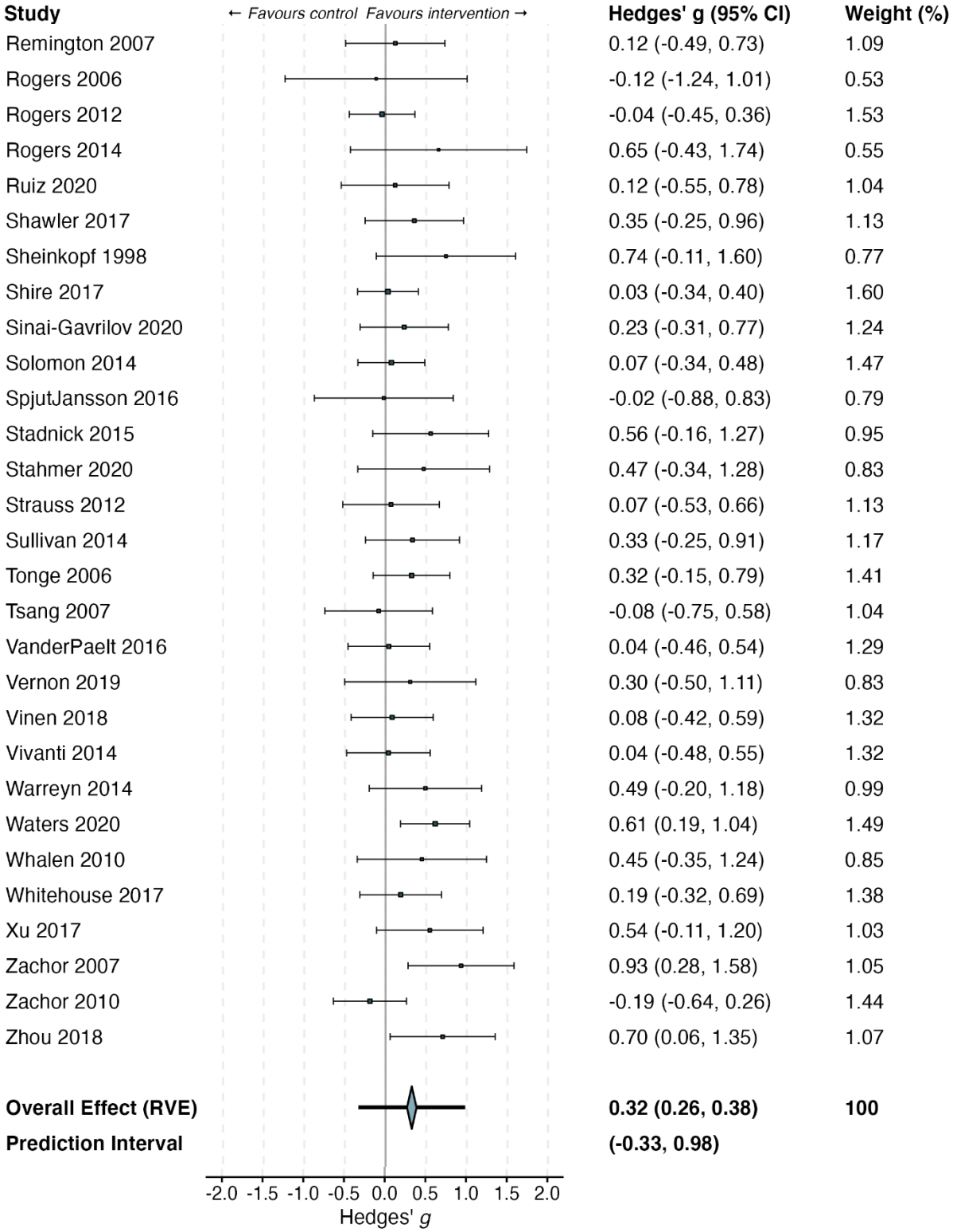


Table B4. Table version of forest plot of all outcome measures.

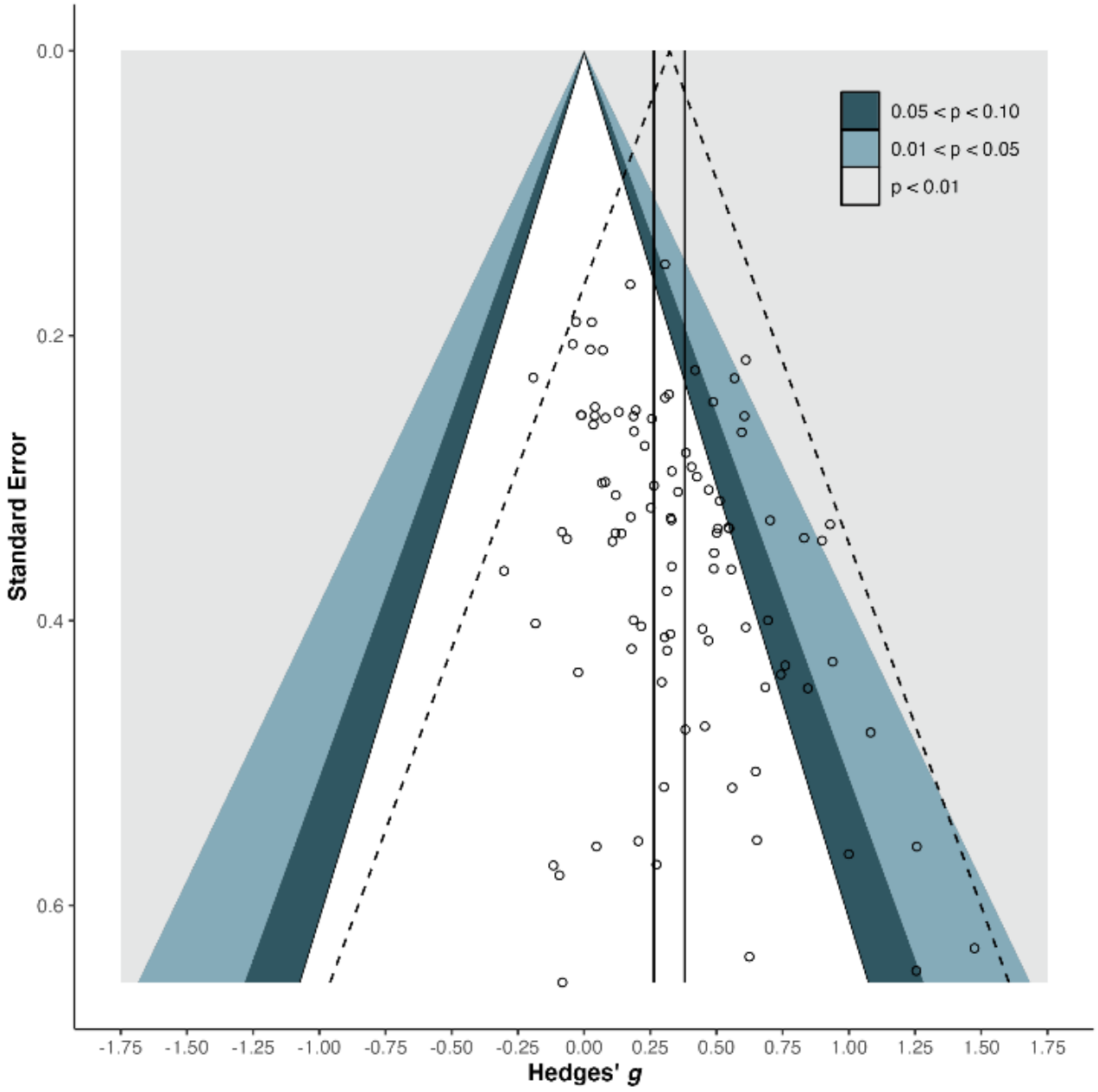
Note: This table presents the information displayed in Figure B2 in an accessible format. Positive Hedges' g values favour the behaviourally based intervention, negative Hedges' g values favour the comparison group.

Study	Hedges' g (95%CI)	Weight (%)
Argumedes 2021	0.22 (-0.58, 1.01)	0.85
Azarbehi 2012	0.30 (-0.71, 1.31)	0.61
Bearss 2015	0.31 (0.01, 0.60)	1.78
Bentenuto 2020	0.11 (-0.57, 0.78)	1.01
Bernard-Opitz 2004	0.62 (-0.62, 1.87)	0.45
Blackman 2020	1.47 (0.24, 2.71)	0.46
Bordini 2020	-0.01 (-0.51, 0.49)	1.33
Boyd 2014	-0.03 (-0.40, 0.34)	1.55
Cariveau 2019	-0.09 (-1.23, 1.04)	0.52
Carr 2007	1.25 (-0.01, 2.52)	0.43
Chang 2016	0.61 (0.10, 1.11)	1.31
Charman 2021	0.20 (-0.30, 0.69)	1.32
Chiang 2016	0.14 (-0.52, 0.81)	1.04
Cohen 2006	0.51 (-0.11, 1.13)	1.1
Coleman 2017	0.46 (-0.47, 1.39)	0.69
Colombi 2018	0.31 (-0.17, 0.78)	1.38
D'Elia 2014	0.33 (-0.38, 1.04)	0.93
Dai 2018	-0.30 (-1.02, 0.41)	0.95
Dawson 2010	0.43 (-0.16, 1.01)	1.13
Dixon 2019	1.08 (0.14, 2.02)	0.68
Drew 2002	0.33 (-0.48, 1.13)	0.83
Duifhuis 2017	0.31 (-0.51, 1.14)	0.78
Eikeseth 2002	0.68 (-0.19, 1.56)	0.75
Eikeseth 2012	0.60 (0.07, 1.12)	1.28
Eldevik 2006	0.69 (-0.09, 1.48)	0.85
Eldevik 2010	0.61 (-0.18, 1.40)	0.85
Eldevik 2012	0.50 (-0.16, 1.16)	1.04
Estes 2014	0.02 (-0.39, 0.43)	1.52
Fava 2011	0.18 (-0.64, 1.00)	0.81
Felzer-Kim 2020	0.56 (-0.45, 1.57)	0.61
Feng 2019	0.04 (-0.45, 0.53)	1.35
Flanagan 2011	0.49 (0.00, 0.97)	1.33
Fox 2018	-0.08 (-1.36, 1.20)	0.42
Frey 2015	-0.06 (-0.74, 0.61)	1.01
Furukawa 2018	0.29 (-0.58, 1.16)	0.74
Gengoux 2019	0.47 (-0.13, 1.07)	1.13
Gengoux 2021	0.26 (-0.33, 0.86)	1.14
Ginn 2017	0.49 (-0.22, 1.20)	0.97

Study	Hedges' g (95%CI)	Weight (%)
Gomes 2019	0.84 (-0.03, 1.72)	0.74
Goods 2013	0.27 (-0.85, 1.39)	0.53
Grahame 2015	0.08 (-0.51, 0.67)	1.17
Grindle 2012	0.76 (-0.09, 1.61)	0.78
Gulsrud 2019	1.00 (-0.11, 2.10)	0.53
Haglund 2021	0.57 (0.12, 1.02)	1.43
Hampton 2020	0.13 (-0.37, 0.63)	1.33
Haraguchi 2020	-0.01 (-0.51, 0.49)	1.32
Hardan 2015	0.41 (-0.17, 0.98)	1.17
Ho 2020	0.19 (-0.60, 0.97)	0.86
Holzinger 2019	0.20 (-0.88, 1.29)	0.54
Howard 2005	0.83 (0.16, 1.50)	0.97
Iadarola 2018	0.17 (-0.15, 0.50)	1.71
Ingersoll 2010	0.94 (0.10, 1.78)	0.79
Johnson 2019	0.33 (-0.32, 0.97)	1.06
Jouen 2017	-0.18 (-0.97, 0.60)	0.87
Kaale 2012	0.26 (-0.25, 0.76)	1.33
Kasari 2006	0.33 (-0.32, 0.98)	1.12
Kasari 2010	0.18 (-0.47, 0.82)	1.06
Kasari 2015	0.42 (-0.02, 0.86)	1.47
Lawton 2012	0.65 (-0.34, 1.64)	0.63
Leaf 2017	1.26 (0.16, 2.35)	0.55
Magiati 2007	0.25 (-0.38, 0.88)	1.1
Manohar 2019	0.38 (-0.17, 0.94)	1.23
Matthews 2018	0.90 (0.22, 1.57)	1.02
Nojiri 2019	0.55 (-0.11, 1.21)	1.05
Nowell 2019	0.38 (-0.55, 1.32)	0.66
Oosterling 2010	0.19 (-0.34, 0.71)	1.29
Pajareya 2011	0.31 (-0.43, 1.06)	0.91
Peters-Scheffer 2013	0.50 (-0.15, 1.16)	1.01
Reitzel 2013	0.05 (-1.05, 1.14)	0.52
Remington 2007	0.12 (-0.49, 0.73)	1.09
Rogers 2006	-0.12 (-1.24, 1.01)	0.53
Rogers 2012	-0.04 (-0.45, 0.36)	1.53
Rogers 2014	0.65 (-0.43, 1.74)	0.55
Ruiz 2020	0.12 (-0.55, 0.78)	1.04
Shawler 2017	0.35 (-0.25, 0.96)	1.13
Sheinkopf 1998	0.74 (-0.11, 1.60)	0.77
Shire 2017	0.03 (-0.34, 0.40)	1.6
Sinai-Gavrilov 2020	0.23 (-0.31, 0.77)	1.24
Solomon 2014	0.07 (-0.34, 0.48)	1.47
SpjutJansson 2016	-0.02 (-0.88, 0.83)	0.79
Stadnick 2015	0.56 (-0.16, 1.27)	0.95

Study	Hedges' g (95%CI)	Weight (%)
Stahmer 2020	0.47 (-0.34, 1.28)	0.83
Strauss 2012	0.07 (-0.53, 0.66)	1.13
Sullivan 2014	0.33 (-0.25, 0.91)	1.17
Tonge 2006	0.32 (-0.15, 0.79)	1.41
Tsang 2007	-0.08 (-0.75, 0.58)	1.04
VanderPaelt 2016	0.04 (-0.46, 0.54)	1.29
Vernon 2019	0.30 (-0.50, 1.11)	0.83
Vinen 2018	0.08 (-0.42, 0.59)	1.32
Vivanti 2014	0.04 (-0.48, 0.55)	1.32
Warreyn 2014	0.49 (-0.20, 1.18)	0.99
Waters 2020	0.61 (0.19, 1.04)	1.49
Whalen 2010	0.45 (-0.35, 1.24)	0.85
Whitehouse 2017	0.19 (-0.32, 0.69)	1.38
Xu 2017	0.54 (-0.11, 1.20)	1.03
Zachor 2007	0.93 (0.28, 1.58)	1.05
Zachor 2010	-0.19 (-0.64, 0.26)	1.44
Zhou 2018	0.70 (0.06, 1.35)	1.07
Overall Effect (RVE)	**0.32 (0.26, 0.38)**	**100**
Prediction Interval	**(-0.33, 0.98)**	**NA**

Figure B3. Funnel plot of all outcomes



B4.2 Autism characteristics

Autism characteristics outcomes were reported in 82 studies. The combined effect size was small and significant (g = 0.32, 95% CI 0.24 to 0.39, $tau^2 = 0.11$; **Figure B4**). The funnel plot did indicate evidence of small study effect (**Figure B5**), which was confirmed through formal testing (Egger's intercept = 1.22, p = 0.002). Adjusting for this effect (imputing 14 studies) resulted in a reduction in effect size (Hedges' g = 0.22, 95%CI 0.15 – 0.29, p < 0.001), although still small and statistically significant.

Figure B4.1. Forest plot of autism characteristic outcomes.

Note: Figures B4.1-B4.3 comprise one figure, displayed across multiple pages to ensure readability. An accessible version of the data displayed in this figure is presented in Table B5 below.

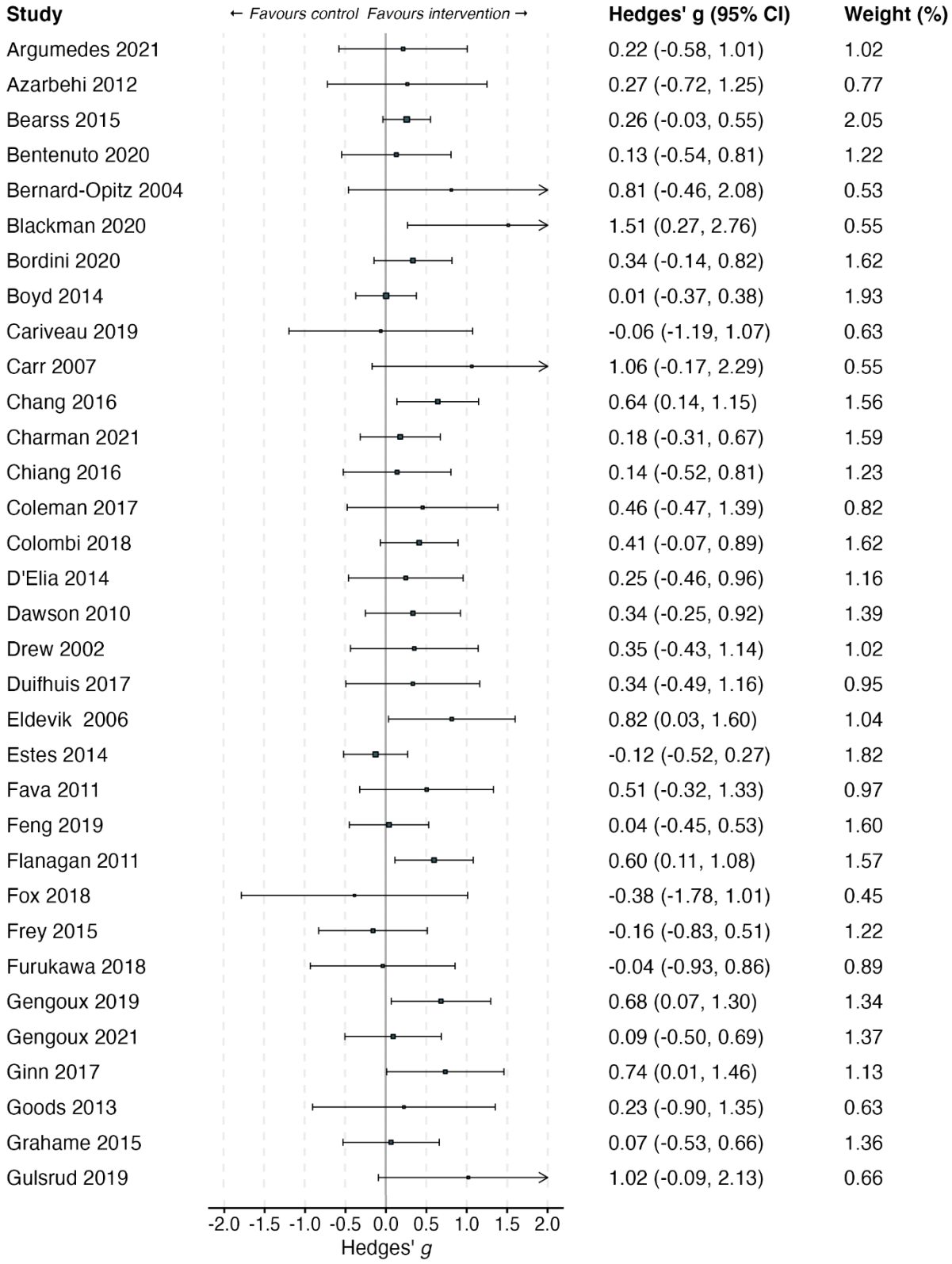


Figure B4.2: Forest plot of autism characteristic outcomes

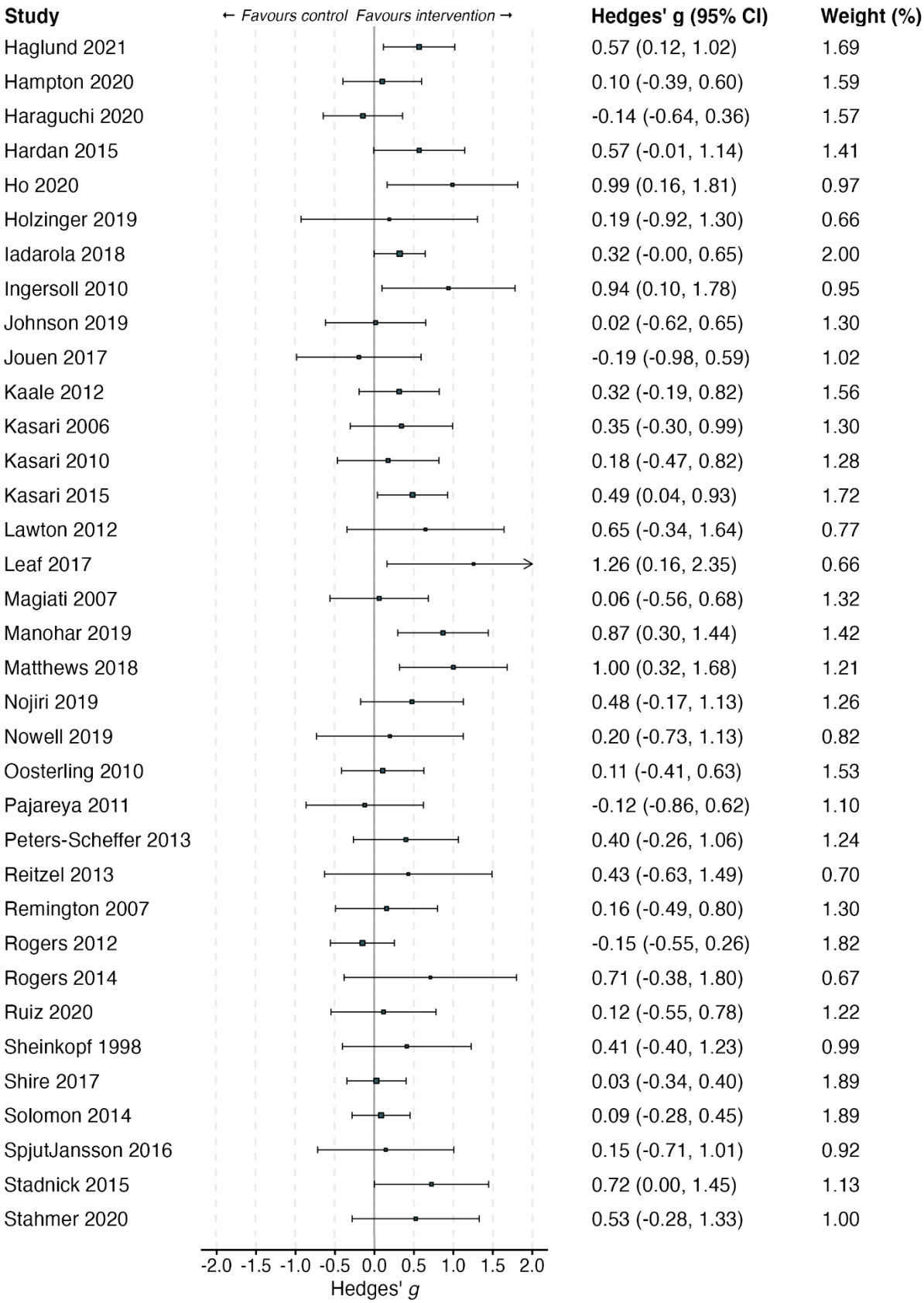


Figure B4.3: Forest plot of autism characteristic outcomes.

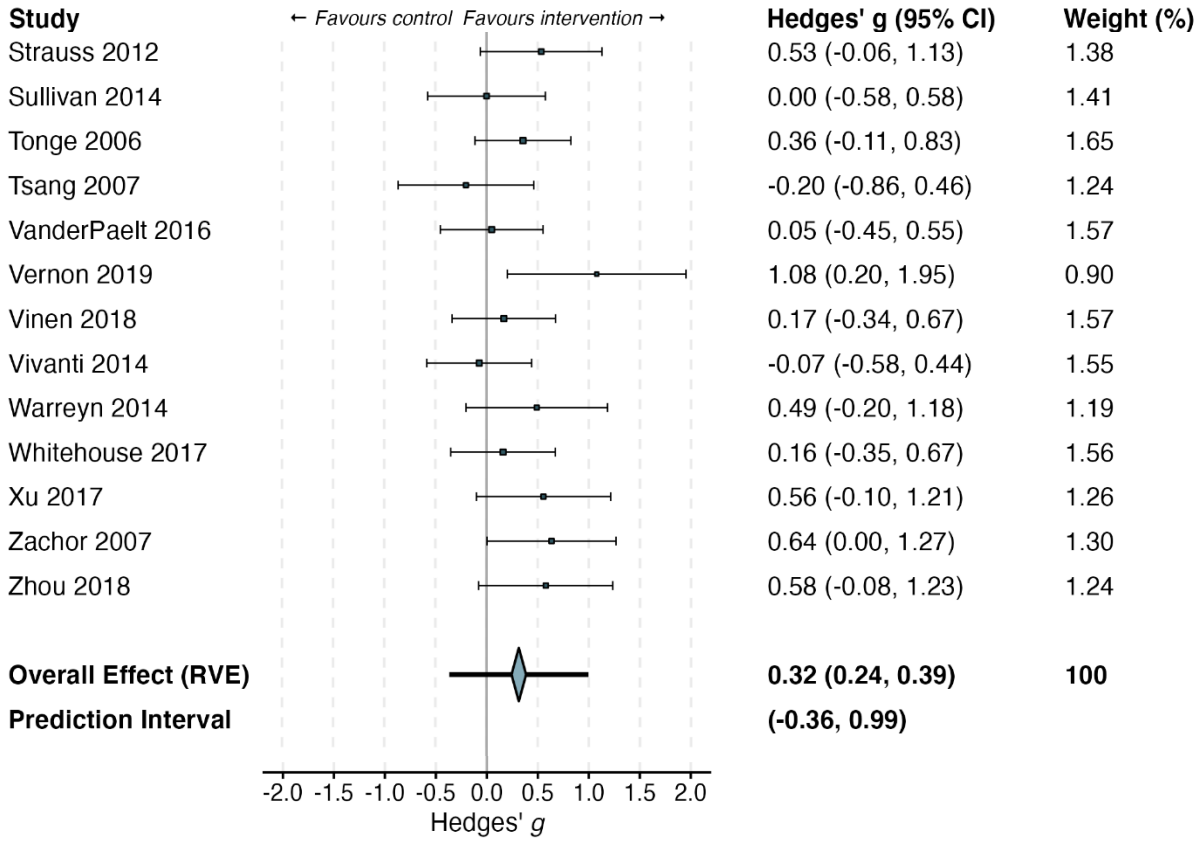


Table B5. Table version of forest plot of autism characteristic outcome measures.

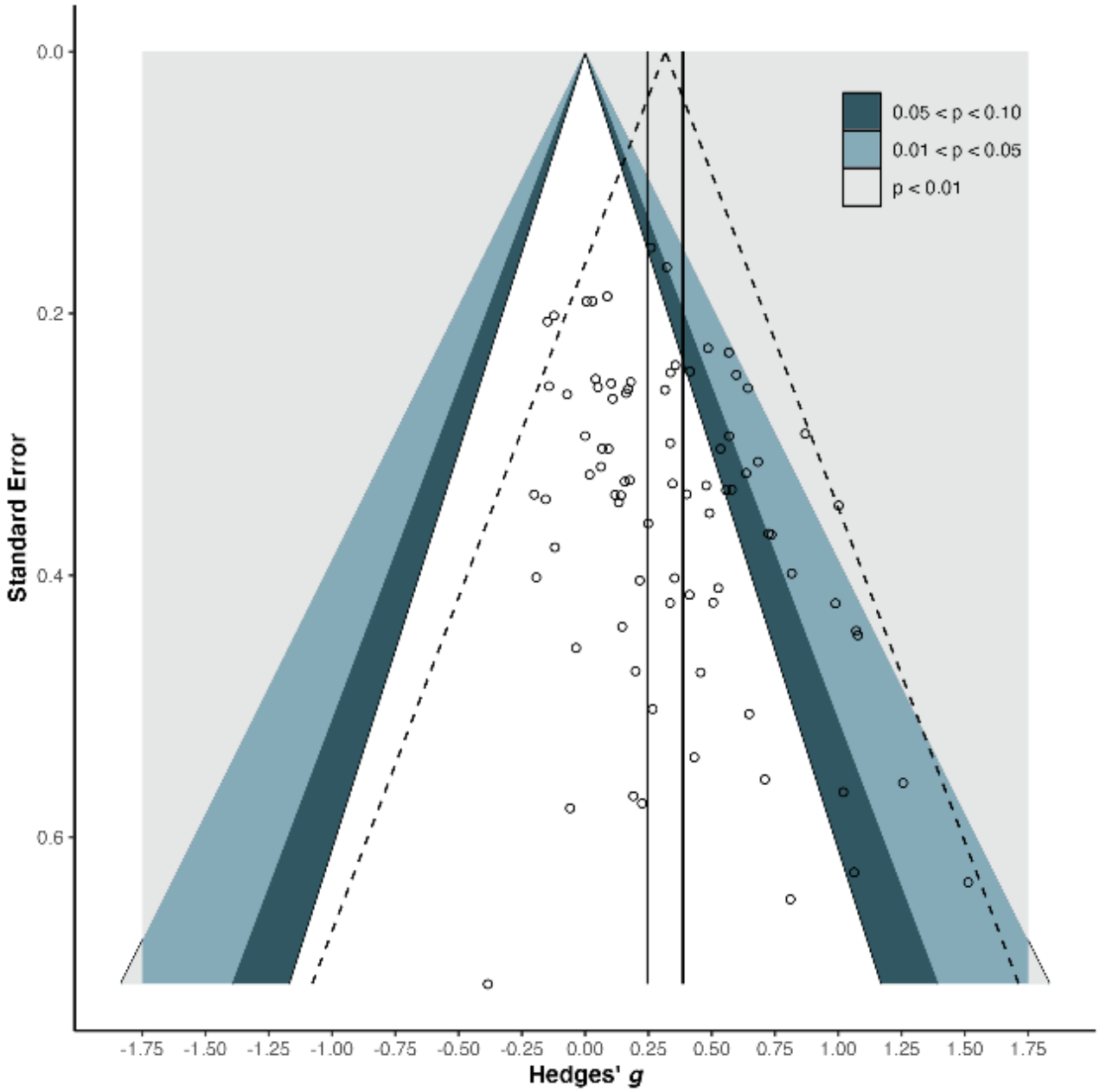
Note: This table presents the information displayed in Figure B4 in an accessible format. Positive Hedges' g values favour the behaviourally based intervention, negative Hedges' g values favour the comparison group.

Study	Hedges' g (95%CI)	Weight (%)
Argumedes 2021	0.22 (-0.58, 1.01)	1.02
Azarbehi 2012	0.27 (-0.72, 1.25)	0.77
Bearss 2015	0.26 (-0.03, 0.55)	2.05
Bentenuto 2020	0.13 (-0.54, 0.81)	1.22
Bernard-Opitz 2004	0.81 (-0.46, 2.08)	0.53
Blackman 2020	1.51 (0.27, 2.76)	0.55
Bordini 2020	0.34 (-0.14, 0.82)	1.62
Boyd 2014	0.01 (-0.37, 0.38)	1.93
Cariveau 2019	-0.06 (-1.19, 1.07)	0.63
Carr 2007	1.06 (-0.17, 2.29)	0.55
Chang 2016	0.64 (0.14, 1.15)	1.56
Charman 2021	0.18 (-0.31, 0.67)	1.59
Chiang 2016	0.14 (-0.52, 0.81)	1.23

Study	Hedges' g (95%CI)	Weight (%)
Coleman 2017	0.46 (-0.47, 1.39)	0.82
Colombi 2018	0.41 (-0.07, 0.89)	1.62
D'Elia 2014	0.25 (-0.46, 0.96)	1.16
Dawson 2010	0.34 (-0.25, 0.92)	1.39
Drew 2002	0.35 (-0.43, 1.14)	1.02
Duifhuis 2017	0.34 (-0.49, 1.16)	0.95
Eldevik 2006	0.82 (0.03, 1.60)	1.04
Estes 2014	-0.12 (-0.52, 0.27)	1.82
Fava 2011	0.51 (-0.32, 1.33)	0.97
Feng 2019	0.04 (-0.45, 0.53)	1.6
Flanagan 2011	0.60 (0.11, 1.08)	1.57
Fox 2018	-0.38 (-1.78, 1.01)	0.45
Frey 2015	-0.16 (-0.83, 0.51)	1.22
Furukawa 2018	-0.04 (-0.93, 0.86)	0.89
Gengoux 2019	0.68 (0.07, 1.30)	1.34
Gengoux 2021	0.09 (-0.50, 0.69)	1.37
Ginn 2017	0.74 (0.01, 1.46)	1.13
Goods 2013	0.23 (-0.90, 1.35)	0.63
Grahame 2015	0.07 (-0.53, 0.66)	1.36
Gulsrud 2019	1.02 (-0.09, 2.13)	0.66
Haglund 2021	0.57 (0.12, 1.02)	1.69
Hampton 2020	0.10 (-0.39, 0.60)	1.59
Haraguchi 2020	-0.14 (-0.64, 0.36)	1.57
Hardan 2015	0.57 (-0.01, 1.14)	1.41
Ho 2020	0.99 (0.16, 1.81)	0.97
Holzinger 2019	0.19 (-0.92, 1.30)	0.66
Iadarola 2018	0.32 (-0.00, 0.65)	2
Ingersoll 2010	0.94 (0.10, 1.78)	0.95
Johnson 2019	0.02 (-0.62, 0.65)	1.3
Jouen 2017	-0.19 (-0.98, 0.59)	1.02
Kaale 2012	0.32 (-0.19, 0.82)	1.56
Kasari 2006	0.35 (-0.30, 0.99)	1.3
Kasari 2010	0.18 (-0.47, 0.82)	1.28
Kasari 2015	0.49 (0.04, 0.93)	1.72
Lawton 2012	0.65 (-0.34, 1.64)	0.77
Leaf 2017	1.26 (0.16, 2.35)	0.66
Magiati 2007	0.06 (-0.56, 0.68)	1.32
Manohar 2019	0.87 (0.30, 1.44)	1.42
Matthews 2018	1.00 (0.32, 1.68)	1.21
Nojiri 2019	0.48 (-0.17, 1.13)	1.26
Nowell 2019	0.20 (-0.73, 1.13)	0.82
Oosterling 2010	0.11 (-0.41, 0.63)	1.53
Pajareya 2011	-0.12 (-0.86, 0.62)	1.1

Study	Hedges' g (95%CI)	Weight (%)
Peters-Scheffer 2013	0.40 (-0.26, 1.06)	1.24
Reitzel 2013	0.43 (-0.63, 1.49)	0.7
Remington 2007	0.16 (-0.49, 0.80)	1.3
Rogers 2012	-0.15 (-0.55, 0.26)	1.82
Rogers 2014	0.71 (-0.38, 1.80)	0.67
Ruiz 2020	0.12 (-0.55, 0.78)	1.22
Sheinkopf 1998	0.41 (-0.40, 1.23)	0.99
Shire 2017	0.03 (-0.34, 0.40)	1.89
Solomon 2014	0.09 (-0.28, 0.45)	1.89
SpjutJansson 2016	0.15 (-0.71, 1.01)	0.92
Stadnick 2015	0.72 (0.00, 1.45)	1.13
Stahmer 2020	0.53 (-0.28, 1.33)	1
Strauss 2012	0.53 (-0.06, 1.13)	1.38
Sullivan 2014	0.00 (-0.58, 0.58)	1.41
Tonge 2006	0.36 (-0.11, 0.83)	1.65
Tsang 2007	-0.20 (-0.86, 0.46)	1.24
VanderPaelt 2016	0.05 (-0.45, 0.55)	1.57
Vernon 2019	1.08 (0.20, 1.95)	0.9
Vinen 2018	0.17 (-0.34, 0.67)	1.57
Vivanti 2014	-0.07 (-0.58, 0.44)	1.55
Warreyn 2014	0.49 (-0.20, 1.18)	1.19
Whitehouse 2017	0.16 (-0.35, 0.67)	1.56
Xu 2017	0.56 (-0.10, 1.21)	1.26
Zachor 2007	0.64 (0.00, 1.27)	1.3
Zhou 2018	0.58 (-0.08, 1.23)	1.24
Overall Effect (RVE)	**0.32 (0.24, 0.39)**	**100**
Prediction Interval	**(-0.36, 0.99)**	**NA**

Figure B5. Funnel plot of autism characteristic outcomes



B4.3 Adaptive functioning

Adaptive functioning outcomes were reported by 47 studies. The combined effect size was small and significant (g = 0.24, 95% CI 0.12 to 0.36, $tau^2 = 0.09$; **Figure B6**). The funnel plot did not indicate evidence of small study effect (**Figure B7**), which was confirmed through formal testing (Egger's intercept = 0.73, p = 0.255).

Figure B6.1. Forest plot of adaptive functioning outcomes

Note: Figures B6.1-B6.2 comprise one figure, displayed across multiple pages to ensure readability. An accessible version of the data displayed in this figure is presented in Table B6 below.

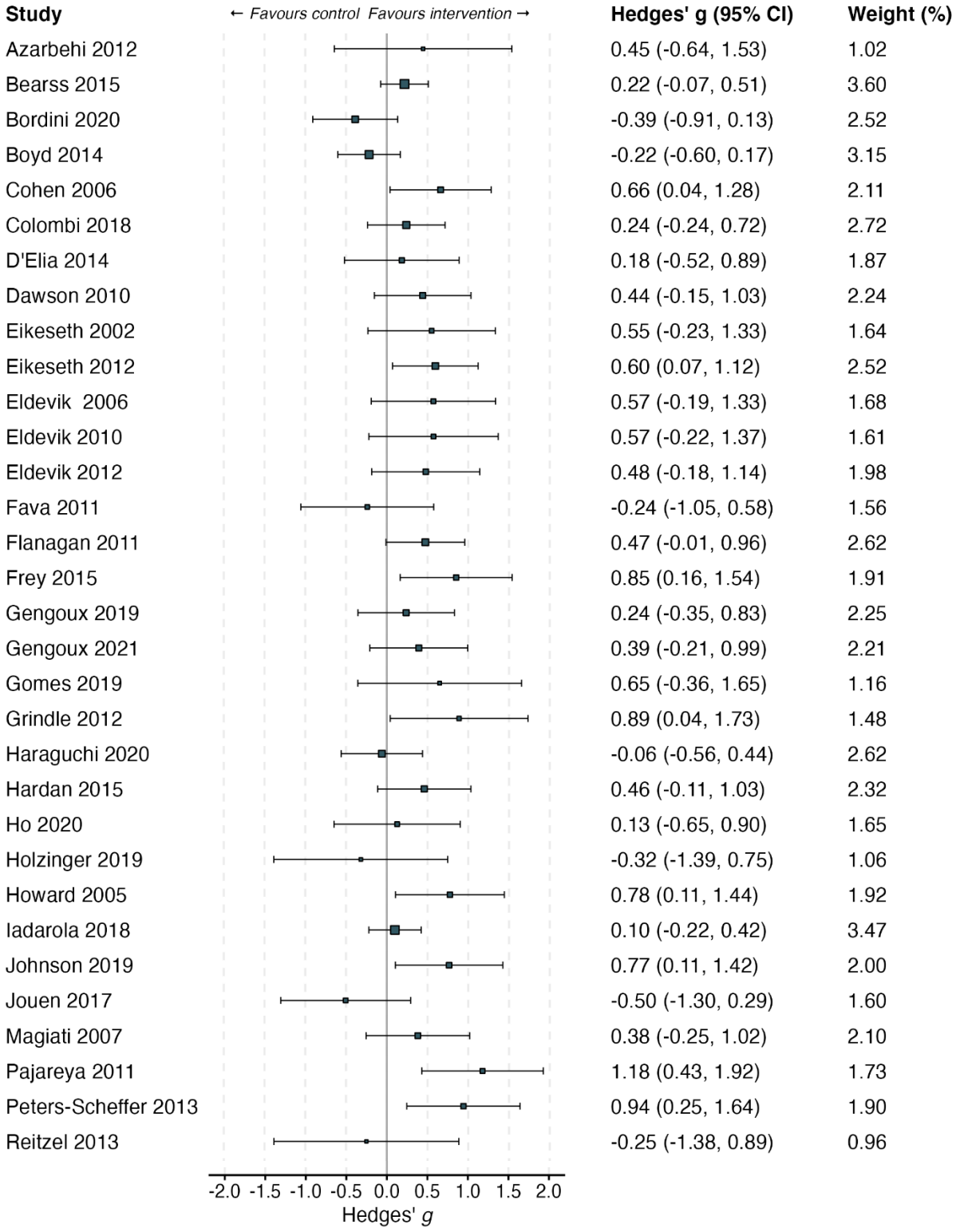


Figure B6.2. Forest plot of adaptive functioning outcomes

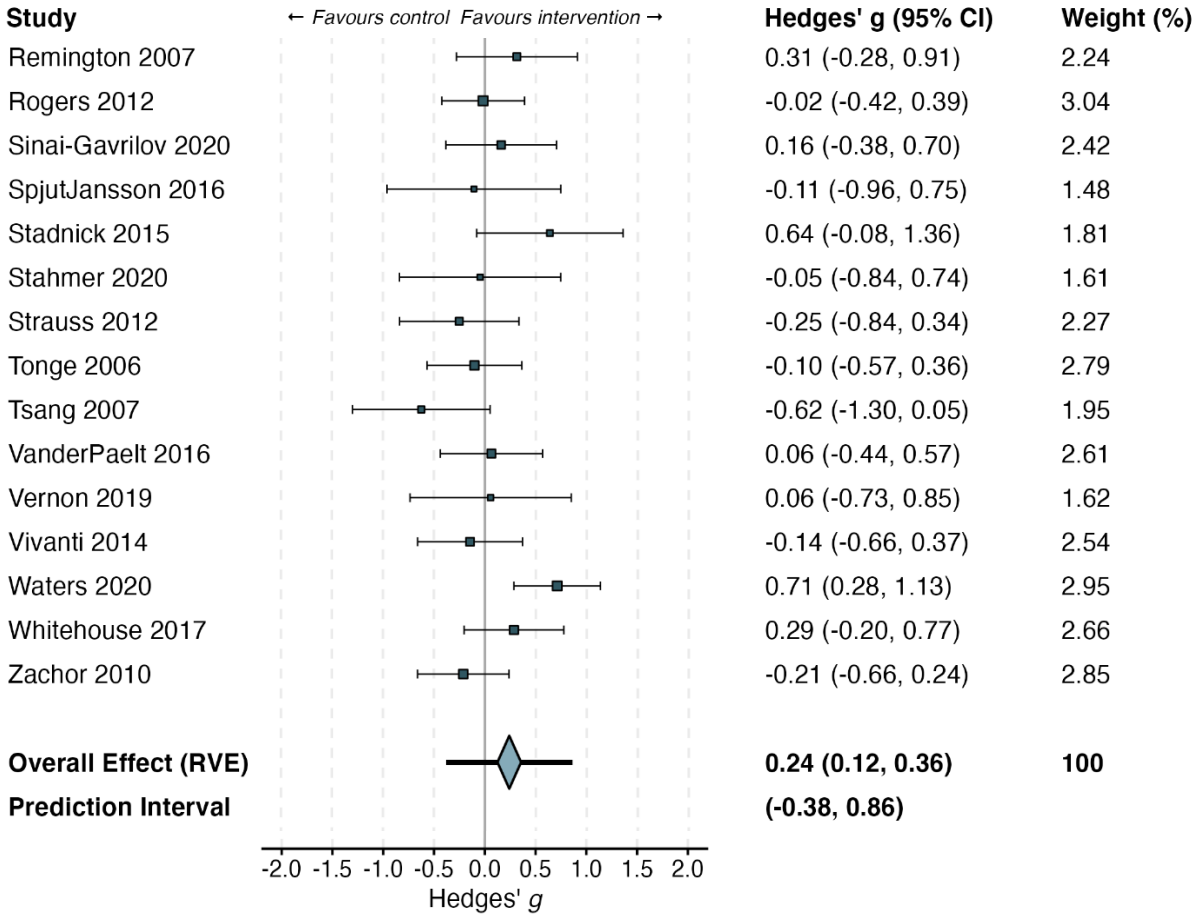


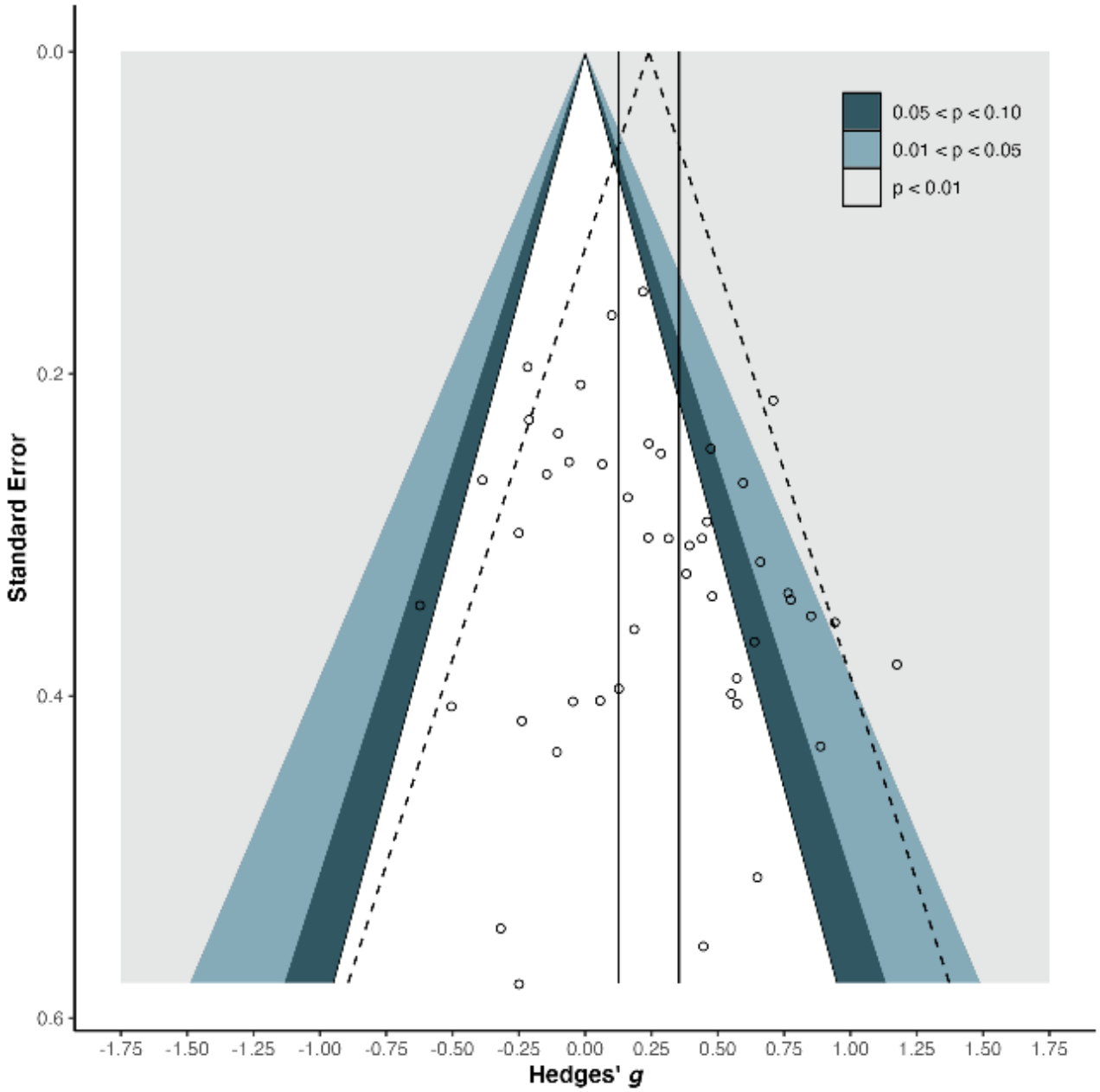
Table B6. Table version of forest plot of adaptive functioning outcome measures.

Note: This table presents the information displayed in Figure B6 in an accessible format. Positive Hedges' g values favour the behaviourally based intervention, negative Hedges' g values favour the comparison group.

Study	Hedges' g (95%CI)	Weight (%)
Azarbehi 2012	0.45 (-0.64, 1.53)	1.02
Bearss 2015	0.22 (-0.07, 0.51)	3.6
Bordini 2020	-0.39 (-0.91, 0.13)	2.52
Boyd 2014	-0.22 (-0.60, 0.17)	3.15
Cohen 2006	0.66 (0.04, 1.28)	2.11
Colombi 2018	0.24 (-0.24, 0.72)	2.72
D'Elia 2014	0.18 (-0.52, 0.89)	1.87
Dawson 2010	0.44 (-0.15, 1.03)	2.24
Eikeseth 2002	0.55 (-0.23, 1.33)	1.64
Eikeseth 2012	0.60 (0.07, 1.12)	2.52
Eldevik 2006	0.57 (-0.19, 1.33)	1.68

Study	Hedges' g (95%CI)	Weight (%)
Eldevik 2010	0.57 (-0.22, 1.37)	1.61
Eldevik 2012	0.48 (-0.18, 1.14)	1.98
Fava 2011	-0.24 (-1.05, 0.58)	1.56
Flanagan 2011	0.47 (-0.01, 0.96)	2.62
Frey 2015	0.85 (0.16, 1.54)	1.91
Gengoux 2019	0.24 (-0.35, 0.83)	2.25
Gengoux 2021	0.39 (-0.21, 0.99)	2.21
Gomes 2019	0.65 (-0.36, 1.65)	1.16
Grindle 2012	0.89 (0.04, 1.73)	1.48
Haraguchi 2020	-0.06 (-0.56, 0.44)	2.62
Hardan 2015	0.46 (-0.11, 1.03)	2.32
Ho 2020	0.13 (-0.65, 0.90)	1.65
Holzinger 2019	-0.32 (-1.39, 0.75)	1.06
Howard 2005	0.78 (0.11, 1.44)	1.92
Iadarola 2018	0.10 (-0.22, 0.42)	3.47
Johnson 2019	0.77 (0.11, 1.42)	2
Jouen 2017	-0.50 (-1.30, 0.29)	1.6
Magiati 2007	0.38 (-0.25, 1.02)	2.1
Pajareya 2011	1.18 (0.43, 1.92)	1.73
Peters-Scheffer 2013	0.94 (0.25, 1.64)	1.9
Reitzel 2013	-0.25 (-1.38, 0.89)	0.96
Remington 2007	0.31 (-0.28, 0.91)	2.24
Rogers 2012	-0.02 (-0.42, 0.39)	3.04
Sinai-Gavrilov 2020	0.16 (-0.38, 0.70)	2.42
SpjutJansson 2016	-0.11 (-0.96, 0.75)	1.48
Stadnick 2015	0.64 (-0.08, 1.36)	1.81
Stahmer 2020	-0.05 (-0.84, 0.74)	1.61
Strauss 2012	-0.25 (-0.84, 0.34)	2.27
Tonge 2006	-0.10 (-0.57, 0.36)	2.79
Tsang 2007	-0.62 (-1.30, 0.05)	1.95
VanderPaelt 2016	0.06 (-0.44, 0.57)	2.61
Vernon 2019	0.06 (-0.73, 0.85)	1.62
Vivanti 2014	-0.14 (-0.66, 0.37)	2.54
Waters 2020	0.71 (0.28, 1.13)	2.95
Whitehouse 2017	0.29 (-0.20, 0.77)	2.66
Zachor 2010	-0.21 (-0.66, 0.24)	2.85
Overall Effect (RVE)	**0.24 (0.12, 0.36)**	**100**
Prediction Interval	**(-0.38, 0.86)**	**NA**

Figure B7. Funnel plot of adaptive functioning outcomes



B4.4 Cognition and language

Cognition and language outcomes were reported by 64 studies. The combined effect size was small and significant (g = 0.30, 95% CI 0.22 to 0.38, $tau^2 = 0.05$; **Figure B8**). The funnel plot did indicate evidence of small study effect (**Figure B9**), which was confirmed through formal testing (Egger's intercept = 1.63, p < 0.001). Adjusting for this effect (imputing 15 studies) resulting in a reduction in effect size (Hedges' g = 0.19, 95%CI 0.10 – 0.28, p < 0.001), although still small and statistically significant.

Figure B8.1. Forest plot of cognition outcomes

Note: Figures B8.1-B8.2 comprise one figure, displayed across multiple pages to ensure readability. An accessible version of the data displayed in this figure is presented in Table B7 below.

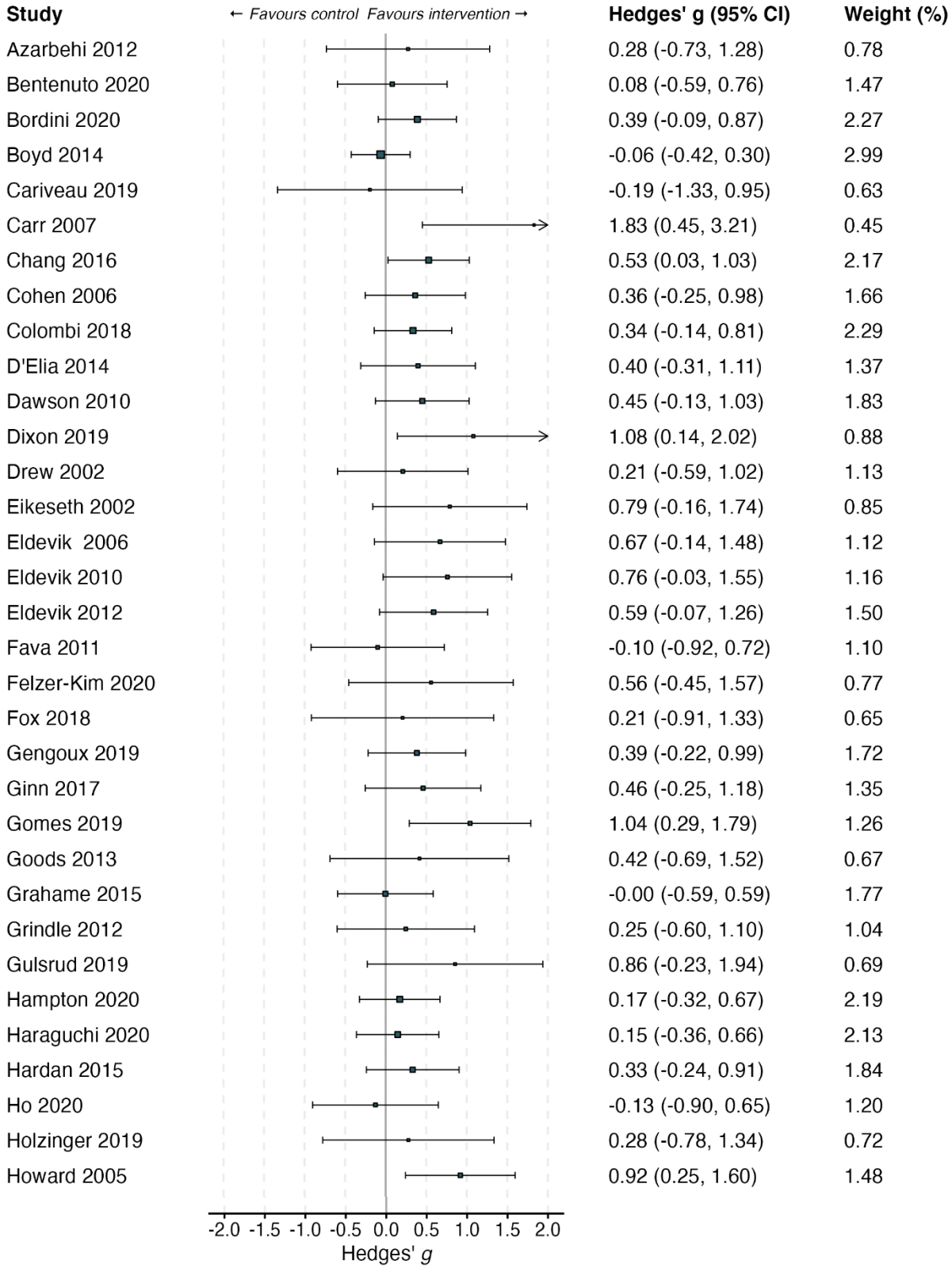


Figure B8.2. Forest plot of cognition outcomes

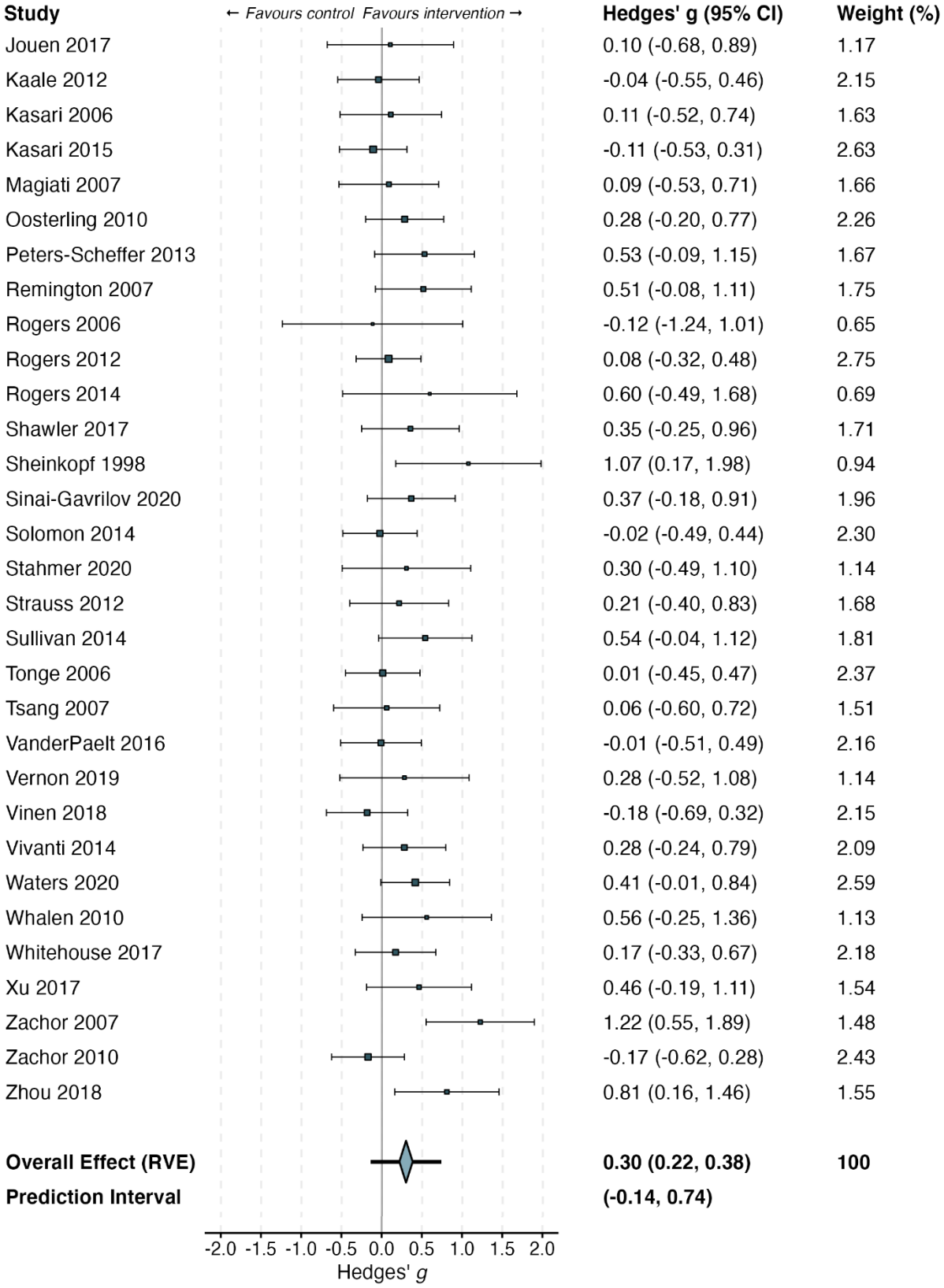


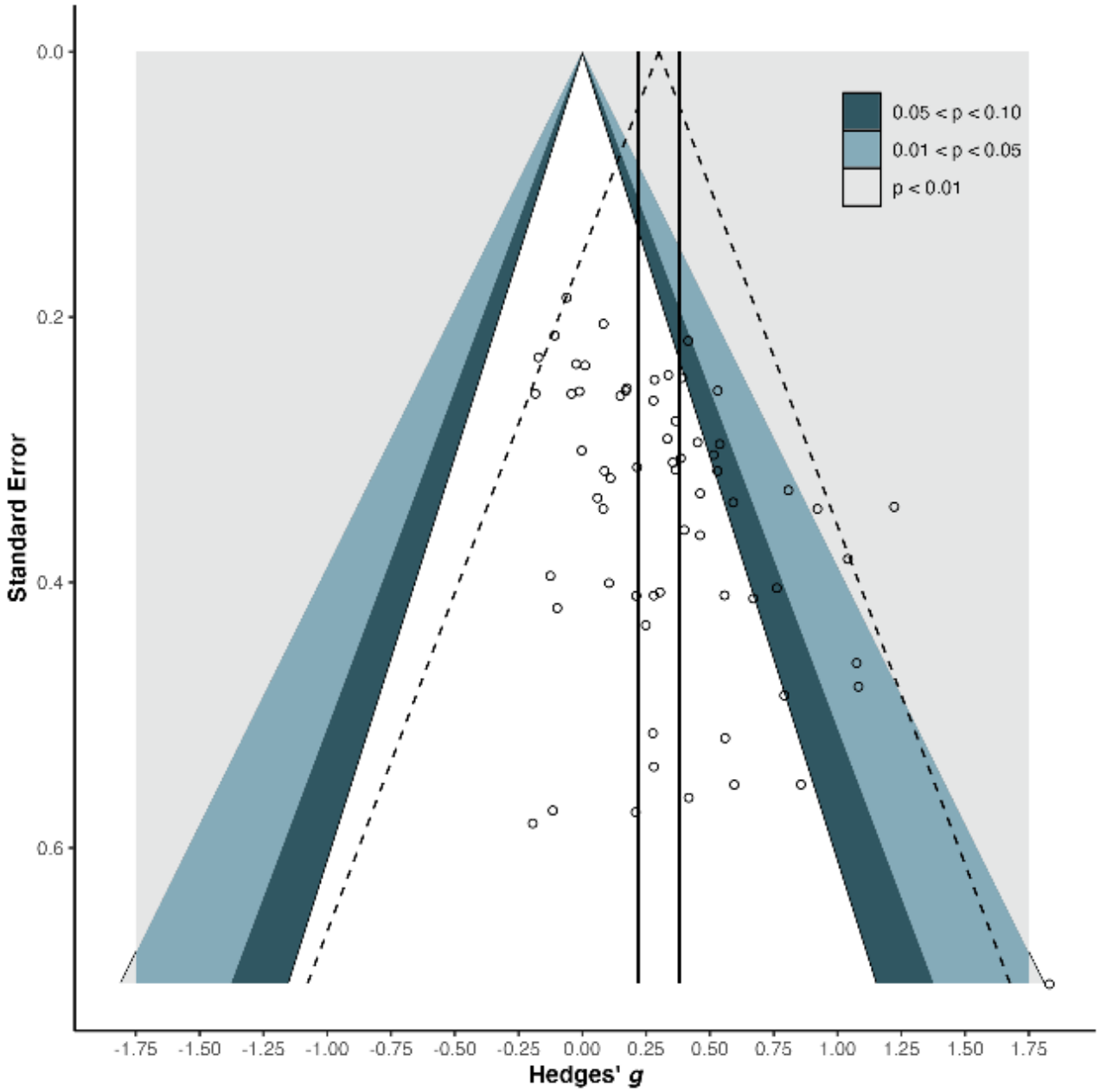
Table B7. Table version of forest plot of cognition and language outcome measures.

Note: This table presents the information displayed in Figure B8 in an accessible format. Positive Hedges' g values favour the behaviourally based intervention, negative Hedges' g values favour the comparison group.

Study	Hedges' g (95%CI)	Weight (%)
Azarbehi 2012	0.28 (-0.73, 1.28)	0.78
Bentenuto 2020	0.08 (-0.59, 0.76)	1.47
Bordini 2020	0.39 (-0.09, 0.87)	2.27
Boyd 2014	-0.06 (-0.42, 0.30)	2.99
Cariveau 2019	-0.19 (-1.33, 0.95)	0.63
Carr 2007	1.83 (0.45, 3.21)	0.45
Chang 2016	0.53 (0.03, 1.03)	2.17
Cohen 2006	0.36 (-0.25, 0.98)	1.66
Colombi 2018	0.34 (-0.14, 0.81)	2.29
D'Elia 2014	0.40 (-0.31, 1.11)	1.37
Dawson 2010	0.45 (-0.13, 1.03)	1.83
Dixon 2019	1.08 (0.14, 2.02)	0.88
Drew 2002	0.21 (-0.59, 1.02)	1.13
Eikeseth 2002	0.79 (-0.16, 1.74)	0.85
Eldevik 2006	0.67 (-0.14, 1.48)	1.12
Eldevik 2010	0.76 (-0.03, 1.55)	1.16
Eldevik 2012	0.59 (-0.07, 1.26)	1.5
Fava 2011	-0.10 (-0.92, 0.72)	1.1
Felzer-Kim 2020	0.56 (-0.45, 1.57)	0.77
Fox 2018	0.21 (-0.91, 1.33)	0.65
Gengoux 2019	0.39 (-0.22, 0.99)	1.72
Ginn 2017	0.46 (-0.25, 1.18)	1.35
Gomes 2019	1.04 (0.29, 1.79)	1.26
Goods 2013	0.42 (-0.69, 1.52)	0.67
Grahame 2015	-0.00 (-0.59, 0.59)	1.77
Grindle 2012	0.25 (-0.60, 1.10)	1.04
Gulsrud 2019	0.86 (-0.23, 1.94)	0.69
Hampton 2020	0.17 (-0.32, 0.67)	2.19
Haraguchi 2020	0.15 (-0.36, 0.66)	2.13
Hardan 2015	0.33 (-0.24, 0.91)	1.84
Ho 2020	-0.13 (-0.90, 0.65)	1.2
Holzinger 2019	0.28 (-0.78, 1.34)	0.72
Howard 2005	0.92 (0.25, 1.60)	1.48
Jouen 2017	0.10 (-0.68, 0.89)	1.17
Kaale 2012	-0.04 (-0.55, 0.46)	2.15
Kasari 2006	0.11 (-0.52, 0.74)	1.63
Kasari 2015	-0.11 (-0.53, 0.31)	2.63
Magiati 2007	0.09 (-0.53, 0.71)	1.66

Study	Hedges' g (95%CI)	Weight (%)
Oosterling 2010	0.28 (-0.20, 0.77)	2.26
Peters-Scheffer 2013	0.53 (-0.09, 1.15)	1.67
Remington 2007	0.51 (-0.08, 1.11)	1.75
Rogers 2006	-0.12 (-1.24, 1.01)	0.65
Rogers 2012	0.08 (-0.32, 0.48)	2.75
Rogers 2014	0.60 (-0.49, 1.68)	0.69
Shawler 2017	0.35 (-0.25, 0.96)	1.71
Sheinkopf 1998	1.07 (0.17, 1.98)	0.94
Sinai-Gavrilov 2020	0.37 (-0.18, 0.91)	1.96
Solomon 2014	-0.02 (-0.49, 0.44)	2.3
Stahmer 2020	0.30 (-0.49, 1.10)	1.14
Strauss 2012	0.21 (-0.40, 0.83)	1.68
Sullivan 2014	0.54 (-0.04, 1.12)	1.81
Tonge 2006	0.01 (-0.45, 0.47)	2.37
Tsang 2007	0.06 (-0.60, 0.72)	1.51
VanderPaelt 2016	-0.01 (-0.51, 0.49)	2.16
Vernon 2019	0.28 (-0.52, 1.08)	1.14
Vinen 2018	-0.18 (-0.69, 0.32)	2.15
Vivanti 2014	0.28 (-0.24, 0.79)	2.09
Waters 2020	0.41 (-0.01, 0.84)	2.59
Whalen 2010	0.56 (-0.25, 1.36)	1.13
Whitehouse 2017	0.17 (-0.33, 0.67)	2.18
Xu 2017	0.46 (-0.19, 1.11)	1.54
Zachor 2007	1.22 (0.55, 1.89)	1.48
Zachor 2010	-0.17 (-0.62, 0.28)	2.43
Zhou 2018	0.81 (0.16, 1.46)	1.55
Overall Effect (RVE)	**0.30 (0.22, 0.38)**	**100**
Prediction Interval	**(-0.14, 0.74)**	**NA**

Figure B9. Funnel plot of cognition and language outcomes



B4.5 Family outcomes

Family outcomes were reported by 20 studies. The combined effect size was small and significant (g = 0.39, 95% CI 0.21 to 0.58, $tau^2 = 0.18$; **Figure B10**). The funnel plot did not indicate evidence of small study effect (**Figure B11**), which was confirmed through formal testing (Egger's intercept = 0.48, p = 0.533).

Figure B10. Forest plot of family outcomes

Note: An accessible version of the data displayed in this figure is presented in Table B8 below.

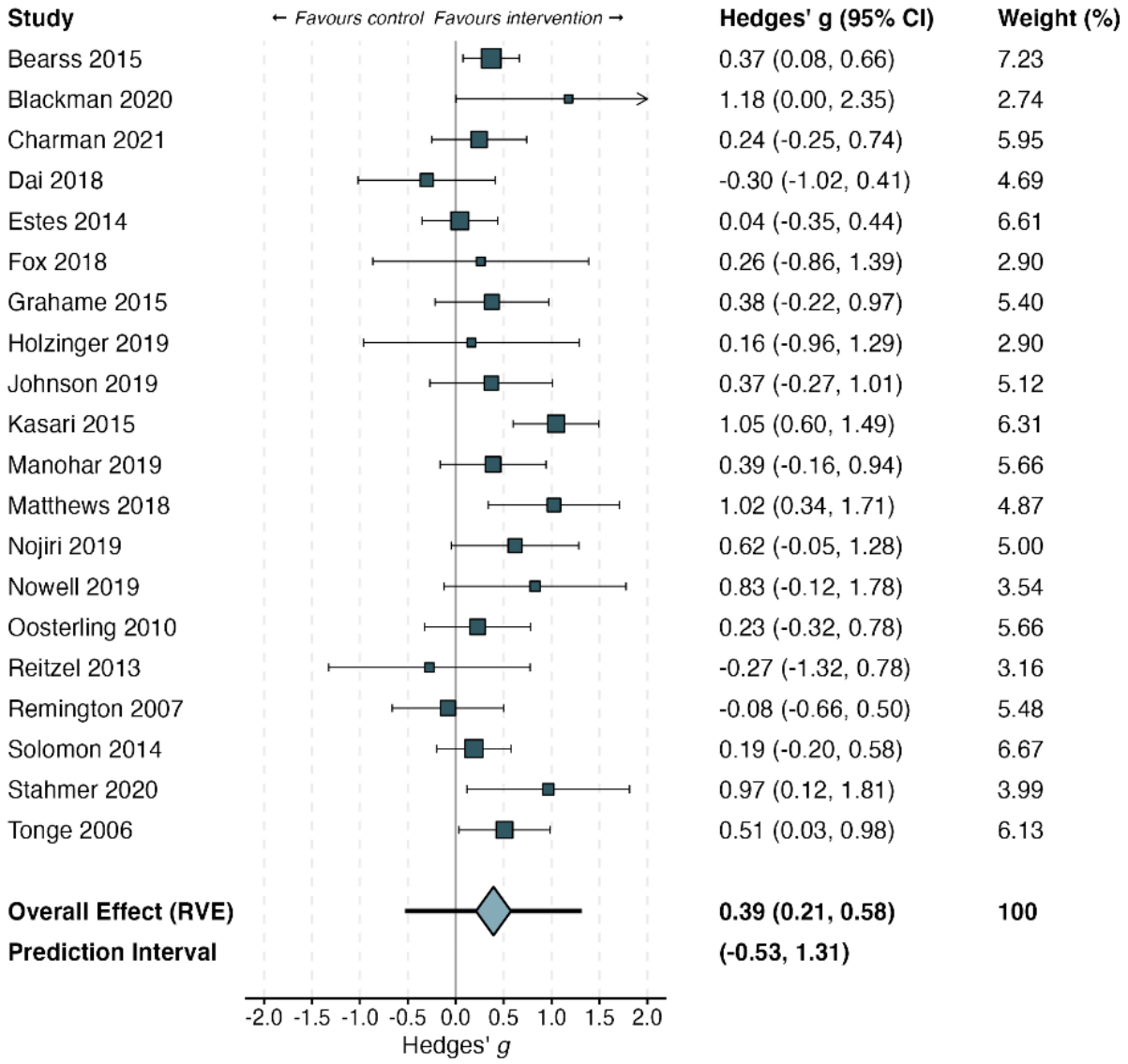


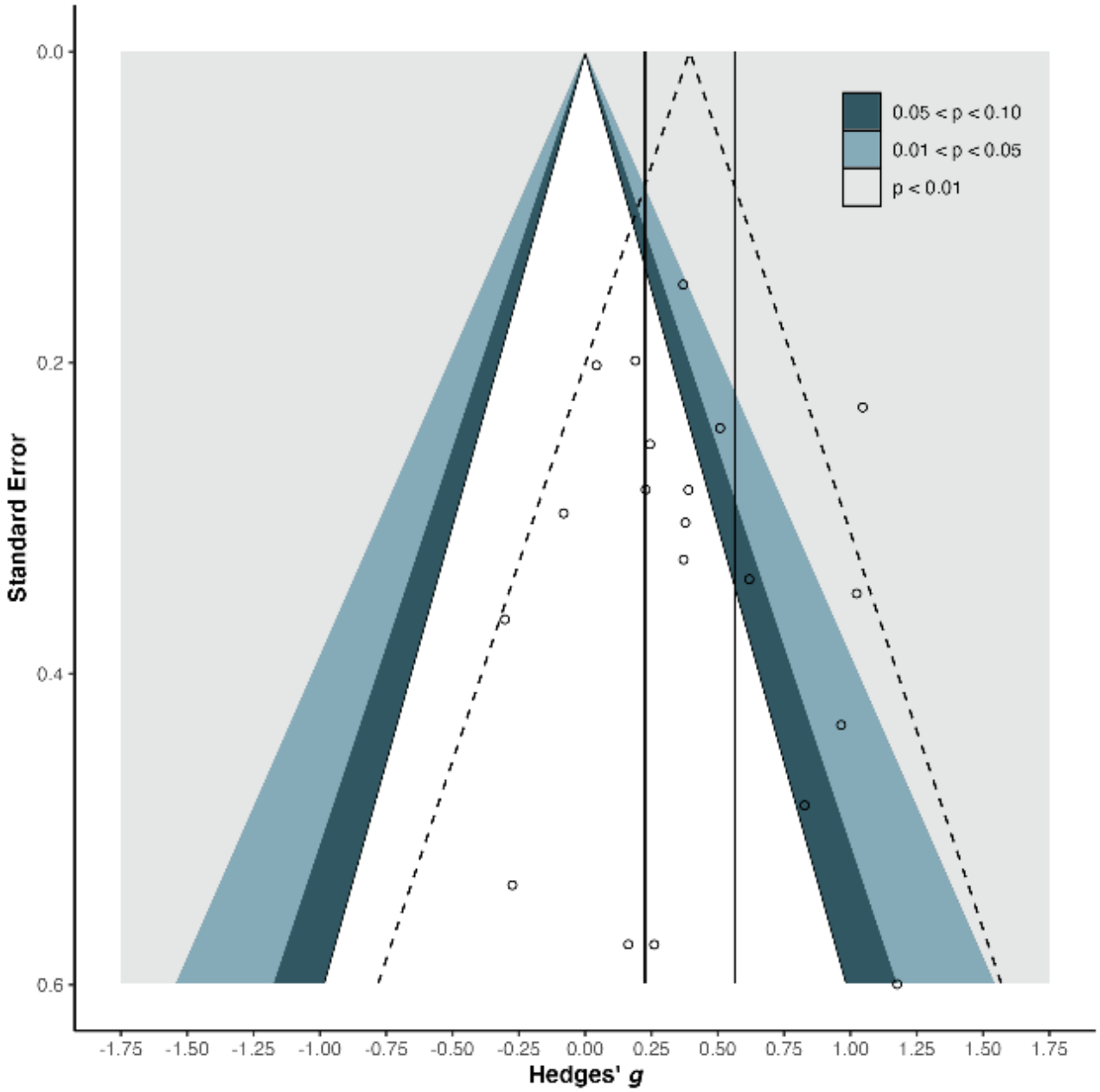
Table B8. Table version of forest plot of family outcome outcome measures.

Note: This table presents the information displayed in Figure B10 in an accessible format. Positive Hedges' g values favour the behaviourally based intervention, negative Hedges' g values favour the comparison group.

Study	Hedges' g (95%CI)	Weight (%)
Bearss 2015	0.37 (0.08, 0.66)	7.23
Blackman 2020	1.18 (0.00, 2.35)	2.74
Charman 2021	0.24 (-0.25, 0.74)	5.95
Dai 2018	-0.30 (-1.02, 0.41)	4.69
Estes 2014	0.04 (-0.35, 0.44)	6.61

Study	Hedges' g (95%CI)	Weight (%)
Fox 2018	0.26 (-0.86, 1.39)	2.9
Grahame 2015	0.38 (-0.22, 0.97)	5.4
Holzinger 2019	0.16 (-0.96, 1.29)	2.9
Johnson 2019	0.37 (-0.27, 1.01)	5.12
Kasari 2015	1.05 (0.60, 1.49)	6.31
Manohar 2019	0.39 (-0.16, 0.94)	5.66
Matthews 2018	1.02 (0.34, 1.71)	4.87
Nojiri 2019	0.62 (-0.05, 1.28)	5
Nowell 2019	0.83 (-0.12, 1.78)	3.54
Oosterling 2010	0.23 (-0.32, 0.78)	5.66
Reitzel 2013	-0.27 (-1.32, 0.78)	3.16
Remington 2007	-0.08 (-0.66, 0.50)	5.48
Solomon 2014	0.19 (-0.20, 0.58)	6.67
Stahmer 2020	0.97 (0.12, 1.81)	3.99
Tonge 2006	0.51 (0.03, 0.98)	6.13
Overall Effect (RVE)	**0.39 (0.21, 0.58)**	**100**
Prediction Interval	**(-0.53, 1.31)**	**NA**

Figure B11. Funnel plot of family outcomes



B4.6 Adverse effects

Adverse effects outcomes were reported by 27 studies. The combined effect size was small and significant (g = 0.24, 95% CI 0.09 to 0.39, $tau^2 = 0.09$; **Figure B12**). The funnel plot did not indicate evidence of small study effect (**Figure B13**), which was confirmed through formal testing (Egger's intercept = 0.96, p = 0.220).

Figure B12. Forest plot of adverse effects outcomes

Note: An accessible version of the data displayed in this figure is presented in Table B9, which follows.

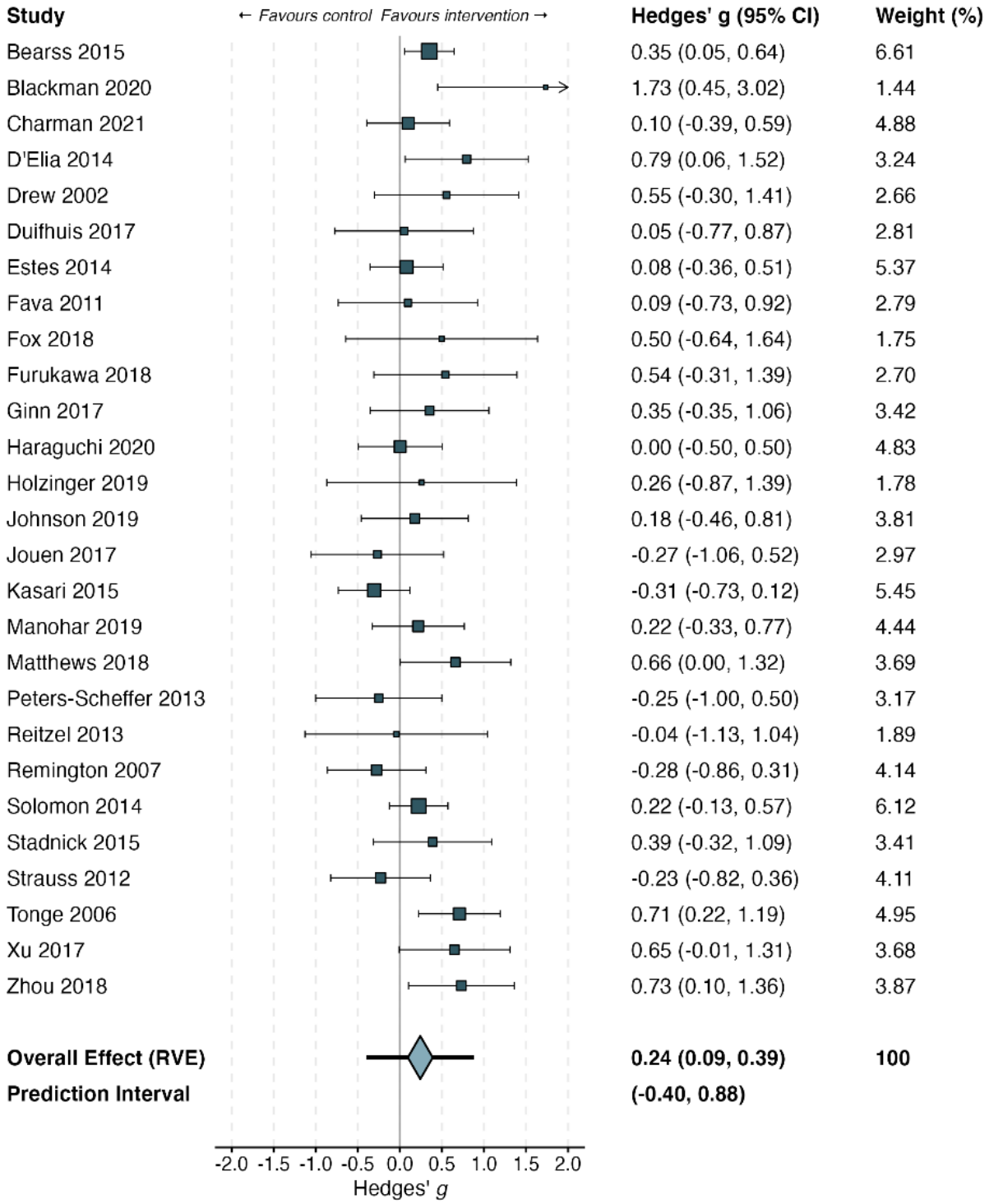
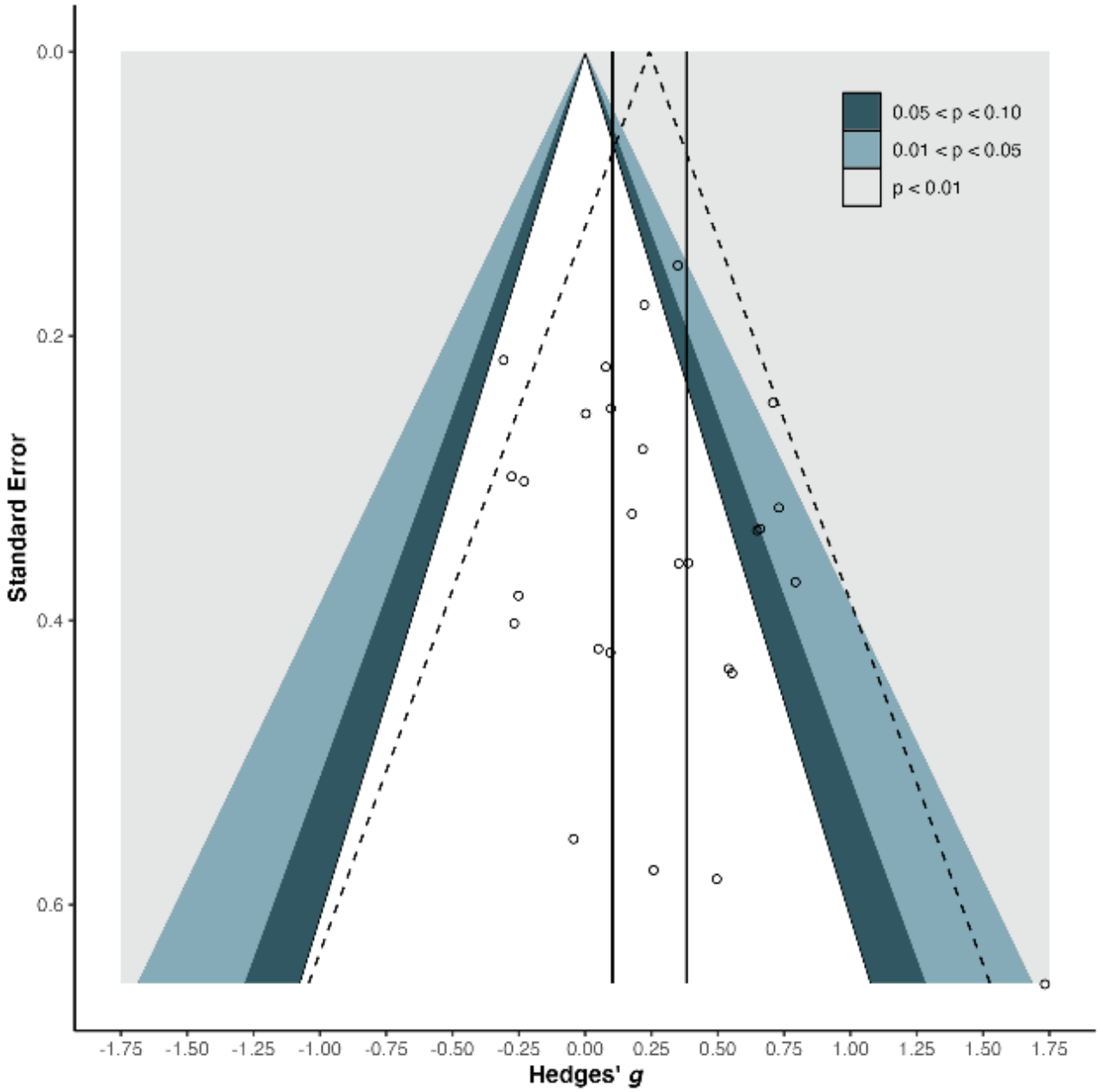


Table B9. Table version of forest plot of adverse effects outcome measures.

Note: This table presents the information displayed in Figure B12 in an accessible format. Positive Hedges' g values favour the behaviourally based intervention, negative Hedges' g values favour the comparison group.

Study	Hedges' g (95%CI)	Weight (%)
Bearss 2015	0.35 (0.05, 0.64)	6.61
Blackman 2020	1.73 (0.45, 3.02)	1.44
Charman 2021	0.10 (-0.39, 0.59)	4.88
D'Elia 2014	0.79 (0.06, 1.52)	3.24
Drew 2002	0.55 (-0.30, 1.41)	2.66
Duifhuis 2017	0.05 (-0.77, 0.87)	2.81
Estes 2014	0.08 (-0.36, 0.51)	5.37
Fava 2011	0.09 (-0.73, 0.92)	2.79
Fox 2018	0.50 (-0.64, 1.64)	1.75
Furukawa 2018	0.54 (-0.31, 1.39)	2.7
Ginn 2017	0.35 (-0.35, 1.06)	3.42
Haraguchi 2020	0.00 (-0.50, 0.50)	4.83
Holzinger 2019	0.26 (-0.87, 1.39)	1.78
Johnson 2019	0.18 (-0.46, 0.81)	3.81
Jouen 2017	-0.27 (-1.06, 0.52)	2.97
Kasari 2015	-0.31 (-0.73, 0.12)	5.45
Manohar 2019	0.22 (-0.33, 0.77)	4.44
Matthews 2018	0.66 (0.00, 1.32)	3.69
Peters-Scheffer 2013	-0.25 (-1.00, 0.50)	3.17
Reitzel 2013	-0.04 (-1.13, 1.04)	1.89
Remington 2007	-0.28 (-0.86, 0.31)	4.14
Solomon 2014	0.22 (-0.13, 0.57)	6.12
Stadnick 2015	0.39 (-0.32, 1.09)	3.41
Strauss 2012	-0.23 (-0.82, 0.36)	4.11
Tonge 2006	0.71 (0.22, 1.19)	4.95
Xu 2017	0.65 (-0.01, 1.31)	3.68
Zhou 2018	0.73 (0.10, 1.36)	3.87
Overall Effect (RVE)	**0.24 (0.09, 0.39)**	**100**
Prediction Interval	**(-0.40, 0.88)**	**NA**

Figure B13. Funnel plot of adverse effects outcomes



B5. Investigating the effect of dose

B5.1 Relationship between dose and efficacy

Table B10 provides model statistics for the linear models which investigate the relationship between dose (monthly and total clinician hours) and efficacy of behaviourally based interventions as compared to a control group. This is reported across all available outcomes, as well as within three outcome domains: autism characteristics, adaptive functioning, and cognition and language.

Table B10.1. Linear model statistics for association between dose (monthly clinician hours) and autism characteristic, adaptive functioning and cognition and language outcomes.

Note: These tables have been grouped together using a number-letter referencing system as they are related.

Dose/outcome domain	N	β	95%CI	p-value
All outcomes	34	0.001	-0.0006 to 0.0031	0.195
Autism characteristics	31	0.001	-0.0022 to 0.0034	0.666
Adaptive functioning	18	0.003	0.0004 to 0.0062	0.025
Cognition and language	21	0.002	0.0002 to 0.0037	0.029

Table B10.2. Linear model statistics for association between dose (total clinician hours) and autism characteristic, adaptive functioning and cognition and language outcomes.

Dose/outcome domain	N	β	95%CI	p-value
All outcomes	34	0.00005	0.000003 to 0.000998	0.036
Autism characteristics	31	-0.00002	-0.0001 to 0.0001	0.700
Adaptive functioning	18	0.00009	0.00002 to 0.00015	0.007
Cognition and language	21	0.00008	0.00001 to 0.00014	0.022

Relationship between dose and change from baseline to follow-up separately within the intervention group and the comparison group

Table B11 provides model statistics for the linear models which investigate the relationship between dose (monthly and total clinician hours) and change from pre- to post-intervention in the group of children who underwent behaviourally based intervention. This is reported across all available outcomes, as well as within three outcome domains: autism characteristics, adaptive functioning, and cognition and language.

Table B11.1. Linear model statistics for association between dose (monthly clinician hours) and autism characteristic, adaptive functioning and cognition and language outcomes for change from pre to post in the intervention group

Note: These tables have been grouped together using a number-letter referencing system as they are related.

Outcome domain	N	β	95%CI	p-value
All outcomes	33	0.004	0.0004 to 0.0083	0.031
Autism characteristics	29	0.003	-0.001 to 0.008	0.115
Adaptive functioning	17	0.007	0.0003 to 0.0134	0.040
Cognition and language	20	0.004	-0.0006 to 0.0092	0.085

Table B11.2. Linear model statistics for association between dose (total clinician hours) and autism characteristic, adaptive functioning and cognition and language outcomes for change from pre to post in the intervention group

Outcome domain	N	β	95%CI	p-value
All outcomes	33	0.00007	0.00003 to 0.00010	<0.001
Autism characteristics	29	0.00001	-0.00007 to 0.00009	0.840
Adaptive functioning	17	0.00005	0.00001 to 0.00010	0.017
Cognition and language	20	0.00012	0.00005 to 0.00019	0.001

All outcomes

The linear and non-linear models of total and monthly clinician-delivered hours of intervention by effect size for the change from baseline to follow-up in the behaviourally based intervention group, with 95% confidence intervals, for all outcomes are shown in **Figure B14**.

Figure B14.1. Linear model of monthly clinician-delivered dose by effect size (Hedges' g) for change in all outcomes from pre- to post-intervention in the intervention group.

Note: These figures have been grouped together using a number-letter referencing system as they are related. Hedges' g > 0 = improvement in outcomes from baseline to follow-up. Hedges' g > 0 = decrease in outcomes from baseline to follow-up.

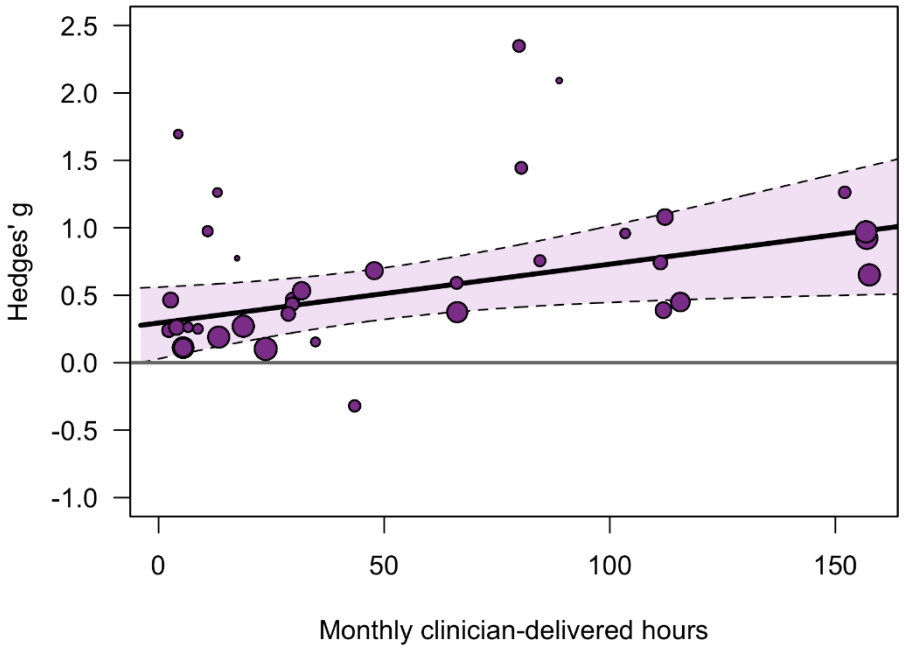


Figure B14.2. Linear model of total clinician-delivered dose by effect size (Hedges' g) for change in all outcomes from pre- to post-intervention in the intervention group.

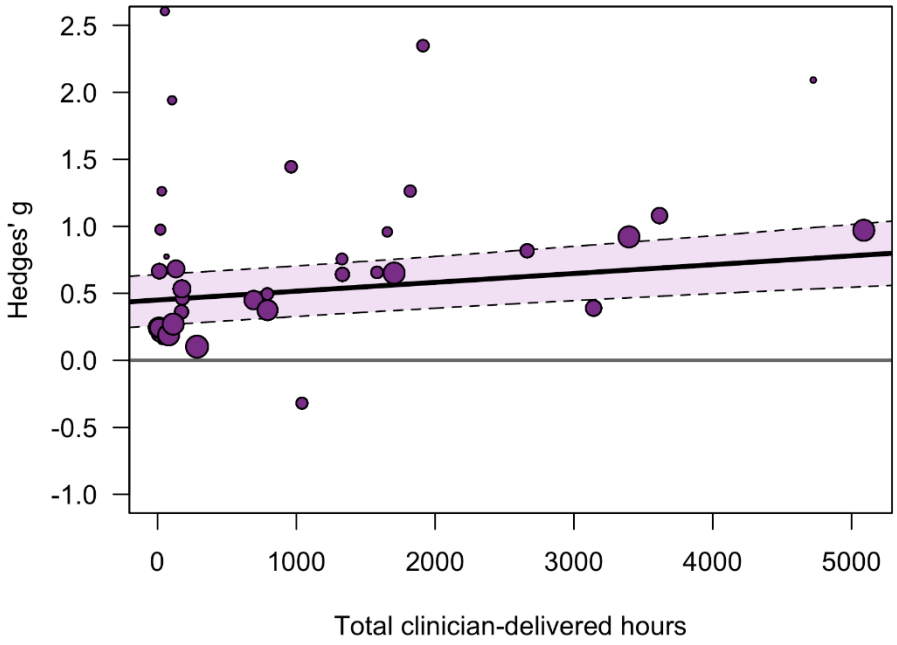


Figure B14.3. Non-linear model of monthly clinician-delivered dose by effect size (Hedges' g) for change in all outcomes from pre- to post-intervention in the intervention group.

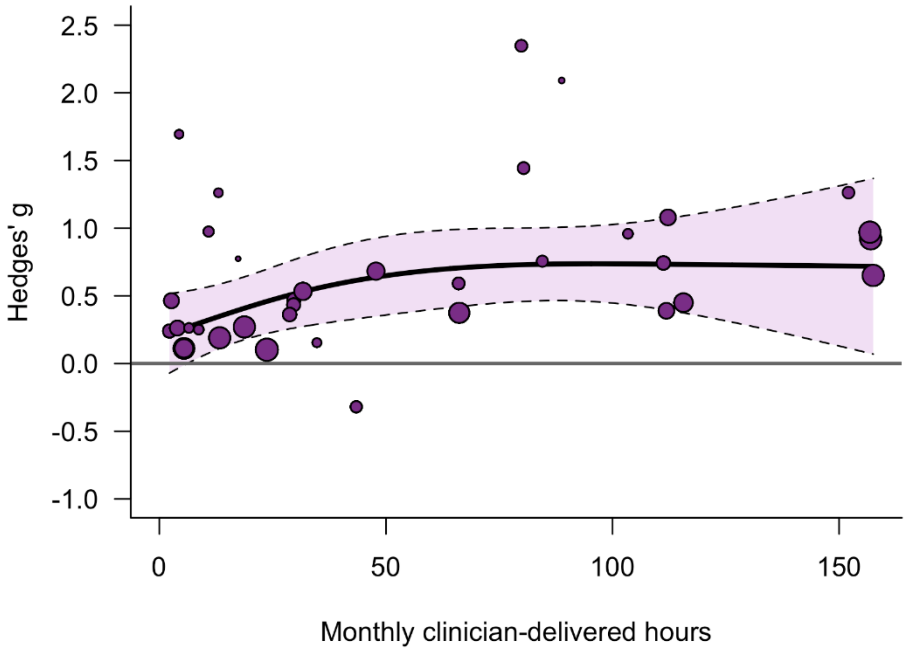
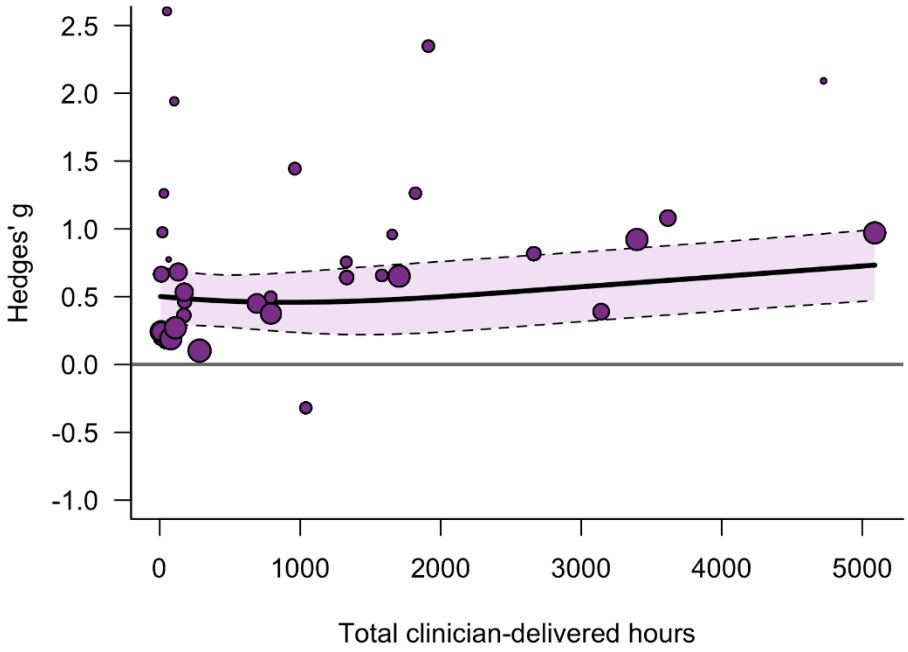


Figure B14.4. Non-linear model of total clinician-delivered dose by effect size (Hedges' g) for change in all outcomes from pre- to post-intervention in the intervention group.



Autism characteristics

The linear and non-linear models of total and monthly clinician-delivered hours of intervention by effect size for the change from baseline to follow-up in the behaviourally based intervention group, with 95% confidence intervals, for autism characteristic outcomes are shown in **Figure B15**.

Figure B15.1. Linear model of monthly clinician-delivered dose by effect size (Hedges' g) for change in autism characteristic outcomes from pre- to post-intervention in the intervention group.

Note: These figures have been grouped together using a number-letter referencing system as they are related. Hedges' g > 0 = improvement in outcomes from baseline to follow-up. Hedges' g > 0 = decrease in outcomes from baseline to follow-up.

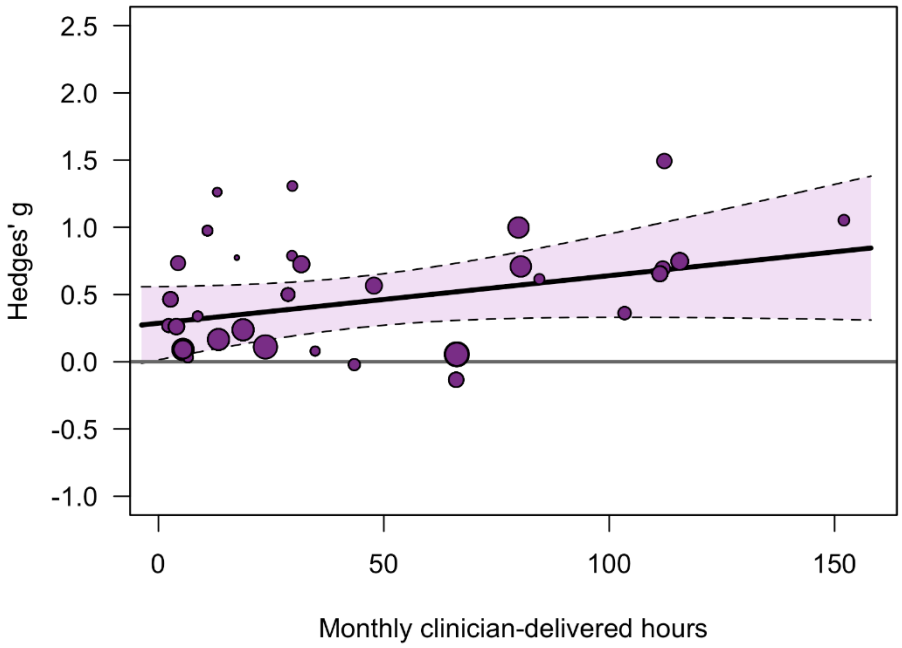


Figure B15.2. Linear model of total clinician-delivered dose by effect size (Hedges' g) for change in autism characteristic outcomes from pre- to post-intervention in the intervention group.

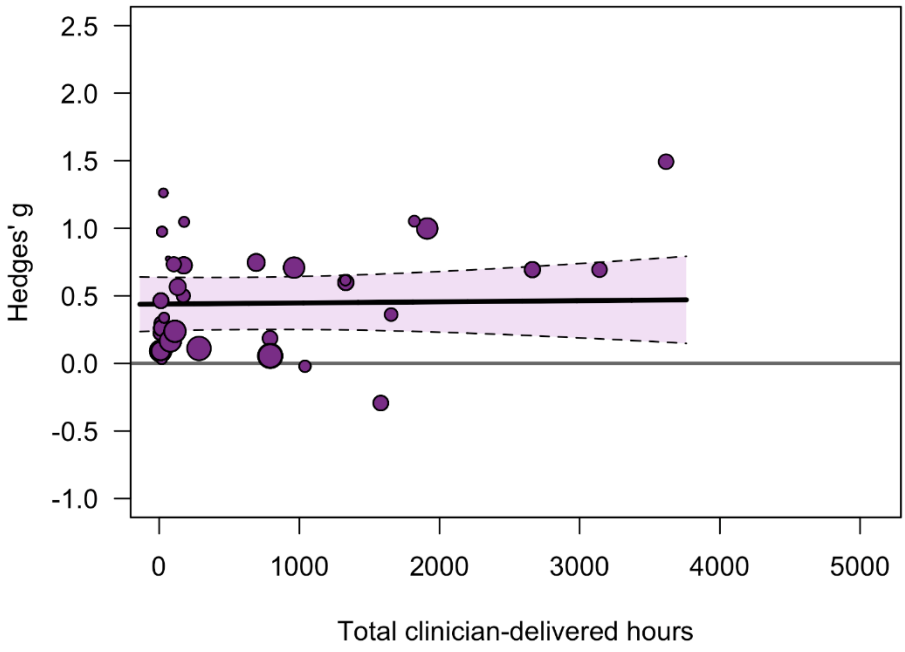


Figure B15.3. Non-linear model of monthly clinician-delivered dose by effect size (Hedges' g) for change in autism characteristic outcomes from pre- to post-intervention in the intervention group.

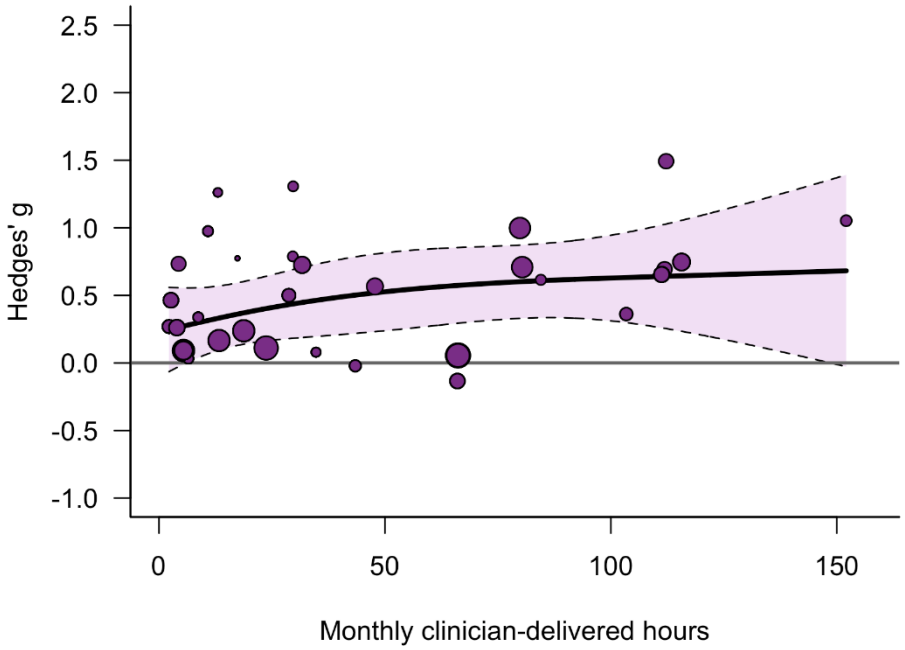
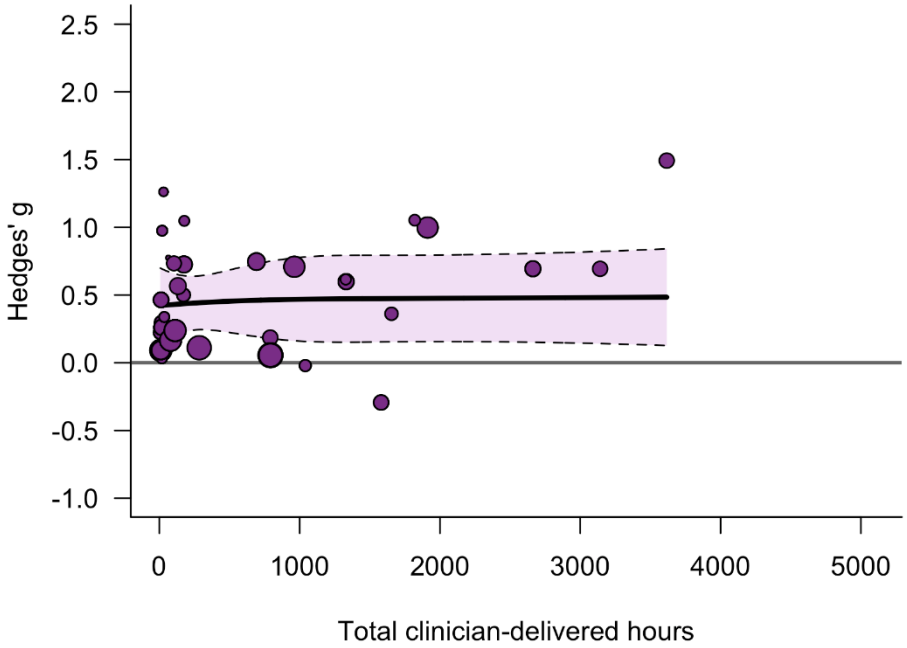


Figure B15.4. Non-linear model of total clinician-delivered dose by effect size (Hedges' g) for change in autism characteristic outcomes from pre- to post-intervention in the intervention group.



Adaptive functioning

The linear and non-linear models of total and monthly clinician-delivered hours of intervention by effect size for the change from baseline to follow-up in the behaviourally based intervention group, with 95% confidence intervals, for adaptive functioning outcomes are shown in **Figure B16**.

Figure B16.1. Linear model of monthly clinician-delivered dose by effect size (Hedges' g) for change in adaptive functioning outcomes from pre- to post-intervention in the intervention group.

Note: These figures have been grouped together using a number-letter referencing system as they are related. Hedges' g > 0 = improvement in outcomes from baseline to follow-up. Hedges' g > 0 = decrease in outcomes from baseline to follow-up.

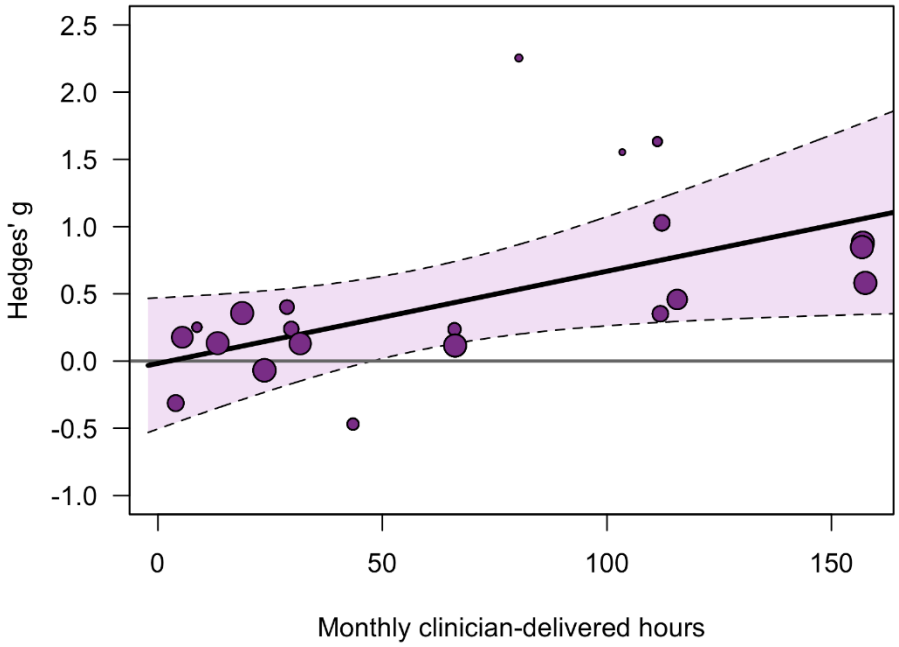


Figure B16.2. Linear model of total clinician-delivered dose by effect size (Hedges' g) for change in adaptive functioning outcomes from pre- to post-intervention in the intervention group.

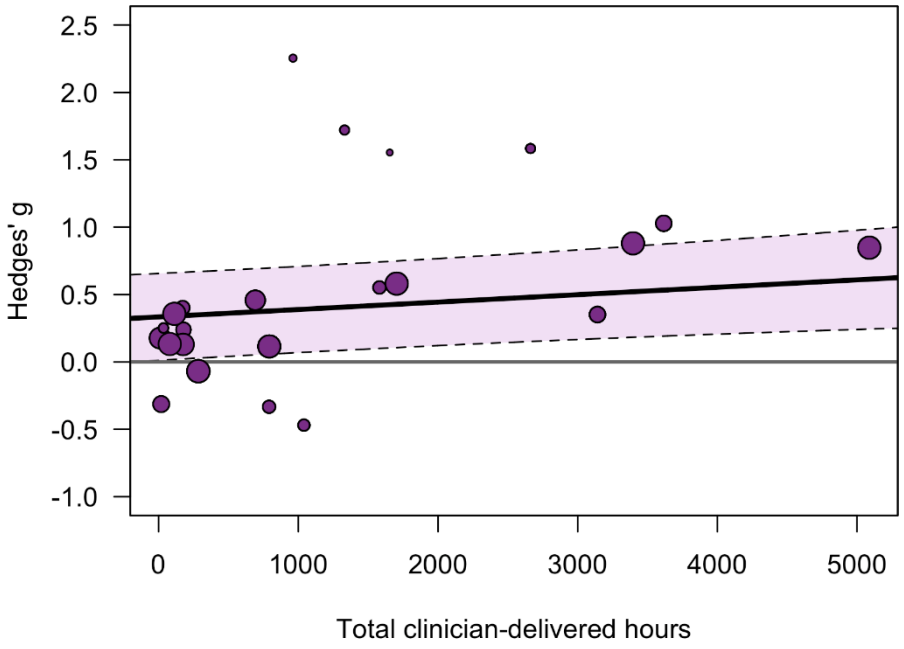


Figure B16.3. Non-linear model of monthly clinician-delivered dose by effect size (Hedges' g) for change in adaptive functioning outcomes from pre- to post-intervention in the intervention group.

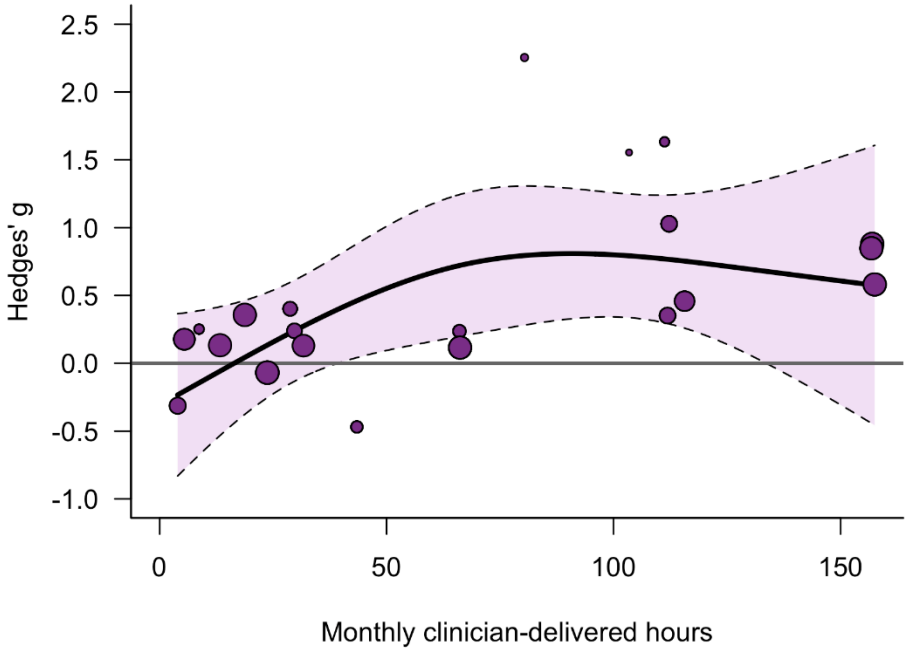
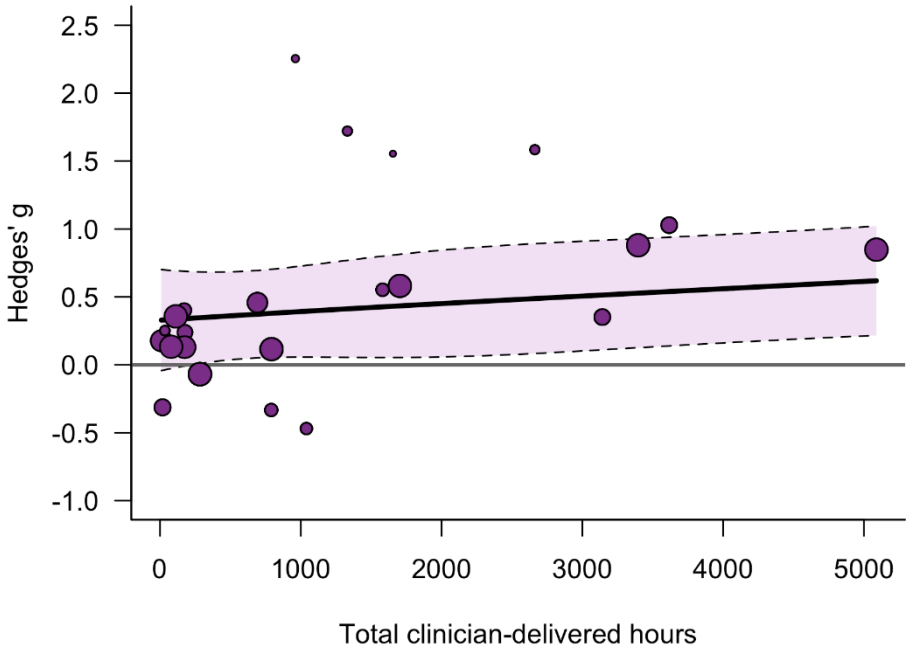


Figure B16.4. Non-linear model of total clinician-delivered dose by effect size (Hedges' g) for change in adaptive functioning outcomes from pre- to post-intervention in the intervention group.



Cognition and language

The linear and non-linear models of total and monthly clinician-delivered hours of intervention by effect size for the change from baseline to follow-up in the behaviourally based intervention group, with 95% confidence intervals, for cognition and language outcomes are shown in **Figure B17**.

Figure B17.1. Linear model of monthly clinician-delivered dose by effect size (Hedges' g) for change in cognition and language outcomes from pre- to post-intervention in the intervention group.

Note: These figures have been grouped together using a number-letter referencing system as they are related. Hedges' g > 0 = improvement in outcomes from baseline to follow-up. Hedges' g > 0 = decrease in outcomes from baseline to follow-up.

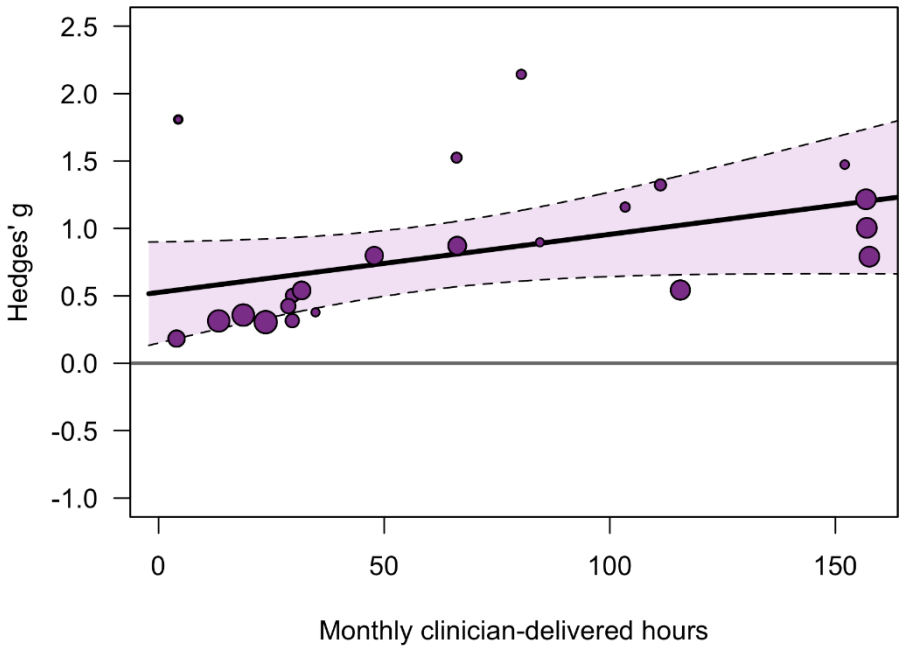


Figure B17.2. Linear model of total clinician-delivered dose by effect size (Hedges' g) for change in cognition and language outcomes from pre- to post-intervention in the intervention group.

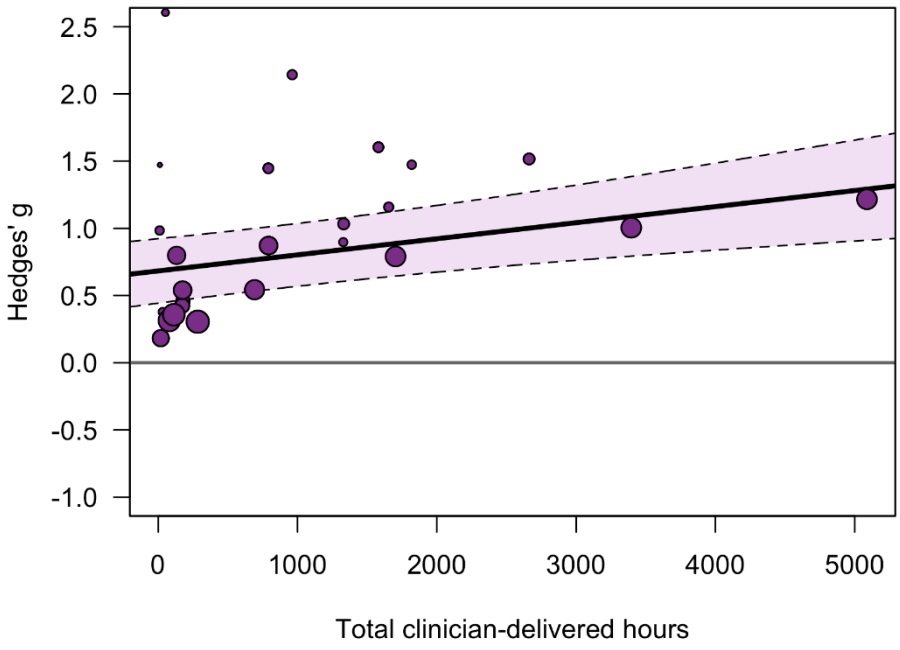


Figure B17.3. Non-linear model of monthly clinician-delivered dose by effect size (Hedges' g) for change in cognition and language outcomes from pre- to post-intervention in the intervention group.

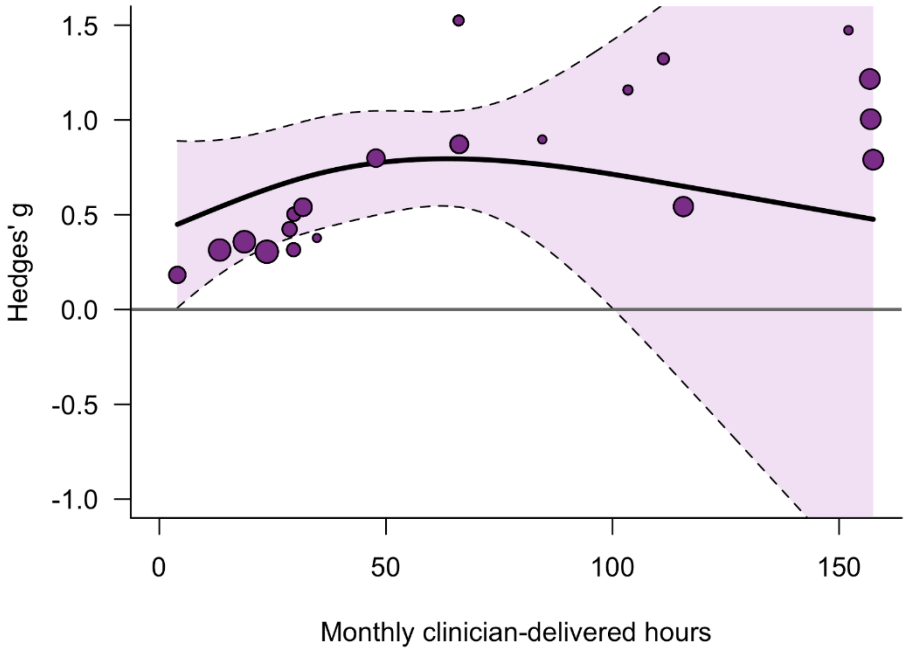
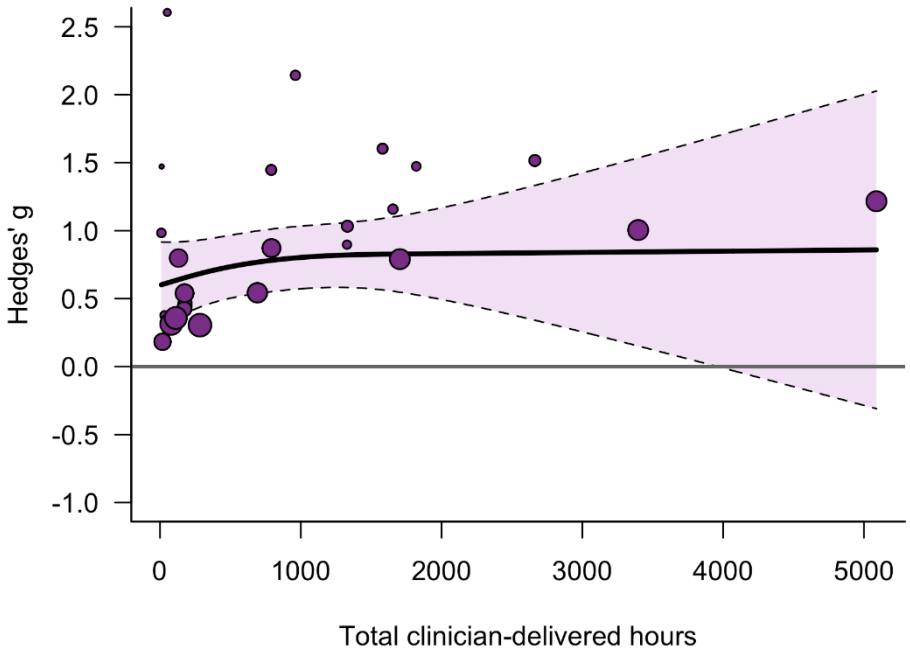


Figure B17.4. Non-linear model of total clinician-delivered dose by effect size (Hedges' g) for change in cognition and language outcomes from pre- to post-intervention in the intervention group.



Comparison group

Table B12 provides model statistics for the linear models which investigate the relationship between dose (monthly and total clinician hours) and change from baseline to follow-up in the group of children who **did not** undergo behaviourally based intervention (i.e., the comparison group). This is reported across all available outcomes, as well as within three outcome domains: autism characteristics, adaptive functioning, and cognition and language.

Table B12.1. Linear model statistics for association between dose (monthly clinician hours) and autism characteristic, adaptive functioning and cognition and language outcomes for change from baseline to follow-up in the comparison group

Note: These tables have been grouped together using a number-letter referencing system as they are related.

Dose/outcome domain	N	β	95%CI	p-value
All outcomes	17	0.0001	-0.006 to 0.006	0.965
Autism characteristics	17	0.0002	-0.006 to 0.006	0.940
Adaptive functioning	12	-0.0028	-0.021 to 0.016	0.767
Cognition and language	12	0.0025	-0.002 to 0.007	0.291

Table B12.2. Linear model statistics for association between dose (total clinician hours) and autism characteristic, adaptive functioning and cognition and language outcomes for change from baseline to follow-up in the comparison group

Dose/outcome domain	N	β	95%CI	p-value
All outcomes	19	0.0001	-0.0001 to 0.0004	0.378
Autism characteristics	18	0.0001	-0.0002 to 0.0004	0.470
Adaptive functioning	12	0.0004	-0.0004 to 0.0012	0.295
Cognition and language	13	0.0002	0.00003 to 0.00035	0.021

All outcomes

The linear and non-linear models of total and monthly clinician-delivered hours of intervention by effect size for the change from baseline to follow-up in the comparison group, with 95% confidence intervals, for all outcomes are shown in **Figure B18**.

Figure B18.1. Linear model of monthly clinician-delivered dose by effect size (Hedges' g) for change in all outcomes from pre- to post-intervention in the comparison group.

Note: These figures have been grouped together using a number-letter referencing system as they are related. Hedges' g > 0 = improvement in outcomes from baseline to follow-up. Hedges' g > 0 = decrease in outcomes from baseline to follow-up.

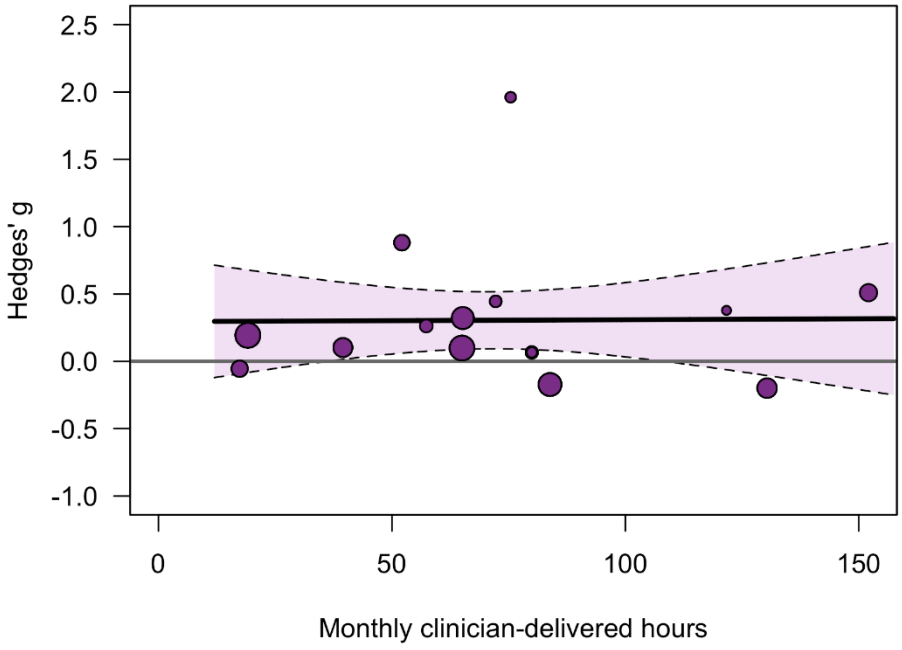


Figure B18.2. Linear model of total clinician-delivered dose by effect size (Hedges' g) for change in all outcomes from pre- to post-intervention in the comparison group.

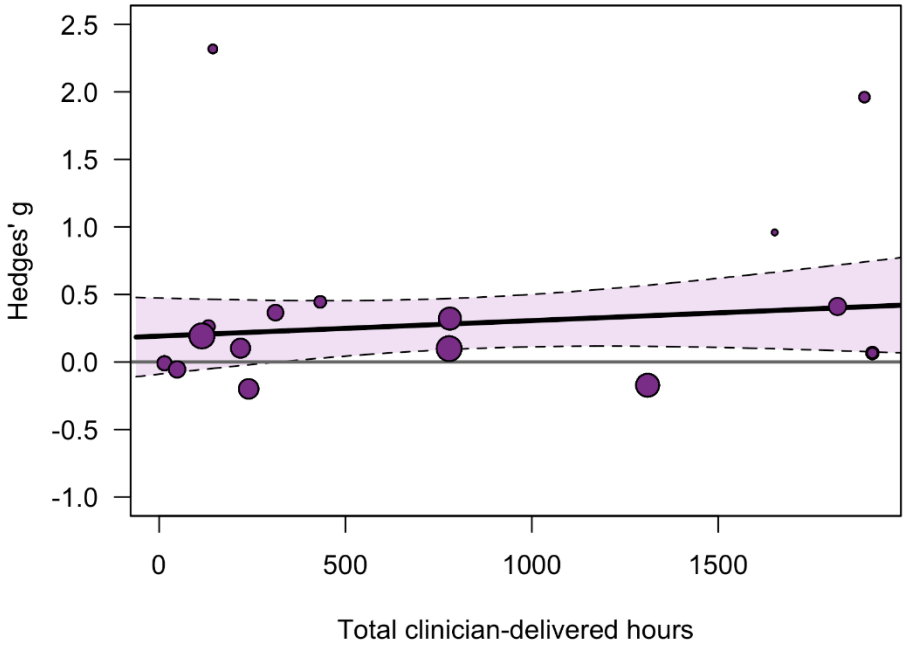


Figure B18.3. Non-linear model of monthly clinician-delivered dose by effect size (Hedges' g) for change in all outcomes from pre- to post-intervention in the comparison group.

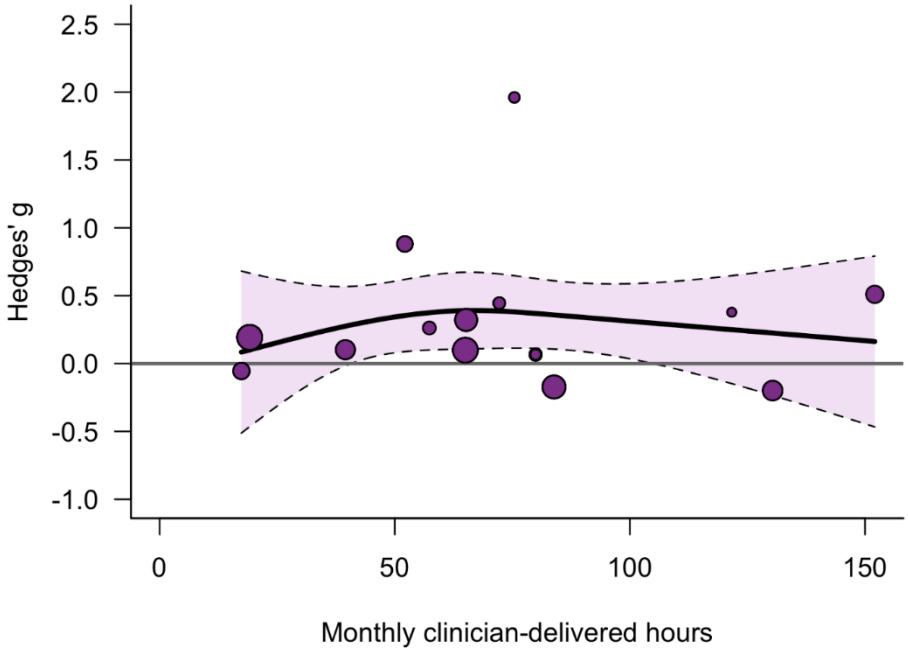
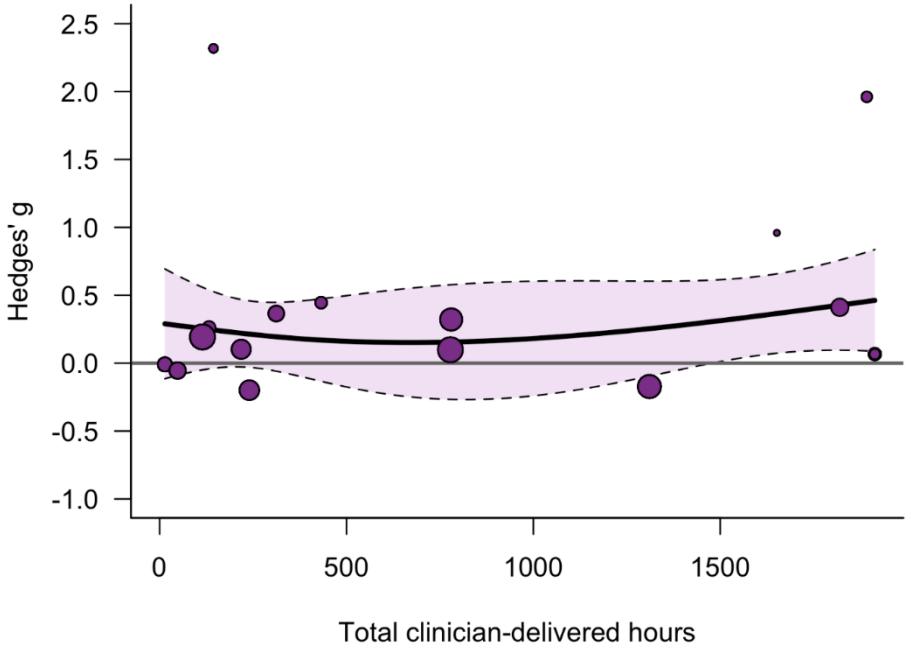


Figure B18.4. Non-linear model of total clinician-delivered dose by effect size (Hedges' g) for change in all outcomes from pre- to post-intervention in the comparison group.



Autism characteristics

The linear and non-linear models of total and monthly clinician-delivered hours of intervention by effect size for the change from baseline to follow-up in the comparison group, with 95% confidence intervals, for autism characteristic outcomes are shown in **Figure B19**.

Figure B19.1. Linear model of monthly clinician-delivered dose by effect size (Hedges' g) for change in autism characteristic outcomes from pre- to post-intervention in the comparison group.

Note: These figures have been grouped together using a number-letter referencing system as they are related. Hedges' g > 0 = improvement in outcomes from baseline to follow-up. Hedges' g > 0 = decrease in outcomes from baseline to follow-up.

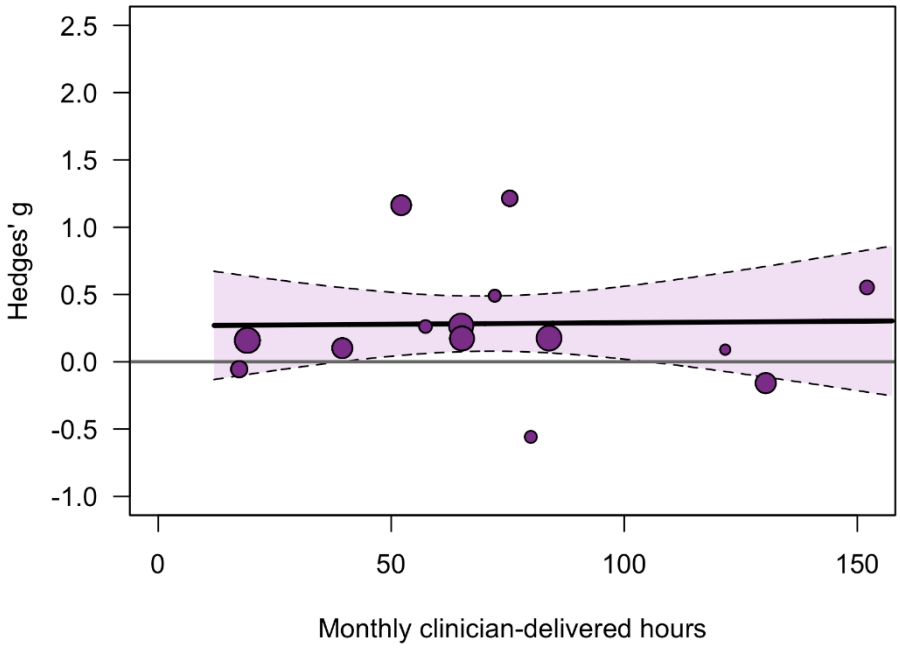


Figure B19.2. Linear model of total clinician-delivered dose by effect size (Hedges' g) for change in autism characteristic outcomes from pre- to post-intervention in the comparison group.

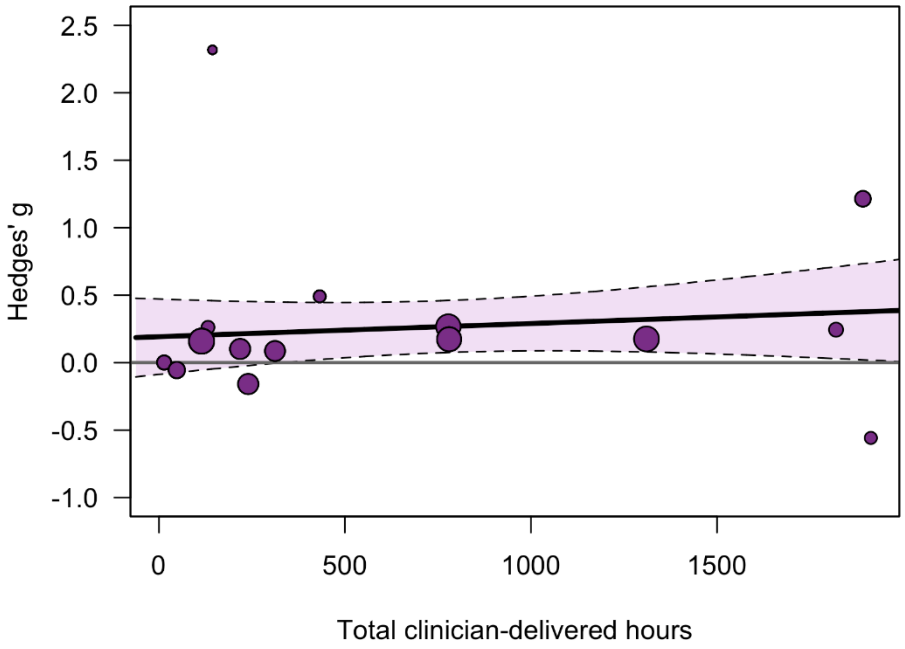


Figure B19.3. Non-linear model of monthly clinician-delivered dose by effect size (Hedges' g) for change in autism characteristic outcomes from pre- to post-intervention in the comparison group.

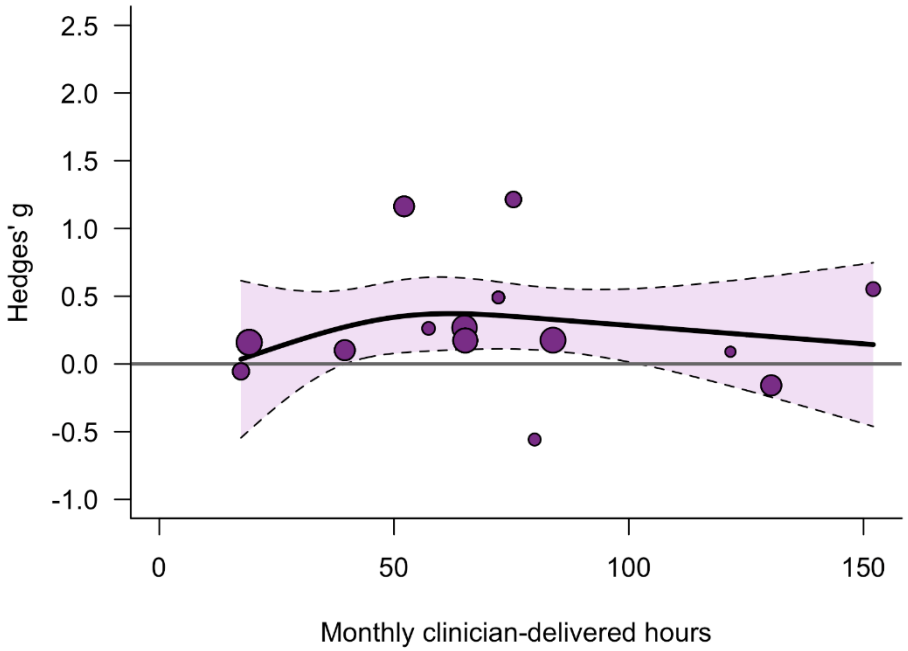
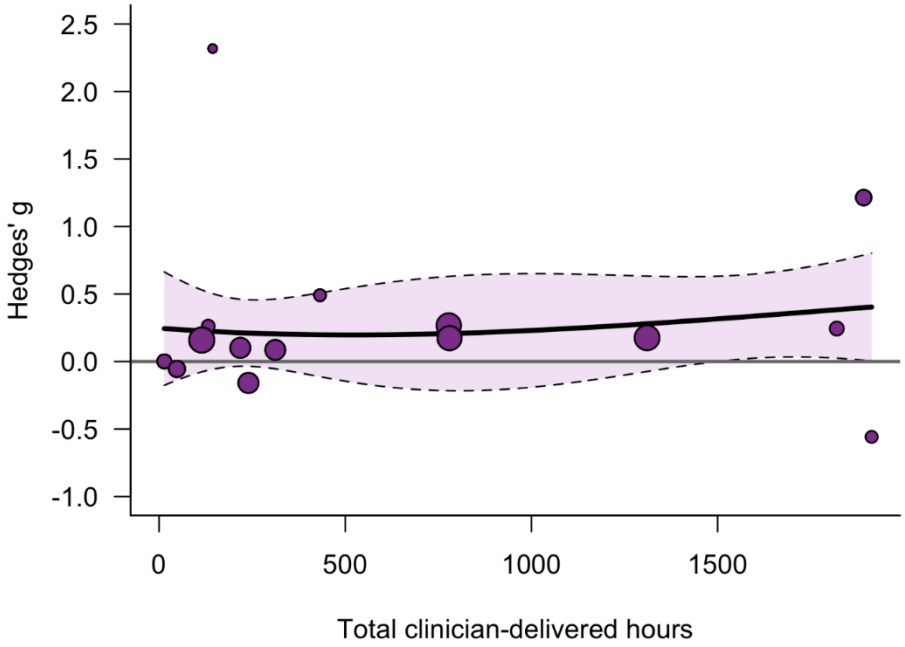


Figure B19.4. Non-linear model of total clinician-delivered dose by effect size (Hedges' g) for change in autism characteristic outcomes from pre- to post-intervention in the comparison group.



Adaptive functioning

The linear and non-linear models of total and monthly clinician-delivered hours of intervention by effect size for the change from baseline to follow-up in the comparison group, with 95% confidence intervals, for adaptive functioning outcomes are shown in **Figure B20**.

Figure B20.1. Linear model of monthly clinician-delivered dose by effect size (Hedges' g) for change in adaptive functioning outcomes from pre- to post-intervention in the comparison group.

Note: These figures have been grouped together using a number-letter referencing system as they are related. Hedges' g > 0 = improvement in outcomes from baseline to follow-up. Hedges' g > 0 = decrease in outcomes from baseline to follow-up.

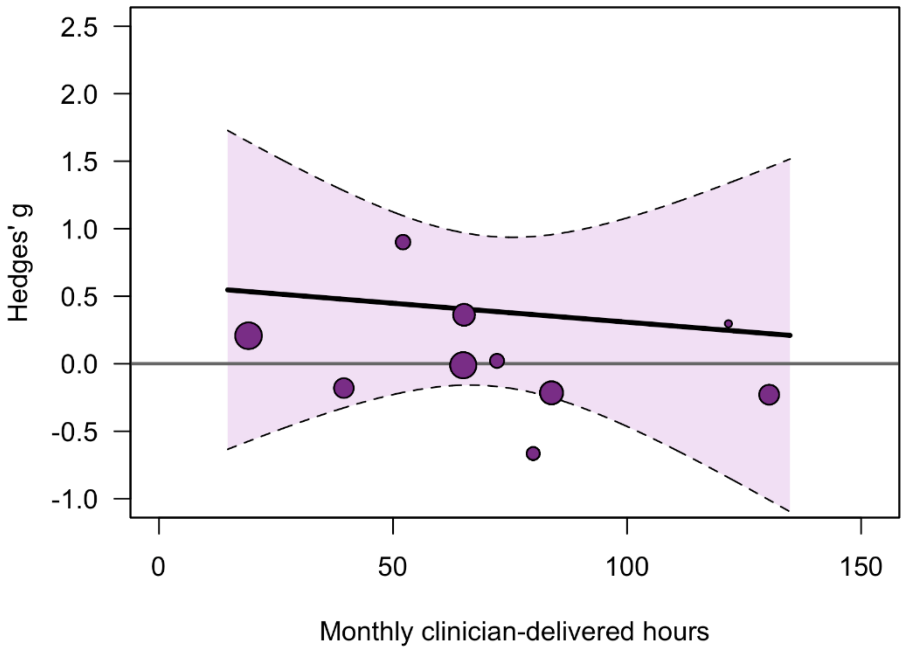


Figure B20.2. Linear model of total clinician-delivered dose by effect size (Hedges' g) for change in adaptive functioning outcomes from pre- to post-intervention in the comparison group.

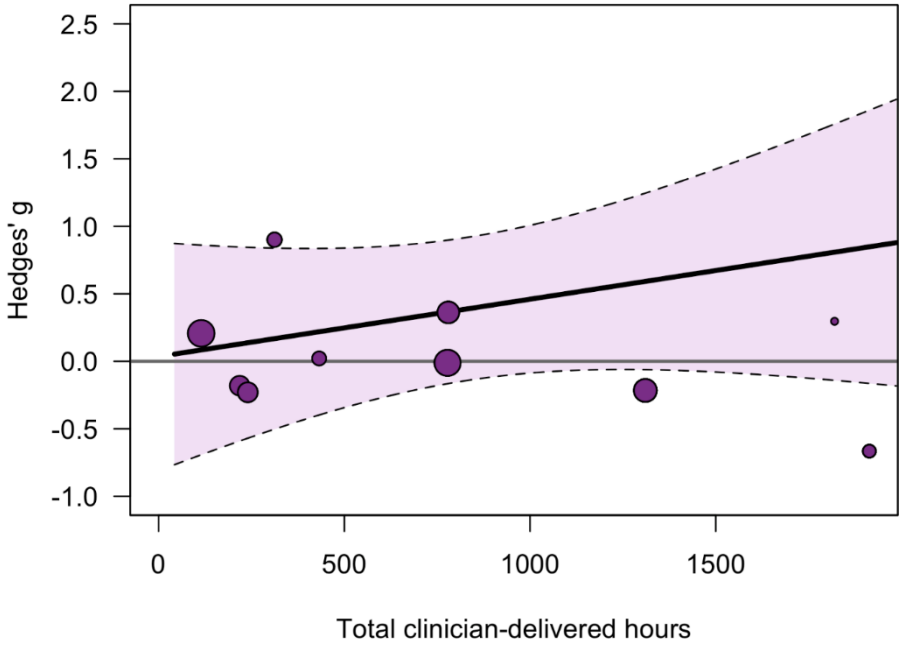


Figure B20.3. Non-linear model of monthly clinician-delivered dose by effect size (Hedges' g) for change in adaptive functioning outcomes from pre- to post-intervention in the comparison group.

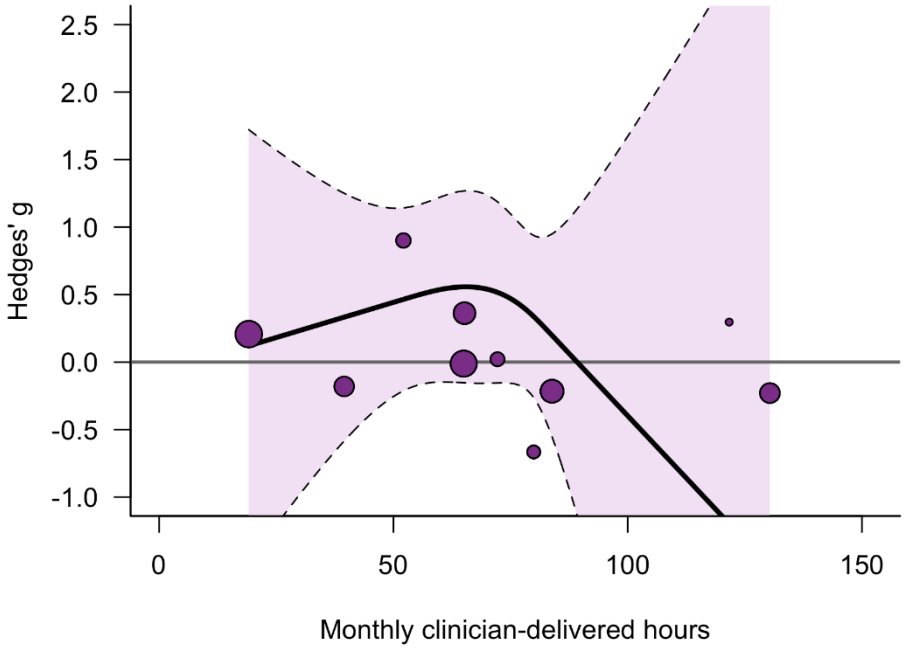
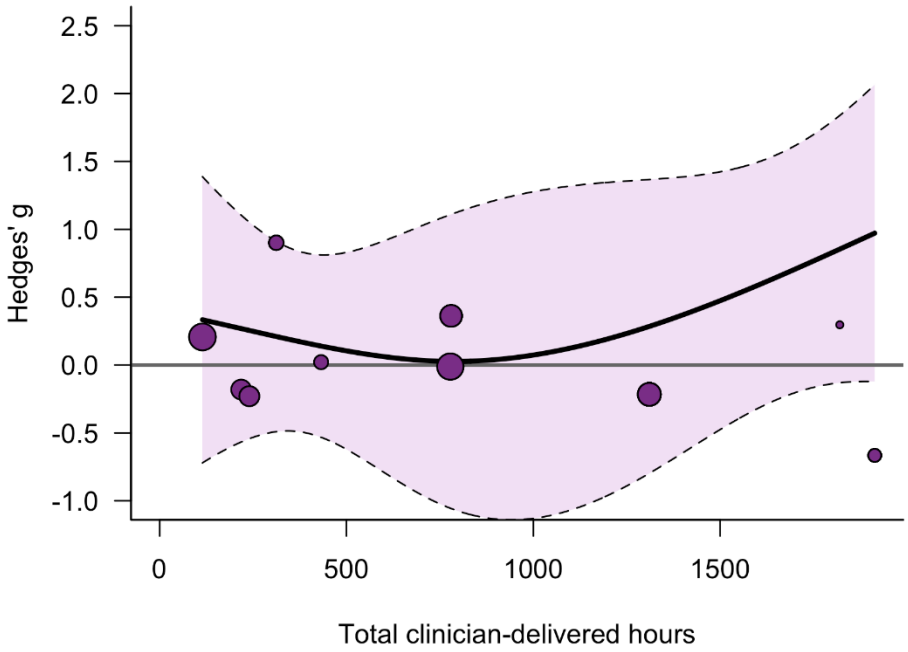


Figure B20.4. Non-linear model of total clinician-delivered dose by effect size (Hedges' g) for change in adaptive functioning outcomes from pre- to post-intervention in the comparison group.



Cognition and language

The linear and non-linear models of total and monthly clinician-delivered hours of intervention by effect size for the change from baseline to follow-up in the comparison group, with 95% confidence intervals, for cognition and language outcomes are shown in **Figure B21**.

Figure B21.1. Linear model of monthly clinician-delivered dose by effect size (Hedges' g) for change in cognition and language outcomes from pre- to post-intervention in the comparison group.

Note: These figures have been grouped together using a number-letter referencing system as they are related. Hedges' g > 0 = improvement in outcomes from baseline to follow-up. Hedges' g > 0 = decrease in outcomes from baseline to follow-up.

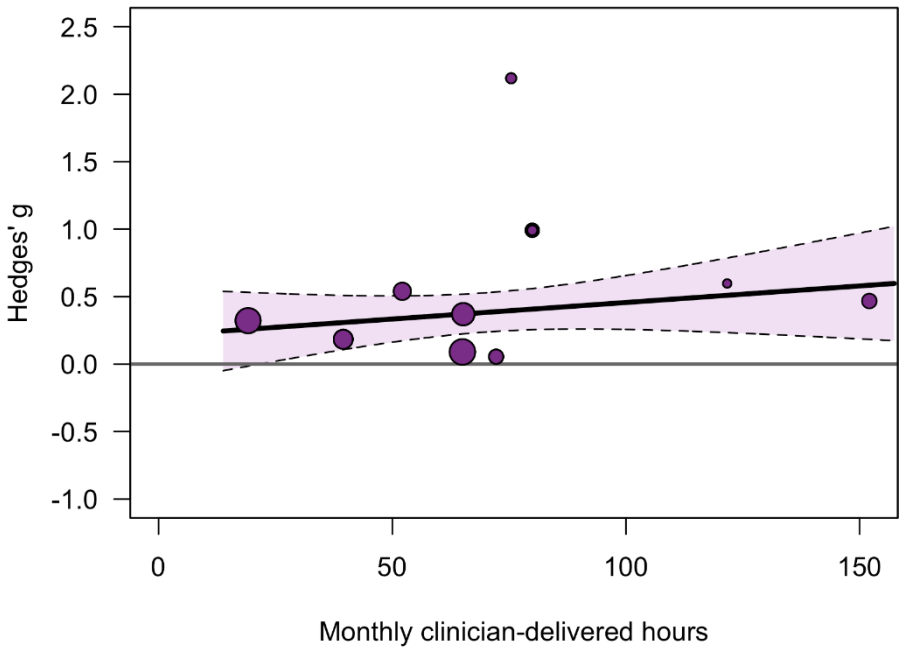


Figure B21.2. Linear model of total clinician-delivered dose by effect size (Hedges' g) for change in cognition and language outcomes from pre- to post-intervention in the comparison group.

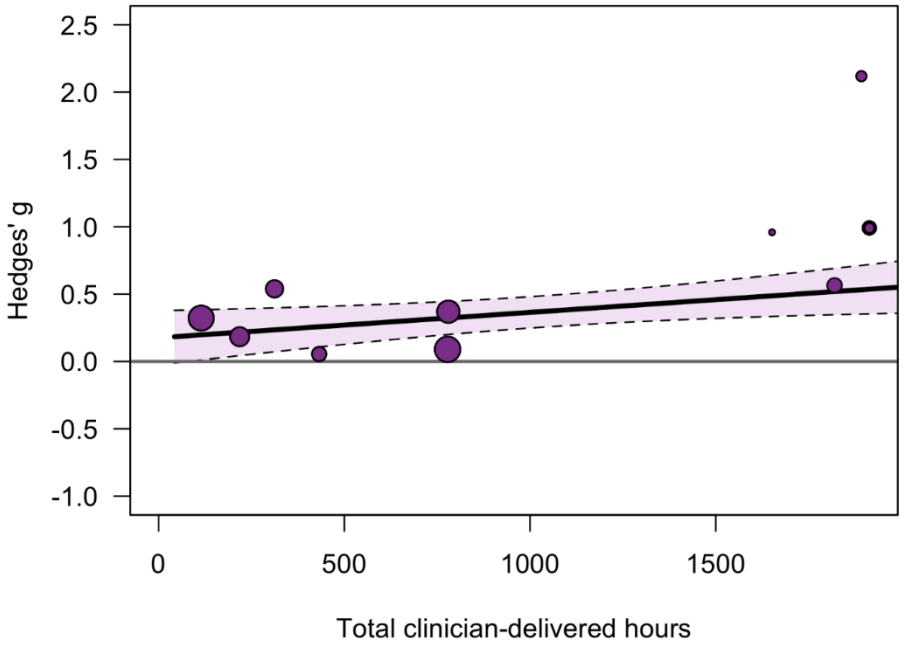


Figure B21.3. Non-linear model of monthly clinician-delivered dose by effect size (Hedges' g) for change in cognition and language outcomes from pre- to post-intervention in the comparison group.

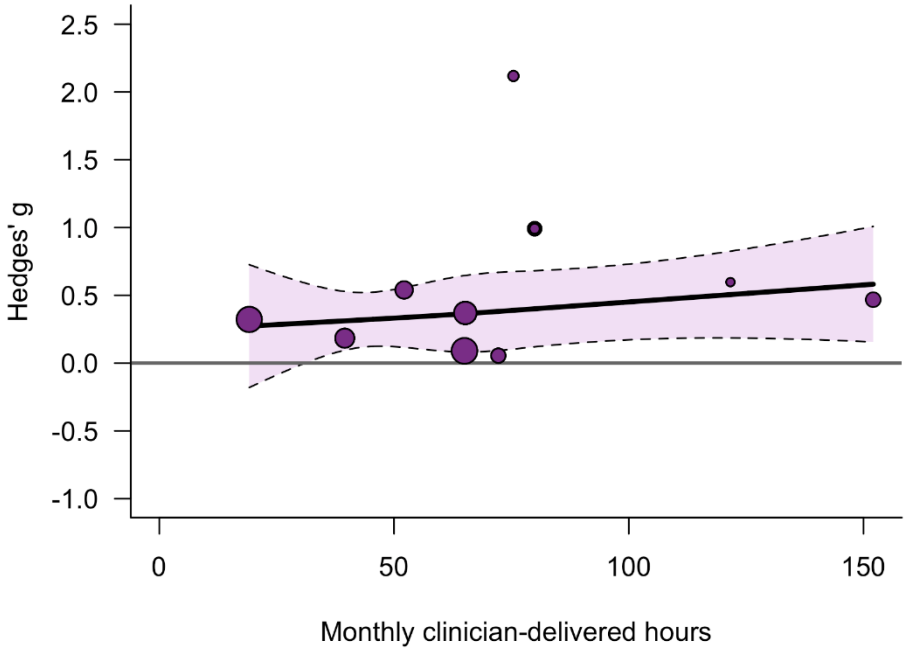
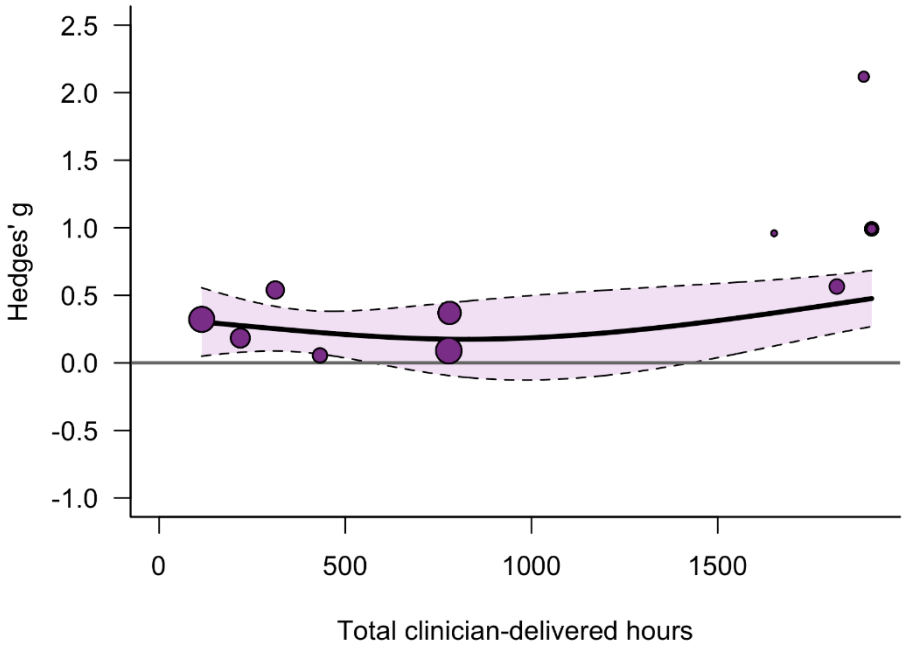


Figure B21.4. Non-linear model of total clinician-delivered dose by effect size (Hedges' g) for change in cognition and language outcomes from pre- to post-intervention in the comparison group.



B6. Investigating the effect of population, intervention, and study design factors on efficacy

Results and forest plots for each subgroup analysis are shown for each outcome domain in **Figures B22-B26**.

Figure B22. Results of subgroup analysis for autism characteristic outcome domain

Note: An accessible version of the data displayed in this figure is presented in Table B13, which follows. The *F* and the *p* statistic are from the Wald-type test. If statistically significant ($p < 0.05$), this indicates that there is a difference in efficacy of the intervention between levels of the subgroup. Tau^2 is a measure of statistical heterogeneity, which gives an estimation of the extent to which an effect estimate is inconsistent across studies.

Subgroup	N studies	Hedges' g (95% CI)	F	p	Tau^2
Parent involvement			0.15	0.930	0.12
Clinician	25	0.33 (0.20 to 0.46)			
Clinician and parent	18	0.30 (0.12 to 0.48)			
Parent delivered	31	0.33 (0.21 to 0.46)			
Teacher delivered	8	0.26 (0.01 to 0.51)			
Age			2.72	0.104	0.11
0-1 years	7	0.15 (-0.15 to 0.44)			
2-4 years	61	0.36 (0.27 to 0.44)			
5-6 years	13	0.21 (0.08 to 0.34)			
Intervention category			1.03	0.503	0.12
Behavioural	30	0.37 (0.25 to 0.49)			
Developmental	13	0.23 (0.09 to 0.37)			
NDBI	31	0.35 (0.22 to 0.49)			
Other	3	0.27 (-0.92 to 1.46)			
TEACCH	4	0.05 (-0.32 to 0.43)			
Technology-based	2	0.02 (-2.19 to 2.23)			
Comparison group			0.56	0.466	0.12
Eclectic	15	0.38 (0.19 to 0.56)			
TAU	69	0.31 (0.23 to 0.39)			
Setting			0.16	0.922	0.12
Community	9	0.28 (0.02 to 0.54)			
Early education	17	0.37 (0.19 to 0.54)			
Health	35	0.31 (0.19 to 0.42)			
Home	25	0.31 (0.18 to 0.45)			
Study design			0.37	0.693	0.12
Cohort	13	0.26 (0.09 to 0.43)			
Non-random	23	0.35 (0.19 to 0.52)			
Random	45	0.32 (0.22 to 0.41)			

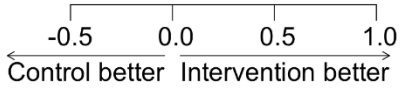


Table B13. Table version of results of subgroup analysis for autism characteristic outcome domain

Note: This table presents the information displayed in Figure B22 in an accessible format. The F and the p statistic are from the Wald-type test. If statistically significant ($p < 0.05$), this indicates that there is a difference in efficacy of the intervention between levels of the subgroup. Tau^2 is a measure of statistical heterogeneity, which gives an estimation of the extent to which an effect estimate is inconsistent across studies. NA = not applicable.

Subgroup	N studies	Hedges' g (95% CI)	F	p	Tau^2
Subgroup: Parent involvement	NA	NA	0.15	0.93	0.12
Clinician	25	0.33 (0.20 to 0.46)	NA	NA	NA
Clinician and parent	18	0.30 (0.12 to 0.48)	NA	NA	NA
Parent delivered	31	0.33 (0.21 to 0.46)	NA	NA	NA
Teacher delivered	8	0.26 (0.01 to 0.51)	NA	NA	NA
Subgroup: Age	NA	NA	2.72	0.104	0.11
0-1 years	7	0.15 (-0.15 to 0.44)	NA	NA	NA
2-4 years	61	0.36 (0.27 to 0.44)	NA	NA	NA
5-6 years	13	0.21 (0.08 to 0.34)	NA	NA	NA
Subgroup: Intervention category	NA	NA	1.03	0.503	0.12
Behavioural	30	0.37 (0.25 to 0.49)	NA	NA	NA
Developmental	13	0.23 (0.09 to 0.37)	NA	NA	NA
NDBI	31	0.35 (0.22 to 0.49)	NA	NA	NA
Other	3	0.27 (-0.92 to 1.46)	NA	NA	NA
TEACCH	4	0.05 (-0.32 to 0.43)	NA	NA	NA
Technology-based	2	0.02 (-2.19 to 2.23)	NA	NA	NA
Subgroup: Comparison group	NA	NA	0.05	0.82	0.12
Eclectic	15	0.38 (0.16 to 0.50)	NA	NA	NA
TAU	69	0.31 (0.23 to 0.39)	NA	NA	NA
Subgroup: Setting	NA	NA	0.16	0.922	0.12
Community	9	0.28 (0.02 to 0.54)	NA	NA	NA
Early education	17	0.37 (0.19 to 0.54)	NA	NA	NA
Health	35	0.31 (0.19 to 0.42)	NA	NA	NA
Home	25	0.31 (0.18 to 0.45)	NA	NA	NA
Subgroup: Study design	NA	NA	0.37	0.693	0.12
Cohort	13	0.26 (0.09 to 0.43)	NA	NA	NA
Non-random	23	0.35 (0.19 to 0.52)	NA	NA	NA
Random	45	0.32 (0.22 to 0.41)	NA	NA	NA

Figure B23. Results of subgroup analysis for adaptive functioning outcome domain

Note: An accessible version of the data displayed in this figure is presented in Table B14, which follows. The *F* and the *p* statistic are from the Wald-type test. If statistically significant ($p < 0.05$), this indicates that there is a difference in efficacy of the intervention between levels of the subgroup. Tau^2 is a measure of statistical heterogeneity, which gives an estimation of the extent to which an effect estimate is inconsistent across studies.

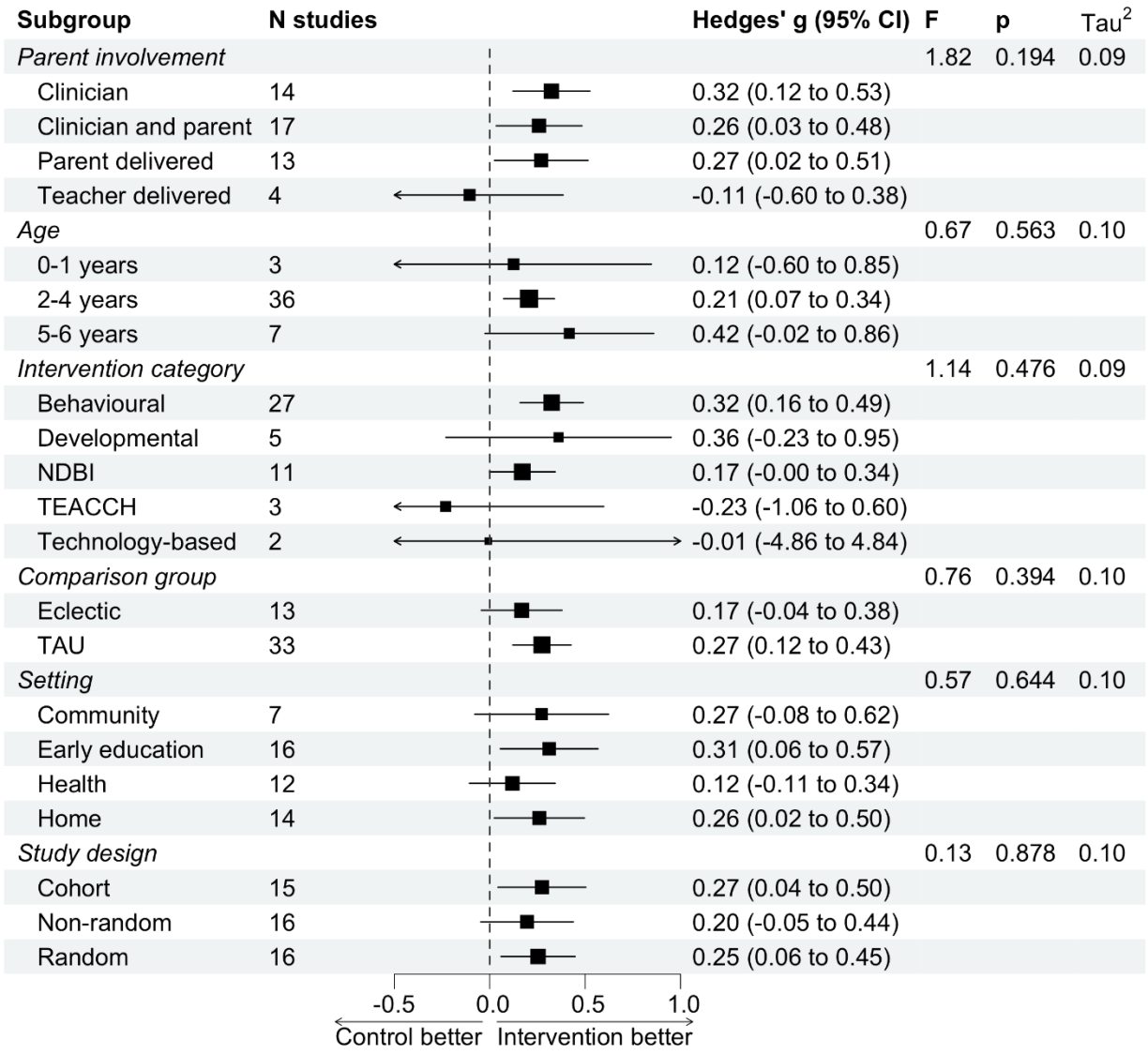


Table B14. Table version of results of subgroup analysis for adaptive functioning outcome domain

Note: This table presents the information displayed in Figure B23 in an accessible format. The F and the p statistic are from the Wald-type test. If statistically significant ($p < 0.05$), this indicates that there is a difference in efficacy of the intervention between levels of the subgroup. Tau^2 is a measure of statistical heterogeneity, which gives an estimation of the extent to which an effect estimate is inconsistent across studies. NA = not applicable.

Subgroup	N studies	Hedges' g (95% CI)	F	p	Tau^2
Subgroup: Parent involvement	NA	NA	1.82	0.194	0.09
Clinician	14	0.32 (0.12 to 0.53)	NA	NA	NA
Clinician and parent	17	0.26 (0.03 to 0.48)	NA	NA	NA
Parent delivered	13	0.27 (0.02 to 0.51)	NA	NA	NA
Teacher delivered	4	-0.11 (-0.60 to 0.38)	NA	NA	NA
Subgroup: Age	NA	NA	0.67	0.563	0.1
0-1 years	3	0.12 (-0.60 to 0.85)	NA	NA	NA
2-4 years	36	0.21 (0.07 to 0.34)	NA	NA	NA
5-6 years	7	0.42 (-0.02 to 0.86)	NA	NA	NA
Subgroup: Intervention category	NA	NA	1.14	0.476	0.09
Behavioural	27	0.32 (0.16 to 0.49)	NA	NA	NA
Developmental	5	0.36 (-0.23 to 0.95)	NA	NA	NA
NDBI	11	0.17 (-0.00 to 0.34)	NA	NA	NA
TEACCH	3	-0.23 (-1.06 to 0.60)	NA	NA	NA
Technology-based	2	-0.01 (-4.86 to 4.84)	NA	NA	NA
Subgroup: Comparison group	NA	NA	0.75	0.396	0.1
Eclectic	14	0.17 (-0.01 to 0.36)	NA	NA	NA
TAU	33	0.27 (0.12 to 0.42)	NA	NA	NA
Subgroup: Setting	NA	NA	0.57	0.644	0.1
Community	7	0.27 (-0.08 to 0.62)	NA	NA	NA
Early education	16	0.31 (0.06 to 0.57)	NA	NA	NA
Health	12	0.12 (-0.11 to 0.34)	NA	NA	NA
Home	14	0.26 (0.02 to 0.50)	NA	NA	NA
Subgroup: Study design	NA	NA	0.13	0.878	0.1
Cohort	15	0.27 (0.04 to 0.50)	NA	NA	NA
Non-random	16	0.20 (-0.05 to 0.44)	NA	NA	NA
Random	16	0.25 (0.06 to 0.45)	NA	NA	NA

Figure B24. Results of subgroup analysis for cognition and language outcome domain

Note: An accessible version of the data displayed in this figure is presented in Table B15, which follows. The *F* and the *p* statistic are from the Wald-type test. If statistically significant ($p < 0.05$), this indicates that there is a difference in efficacy of the intervention between levels of the subgroup. Tau^2 is a measure of statistical heterogeneity, which gives an estimation of the extent to which an effect estimate is inconsistent across studies.

Subgroup	N studies	Hedges' g (95% CI)	F	p	Tau^2
Parent involvement			0.99	0.414	0.05
Clinician	20	0.41 (0.20 to 0.63)			
Clinician and parent	20	0.32 (0.18 to 0.45)			
Parent delivered	18	0.22 (0.07 to 0.36)			
Teacher delivered	7	0.22 (-0.06 to 0.51)			
Age			0.54	0.605	0.05
0-1 years	6	0.32 (0.05 to 0.60)			
2-4 years	49	0.28 (0.18 to 0.37)			
5-6 years	7	0.47 (0.01 to 0.94)			
Intervention category			1.22	0.475	0.04
Behavioural	28	0.41 (0.25 to 0.57)			
Developmental	7	0.06 (-0.08 to 0.21)			
NDBI	23	0.29 (0.17 to 0.41)			
Other	2	0.21 (-2.79 to 3.21)			
TEACCH	3	0.11 (-0.52 to 0.73)			
Technology-based	3	0.25 (-0.29 to 0.79)			
Comparison group			0.86	0.365	0.05
Eclectic	16	0.23 (0.01 to 0.46)			
TAU	46	0.34 (0.25 to 0.43)			
Setting			0.61	0.619	0.05
Community	6	0.17 (-0.13 to 0.46)			
Early education	20	0.37 (0.18 to 0.56)			
Health	21	0.29 (0.13 to 0.45)			
Home	18	0.30 (0.15 to 0.45)			
Study design			2.16	0.131	0.05
Cohort	16	0.33 (0.08 to 0.58)			
Non-random	20	0.39 (0.25 to 0.52)			
Random	28	0.22 (0.13 to 0.32)			

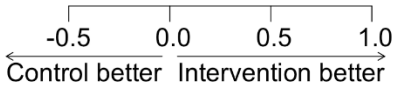


Table B15. Table version of results of subgroup analysis for cognition and language outcome domain

Note: This table presents the information displayed in Figure B24 in an accessible format. The F and the p statistic are from the Wald-type test. If statistically significant ($p < 0.05$), this indicates that there is a difference in efficacy of the intervention between levels of the subgroup. Tau^2 is a measure of statistical heterogeneity, which gives an estimation of the extent to which an effect estimate is inconsistent across studies. NA = not applicable.

Subgroup	N studies	Hedges' g (95% CI)	F	p	Tau^2
Subgroup: Parent involvement	NA	NA	0.99	0.414	0.05
Clinician	20	0.41 (0.20 to 0.63)	NA	NA	NA
Clinician and parent	20	0.32 (0.18 to 0.45)	NA	NA	NA
Parent delivered	18	0.22 (0.07 to 0.36)	NA	NA	NA
Teacher delivered	7	0.22 (-0.06 to 0.51)	NA	NA	NA
Subgroup: Age	NA	NA	0.54	0.605	0.05
0-1 years	6	0.32 (0.05 to 0.60)	NA	NA	NA
2-4 years	49	0.28 (0.18 to 0.37)	NA	NA	NA
5-6 years	7	0.47 (0.01 to 0.94)	NA	NA	NA
Subgroup: Intervention category	NA	NA	1.22	0.475	0.04
Behavioural	28	0.41 (0.25 to 0.57)	NA	NA	NA
Developmental	7	0.06 (-0.08 to 0.21)	NA	NA	NA
NDBI	23	0.29 (0.17 to 0.41)	NA	NA	NA
Other	2	0.21 (-2.79 to 3.21)	NA	NA	NA
TEACCH	3	0.11 (-0.52 to 0.73)	NA	NA	NA
Technology-based	3	0.25 (-0.29 to 0.79)	NA	NA	NA
Subgroup: Comparison group	NA	NA	0.62	0.439	0.05
Eclectic	16	0.23 (0.01 to 0.46)	NA	NA	NA
TAU	48	0.34 (0.25 to 0.43)	NA	NA	NA
Subgroup: Setting	NA	NA	0.61	0.619	0.05
Community	6	0.17 (-0.13 to 0.46)	NA	NA	NA
Early education	20	0.37 (0.18 to 0.56)	NA	NA	NA
Health	21	0.29 (0.13 to 0.45)	NA	NA	NA
Home	18	0.30 (0.15 to 0.45)	NA	NA	NA
Subgroup: Study design	NA	NA	2.16	0.131	0.05
Cohort	16	0.33 (0.08 to 0.58)	NA	NA	NA
Non-random	20	0.39 (0.25 to 0.52)	NA	NA	NA
Random	28	0.22 (0.13 to 0.32)	NA	NA	NA

Figure B25. Results of subgroup analysis for family outcomes domain

Note: An accessible version of the data displayed in this figure is presented in Table B16, which follows. The *F* and the *p* statistic are from the Wald-type test. If statistically significant ($p < 0.05$), this indicates that there is a difference in efficacy of the intervention between levels of the subgroup. Tau^2 is a measure of statistical heterogeneity, which gives an estimation of the extent to which an effect estimate is inconsistent across studies.

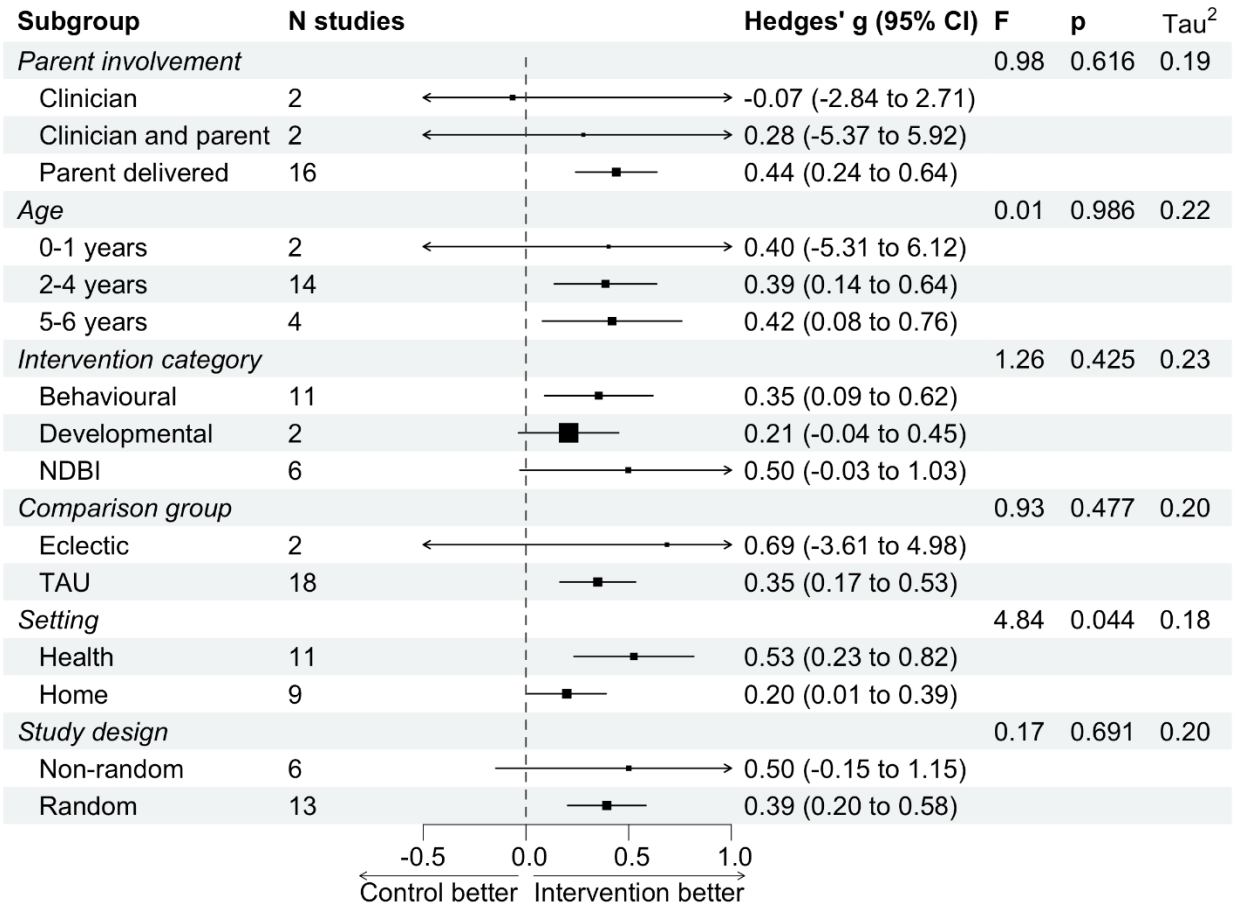


Table B16. Table version of results of subgroup analysis for family outcome domain

Note: This table presents the information displayed in Figure B25 in an accessible format. The F and the p statistic are from the Wald-type test. If statistically significant ($p < 0.05$), this indicates that there is a difference in efficacy of the intervention between levels of the subgroup. Tau^2 is a measure of statistical heterogeneity, which gives an estimation of the extent to which an effect estimate is inconsistent across studies. NA = not applicable.

Subgroup	N studies	Hedges' g (95% CI)	F	p	Tau^2
Subgroup: Parent involvement	NA	NA	0.98	0.616	0.19
Clinician	2	-0.07 (-2.84 to 2.71)	NA	NA	NA
Clinician and parent	2	0.28 (-5.37 to 5.92)	NA	NA	NA
Parent delivered	16	0.44 (0.24 to 0.64)	NA	NA	NA
Subgroup: Age	NA	NA	0.01	0.986	0.22
0-1 years	2	0.40 (-5.31 to 6.12)	NA	NA	NA
2-4 years	14	0.39 (0.14 to 0.64)	NA	NA	NA
5-6 years	4	0.42 (0.08 to 0.76)	NA	NA	NA
Subgroup: Intervention category	NA	NA	1.26	0.425	0.23
Behavioural	11	0.35 (0.09 to 0.62)	NA	NA	NA
Developmental	2	0.21 (-0.04 to 0.45)	NA	NA	NA
NDBI	6	0.50 (-0.03 to 1.03)	NA	NA	NA
Subgroup: Comparison group	NA	NA	0.93	0.477	0.2
Eclectic	2	0.69 (-3.61 to 4.98)	NA	NA	NA
TAU	18	0.35 (0.17 to 0.53)	NA	NA	NA
Subgroup: Setting	NA	NA	4.84	0.044	0.18
Health	11	0.53 (0.23 to 0.82)	NA	NA	NA
Home	9	0.20 (0.01 to 0.39)	NA	NA	NA
Subgroup: Study design	NA	NA	0.17	0.691	0.2
Non-random	6	0.50 (-0.15 to 1.15)	NA	NA	NA
Random	13	0.39 (0.20 to 0.58)	NA	NA	NA

Figure B26. Results of subgroup analysis for adverse effects outcome domain

Note: An accessible version of the data displayed in this figure is presented in Table B17, which follows. The *F* and the *p* statistic are from the Wald-type test. If statistically significant ($p < 0.05$), this indicates that there is a difference in efficacy of the intervention between levels of the subgroup. Tau^2 is a measure of statistical heterogeneity, which gives an estimation of the extent to which an effect estimate is inconsistent across studies.

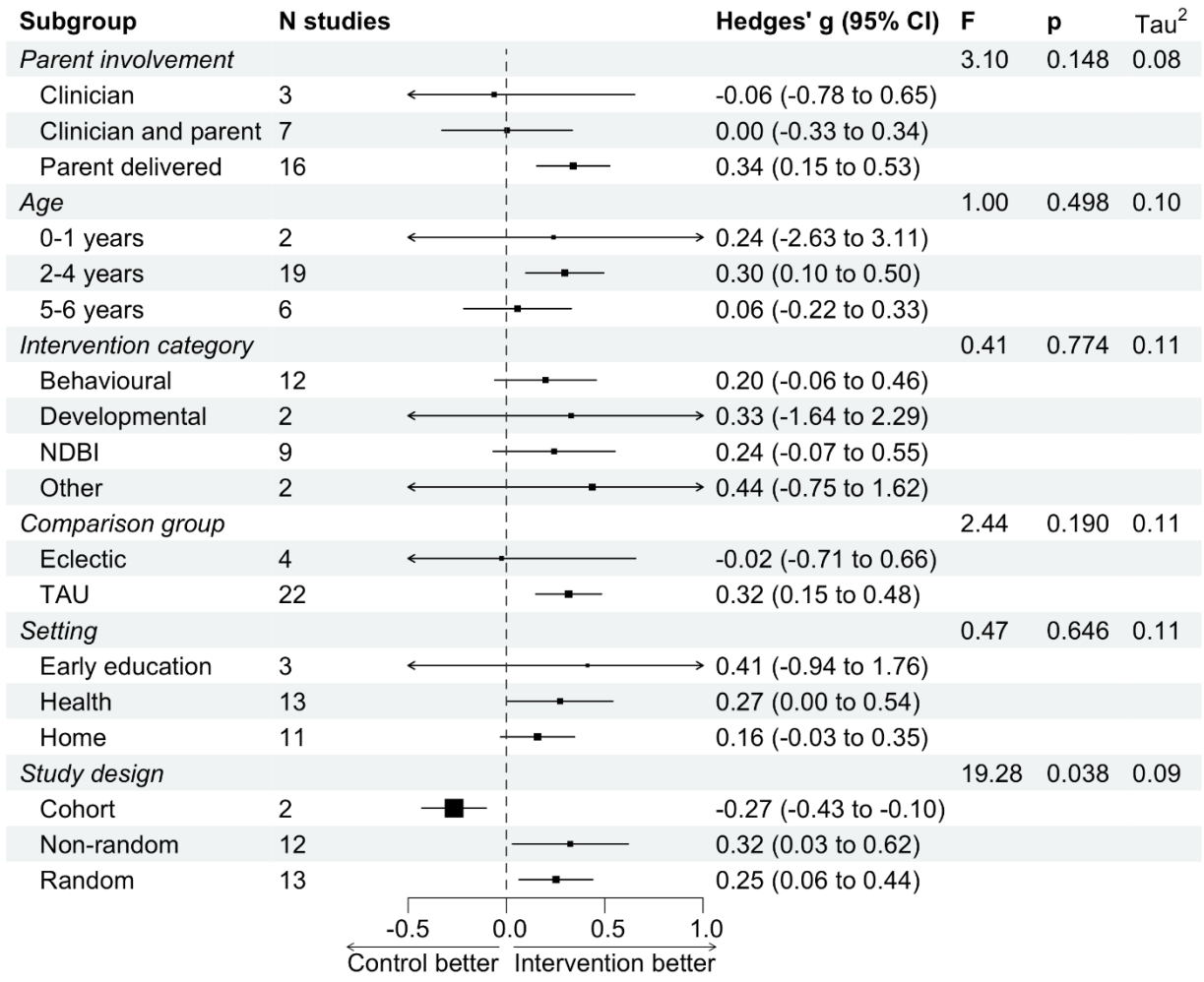


Table B17. Table version of results of subgroup analysis for adverse effects outcome domain

Note: This table presents the information displayed in Figure B26 in an accessible format. The F and the p statistic are from the Wald-type test. If statistically significant ($p < 0.05$), this indicates that there is a difference in efficacy of the intervention between levels of the subgroup. Tau^2 is a measure of statistical heterogeneity, which gives an estimation of the extent to which an effect estimate is inconsistent across studies. NA = not applicable.

Subgroup	N studies	Hedges' g (95% CI)	F	p	Tau^2
Subgroup: Parent involvement	NA	NA	3.1	0.148	0.08
Clinician	3	-0.06 (-0.78 to 0.65)	NA	NA	NA
Clinician and parent	7	0.00 (-0.33 to 0.34)	NA	NA	NA
Parent delivered	16	0.34 (0.15 to 0.53)	NA	NA	NA
Subgroup: Age	NA	NA	1	0.498	0.1
0-1 years	2	0.24 (-2.63 to 3.11)	NA	NA	NA
2-4 years	19	0.30 (0.10 to 0.50)	NA	NA	NA
5-6 years	6	0.06 (-0.22 to 0.33)	NA	NA	NA
Subgroup: Intervention category	NA	NA	0.41	0.774	0.11
Behavioural	12	0.20 (-0.06 to 0.46)	NA	NA	NA
Developmental	2	0.33 (-1.64 to 2.29)	NA	NA	NA
NDBI	9	0.24 (-0.07 to 0.55)	NA	NA	NA
Other	2	0.44 (-0.75 to 1.62)	NA	NA	NA
Subgroup: Comparison group	NA	NA	2.15	0.215	0.09
Eclectic	4	-0.00 (-0.65 to 0.65)	NA	NA	NA
TAU	23	0.30 (0.14 to 0.46)	NA	NA	NA
Subgroup: Setting	NA	NA	0.47	0.646	0.11
Early education	3	0.41 (-0.94 to 1.76)	NA	NA	NA
Health	13	0.27 (0.00 to 0.54)	NA	NA	NA
Home	11	0.16 (-0.03 to 0.35)	NA	NA	NA
Subgroup: Study design	NA	NA	19.28	0.038	0.09
Cohort	2	-0.27 (-0.43 to -0.10)	NA	NA	NA
Non-random	12	0.32 (0.03 to 0.62)	NA	NA	NA
Random	13	0.25 (0.06 to 0.44)	NA	NA	NA

Appendix References

Argumedes, M., Lanovaz, M. J., Larivee, S., & Giannakakos, A. R. (2021). Using the Prevent-Teach-Reinforce model to reduce challenging behaviors in children with autism spectrum disorder in home settings: A feasibility study. *Research in Autism Spectrum Disorders, 86 (no pagination)*. doi:http://dx.doi.org/10.1016/j.rasd.2021.101804

Arora, T. (2008). Perseveration in young children with autism and the association with joint attention. *Dissertation Abstracts International Section A: Humanities and Social Sciences, 69*(6-A), 2216.

Azarbehi, A. C. (2012). The effectiveness of early intervention programs for children with autism: A one-year follow-up study of Intensive Behavioural Intervention versus preschool integration. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 73*(2-B), 1290.

Barrett, A. C., Vernon, T. W., McGarry, E. S., Holden, A. N., Bradshaw, J., Ko, J. A., . . . German, T. C. (2020). Social responsiveness and language use associated with an enhanced PRT approach for young children with ASD: results from a pilot RCT of the PRISM model. *Research in autism spectrum disorders, 71*.

Bearss, K., Johnson, C., Smith, T., Lecavalier, L., Swiezy, N., Aman, M., . . . Scahill, L. (2015). Effect of parent training vs parent education on behavioral problems in children with autism spectrum disorder: a randomized clinical trial. *JAMA, 313*(15), 1524-1533. doi:https://dx.doi.org/10.1001/jama.2015.3150

Bentenuto, A., Bertamini, G., Perzolli, S., & Venuti, P. (2020). Changes in Developmental Trajectories of Preschool Children with Autism Spectrum Disorder during Parental Based Intensive Intervention. *Brain Sciences, 10*(5), 12. doi:https://dx.doi.org/10.3390/brainsci10050289

Bernard-Opitz, V., Ing, S., & Kong, T. Y. (2004). Comparison of behavioural and natural play interventions for young children with autism. *Autism, 8*(3), 319-333.

Blackman, A. L., Jimenez-Gomez, C., & Shvarts, S. (2020). Comparison of the efficacy of online versus in-vivo behavior analytic training for parents of children with autism spectrum disorder. *Behavior Analysis: Research and Practice, 20*(1), 13-23. doi:http://dx.doi.org/10.1037/bar0000163

Bordini, D., Paula, C., Cunha, G., Caetano, S., Bagaiolo, L., Ribeiro, T., . . . Brunoni, D. (2020). A randomised clinical pilot trial to test the effectiveness of parent training with video modelling to improve functioning and symptoms in children with autism spectrum disorders and intellectual disability. *Journal of Intellectual Disability Research, 64*(8), 629-643.

Boyd, B. A., Hume, K., McBee, M. T., Alessandri, M., Gutierrez, A., Johnson, L., . . . Odom, S. L. (2014). Comparative efficacy of LEAP, TEACCH and non-model-specific special education programs for preschoolers with autism spectrum disorders. *Journal of Autism & Developmental Disorders, 44*(2), 366-380. doi:https://dx.doi.org/10.1007/s10803-013-1877-9

Cariveau, T., Shillingsburg, M. A., Alamoudi, A., Thompson, T., Bartlett, B., Gillespie, S., & Scahill, L. (2019). Brief Report: Feasibility and Preliminary Efficacy of a Behavioral Intervention for Minimally Verbal Girls with Autism Spectrum Disorder. *Journal of Autism & Developmental Disorders, 49*(5), 2203-2209. doi:https://dx.doi.org/10.1007/s10803-018-03872-3

Carr, D., & Felce, J. (2007). "Brief report: increase in production of spoken words in some children with autism after PECS teaching to Phase III". *Journal of Autism & Developmental Disorders, 37*(4), 780-787.

Chang, Y. C., Shire, S. Y., Shih, W., Gelfand, C., & Kasari, C. (2016). Preschool Deployment of Evidence-Based Social Communication Intervention: JASPER in the Classroom. *Journal of Autism & Developmental Disorders, 46*(6), 2211-2223. doi:https://dx.doi.org/10.1007/s10803-016-2752-2

Charman, T., Palmer, M., Stringer, D., Hallett, V., Mueller, J., Romeo, R., . . . Simonoff, E. (2021). A Novel Group Parenting Intervention for Emotional and Behavioral Difficulties in Young

Autistic Children: autism Spectrum Treatment and Resilience (ASTAR): a Randomized Controlled Trial. *Journal of the American Academy of Child and Adolescent Psychiatry*.
Chiang, C. H., Chu, C. L., & Lee, T. C. (2016). Efficacy of caregiver-mediated joint engagement intervention for young children with autism spectrum disorders. *Autism, 20*(2), 172-182. doi:https://dx.doi.org/10.1177/1362361315575725
Cohen, H., Amerine-Dickens, M., & Smith, T. (2006). Early intensive behavioral treatment: replication of the UCLA model in a community setting. *Journal of Developmental & Behavioral Pediatrics, 27*(2 Suppl), S145-155.
Coleman, B. (2017). The effect of parent implemented interventions for Autism on parent ratings of maladaptive behavior. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 77*(12-B(E)), No Pagination Specified.
Colombi, C., Narzisi, A., Ruta, L., Cigala, V., Gagliano, A., Pioggia, G., . . . Prima Pietra, T. (2018). Implementation of the Early Start Denver Model in an Italian community. *Autism, 22*(2), 126-133. doi:https://dx.doi.org/10.1177/1362361316665792
Coman, D. C. (2014). The role of teacher commitment and burnout in predicting outcomes of preschoolers with autism spectrum disorders: A multilevel structural equation approach. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 74*(9-B(E)), No Pagination Specified.
D'Elia, L., Valeri, G., Sonnino, F., Fontana, I., Mammone, A., & Vicari, S. (2014). A longitudinal study of the teacch program in different settings: the potential benefits of low intensity intervention in preschool children with autism spectrum disorder. *Journal of Autism & Developmental Disorders, 44*(3), 615-626. doi:https://dx.doi.org/10.1007/s10803-013-1911-y
Dai, Y. G., Brennan, L., Como, A., Hughes-Lika, J., Dumont-Mathieu, T., Rathwell, I. C., . . . Fein, D. A. (2018). A Video Parent-Training Program for Families of Children with Autism Spectrum Disorder in Albania. *Research in Autism Spectrum Disorders, 56*, 36-49. doi:https://dx.doi.org/10.1016/j.rasd.2018.08.008
Dawson, G., Rogers, S., Munson, J., Smith, M., Winter, J., Greenson, J., . . . Varley, J. (2010). Randomized, controlled trial of an intervention for toddlers with autism: the Early Start Denver Model. *Pediatrics, 125*(1), e17-23. doi:https://dx.doi.org/10.1542/peds.2009-0958
Dimachkie, A. M. (2021). Compliance in toddlers with autism spectrum disorder: Exploring the effects of treatment and caregiver requests. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 82*(3-B), No Pagination Specified.
Dixon, M. R., Paliliunas, D., Barron, B. F., Schmick, A. M., & Stanley, C. R. (2019). Randomized controlled trial evaluation of aba content on iq gains in children with autism. *Journal of Behavioral Education*, No Pagination Specified. doi:http://dx.doi.org/10.1007/s10864-019-09344-7
Drew, A., Baird, G., Baron-Cohen, S., Cox, A., Slonims, V., Wheelwright, S., . . . Charman, T. (2002). A pilot randomised control trial of a parent training intervention for pre-school children with autism. Preliminary findings and methodological challenges. *European Child & Adolescent Psychiatry, 11*(6), 266-272.
Duifhuis, E. A., den Boer, J. C., Doornbos, A., Buitelaar, J. K., Oosterling, I. J., & Klip, H. (2017). The Effect of Pivotal Response Treatment in Children with Autism Spectrum Disorders: A Non-randomized Study with a Blinded Outcome Measure. *Journal of Autism & Developmental Disorders, 47*(2), 231-242. doi:https://dx.doi.org/10.1007/s10803-016-2916-0
Eikeseth, S., Klintwall, L., Jahr, E., & Karlsson, P. (2012). Outcome for children with autism receiving early and intensive behavioral intervention in mainstream preschool and kindergarten settings. *Research in autism spectrum disorders, 6*(2), 829-835.
Eikeseth, S., Smith, T., Jahr, E., & Eldevik, S. (2002). Intensive behavioral treatment at school for 4- to 7-year-old children with autism. A 1-year comparison controlled study. *Behavior Modification, 26*(1), 49-68.
Elder, L. M. (2012). The relationship between parent training and parent-child interaction in autism. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 73*(2-B), 1246.

Eldevik, S., Eikeseth, S., Jahr, E., & Smith, T. (2006). Effects of low-intensity behavioral treatment for children with autism and mental retardation. *Journal of Autism & Developmental Disorders, 36*(2), 211-224.

Eldevik, S., Hastings, R. P., Jahr, E., & Hughes, J. C. (2012). Outcomes of behavioral intervention for children with autism in mainstream pre-school settings. *Journal of Autism & Developmental Disorders, 42*(2), 210-220. doi:https://dx.doi.org/10.1007/s10803-011-1234-9

Eldevik, S., Jahr, E., Eikeseth, S., Hastings, R. P., & Hughes, C. J. (2010). Cognitive and adaptive behavior outcomes of behavioral intervention for young children with intellectual disability. *Behavior Modification, 34*(1), 16-34. doi:https://dx.doi.org/10.1177/0145445509351961

Estes, A., Munson, J., Rogers, S. J., Greenson, J., Winter, J., & Dawson, G. (2015). Long-Term Outcomes of Early Intervention in 6-Year-Old Children With Autism Spectrum Disorder. *Journal of the American Academy of Child & Adolescent Psychiatry, 54*(7), 580-587. doi:https://dx.doi.org/10.1016/j.jaac.2015.04.005

Estes, A., Vismara, L., Mercado, C., Fitzpatrick, A., Elder, L., Greenson, J., . . . Rogers, S. (2014). The impact of parent-delivered intervention on parents of very young children with autism. *Journal of Autism & Developmental Disorders, 44*(2), 353-365. doi:https://dx.doi.org/10.1007/s10803-013-1874-z

Fava, L., Strauss, K., Valeri, G., D'Elia, L., Arima, S., & Vicari, S. (2011). The effectiveness of a cross-setting complementary staff- and parent-mediated early intensive behavioral intervention for young children with ASD. *Research in Autism Spectrum Disorders, 5(4)*, 1479-1492. doi:http://dx.doi.org/10.1016/j.rasd.2011.02.009

Felzer-Kim, I. T., & Hauck, J. L. (2020). How Much Instructional Time Is Necessary? Mid-intervention Results of Fundamental Movement Skills Training Within ABA Early Intervention Centers. *Frontiers in Integrative Neuroscience, 14*, 24. doi:https://dx.doi.org/10.3389/fnint.2020.00024

Feng, J. Y., Li, H. H., Shan, L., Wang, B., Jia, F. Y., & Du, L. (2019). Clinical effect of vitamin D3 combined with the Early Start Denver Model in the treatment of autism spectrum disorder in toddlers. [Chinese]. [D<inf>3</inf>.]. *Chinese Journal of Contemporary Pediatrics, 21(4)*, 337-341. doi:http://dx.doi.org/10.7499/j.issn.1008-8830.2019.04.007

Flanagan, H. E. (2011). The impact of community-based intensive behavioural intervention. *Dissertation Abstracts International Section A: Humanities and Social Sciences, 72*(1-A), 73.

Fox, S. A. (2018). An early start denver model-based group intervention for parents of very young children diagnosed with or at risk for autism spectrum disorder. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 79*(1-B(E)), No Pagination Specified.

Frey, A. J., Small, J. W., Feil, E. G., Seeley, J. R., Walker, H. M., & Forness, S. (2015). First Step to Success: Applications to Preschoolers at Risk of Developing Autism Spectrum Disorders. *Education & Training in Autism & Developmental Disabilities, 50*(4), 397-407.

Furukawa, K., Okuno, H., Mohri, I., Nakanishi, M., Eyberg, S. M., & Sakai, S. (2018). Effectiveness of child-directed interaction training for young Japanese children with autism spectrum disorders. *Child & Family Behavior Therapy, 40*(2), 166-186. doi:http://dx.doi.org/10.1080/07317107.2018.1477344

Gengoux, G. W., Abrams, D. A., Schuck, R., Millan, M. E., Libove, R., Ardel, C. M., . . . Hardan, A. Y. (2019). A pivotal response treatment package for children with autism spectrum disorder: An RCT. *Pediatrics Vol 144(3), 2019, ArtID e20190178, 144*(3).

Gengoux, G. W., Schwartzman, J. M., Millan, M. E., Schuck, R. K., Ruiz, A. A., Weng, Y., . . . Hardan, A. Y. (2021). Enhancing Social Initiations Using Naturalistic Behavioral Intervention: Outcomes from a Randomized Controlled Trial for Children with Autism. *Journal of Autism & Developmental Disorders, 51*(10), 3547-3563. doi:https://dx.doi.org/10.1007/s10803-020-04787-8

Ginn, N. C., Clionsky, L. N., Eyberg, S. M., Warner-Metzger, C., & Abner, J.-P. (2017). Child-directed interaction training for young children with autism spectrum disorders: Parent and child outcomes. *Journal of Clinical Child & Adolescent Psychology, 46*(1), 101-109.

Gomes, C. G. S., de Souza, D. d. G., Silveira, A. D., Rates, A. C., de Castro Paiva, G. C., & de Castro, N. P. (2019). Effects of intensive behavioral intervention through training of caregivers of children with autism. *Psicologia: Teoria e Pesquisa Vol 35 2019, ArtID e3523, 35*. doi:http://dx.doi.org/10.1590/0102.3772e3523

Goods, K. S., Ishijima, E., Chang, Y. C., & Kasari, C. (2013). Preschool based JASPER intervention in minimally verbal children with autism: pilot RCT. *Journal of Autism & Developmental Disorders, 43*(5), 1050-1056. doi:https://dx.doi.org/10.1007/s10803-012-1644-3

Grahame, V., Brett, D., Dixon, L., McConachie, H., Lowry, J., Rodgers, J., . . . Le Couteur, A. (2015). Managing repetitive behaviours in young children with autism spectrum disorder (ASD): pilot randomised controlled trial of a new parent group intervention. *Journal of autism and developmental disorders, 45*, 3168-3182.

Grindle, C. F., Hastings, R. P., Saville, M., Hughes, J. C., Huxley, K., Kovshoff, H., . . . Remington, B. (2012). Outcomes of a behavioral education model for children with autism in a mainstream school setting. *Behavior Modification, 36*(3), 298-319. doi:https://dx.doi.org/10.1177/0145445512441199

Gulsrud, A., Carr, T., Williams, J., Panganiban, J., Jones, F., Kimbrough, J., . . . Kasari, C. (2019). Developmental screening and early intervention in a childcare setting for young children at risk for autism and other developmental delays: A feasibility trial. *Autism research : Official Journal of the International Society for Autism Research, 12*(9), 1423-1433. doi:https://dx.doi.org/10.1002/aur.2160

Gulsrud, A. C., Hellemann, G., Shire, S., & Kasari, C. (2016). Isolating active ingredients in a parent-mediated social communication intervention for toddlers with autism spectrum disorder. *Journal of Child Psychology & Psychiatry & Allied Disciplines, 57*(5), 606-613. doi:https://dx.doi.org/10.1111/jcpp.12481

Haglund, N., Dahlgren, S., Rastam, M., Gustafsson, P., & Kallen, K. (2021). Improvement of Autism Symptoms After Comprehensive Intensive Early Interventions in Community Settings. *Journal of the American Psychiatric Nurses Association, 27*(6), 483-495. doi:https://dx.doi.org/10.1177/1078390320915257

Hampton, L. H., Kaiser, A. P., & Fuller, E. A. (2020). Multi-component communication intervention for children with autism: A randomized controlled trial. *Autism, 24*(8), 2104-2116. doi:https://dx.doi.org/10.1177/1362361320934558

Haraguchi, H., Yamaguchi, H., Miyake, A., Tachibana, Y., Stickley, A., Horiguchi, M., . . . Kamio, Y. (2020). One-year outcomes of low-intensity behavioral interventions among Japanese preschoolers with autism spectrum disorders: Community-based study. *Research in Autism Spectrum Disorders, 76 (no pagination)*. doi:http://dx.doi.org/10.1016/j.rasd.2020.101556

Hardan, A. Y., Gengoux, G. W., Berquist, K. L., Libove, R. A., Ardel, C. M., Phillips, J., . . . Minjarez, M. B. (2015). A randomized controlled trial of Pivotal Response Treatment Group for parents of children with autism. *Journal of Child Psychology & Psychiatry & Allied Disciplines, 56*(8), 884-892. doi:https://dx.doi.org/10.1111/jcpp.12354

Ho, M. H., & Lin, L. Y. (2020). Efficacy of parent-training programs for preschool children with autism spectrum disorder: A randomized controlled trial. *Research in Autism Spectrum Disorders, 71 (no pagination)*. doi:http://dx.doi.org/10.1016/j.rasd.2019.101495

Holzinger, D., Laister, D., Vivanti, G., Barbaresi, W. J., & Fellinger, J. (2019). Feasibility and Outcomes of the Early Start Denver Model Implemented with Low Intensity in a Community Setting in Austria. *Journal of Developmental & Behavioral Pediatrics, 40*(5), 354-363. doi:https://dx.doi.org/10.1097/DBP.0000000000000675

Howard, J. S., Sparkman, C. R., Cohen, H. G., Green, G., & Stanislaw, H. (2005). A comparison of intensive behavior analytic and eclectic treatments for young children with autism. *Research in Developmental Disabilities, 26*(4), 359-383.

Howard, J. S., Stanislaw, H., Green, G., Sparkman, C. R., & Cohen, H. G. (2014). Comparison of behavior analytic and eclectic early interventions for young children with autism after three years. *Research in Developmental Disabilities, 35*(12), 3326-3344. doi:https://dx.doi.org/10.1016/j.ridd.2014.08.021

Iadarola, S., Levato, L., Harrison, B., Smith, T., Lecavalier, L., Johnson, C., . . . Scahill, L. (2018). Teaching Parents Behavioral Strategies for Autism Spectrum Disorder (ASD): Effects on Stress, Strain, and Competence. *Journal of Autism & Developmental Disorders, 48*(4), 1031-1040. doi:https://dx.doi.org/10.1007/s10803-017-3339-2

Iadarola, S., Shih, W., Dean, M., Blanch, E., Harwood, R., Hetherington, S., . . . Smith, T. (2018). Implementing a Manualized, Classroom Transition Intervention for Students With ASD in Underresourced Schools. *Behavior Modification, 42*(1), 126-147. doi:https://dx.doi.org/10.1177/0145445517711437

Ingersoll, B. (2010). Pilot randomized controlled trial of Reciprocal Imitation Training for teaching elicited and spontaneous imitation to children with autism. *Journal of Autism & Developmental Disorders, 40*(9), 1154-1160. doi:https://dx.doi.org/10.1007/s10803-010-0966-2

Ingersoll, B. (2012). Brief report: effect of a focused imitation intervention on social functioning in children with autism. *Journal of Autism & Developmental Disorders, 42*(8), 1768-1773. doi:https://dx.doi.org/10.1007/s10803-011-1423-6

Johnson, C. R., Brown, K., Hyman, S. L., Brooks, M. M., Aponte, C., Levato, L., . . . Smith, T. (2019). Parent Training for Feeding Problems in Children With Autism Spectrum Disorder: initial Randomized Trial. *Journal of pediatric psychology, 44*(2), 164-175.

Jouen, A. L., Narzisi, A., Xavier, J., Tilmont, E., Bodeau, N., Bono, V., . . . Group, M. S. (2017). GOLIAH (Gaming Open Library for Intervention in Autism at Home): a 6-month single blind matched controlled exploratory study. *Child & Adolescent Psychiatry & Mental Health [Electronic Resource], 11*, 17. doi:https://dx.doi.org/10.1186/s13034-017-0154-7

Kaale, A., Fagerland, M. W., Martinsen, E. W., & Smith, L. (2014). Preschool-based social communication treatment for children with autism: 12-month follow-up of a randomized trial. *Journal of the American Academy of Child & Adolescent Psychiatry, 53*(2), 188-198. doi:https://dx.doi.org/10.1016/j.jaac.2013.09.019

Kaale, A., Smith, L., & Sponheim, E. (2012). A randomized controlled trial of preschool-based joint attention intervention for children with autism. *Journal of Child Psychology & Psychiatry & Allied Disciplines, 53*(1), 97-105. doi:https://dx.doi.org/10.1111/j.1469-7610.2011.02450.x

Kasari, C., Freeman, S., & Paparella, T. (2006). Joint attention and symbolic play in young children with autism: a randomized controlled intervention study. *Journal of Child Psychology & Psychiatry & Allied Disciplines, 47*(6), 611-620.

Kasari, C., Gulsrud, A., Paparella, T., Hellemann, G., & Berry, K. (2015). Randomized comparative efficacy study of parent-mediated interventions for toddlers with autism. *Journal of Consulting & Clinical Psychology, 83*(3), 554-563. doi:https://dx.doi.org/10.1037/a0039080

Kasari, C., Gulsrud, A. C., Wong, C., Kwon, S., & Locke, J. (2010). Randomized controlled caregiver mediated joint engagement intervention for toddlers with autism. *Journal of Autism & Developmental Disorders, 40*(9), 1045-1056. doi:https://dx.doi.org/10.1007/s10803-010-0955-5

Kasari, C., Paparella, T., Freeman, S., & Jahromi, L. B. (2008). Language outcome in autism: randomized comparison of joint attention and play interventions. *Journal of consulting and clinical psychology, 76*(1), 125.

Kovshoff, H., Hastings, R. P., & Remington, B. (2011). Two-year outcomes for children with autism after the cessation of early intensive behavioral intervention. *Behavior Modification, 35*(5), 427-450. doi:https://dx.doi.org/10.1177/0145445511405513

Lawton, K., & Kasari, C. (2012a). Brief report: longitudinal improvements in the quality of joint attention in preschool children with autism. *Journal of Autism & Developmental Disorders, 42*(2), 307-312. doi:https://dx.doi.org/10.1007/s10803-011-1231-z

Lawton, K., & Kasari, C. (2012b). Teacher-implemented joint attention intervention: pilot randomized controlled study for preschoolers with autism. *Journal of Consulting & Clinical Psychology, 80*(4), 687-693. doi:https://dx.doi.org/10.1037/a0028506

Leaf, J. B., Leaf, J. A., Milne, C., Taubman, M., Oppenheim-Leaf, M., Torres, N., . . . Autism Partnership, F. (2017). An Evaluation of a Behaviorally Based Social Skills Group for

Individuals Diagnosed with Autism Spectrum Disorder. *Journal of Autism & Developmental Disorders, 47*(2), 243-259. doi:https://dx.doi.org/10.1007/s10803-016-2949-4

Magiati, I., Charman, T., & Howlin, P. (2007). A two-year prospective follow-up study of community-based early intensive behavioural intervention and specialist nursery provision for children with autism spectrum disorders. *Journal of Child Psychology & Psychiatry & Allied Disciplines, 48*(8), 803-812.

Manohar, H., Kandasamy, P., Chandrasekaran, V., & Rajkumar, R. P. (2019). Brief Parent-Mediated Intervention for Children with Autism Spectrum Disorder: a Feasibility Study from South India. *Journal of autism and developmental disorders*.

Matthews, N. L., Orr, B. C., Harris, B., McIntosh, R., Openden, D., & Smith, C. J. (2018). Parent and child outcomes of JumpStartTM, an education and training program for parents of children with autism spectrum disorder. *Research in Autism Spectrum Disorders, 56*, 21-35. doi:http://dx.doi.org/10.1016/j.rasd.2018.08.009

Nojiri, J., & Yanagawa, T. (2019). [Effects of the Stepping Stones Triple P for mothers of pre-school children with suspected Autistic Spectrum Disorder]. *Nippon Koshu Eisei Zasshi - Japanese Journal of Public Health, 66*(5), 237-245. doi:https://dx.doi.org/10.11236/jph.66.5_237

Nowell, S. W., Watson, L. R., Boyd, B., & Klinger, L. G. (2019). Efficacy Study of a Social Communication and Self-Regulation Intervention for School-Age Children With Autism Spectrum Disorder: A Randomized Controlled Trial. *Language, Speech & Hearing Services in the Schools, 50*(3), 416-433. doi:https://dx.doi.org/10.1044/2019_LSHSS-18-0093

Oosterling, I., Visser, J., Swinkels, S., Rommelse, N., Donders, R., Woudenberg, T., . . . Buitelaar, J. (2010). Randomized controlled trial of the focus parent training for toddlers with autism: 1-year outcome. *Journal of Autism & Developmental Disorders, 40*(12), 1447-1458. doi:https://dx.doi.org/10.1007/s10803-010-1004-0

Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., . . . McKenzie, J. E. (2021). PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ, 372*, n160. doi:10.1136/bmj.n160

Pajareya, K., & Nopmaneejumruslers, K. (2011). A pilot randomized controlled trial of DIR/FloortimeTM parent training intervention for pre-school children with autistic spectrum disorders. *Autism, 15(5)*, 563-577. doi:http://dx.doi.org/10.1177/1362361310386502

Peters-Scheffer, N., Didden, R., Mulders, M., & Korzilius, H. (2013). Effectiveness of low intensity behavioral treatment for children with autism spectrum disorder and intellectual disability. *Research in Autism Spectrum Disorders, 7(9)*, 1012-1025. doi:http://dx.doi.org/10.1016/j.rasd.2013.05.001

Reitzel, J., Summers, J., Lorv, B., Szatmari, P., Zwaigenbaum, L., Georgiades, S., & Duku, E. (2013). Pilot randomized controlled trial of a Functional Behavior Skills Training program for young children with Autism Spectrum Disorder who have significant early learning skill impairments and their families. *Research in autism spectrum disorders, 7*(11), 1418-1432.

Remington, B., Hastings, R. P., Kovshoff, H., degli Espinosa, F., Jahr, E., Brown, T., . . . Ward, N. (2007). Early intensive behavioral intervention: outcomes for children with autism and their parents after two years. *American Journal of Mental Retardation, 112*(6), 418-438.

Rodgers, M., Marshall, D., Simmonds, M., Le Couteur, A., Biswas, M., Wright, K., . . . Hodgson, R. (2020). Interventions based on early intensive applied behaviour analysis for autistic children: a systematic review and cost-effectiveness analysis. *Health Technology Assessment, 24*(35), 1-306. doi:10.3310/hta24350

Rogers, S. J., Estes, A., Lord, C., Munson, J., Rocha, M., Winter, J., . . . Talbott, M. (2019). A Multisite Randomized Controlled Two-Phase Trial of the Early Start Denver Model Compared to Treatment as Usual. *Journal of the american academy of child and adolescent psychiatry*.

Rogers, S. J., Estes, A., Lord, C., Vismara, L., Winter, J., Fitzpatrick, A., . . . Dawson, G. (2012). Effects of a brief Early Start Denver model (ESDM)-based parent intervention on toddlers at risk for autism spectrum disorders: a randomized controlled trial. *Journal of the American*

Academy of Child & Adolescent Psychiatry, 51(10), 1052-1065. doi:https://dx.doi.org/10.1016/j.jaac.2012.08.003

Rogers, S. J., Hayden, D., Hepburn, S., Charlifue-Smith, R., Hall, T., & Hayes, A. (2006). Teaching young nonverbal children with autism useful speech: a pilot study of the Denver Model and PROMPT interventions. *Journal of Autism & Developmental Disorders, 36*(8), 1007-1024.

Rogers, S. J., Vismara, L., Wagner, A. L., McCormick, C., Young, G., & Ozonoff, S. (2014). Autism treatment in the first year of life: a pilot study of infant start, a parent-implemented intervention for symptomatic infants. *Journal of Autism & Developmental Disorders, 44*(12), 2981-2995. doi:https://dx.doi.org/10.1007/s10803-014-2202-y

Ruiz, A. A. (2020). Motivational behavioral group treatment for social deficits in preschool children with autism: Impact on talkativeness with typically developing peers. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 81*(6-B), No Pagination Specified.

Scahill, L., Bearss, K., Lecavalier, L., Smith, T., Swiezy, N., Aman, M. G., . . . Johnson, C. (2016). Effect of Parent Training on Adaptive Behavior in Children With Autism Spectrum Disorder and Disruptive Behavior: Results of a Randomized Trial. *Journal of the American Academy of Child and Adolescent Psychiatry, 55(7)*, 602-609.e603. doi:http://dx.doi.org/10.1016/j.jaac.2016.05.001

Schlink, A., Williams, J., Pizzano, M., Gulsrud, A., & Kasari, C. (2022). Parenting stress in caregiver-mediated interventions for toddlers with autism: An application of quantile regression mixed models. *Autism Research, 15*(2), 353-365.

Shawler, P. (2017). Does early intervention reduce the risk of future emotional and behavioral problems in children with autism spectrum disorder. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 78*(4-B(E)), No Pagination Specified.

Sheinkopf, S. J., & Siegel, B. (1998). Home-based behavioral treatment of young children with autism. *Journal of Autism & Developmental Disorders, 28*(1), 15-23.

Shire, S. Y., Chang, Y. C., Shih, W., Bracaglia, S., Kodjoe, M., & Kasari, C. (2017). Hybrid implementation model of community-partnered early intervention for toddlers with autism: a randomized trial. *Journal of Child Psychology & Psychiatry & Allied Disciplines, 58*(5), 612-622. doi:https://dx.doi.org/10.1111/jcpp.12672

Sinai-Gavrilov, Y., Gev, T., Mor-Snir, I., Vivanti, G., & Golan, O. (2020). Integrating the Early Start Denver Model into Israeli community autism spectrum disorder preschools: Effectiveness and treatment response predictors. *Autism, 24*(8), 2081-2093. doi:https://dx.doi.org/10.1177/1362361320934221

Solomon, R., Van Egeren, L. A., Mahoney, G., Huber, M. S. Q., & Zimmerman, P. (2014). PLAY Project Home Consultation intervention program for young children with autism spectrum disorders: A randomized controlled trial. *Journal of Developmental and Behavioral Pediatrics, 35*(8), 475.

Spjut Jansson, B., Miniscalco, C., Westerlund, J., Kantzer, A. K., Fernell, E., & Gillberg, C. (2016). Children who screen positive for autism at 2.5 years and receive early intervention: a prospective naturalistic 2-year outcome study. *Neuropsychiatric Disease & Treatment, 12*, 2255-2263. doi:https://dx.doi.org/10.2147/NDT.S108899

Stadnick, N. A., Stahmer, A., & Brookman-Frazee, L. (2015). Preliminary Effectiveness of Project ImPACT: A Parent-Mediated Intervention for Children with Autism Spectrum Disorder Delivered in a Community Program. *Journal of Autism & Developmental Disorders, 45*(7), 2092-2104. doi:https://dx.doi.org/10.1007/s10803-015-2376-y

Stahmer, A. C., Rieth, S. R., Dickson, K. S., Feder, J., Burgeson, M., Searcy, K., & Brookman-Frazee, L. (2020). Project ImPACT for Toddlers: Pilot outcomes of a community adaptation of an intervention for autism risk. *Autism, 24*(3), 617-632. doi:https://dx.doi.org/10.1177/1362361319878080

Sterne, J. A. C., Hernán, M. A., Reeves, B. C., Savović, J., Berkman, N. D., Viswanathan, M., . . . Higgins, J. P. T. (2016). ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ, 355*, i4919. doi:10.1136/bmj.i4919

Sterne, J. A. C., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., . . . Higgins, J. P. T. (2019). RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ, 366*, l4898. doi:10.1136/bmj.l4898

Strauss, K., Vicari, S., Valeri, G., D'Elia, L., Arima, S., & Fava, L. (2012). Parent inclusion in Early Intensive Behavioral Intervention: the influence of parental stress, parent treatment fidelity and parent-mediated generalization of behavior targets on child outcomes. *Research in Developmental Disabilities, 33*(2), 688-703. doi:https://dx.doi.org/10.1016/j.ridd.2011.11.008

Sullivan, K. (2014). The early start denver model: Outcomes and moderators of an intervention for toddlers with autism. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 75*(2-B(E)), No Pagination Specified.

Tonge, B., Brereton, A., Kiomall, M., Mackinnon, A., King, N., & Rinehart, N. (2006). Effects on parental mental health of an education and skills training program for parents of young children with autism: a randomized controlled trial. *Journal of the American Academy of Child & Adolescent Psychiatry, 45*(5), 561-569. doi:https://dx.doi.org/10.1097/01.chi.0000205701.48324.26

Tonge, B., Brereton, A., Kiomall, M., Mackinnon, A., & Rinehart, N. J. (2014). A randomised group comparison controlled trial of 'preschoolers with autism': a parent education and skills training intervention for young children with autistic disorder. *Autism, 18*(2), 166-177. doi:https://dx.doi.org/10.1177/1362361312458186

Tsang, S. K., Shek, D. T., Lam, L. L., Tang, F. L., & Cheung, P. M. (2007). Brief report: application of the TEACCH program on Chinese pre-school children with autism--Does culture make a difference? *Journal of Autism & Developmental Disorders, 37*(2), 390-396.

Van der Paelt, S., Warreyn, P., & Roeyers, H. (2016). Effect of community interventions on social-communicative abilities of preschoolers with autism spectrum disorder. *Developmental neurorehabilitation, 19*(3), 162-174. doi:https://dx.doi.org/10.3109/17518423.2014.933983

Vernon, T. W., Holden, A. N., Barrett, A. C., Bradshaw, J., Ko, J. A., McGarry, E. S., . . . German, T. C. (2019). A Pilot Randomized Clinical Trial of an Enhanced Pivotal Response Treatment Approach for Young Children with Autism: The PRISM Model. *Journal of Autism & Developmental Disorders, 49*(6), 2358-2373. doi:https://dx.doi.org/10.1007/s10803-019-03909-1

Vinen, Z., Clark, M., Paynter, J., & Dissanayake, C. (2018). School Age Outcomes of Children with Autism Spectrum Disorder Who Received Community-Based Early Interventions. *Journal of Autism & Developmental Disorders, 48*(5), 1673-1683. doi:https://dx.doi.org/10.1007/s10803-017-3414-8

Vivanti, G., Paynter, J., Duncan, E., Fothergill, H., Dissanayake, C., Rogers, S. J., & Victorian, A. T. (2014). Effectiveness and feasibility of the early start denver model implemented in a group-based community childcare setting. *Journal of Autism & Developmental Disorders, 44*(12), 3140-3153. doi:https://dx.doi.org/10.1007/s10803-014-2168-9

Warreyn, P., & Roeyers, H. (2014). See what I see, do as I do: promoting joint attention and imitation in preschoolers with autism spectrum disorder. *Autism, 18*(6), 658-671. doi:https://dx.doi.org/10.1177/1362361313493834

Waters, C. F., Amerine Dickens, M., Thurston, S. W., Lu, X., & Smith, T. (2020). Sustainability of Early Intensive Behavioral Intervention for Children With Autism Spectrum Disorder in a Community Setting. *Behavior Modification, 44*(1), 3-26. doi:https://dx.doi.org/10.1177/0145445518786463

Whalen, C., Moss, D., Ilan, A. B., Vaupel, M., Fielding, P., Macdonald, K., . . . Symon, J. (2010). Efficacy of TeachTown: basics computer-assisted intervention for the Intensive Comprehensive Autism Program in Los Angeles Unified School District. *Autism, 14*(3), 179-197.

Whitehouse, A., Varcin, K., Waddington, H., Sulek, R., Bent, C., Ashburner, J.,, & Eapen, V., Goodall, E., Hudry, K., Roberts, J., Silove, N., Trembath, D. (2020). *Interventions for children on the autism spectrum: A synthesis of research evidence*. Retrieved from Brisbane: https://www.autismcrc.com.au/interventions-evidence

Whitehouse, A. J. O., Granich, J., Alvares, G., Busacca, M., Cooper, M. N., Dass, A., . . . Anderson, A. (2017). A randomised controlled trial of an iPad-based application to complement early behavioural intervention in Autism Spectrum Disorder. *Journal of Child Psychology & Psychiatry & Allied Disciplines, 58*(9), 1042-1052. doi:https://dx.doi.org/10.1111/jcpp.12752

Xu, Y., Yang, J., Yao, J., Chen, J., Zhuang, X., Wang, W., . . . Lee, G. T. (2018). A pilot study of a culturally adapted early intervention for young children with autism spectrum disorders in China. *Journal of Early Intervention, 40*(1), 52-68. doi:http://dx.doi.org/10.1177/1053815117748408

Xu, Y., Yao, J., & Yang, J. (2017). Application of early start Denver model for early intervention on autistic children. *Chinese Journal of Clinical Psychology, 25*(1), 188-191.

Zachor, D. A., & Ben Itzchak, E. (2010). Treatment approach, autism severity and intervention outcomes in young children. *Research in Autism Spectrum Disorders, 4(3)*, 425-432. doi:http://dx.doi.org/10.1016/j.rasd.2009.10.013

Zachor, D. A., Ben-Itzchak, E., Rabinovich, A. L., & Lahat, E. (2007). Change in autism core symptoms with intervention. *Research in Autism Spectrum Disorders, 1(4)*, 304-317. doi:http://dx.doi.org/10.1016/j.rasd.2006.12.001

Zhou, B., Xu, Q., Li, H., Zhang, Y., Wang, Y., Rogers, S. J., & Xu, X. (2018). Effects of Parent-Implemented Early Start Denver Model Intervention on Chinese Toddlers with Autism Spectrum Disorder: A Non-Randomized Controlled Trial. *Autism research : Official Journal of the International Society for Autism Research, 11*(4), 654-666. doi:https://dx.doi.org/10.1002/aur.1917