

A systematic review of fidelity measurements in complex interventions for people with intellectual disabilities and behaviours that challenge

Journal:	<i>Advances in Mental Health and Intellectual Disabilities</i>
Manuscript ID	AMHID-09-2018-0040.R1
Manuscript Type:	Research Paper
Keywords:	Fidelity, Complex intervention, Intellectual disability, Behaviours that challenge, Positive Behaviour Support, Randomised controlled trial

Introduction

There are approximately 26,500 adults with intellectual disabilities (ID) and behaviours that challenge in England (Emerson et al, 2011). Behaviours that challenge are enduring and often have serious negative consequences for the person (e.g. administration of medication, abuse or in-patient admissions) and for those who support them (e.g. physical/mental health problems or high job turnover) (Emerson and Enfield, 2011; Allen, 2009; Sturmey, 2009).

Currently, Positive Behaviour Support (PBS) is considered the best practice for supporting individuals with ID who engage in behaviours that challenge (Department of Health, 2014; British Psychological Society, 2018; National Institute for Clinical Excellence, 2015). PBS is a multicomponent biopsychosocial intervention that aims to improve individuals' quality of life and eventually reduce the display of behaviours that challenge (PBS competence framework, 2015). PBS evolved from Applied Behaviour Analysis (ABA) that forms a central core of PBS and provides the framework and techniques for understanding and reducing behaviours that challenge (Allen, 2009). Behaviour support plans are developed as part of PBS to provide a step by step guide to carers of how to support an individual with ID effectively (Gore et al, 2015; Allen et al, 2005; LaVigna and Willis, 2005).

There is evidence from systematic and meta-analytic reviews of studies using N=1 designs that interventions based on PBS can produce reductions in behaviours that challenge up to 50% (Carr et al, 1999; Dunlap et al, 2008; LaVigna and Willis, 2012). One UK-based RCT showed that behaviours that challenge had reduced by 43% after the implementation of PBS (Hassiotis et al, 2009). However, only a small proportion of those who display behaviours that challenge receive behavioural support in England (Emerson et al, 2000; Allen et al, 2005). Possible reasons as to why behavioural support is not widely available are: there is a lack of services available in the UK based on PBS principles, small numbers of staff

1
2 members are trained in PBS and there are long waiting lists in the services that provides
3
4 PBS/ABA (Toogood, 2015; Allen et al, 2005; Beadle et al, 2006). Davidson et al (2015)
5
6 found that the current provision of peripatetic support for individuals with behaviours that
7
8 challenge in the UK is very low; only 46 services identified themselves as such and out of
9
10 these only 47% had reported providing support for behaviours that challenge based on PBS
11
12 principles. These findings are in contract to the main focus of PBS which aims to apply
13
14 research-validated principles in real-life settings where individuals with ID live (Carr et al,
15
16 2002).
17
18
19
20
21

22
23 Fidelity data in clinical trials of PBS could provide useful insights how to improve the
24
25 application of PBS in everyday settings (Boutron et al, 2008; Craig et al, 2008). Fidelity,
26
27 which is also known as integrity or adherence, refers to the extent an intervention is
28
29 implemented and delivered as indicated by the intervention developers (Moore et al, 2015;
30
31 Craig et al, 2008). Fidelity measurements are particularly important for complex
32
33 interventions because there are multiple factors that affect its delivery and efficacy (Craig et
34
35 al, 2008). According to the Medical Research Council (MRC) guidelines, complex
36
37 interventions are defined as having multiple active components that influence outcomes,
38
39 multiple agents with a range of different skills are involved in the delivery and the
40
41 intervention is delivered in real life contexts (Craig et al, 2008).
42
43
44
45
46
47

48
49 PBS is a complex intervention and as such has multiple interacting elements that are
50
51 important for the reduction of behaviours that challenge, it is constantly adapted to take into
52
53 account the service users' strengths and targets the person's environment (Carr et al, 2002).
54
55 Therefore, it is plausible that during the implementation of PBS several elements of this
56
57 intervention might be omitted or added (Carr et al, 2002). Besides, multiple other variables
58
59 (such as therapists' competence, skills, motivation, time commitments, goodness of fit,
60

management support, staff-turnover rates) also impact its delivery and implementation

(Hassiotis et al, 2018). Examining fidelity data reported in PBS trials are not only necessary for the internal and external validity of the study but it also provides information about essential elements of PBS that need to be implemented as identified in the protocol and of those that can be adapted without decreasing its effectiveness (Carroll et al, 2007); it highlights the aspects of PBS that service providers find difficult to deliver, pointing to modifications in the administration of PBS and what additional training is needed in order to improve its implementation (Spee et al, 2014; Mowbray et al, 2003).

Systematic reviews report that fidelity measurements in the behavioural analysis literature are uncommon (McIntyre et al, 2007; Sanetti et al, 2012). The McIntyre's et al (2007) review examined fidelity measurements reported in school-based experimental (within-group or single-case design) studies with children published in the Journal of Applied Behavior Analysis between 1991 to 2005. Researchers found that only 39% of studies provided any mention of fidelity assessments; 30% reported quantitative data (in the form of percentage of implementation) and 9% of studies mentioned that fidelity measurements were completed but quantitative data were not provided in the papers (McIntyre et al, 2007). Another review (Sanetti et al, 2012) found similar findings: 42% of studies with children published in the Journal of Positive Behavior Interventions between 1999 and 2009 included quantitative fidelity data and 7% of studies mentioned fidelity assessments without providing quantitative data (McIntyre et al, 2007; Sanetti et al, 2012). These reviews indicate that fidelity measurements in behavioural literature are not new but the low rates of reporting fidelity measurements in the studies over the years remain stable. Nevertheless, these systematic reviews examined fidelity in single-case design studies published in only one journal and often excluding people with ID from their samples. Different trends of fidelity measurements

1
2 might be found in the literature using different inclusion criteria, considering more than one
3
4 medical journal and including people with ID.
5
6
7

8
9 With this in mind, the primary aim of this study was to carry out a systematic review
10
11 examining the approaches used to measure fidelity in randomised controlled trials of complex
12
13 interventions for behaviours that challenge which are based on PBS/ABA principles.
14
15
16
17

18 **Methods**

21 **Search strategy**

22
23 Searchers of the electronic databases MEDLINE, Embase, PsycINFO, Web of Science and
24
25 CINAHL Plus were conducted from April 2017 to June 2017. The first 50 pages of grey
26
27 literature were also searched and sources included: Research gate, Google scholar, Google,
28
29 NICE Evidence Search, The King's Fund, Zetoc, Proquest, WorldCat, OpenGrey and Clinical
30
31 Trials.gov. We also hand searched the reference lists of the studies included in the review.
32
33 The search terms related to intellectual disability, behaviours that challenge, fidelity and
34
35 randomised controlled trial were combined using the Boolean operator "AND" (see
36
37 supplementary material 1 for the full search strategy).
38
39
40
41
42
43
44

45 **Inclusion criteria**

46
47 *Participants:* children or adults with mild, moderate, severe or profound ID and behaviours
48
49 that challenge. ID had to be explicitly identified in the studies such as an IQ below 70
50
51 alongside a standardised measure or diagnostic criteria; behaviours that challenge also had to
52
53 be measured by standardised questionnaires.
54
55
56
57
58
59
60

1
2 Studies, which had participants with other neurodevelopmental comorbidities in addition to
3
4 ID and behaviours that challenge (e.g. Autistic Spectrum Disorders), were included only if at
5
6 least 50% of the sample had ID.
7

8
9 *Intervention:* complex interventions for behaviours that challenge which are based on PBS or
10
11 ABA approaches (e.g. pivotal response training or differential reinforcement). Interventions
12
13 that did not target behaviours that challenge were excluded.
14

15
16 *Comparison:* any control group (e.g. treatment as usual, no treatment, waitlist control or any
17
18 alternative management strategy).
19

20
21 *Outcomes:* studies had to include a statement about measuring and reporting fidelity. The
22
23 approach proposed by Dane and Schneider (1998) was used to appraise the dimensions of
24
25 fidelity reported in the studies (see Table1 for definitions).
26

27
28 *Study design:* randomised controlled trials.
29

30
31 *Publication:* published articles were limited to the English language covering the period from
32
33 January 1990 to Jun 2017.
34
35

36 **Table 1: Fidelity components (based on Dane and Schneider, 1998)**
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Study selection

Titles and abstracts were screened by one reviewer (LP) against the inclusion criteria. Full text articles of potentially eligible studies were obtained and assessed by two reviewers (LP and AH). Numbers of excluded studies and its reasons were recorded. Any disagreements or discrepancies about included studies were resolved by a consensus and discussion with a third reviewer (AA). Two reviewers (LP and AH) then independently extracted data into a standardised form that was developed for this study.

Risk of bias in included studies

The Cochrane risk of bias tool (Higgins and Green, 2011) was used independently by two reviewers (LP and AA) in order to detect risk of bias in each study. Reviewers rated risk of bias as low (+), high (-) or unclear (?) for each of the seven domains: allocation, selection, performance, detection, attrition, reporting and other biases (see Table 2 for definitions).

Table 2: Risk of bias (Higgins and Green, 2011)

Results

Electronic and hand searches identified 3482 records. Figure 1 shows a flow diagram of the study selection process for this review. After removal of duplicates, 3439 unique articles were screened and 36 full-text papers of potentially relevant articles were retrieved. Thirty-one articles did not meet inclusion criteria (see supplementary material 2 for a list of excluded articles). Thus, five articles were included in the review. Table 3 shows an overview of the studies.

Figure 1: A PRISMA flow chart of study selection process.

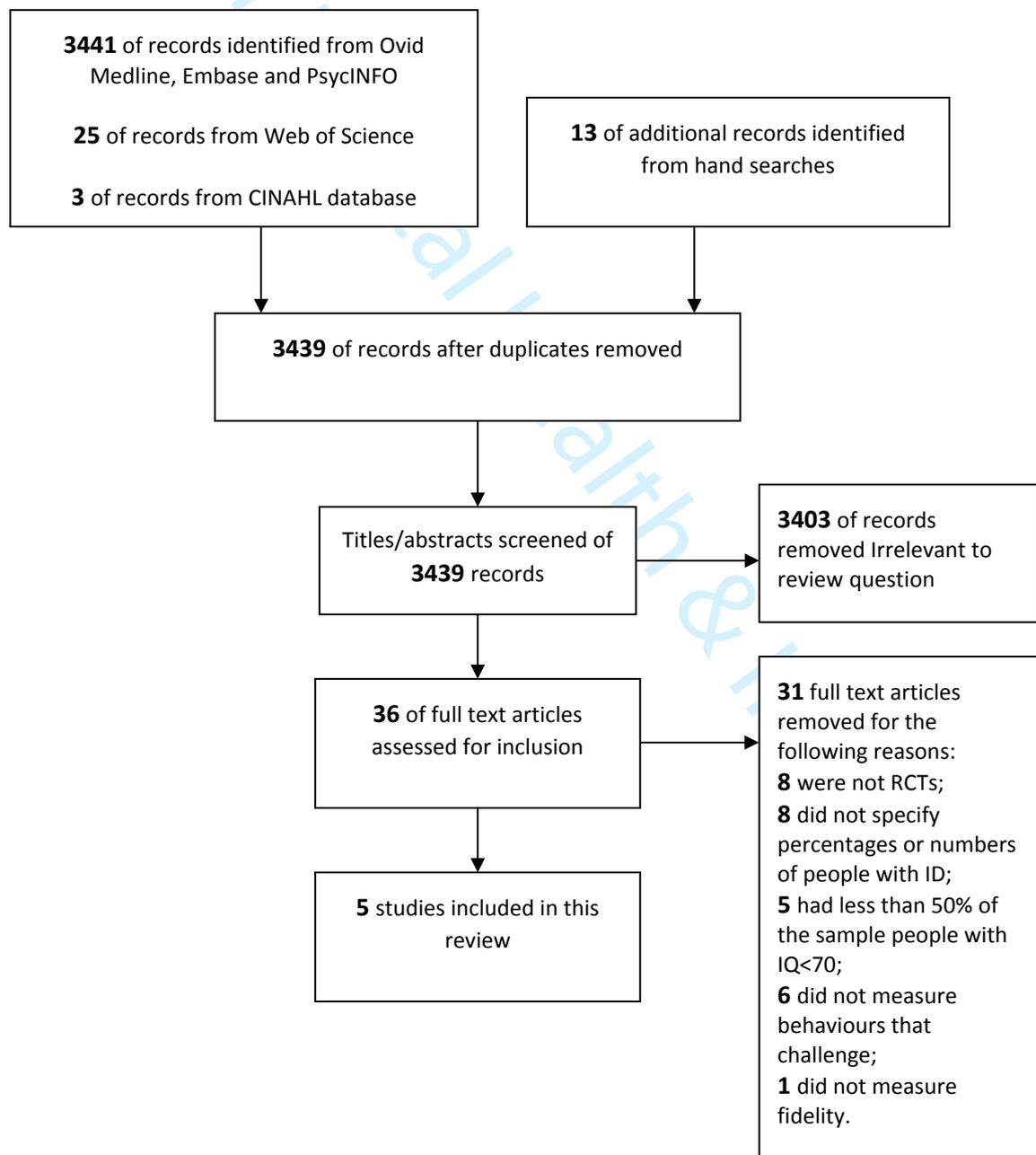


Table 3: Characteristics of the included studies

Advances in Mental Health & Intellectual

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46

Table 4: Elements of fidelity reported in the included studies

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46

Overview of the included studies

Study design and location

The current review identified five small sample studies of which one was randomised controlled trial (Singh et al, 2016) and four were pilot clinical trials (Ingersoll et al, 2016; Johnson et al, 2013; Hassiotis et al, 2009; Reitzel et al, 2013). Studies were published between 2009 and 2016 in the UK, USA and Canada. Participants were recruited from the following settings: residential homes, inpatient, community intellectual disabilities service and education. An overview of studies is provided in *Table 3*.

Participants

The total number of participants of the five included studies was 178. The number of participants in the studies ranged from 15 to 45. Majority of participants in the included studies were males apart from one study which did not specify numbers of males and females in their sample (Reitzel et al, 2013). Two studies focused on adults (Singh et al, 2016; Hassiotis et al, 2009), the other two on children (Johnson et al, 2013; Reitzel et al, 2013) and one on adolescents (Ingersoll et al, 2016). Participants in the included studies had various levels of ID (mild to severe) and behaviours that challenge. Three studies had participants with Autistic Spectrum Disorders or Pervasive Developmental Disorders, Not Otherwise Specified in addition to ID (Ingersoll et al, 2016; Johnson et al, 2013; Reitzel et al, 2013).

Interventions

All interventions were for behaviours that challenge based on ABA or PBS principles and had some overlap in terms of the content of techniques used (e.g. functional assessments, antecedent manipulation, data-collection, extinction, reinforcement and skills teaching).

One study examined individual reciprocal imitation training in addition to treatment as usual against treatment as usual which consisted of PBS intervention and individual education sessions (Ingersoll et al, 2016). Singh et al (2016) compared mindfulness-based PBS training

1
2 to a control group which received PBS training alone. Johnson et al (2013) examined
3
4 behavioural parent training in addition to psychoeducational programme against treatment as
5
6 usual (which consisted of psychoeducational programme). Hassiotis et al (2009) assessed
7
8 behaviour therapy (applied behaviour analysis and positive behavioural support) against
9
10 standard care alone. Reitzel et al (2013) study compared functional behaviour skills training
11
12 to a waitlist-control group.
13
14

16 **Outcomes**

17
18 See Table 4 for the characteristics of fidelity elements reported in the studies.
19
20
21
22

23 *Adherence of implementation*

24
25 Steps taken to measure adherence were inconsistently reported in the studies. Only two
26
27 studies provided detailed evaluations of whether an intervention was delivered as described in
28
29 the protocol which included: records of methodological techniques used, interrater-reliability
30
31 for the adherence measures and quantitative adherence data (Ingersoll et al, 2016; Johnson et
32
33 al, 2013).
34
35
36
37
38

39
40 In Ingersoll's et al (2016) study six randomly selected intervention sessions were assessed by
41
42 researchers using fidelity checklists. The agreement between researchers for fidelity
43
44 checklists was 0.96 which was calculated on 33% of the sessions using intraclass correlation
45
46 coefficients. Average fidelity rating for the intervention was 91.21% (range 90.37–100%)
47
48 indicating that teachers implemented the reciprocal imitation intervention, on average, as
49
50 described by the protocol. However, there was no information provided in the paper what
51
52 these fidelity checklists were capturing.
53
54
55
56
57

58 Johnson's et al (2013) study used both fidelity checklists and 10% of randomly selected
59
60 videotaped sessions to assess the therapists' adherence. This study was the only one that

1 provided fidelity definition that authors are using in the paper. The fidelity checklist included
2 therapists' adherence to the protocol (whether all essential elements were provided) and
3
4 parents' adherence to the intervention (whether parents completed homework or implemented
5
6 suggested strategies). Therapists were asked to use fidelity checklists after each session and
7
8 rate themselves on a scale 0 to 2 whether they achieved the goals of the intervention,
9
10 indicating 0 if the goals were not achieved, 1 if goals were partially achieved or 2 if goals
11
12 were achieved. Additionally, all intervention sessions were videotaped and 10% of the
13
14 sessions were randomly selected and assessed by independent observers. Average Interrater-
15
16 reliability for the randomly selected sessions was 98.2% (range 81.8–100%) and for parent
17
18 adherence was 91.8% (range 66.7–100%). Mean fidelity rating for the Behavioural Parent
19
20 Intervention was 98% (range 83- 100%) and for parent adherence 93% (range 75–100%).
21
22
23
24
25
26
27
28

29 Partially reported adherence was found in Hassiotis's et al (2009) and Reitzel's et al (2013)
30 studies. Hassiotis et al (2009) indicated that fidelity for implementation of the intervention
31 was high without mentioning methodological techniques used and quantitative data. Reitzel
32 et al (2013) study only reported that staff completed fidelity checklists without providing
33
34 quantitative data in the paper.
35
36
37
38
39
40
41
42

43 *Dose*

44 The amount of an intervention delivered was reported in detail as all five studies provided
45 details of the frequency and duration of an intervention (Ingersoll et al, 2016; Johnson et al,
46
47 2013; Hassiotis et al, 2009; Reitzel et al, 2013; Singh et al, 2016). However, only one study
48
49 provided the number of sessions that participants had actually completed (Johnson et al,
50
51 2013). In this study, self-reports revealed that parents' attendance to the intervention sessions
52
53 was 97.3%.
54
55
56
57
58
59
60

Quality of delivery

The most frequently reported measure of assurance of the quality of delivery was training in the intervention techniques and choice of training providers based on qualifications and experience (all five studies). However, only Reitzel et al (2013) study examined parents' knowledge of ABA principles and sense of competence before and after the training. None of the other included studies examined whether staff or family members acquired necessary knowledge and skills after the training.

Ingersoll et al (2016) and Johnson et al (2013) had also mentioned that supervision was available for intervention providers. Training fidelity checklists (whether the training was provided as indicated in the protocol) only mentioned in Ingersoll's et al (2016) and Singh's et al (2016) studies. Even though Singh et al (2016) videotaped 10 random parts of the training and reported quantitative training fidelity data (which was indicated as 100%), only one person was responsible for rating the quality of the training and it was not reported whether this assessor was independent of the study team or not. Thus, it is unclear whether subjective opinion biased the results.

Responsiveness

Measurement of participants' engagement into the content of the intervention and their perceptions about the usefulness of an intervention were rarely reported. Two studies out of five reported participants' responsiveness. Ingersoll et al (2016) study explored intervention providers' views about the benefits of the intervention and whether participants enjoyed the sessions. However, participants' views about the usefulness of the intervention have not been explored directly in this study. In Johnson's et al (2013) study parents received satisfaction questionnaires to rate how useful they found teaching goals and specific intervention strategies.

Program differentiation

None of the included studies carried out a component analysis or qualitative research to examine which elements of an intervention are essential and the most effective.

Risk of bias in the included studies

See Figure 2 (below) for a summary of risk of bias in the included studies and also supplementary material 3 for reviewers' judgments of risk of bias. Overall, the included studies were found to be at high risk of reporting bias especially for these domains: allocation concealment, blinding, incomplete outcome data, selective reporting and other bias.

Four out of five studies had reported specific mechanisms used to produce random sequence generation (Ingersoll et al, 2016; Singh et al, 2016; Hassiotis et al, 2009; Reitzel et al, 2013).

Two studies had provided information about allocation concealment (Singh et al, 2016; Johnson et al, 2013). Blinding of participants, personnel and outcome assessors were stated in one study (Reitzel et al, 2013). Missing outcome data and its reasons were reported in four studies (Ingersoll et al, 2016; Singh et al, 2016; Hassiotis et al, 2009; Reitzel et al, 2013).

Selective reporting was evident in two studies which indicated that fidelity measurements were completed but quantitative data were not provided (Hassiotis et al, 2009; Reitzel et al, 2013). All five studies had other potential threats to validity which were: small sample sizes reported in all studies (range 15 to 45 participants), longer sessions in the treatment condition compared to the control condition in one study (Ingersoll et al, 2016), administration of reciprocal imitation training in addition to PBS plans which also target communication as a part of a plan (Ingersoll et al, 2016) and participants' baseline characteristics unbalanced between the groups (Reitzel et al (2013).

Figure 2: Summary of risk of bias in the included studies

	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias
Hassiotis 2009	+	+	?	?	+	-	-
Ingersoll 2016	+	-	-	-	+	+	-
Johnson 2013	?	?	?	-	-	+	-
Reitzel 2013	+	+	+	+	-	-	-
Singh 2016	+	?	?	?	+	+	?

Discussion

The review presents data from five RCTs that had reported measurements of fidelity in complex interventions for behaviours that challenge based on PBS/ABA principles. We found variable and inconsistent descriptions of how fidelity was measured and reported in the studies. The most frequently provided fidelity elements found in four out of five studies were adherence of implementation, dose and some aspects of quality of delivery (Ingersoll et al, 2016; Johnson et al, 2013; Hassiotis et al, 2009; Reitzel et al, 2013). However, there were considerable differences between studies about how these fidelity elements were analysed and reported.

1
2
3
4 Only two studies in the review had provided detailed accounts on how adherence of
5
6 implementation was measured reporting methodological techniques, interrater reliability and
7
8 quantitative fidelity data (Ingersoll et al, 2016; Johnson et al, 2013). By contrast, the other
9
10 two studies only mentioned that adherence of implementation was high without providing
11
12 measurement techniques used and quantitative data (Hassiotis et al, 2009; Reitzel et al, 2013).
13
14 Even though frequency and duration of the intervention sessions were reported quite well,
15
16 only one study indicated the number of sessions that participants had actually completed
17
18 (Johnson et al, 2013). Only Reitzel et al (2013) examined whether staff and family members
19
20 acquired skills and knowledge after the training. One study in the review only mentioned
21
22 steps taken to measure training fidelity (whether training was provided as indicated in the
23
24 protocol) and did not report any other fidelity elements (Singh et al, 2016). None of the
25
26 studies assessed program differentiation to examine which components of the intervention are
27
28 essential and the most effective. Only one study reported the definition of fidelity that authors
29
30 are using in the paper (Johnson et al, 2013).
31
32
33
34
35
36
37
38

39 Besides inconsistent fidelity measurements found in the review, there is room for
40
41 improvement for studies which had reported some fidelity measurements. Ingersoll et al
42
43 (2016) study only used one methodological technique to measure adherence of
44
45 implementation and it is accepted that several techniques need to be applied to provide
46
47 reliable data (Spilane et al, 2007). Even though Johnson et al (2013) used several methods to
48
49 collect quantitative data, intervention providers rated themselves using fidelity checklists
50
51 whether they had implemented the intervention as indicated in the protocol. Therapists may
52
53 be afraid to disclose implementation difficulties and may not respond truthfully leading to
54
55 inaccurate fidelity data. Moreover, Singh et al (2016) reported that training fidelity was 100%
56
57 suggesting that training was provided exactly as indicated in the protocol. However, only one
58
59
60

1
2 observer was involved in the measurement of training fidelity and these findings should be
3
4 interpreted with caution.
5

6
7
8
9 We found a paucity of research measuring fidelity in complex interventions for behaviours
10 that challenge based on PBS/ABA principles and this finding is in line with previous
11 literature (McIntyre et al, 2007; Sanetti et al, 2012). Our review included five papers that met
12 the inclusion criteria and this number is lower than McIntyre's et al (2007) and Sanetti's et al
13 (2012) studies found (46 and 35 studies respectively). However, these reviews used different
14 inclusion criteria and involved studies with designs other than RCTs. The limited number of
15 RCTs found in the review reflects a general lack of RCTs including people with intellectual
16 disabilities (Beavis et al, 2007; Hassiotis, 2009).
17
18
19
20
21
22
23
24
25
26
27
28

29 Reasons for poor fidelity measurements in the behavioural literature are unclear. One
30 possible reason could be that there are multiple definitions of fidelity used in the literature
31 and no consensus on which fidelity elements have to be assessed in complex interventions for
32 behaviours that challenge (Bellg et al, 2004). In the review, we used the most commonly
33 applied fidelity framework across different disciplines (Dane and Schneider, 1998), to
34 appraise fidelity elements (Carroll et al, 2007; Century et al, 2010). However, it is unclear
35 whether all these five fidelity elements indicated in the framework are equally important and
36 need to be evaluated in complex interventions for behaviours that challenge in order to
37 provide a comprehensive fidelity measurement. The CONSORT and MRC guidelines on
38 reporting findings from clinical trials indicate the importance of measuring fidelity in clinical
39 trials but there is no recommendation how to do so (Boutron et al, 2008; Craig et al, 2008). A
40 standardised fidelity checklist that could be easily applied in different settings does not exist
41 and usually such scales are developed for each study (Spilane et al, 2007).
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Strengths and limitations

The strengths of this study include a systematic review of fidelity measurements in complex interventions for behaviours that challenge in people with ID that has not been previously conducted in this area. We used a comprehensive search strategy and reviewed grey literature such as conference reports and clinical trial registration websites to find unpublished studies.

Nevertheless, the findings of this study should be interpreted with caution. Firstly, the review included studies published only in the English language and there is a risk that we did not capture all relevant studies. Secondly, we reviewed fidelity measurements in a variety of different behavioural interventions for behaviours that challenge which are mainly based on PBS or ABA principles and this mixture of studies reflects the lack of RCTs carried out including people with behaviours that challenge. Thus, we might focused on the fidelity elements that may not be equally important to look at ABA and PBS interventions. Thirdly, we could not do a meta-analysis because the fidelity measurements reported in the studies were inconsistent and only two studies had provided quantitative data (Ingersoll et al, 2016; Johnson et al, 2013).

Implications for research and practice

Our review confirmed the limited number of RCTs examining complex interventions for behaviours that challenge based on PBS/ABA principles despite it being an important criterion of the study's internal and external validity. Fidelity is being seen as the degree to which the intervention was implemented as indicated in the protocol (Moore et al, 2015). However, for complex interventions nonspecific therapy factors such as therapist's competence, skills, motivation, communication style may play an important part how an intervention is implement and might need to be considered as well (Chatoor and Kurpnick, 2001).

1
2 We therefore suggest that future studies should clarify which fidelity elements have to be
3
4 assessed in complex interventions for behaviours that challenge in people with ID. This
5
6 information might ease the current confusion in the literature and improve the measurements
7
8 of fidelity in the clinical trials. The development of a standardised approach for measuring
9
10 fidelity in such interventions for behaviours that challenge could be the next step aiming to
11
12 increase fidelity measurements in the trials. Our findings from the review could be used as a
13
14 starting point for researchers to look at methodological techniques reported in complex
15
16 interventions in this population and to see if any of the helpful techniques could be used.
17
18 Besides, the Consort and MRC guidelines could establish the best practice for how to
19
20 measure fidelity in complex interventions which might provide clarity for many researchers
21
22 how to measure fidelity in PBS/ABA trials.
23
24
25
26
27
28

29
30 PBS is a complex intervention and as such, it is imperative that it is subject to the same
31
32 quality assurance as other interventions. Given that its delivery can be variable, meaning that
33
34 therapists may implement deviations that are likely to compromise its delivery, it is necessary
35
36 to develop a standardised approach for measuring fidelity in research (Denne et al, 2013).
37
38 PBS is about applying research-validated interventions into everyday settings (Carr et al,
39
40 2002). However, without fidelity data widespread PBS application into real-world contexts is
41
42 challenging.
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

References

References of studies included in the review

Hassiotis, A, Robotham, D, Canagasabey, A, Romeo, R, Langridge, D, Blizard, R, Murad, S, and King, M. (2009), 'Randomized, single-blind, controlled trial of a specialist behavior therapy team for challenging behavior in adults with intellectual disabilities', *American Journal of Psychiatry*, vol. 166, no. 11, pp. 1278–1285.

Ingersoll, B, Berger, N, Carlsen, D and Hamlin, T (2016), 'Improving social functioning and challenging behaviors in adolescents with ASD and significant ID: A randomized pilot feasibility trial of reciprocal imitation training in a residential setting', *Developmental Neurorehabilitation*, vol. 20, no. 4, pp. 1–10.

Johnson, C. R. et al. (2013) 'Behavioral parent training to address sleep disturbances in young children with autism spectrum disorder: A pilot trial', *Sleep Medicine*. Elsevier B.V., 14(10), pp. 995–1004. doi: 10.1016/j.sleep.2013.05.013.

Reitzel, J, Summers, J, Lorv, B, Szatmari, P, Zwaigenbaum, L, Georgiades, S and 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*, vol. 7, no. 11, pp. 1418–1432.

Singh, NN, Lancioni, GE, Karazsia, BT, Chan, J and Winton, ASW (2016), 'Effectiveness of caregiver training in Mindfulness-Based Positive Behavior Support (MBPBS) vs. Training-as-Usual (TAU): A randomized controlled trial', *Frontiers in Psychology*, vol. 7, no. 1, pp. 1–13.

Additional references

Allen, D (2009), 'Positive behavioural support as a service system for people with challenging behaviour', *Psychiatry*, vol. 8, no. 10, pp. 408–411.

Allen, D, James, W, Evans, J, Hawkins, S and Jenkins, R (2005), 'Positive behavioural support: definition, current status and future directions', *Tizard Learning Disability Review*, vol. 10, no. 2, pp. 4–11.

Beadle-Brown J, Mansell J, Whelton B and Hutchinson, A (2006), 'People with intellectual disability in 'out-of-area' residential placements: 2. Reasons for and effects of placement', *Journal of Intellectual Disability Research*, vol. pp. 845–856.

Beavis, J, Kerr, M, Marson, AG (2007), 'Non-pharmacological interventions for epilepsy in people with intellectual disabilities', *Cochrane Database of Systematic Reviews*, vol. 4, pp.1-18

Bellg, AJ, Borrelli, B, Resnick, B, Hecht, J, Minicucci, DS, Ory, M, Ogedegbe, G, Orwig, D, Ernst, D and Czajkowski, S, (2004), 'Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH Behavior Change Consortium', *Health Psychology*, vol. 23, no. 5, pp. 443-451.

Boutron, I, Moher, D, Altman, DG, Schulz, KF and Ravaud, P, (2008), 'Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration', *Annals of Internal Medicine*, vol. 148, no. 4, pp. 295–309.

British Psychological Society (2018), *Positive Behaviour Support (PBS)*, Leicester: British Psychological Society.

1
2
3 Carr, EG, Dunlap, G, Horner, RH, Koegel, RL, Turnbull, AP, Sailor, W and Fox, L, (2002),
4 'Positive behavior support: Evolution of an applied science', *Journal of Positive Behavior*
5
6 *Interventions*, vol. 4, no. 1, pp. 4–16.
7
8
9

10
11 Carr, EG, Horner, RH, Turnbull, AP, Marquis, JG, McLaughlin, DM and McAtee, ML,
12
13 (1999), *Positive behavioral support for people with developmental disabilities: A research*
14
15 *synthesis*. Washington, DC: American Association on Mental Retardation.
16
17

18
19 Carroll, C, Patterson, M, Wood, S, Booth, A, Rick, J and Balain, S, (2007), 'A conceptual
20
21 framework for implementation fidelity', *Implementation Science*, vol. 2, no. 1, pp. 40-47.
22
23

24
25 Century, J, Rudnick, M, and Freeman, C, (2010), 'A framework for measuring fidelity of
26
27 implementation: A foundation for shared language and accumulation of knowledge',
28
29 *American Journal of Evaluation*, vol. 31, no. 2, pp. 199-218.
30
31

32
33 Chatoor, I. and Kurpnick, J, (2001), 'The role of non-specific factors in treatment outcome of
34
35 psychotherapy studies', *European Child & Adolescent Psychiatry*, vo. 10, no. 1, pp. 19-25.
36
37

38
39 Craig, P, Dieppe, P, Macintyre, S, Michie, S, Nazareth, I, Petticrew, M, (2008), 'Developing
40
41 and evaluating complex interventions: the new Medical Research Council guidance', *British*
42
43 *Medical Journal*, vol. 337, no. 7676, pp. 979-983.
44
45

46
47 Dane, AV and Schneider, BH, (1998), 'Program integrity in primary and early secondary
48
49 prevention: Are implementation effects out of control?' *Clinical Psychology Review*, vol. 18,
50
51 no. 1, pp. 23-45.
52
53

54
55 Davison, S, McGill, P, Baker, P and Allen, D et al, (2015), 'A national UK survey of
56
57 peripatetic support teams for children and adults with intellectual and developmental
58
59
60

1
2
3 disability who display challenging behaviour', *International Journal of Positive Behavioural*
4 *Support*, vol. 5, no. 1, pp. 26-33.

5
6
7 Davison, S, McGill, P, Baker, PA and Allen, D, (2015), 'A national UK survey of peripatetic
8 support teams for children and adults with intellectual and developmental disability who
9 display challenging behaviour'. *International Journal of Positive Behavioural Support*, vol.
10 5, no.1, pp. 26-33.

11
12
13 Denne, LD, Noone, SJ, Gore, NJ, Toogood, S, Hughes, JC, Hastings, RP, Allen, D, Baker, P
14 and McGill, P, (2013), 'Developing a core competencies framework for positive behavioural
15 support: Issues and recommendations', *International Journal of Positive Behaviour Support*,
16 vol.3, no. 2, pp. 24-31.

17
18
19 Department of Health (2014). *Positive and proactive care: Reducing the need for restrictive*
20 *interventions*. London: Department of Health.

21
22
23 Donnellan, AM, LaVigna, GW, Negri-Shoultz, N and Fassbender, LL (1988), *Progress*
24 *Without Punishment: effective approaches for learners with behavior problems*. New York:
25 Teachers College Press

26
27
28 Dunlap, G, Carr, EG, Horner, RH, Zarcone, JR and Schawrtz, I (2008), 'Positive Behavior
29 Support and Applied Behaviour Analysis: A Familial Alliance', *Behavior Modification*, vol.
30 32, no. 5, pp. 682–698.

31
32
33 Emerson E, Hatton C, Robertson J, Roberts H, Baines S, Glover G (2011), *People with*
34 *Learning Disabilities in England 2010*, Durham: Improving Health and Lives Learning
35 Disabilities Observatory.

36
37
38 Emerson, E and Einfeld, SL (2011), *Challenging Behaviour*, Cambridge, UK: Cambridge
39 University Press.

1
2
3 Emerson, E, Robertson, J, Gregory, N, Hatton, C, Kessissoglou, S and Hallam, A, (2000),
4 ‘Treatment and management of challenging behaviour in residential settings’, *Journal of*
5
6 *Applied Research in Intellectual Disabilities*, vol. 13, no. 1, pp. 197–215.
7
8
9

10
11 Gore, NJ, McGill, P, Toogood, S, Allen, D, Hughes, JC, Baker, P, Hastings, RP, Noone, SJ
12
13 and Denne, L. D. (2015) ‘Definition and scope for positive behaviour support’, *International*
14
15 *Journal of Positive Behaviour Support*, vol. 3, no. 2, pp.14-23.
16
17

18
19 Hassiotis A, Poppe M, Strydom A, Vickerstaff V, Hall I, Crabtree J, et al, (2018), ‘Positive
20
21 behaviour support training for staff for treating challenging behaviour in people with
22
23 intellectual disabilities: a cluster RCT’, *Health Technology Assessment*, vol. 22, no. 15.
24
25

26
27 Hassiotis, A. (2009), ‘Research in mental health learning disabilities: present challenges and
28
29 future drivers’, *Psychiatry*, vol. 8, no. 11, pp. 457-460.
30

31
32 Hassiotis, A. et al. (2018) ‘Clinical outcomes of staff training in positive behaviour support to
33
34 reduce challenging behaviour in adults with intellectual disability: Cluster randomised
35
36 controlled trial’, *British Journal of Psychiatry*, 212(3), pp. 161–168. doi:
37
38 10.1192/bjp.2017.34.
39

40
41 Higgins, JPT and Green, S (editors) (2011) *Cochrane Handbook of Systematic Reviews of*
42
43 *Interventions Version 5.1.0 [updated March 2011]*. The Cochrane Collaboration.
44
45

46
47 LaVigna, GW and Willis, TJ (2005), ‘A positive behavioural support model for breaking the
48
49 barriers to social and community inclusion’, *Tizard Learning Disability Review*, vol. 10, no.
50
51 2, pp.16-23.
52

53
54 LaVigna, GW and Willis, TJ (2012), ‘The efficacy of positive behavioural support with the
55
56 most challenging behaviour: the evidence and its implications’, *Journal of Intellectual and*
57
58 *Developmental Disability*, vol. 37, no. 3, pp. 185–195.
59
60

1
2
3 McClean, B and Grey, I (2012), 'A component analysis of positive behaviour support plans',
4
5 *Journal of Intellectual and Developmental Disability*, vol. 37, no. 3, pp. 221-231.
6
7

8
9 McIntyre, LL, Gresham, FM, DiGennaro, FD and Reed, DR, (2007), 'Treatment Integrity of
10
11 School-Based Interventions with Children in the Journal of Applied Behavior Analysis 1991-
12
13 2005', *Journal of Applied Behavior Analysis*, vol. 40, no. 4, pp. 659-672.
14
15

16
17 Moore, GF, Audrey, S, Barker, M, Bond, L, Bonell, C, Hardeman, W, Moore, L, Cathain,
18
19 AO, Tinati, T, Wight, D and Baird, J, (2015), 'Process evaluation of complex interventions :
20
21 Medical Research Council guidance', *British Medical Journal*, vol. 350, no. 7676, pp. 1-7.
22
23

24
25 Mowbray, CT, Holter, MC, Teague, GB and Bybee, D, (2003), 'Fidelity criteria:
26
27 development, measurement, and validation', *American Journal of Evaluation*, vol. 20, no. 3,
28
29 pp. 315-340.
30
31

32
33 National Institute for Clinical Excellence (NICE), (2105), *Challenging behaviour and*
34
35 *learning disabilities: Prevention and interventions for people with learning disabilities whose*
36
37 *behaviour challenges (NG11)*. London: NICE
38
39

40
41 Positive Behavioural Support (PBS) Coalition UK, (2015), *Positive behavioural support: A*
42
43 *competence framework*.
44

45
46 Sanetti, LMH, Dobby, LM and Gritter, KL, (2012), 'Treatment Integrity of Interventions
47
48 With Children in the Journal of Positive Behavior Interventions From 1999 to 2009', *Journal*
49
50 *of Positive Behavior Interventions*, vol. 14, no. 1, pp. 29-46.
51
52

53
54 Spee, GA, Polo, AJ and Budd, KS, (2014), 'Establishing Treatment Fidelity in Evidence-
55
56 Based Parent Training Programs for Externalizing Disorders in Children and Adolescents',
57
58 *Clinical Child and Family Psychology Review*, vol.17, no.3, pp.230-247.
59
60

1
2
3 Spillane, V, Byrne, MC, Byrne, M, Leathem, CS, Malley, MO and Margaret, E, (2007),
4
5 'Monitoring treatment fidelity in a randomized controlled trial of a complex intervention',
6
7
8 *Journal of Advanced Nursing*, vol. 60, no. 3, pp.343–352.
9

10
11 Sturme, P (2009), 'Restraint, seclusion and PRN medication in English services for people
12
13 with learning disabilities administered by the National Health Service: An analysis of the
14
15 2007 National Audit Survey', *Journal of Applied Research in Intellectual Disabilities*, vol.
16
17 22, no. 2, pp. 140-144.
18
19

20
21 Toogood, S, Saville, MH, McLennan, K, McWade, P, Morgan, G and Welch, C, (2015),
22
23 'Providing positive behavioural support services: specialist challenging behaviour support
24
25 teams', *International Journal of Positive Behavioural Support*, vol. 5, no. 1, pp. 6-15.
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1: Fidelity components (based on Dane and Schneider, 1998)

Fidelity elements	Description
Adherence	The extent to which an intervention content, methods and activities are implemented as described in the protocol. For example, if an intervention provider covered 15 out of the 30 content areas of an intervention, the adherence to the intervention would be 50%.
Dose (also known as exposure)	The amount of an intervention received by participants compared to the amount of the intervention indicated in the protocol. Dose can include number of sessions completed, frequency and duration of sessions. Ideally, dose should be measured using a combination of different measurements such as attendance logs or checklists, self-reports completed by intervention providers and observations of randomly selected sessions in order to assess the proportion of an intervention delivered (Spillane et al, 2007).
Quality of delivery	The measurement of the effort required to deliver an intervention. Aspect of quality of delivery can include: the assurance of quality of training, materials and support to those delivering an intervention, provider confidence and interaction style.
Responsiveness	The degree to which participants engage in the activities and content of an intervention. This can include: participants' interest in the program, perceived benefits of an intervention or engagement in activities.
Program differentiation	The degree to which unique components of an intervention are differentiated from each other and from other interventions. Component analysis can help to determine which elements of an intervention are essential and which elements of an implemented intervention cannot be incorporated in the control condition.

Table 2: Risk of bias (Higgins and Green, 2011)

Risk of bias
1. Allocation bias was examined by looking at whether the method used to assign participants to the groups was truly random (for example, a flip of coin or computerised sequences).
2. Selection bias was examined by looking at whether anyone in the study could predict allocation to the treatment group and attempts involved to conceal it.
3. Performance bias was examined by looking at whether intervention providers and participants were blinded to treatment allocation.
4. Detection bias was examined by looking at whether outcome assessors were blinded to treatment allocation.
5. Attrition bias was examined by looking at proportion of people in both conditions that stopped having treatment and reasons for it (drop-outs and withdrawal).
6. Reporting bias was examined by looking at whether more outcomes were measured than were reported.
7. Other potential threats to validity not addressed by these six domains were stated in this section.

Table 3: Characteristics of the included studies

Study, location and design	Sample	ID level and its measure	CB and its measure	Intervention	Comparison
<p>Hassiotis et al, 2009; UK, Pilot RCT</p>	<p>Community ID service; 63 adults with severe ID and CB.</p> <p><u>Intervention group:</u> 32 participants (19 males and 13 females) Mean age= 39.6 years, SD=15.5.</p> <p><u>Control group:</u> 31 participants (18 males and 13 females) Mean age= 41.3 years, SD=14.5.</p>	<p><u>Intervention group:</u> 20 (62.5%) participants with mild/moderate ID and 12 (37.5%) with severe/profound ID.</p> <p><u>Control group:</u> 22 (71%) participants with mild/moderate ID and 9 (29%) with severe/profound ID. Diagnostic measures used were not identified.</p>	<p><u>Intervention group:</u> Median ABC=36</p> <p><u>Control group:</u> Median ABC= 47</p> <p>Measured using the ABC</p>	<p>Behaviour therapy team (applied behaviour analysis and positive behavioural support) in addition to TAU</p>	<p>TAU (standard care)</p>
<p>Ingersoll et al, 2016; USA, Pilot RCT</p>	<p>Residential homes; 19 adolescents with ASD, severe ID and CB.</p> <p><u>Intervention group:</u> 10 participants (9 males and 1 female) Mean age=16.26 years, SD=3.15.</p> <p><u>Control group:</u> 9 participants (7 males 2 females) Mean age =16.90 years, SD=2.53.</p>	<p><u>ID level:</u> severe;</p> <p><u>Treatment group:</u> Mean IQ =48.50 (SD=13.42)</p> <p><u>Control group:</u> Mean IQ=47.43 (SD=8.48)</p> <p>Measured using WAIS-IV.</p>	<p><u>Treatment group:</u> Mean ABC= 43.70 (SD=9.50)</p> <p><u>Control group:</u> Mean ABC= 40.88 (SD= 6.92).</p> <p>Measured using the ABC-R.</p>	<p>Individual Reciprocal Imitation Training in addition to TAU (education, PBS plans, medication)</p>	<p>Individual sessions with teachers in addition to TAU</p>
<p>Johnson et al, 2013; USA, Pilot RCT</p>	<p>Inpatient; 33 children with ASD or PDD-NOS and sleep problems.</p> <p><u>Intervention group:</u> 15 participants (11 boys and 4 girls) Mean age= 3.5 years, SD= 0.98.</p> <p><u>Control group:</u> 18 participants (15 boys and 3 girls) Mean age=3.6 years, SD=1.12.</p> <p>Diagnostic measures used: ADI-R and ADOS.</p>	<p><u>Intervention group:</u> 10 out of 15 children with lower cognitive scores (≤ 70 standard score) Mean standard cognitive score=65.73, SD=17.23.</p> <p><u>Control condition:</u> 10 out of 18 children with lower cognitive scores (≤ 70 standard score) Mean standard cognitive score= 68.11, SD=17.48.</p> <p>Measured using the MEIS scale or the SIIS scale.</p>	<p>All participants had sleep problems (e.g. bedtime resistance, night awakening) based on parents' reports.</p>	<p>Behavioural parent training in addition to TAU (psychoeducational programme)</p>	<p>TAU (psychoeducational programme)</p>

Study, location and design	Sample	ID level and its measure	CB and its measure	Intervention	Comparison
Reitzel et al, 2013; Canada, Pilot RCT	Autism Intervention Program; 15 children with ASD and early learning difficulties measured using the ELM. <u>Intervention group:</u> 7 participants (Mean age= 62.8 months, SD= 16.3) <u>Control group:</u> 8 participants (Mean age= 57.9 months, SD=13.1).	Mean IQ= 37.1 (SD=12.4); mental age one of third below their chronological age. Scores were obtained using the Early Learning Measure.	<u>Treatment group:</u> Mean DBS score (CB measure) =21.7 (SD=17.2). <u>Control group:</u> Mean DBS score = 21.4 (SD=24.7).	Functional Behaviour Skills Training in addition to TAU	Waitlist-control group
Singh et al, 2016; USA, RCT	Residential homes; 48 adults with severe or profound ID and CB. <u>Intervention group:</u> 24 participants (16 males 8 females) Mean age= 39.21 years, SD=7.61. <u>Control group:</u> 24 participants (16 males and 8 females) Mean age=42.33 years, SD=9.22.	<u>Intervention group:</u> 9 with severe ID and 25 with profound ID. <u>Control group:</u> 7 with severe ID and 17 with profound ID. Diagnostic criteria not specified.	All individuals exhibited CB. <u>Treatment group:</u> 18 (75%) had behaviour plans for aggressive behaviours. <u>Control group:</u> 16 (67%) had behaviour plans for aggressive behaviours. Standardised measures to assess CB were not reported.	Mindfulness based PBS training (MBPBS)	PBS training

Note: RCT= Randomised Controlled Trial; ASD= Autistic Spectrum Disorder; ID= Intellectual Disabilities; CB= Challenging Behaviour; WAIS-IV= Wechsler Adult Intelligence Scale-Fourth Edition; ABC-R= Aberrant Behaviour Checklist-Residential; TAU= treatment as usual; PDD-NOS= Pervasive Developmental Disorder, not otherwise specified; MEIS= Multifactor Emotional Intelligence Scale; SIIS= Stanford-Binet Intelligence Scale; ADI-R= Autism Diagnostic Interview Revised Version; ADOS= Autism Diagnostic Observation Schedule; ELM= Early Learning Measure; DBS= Developmental Behavior Checklist – Parent/Carer Version.

Table 4: Elements of fidelity reported in the included studies

Study	Elements of Fidelity (Based on Dane and Schneider, 1998)				
	Adherence	Dose	Quality of delivery	Responsiveness	Program differentiation
Hassiotis et al, 2009	<p>Reported that fidelity was 'high'. Methodological techniques used to measure fidelity not reported.</p> <p>Quantitative data: not reported.</p>	<p>Frequency and duration: Mean contacts=9 (SD=7); Mean duration= 6 months.</p> <p>Number of sessions completed: not reported.</p>	<p>The intervention was provided by the behaviour therapy team.</p>	<p>Not reported.</p>	<p>Not reported.</p>
Ingersoll et al, 2016	<p>Six sessions randomly selected and assessed using fidelity checklists by the research team.</p> <p>Quantitative data: Mean fidelity rating for RIT=91.21% (range: 90.37–100%).</p> <p>Interrater reliability for fidelity checklists was calculated on 33% of the sessions using ICCs. ICCs=.96 for fidelity of the RIT sessions.</p>	<p>Frequency and duration: 10 min sessions 2 per day, 2–3 days per week.</p> <p>Number of sessions completed: not reported.</p>	<p>Training (lasted 2 weeks) and supervision was provided for intervention providers by a supervisor who had previously established fidelity in RIT.</p> <p>RIT fidelity checklists were used to assess whether intervention providers learnt RIT and able to implement it.</p> <p>Quantitative data: not reported.</p>	<p>Questionnaires administered to teachers to assess benefits and participants' enjoyment of the intervention.</p> <p>One participant often wanted to finish sessions earlier which might indicate that he found the intervention unpleasant.</p>	<p>Not reported.</p>
Johnson et al, 2013	<p>Used two fidelity measures: checklists and videotapes.</p> <p>Fidelity checklists were completed by therapists after each session. Therapists rated themselves on a scale 0 to 2 whether they achieved the goals of the intervention. The fidelity checklist also included assessment of parents' adherence to the treatment (e.g. whether parents completed homework or implemented new strategies).</p> <p>All intervention sessions were videotaped and 10 % of them were randomly selected and assessed by independent observers.</p> <p>Quantitative data: the fidelity for the intervention=98% (range, 83–100) and parent adherence= 93% (range, 75–100).</p>	<p>Frequency and duration: five individual sessions (60-90 minutes).</p> <p>Number of sessions completed: self-reports indicated that the 15 families in the intervention group attended 73 of the 75 possible sessions (97.3%).</p>	<p>Training was administered by two master-trained doctoral students or one senior doctoral-trained behaviour analyst. Training providers were regularly observed by an investigator and feedbacks were provided.</p> <p>Quantitative data: not reported.</p>	<p>Satisfaction questionnaires were administered to parents to rate the usefulness of the teaching and specific elements of the intervention. Higher scores indicated greater satisfaction.</p> <p>Quantitative data: Mean parent satisfaction with the BPT= 90% (range, 71–100%).</p>	<p>Not reported.</p>

Study	Adherence	Dose	Quality of delivery	Responsiveness	Program differentiation
Singh et al, 2016	Not reported.	<p>Frequency and duration: On average, the caregivers meditated for 89% (range 0-96%) of the 40 days for between 25-40 minutes daily.</p> <p>Number of sessions completed: not reported.</p>	<p>Training was presented in 3 parts over a 10 week period. A training provider was experienced behavioural analyst and mindfulness practitioner. Ten randomly selected training segments were videotaped and one observer (qualified in behaviour analysis and meditation) rated the fidelity of the training.</p> <p>Quantitative data: Mean fidelity rating for the training= 100%.</p>	Not reported.	Not reported.

Note: ICCs=intraclass correlation coefficients; RIT= Reciprocal Imitation Training; ABA=Applied Behaviour Analysis; IBI= Intensive Behaviour Intervention.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

Supplementary material

Supplementary material 1: Search strategy

Ovid Medline, Embase and PsycINFO databases were searched together on 22 March 2017.

Web of Science and CINAHL Plus databases were searched separately on 15 June 2017 using the same search strategy and terms as with Medline, Embase and PsycINFO databases.

The following search terms were used:

1. "learning disab*"
2. "learning difficult*"
3. "learning disorder"
4. "learning impair*"
5. "intellectual* disab*"
6. "intellectual* impair*"
7. "intellectual dysfunction"
8. "development* disab*"
9. "development* disorder*"
10. "development* impair*"
11. "intellectual developmental disorder"
12. "mental* deficien*"
13. "mental* retard*"
14. "mental* handicap*"
15. "mental* disab*"
16. "mental insufficiency"
17. "mental* impair*"
18. "mental* challenged"
19. "intelligence"
20. IQ
21. "subaverage intelligence"
22. 1 OR 2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 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
23. "challenging behavio*"
24. "aberrant behavio*"
25. "maladaptive behavio*"
26. "problem behavio*"
27. "behavio* problems"
28. self-injury
29. self-harm
30. "self-injurious behavio*"
31. stereotypy
32. "stereotyped behavio*"
33. "repetitive behavio*"
34. aggress*
35. "destructive behavio*"

- 1
- 2
- 3
- 4 36. "property destruction"
- 5 37. "disruptive behavio**"
- 6 38. 23 OR 24 OR 25 OR 26 OR 27 OR 28 OR 29 OR 30 OR 31 OR 32 OR 33 OR 34 OR
- 7 35 OR 36 OR 37 OR 38
- 8 39. "positive behav* support"
- 9 40. PBS
- 10 41. "applied behav* analysis"
- 11 42. ABA
- 12 43. "functional analysis"
- 13 44. "functional assessment"
- 14 45. "behavio* intervention*"
- 15 46. "behavio* technique*
- 16 47. "behavio* treatment*"
- 17 48. "behavio* support"
- 18 49. "behavio* modification"
- 19
- 20 50. 39 OR 40 OR 41 OR 42 OR 43 OR 44 OR 45 OR 46 OR 47 OR 48 OR 49
- 21
- 22 51. fidelity
- 23 52. treatment
- 24 53. intervention*
- 25 54. implement*
- 26 55. therap*
- 27 56. clinician
- 28 57. 52 OR 53 OR 54 OR 55 OR 56
- 29 58. 51 AND 57
- 30 59. Integrity
- 31 60. intervention*
- 32 61. treatment
- 33 62. program
- 34 63. 60 OR 61 OR 62
- 35 64. 59 AND 63
- 36 65. Adherence
- 37 66. differentiation
- 38 67. compliance
- 39 68. clinician
- 40 69. therap*
- 41 70. 65 OR 66 Or 67
- 42 71. 68 OR 69
- 43 72. 70 AND 71
- 44 73. 58 OR 64 OR 72
- 45 74. "randomised controlled trial"
- 46 75. "randomized controlled trial"
- 47 76. RCT
- 48 77. "random allocation"
- 49 78. randomisation
- 50 79. randomization
- 51 80. intervention
- 52 81. "clinical trial"
- 53 82. "control* trial*"
- 54 83. "randomly allocated"
- 55 84. "single blind*"
- 56
- 57
- 58
- 59
- 60

- 1
 - 2
 - 3
 - 4
 - 5
 - 6
 - 7
 - 8
 - 9
 - 10
 - 11
 - 12
 - 13
 - 14
 - 15
 - 16
 - 17
 - 18
 - 19
 - 20
 - 21
 - 22
 - 23
 - 24
 - 25
 - 26
 - 27
 - 28
 - 29
 - 30
 - 31
 - 32
 - 33
 - 34
 - 35
 - 36
 - 37
 - 38
 - 39
 - 40
 - 41
 - 42
 - 43
 - 44
 - 45
 - 46
 - 47
 - 48
 - 49
 - 50
 - 51
 - 52
 - 53
 - 54
 - 55
 - 56
 - 57
 - 58
 - 59
 - 60
- 85. 74 OR 75 OR 76 OR 77 Or 78 OR 79 OR 80 OR 81 OR 82 OR 83 OR 84
 - 86. 22 AND 38 AND 50 AND 73 AND 85
 - 87. Limit 86 to English Language
 - 88. Limit 87 to human
 - 89. Limit 88 to year= "1990-Current"
 - 90. Remove duplicates from 89

Supplementary material 2: Excluded studies

33 articles were excluded for the following reasons:

- Articles were not randomised control trials (N=8)
- Articles did not specify percentages of participants with intellectual disabilities in the sample (N=8)
- Interventions were based not on ABA or PBS principles (N=5)
- Interventions did not focus on reducing challenging behaviour (N=6)
- Results of fidelity measures have not been reported (N=1)
- More than 50% of sample had participants with IQ above 70 (N=3)

The following studies were excluded because they were not randomised control trials

(N=8):

1. Allen, D, Lowe, K, Baker, P, Dench, C, Hawkins, S, Jones, E and James, W (2009) 'Assessing the effectiveness of positive behavioural support : The - CPO Project', *International Journal of Positive Behavioural Support*, vol.1, no.1, pp. 14–23.
2. Brookman-Frazer, LI, Drahotka, A and Stadnick, N (2012) 'Training community mental health therapists to deliver a package of evidence-based practice strategies for school-age children with autism spectrum disorders: A pilot study', *Journal of Autism and Developmental Disorders*, vol. 4, no. 8, pp. 1651–1661.
3. Farkas, MS, Simonsen, B, Migdole, S, Donovan, ME, Clemens, K and Cicchese, V (2012) 'Schoolwide positive behavior support in an alternative school setting: An evaluation of fidelity, outcomes, and social validity of tier 1 implementation', *Journal of Emotional and Behavioral Disorders*, vol. 20, no. 4, pp. 275–288.
4. Fava, L, Strauss, K, Valeri, G, D'Elia, L, Arima, S and 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*, vol. 5, no. 4, pp. 1479–1492.

5. Feldman, MA and Werner, SE (2002) ‘Collateral effects of behavioral parent training on families of children with developmental disabilities and severe behavior disorders’, *Behavioral Interventions*, vol. 17, no. 2, pp. 75–83.
6. Hansen, B, Orton, E, Adams, C, Knecht, L, Rindlisbaker, S, Jurtoski, F and Trajkovski, V (2017) ‘A Pilot Study of a Behavioral Parent Training in the Republic of Macedonia’, *Journal of Autism and Developmental Disorders*, vol. 47, no. 6, pp. 1878–1889.
7. Peters-Scheffer, N, Didden, R, Mulders, M and Korzilius, H (2013) ‘Effectiveness of low intensity behavioral treatment for children with autism spectrum disorder and intellectual disability’, *Research in Autism Spectrum Disorders*, vol. 7, no. 9, pp. 1012–1025.
8. Pritchard, D, Hoerger, M, Penney, H, Eiri, L, Hellowell, L, Fothergill, S and Mace, FC (2017) ‘Training Staff to Avoid Problem Behavior Related to Restricting Access to Preferred Activities’, *Behavior Analysis in Practice*, vol. 10, no. 1, pp. 92–95.

The following studies were excluded because percentages of participants with intellectual disabilities were not specified in the sample (N=8):

1. Dawson, G, Rogers, S, Munson, J, Smith, M, Winter, J, Greenson, J, Donaldson, A and Varley, J (2010) ‘Randomized, Controlled Trial of an Intervention for Toddlers With Autism: *The Early Start Denver Model*’, *Pediatrics*, vol. 125, no. 1, pp. e17–e23.
2. Durand, VM, Hieneman, M, Clarke, S, Wang, M and Rinaldi, ML (2013) ‘Positive Family Intervention for Severe Challenging Behavior’, *Journal of Positive Behavior Interventions*, vol. 15, no. 3, pp. 133–143.

- 1
2
3 3. Grahame, V, Brett, D, Dixon, L, McConachie, H, Lowry, J, Rodgers, J, Steen, N and
4
5 Le Couteur, A (2015) 'Managing Repetitive Behaviours in Young Children with
6
7 Autism Spectrum Disorder (ASD): Pilot Randomised Controlled Trial of a New
8
9 Parent Group Intervention', *Journal of Autism and Developmental Disorders*, vol. 45,
10
11 no. 10, pp. 3168–3182.
12
13
14
- 15
16 4. Griffin, C, Guerin, S, Sharry, J and Drumm, M (2010) 'A multicentre controlled study
17
18 of an early intervention parenting programme for young children with behavioural and
19
20 developmental difficulties', *International Journal of Clinical and Health Psychology*,
21
22 vol. 10, no. 2, pp. 279–294.
23
24
25
- 26
27 5. Hemmeter, ML, Snyder, PA, Fox, L and Algina, J (2016) 'Evaluating the
28
29 Implementation of the Pyramid Model for Promoting Social-Emotional Competence
30
31 in Early Childhood Classrooms', *Topics in Early Childhood Special Education*, vol.
32
33 36, no. 3, pp. 133–146.
34
35
- 36
37 6. Rogers, SJ, Estes, A, Lord, C, Vismara, L, Winter, J, Fitzpatrick, A, Guo, M and
38
39 Dawson, G (2012) 'Effects of a brief early start Denver model (ESDM)-based parent
40
41 intervention on toddlers at risk for autism spectrum disorders: A randomized
42
43 controlled trial', *Journal of the American Academy of Child and Adolescent
44
45 Psychiatry*, vol. 51, no. 10, pp. 1052–1065.
46
47
- 48
49 7. Strain, PS and Bovey, EH (2011) 'Randomized, Controlled Trial of the LEAP Model
50
51 of Early Intervention for Young Children With Autism Spectrum Disorders', *Topics
52
53 in Early Childhood Special Education*, vol. 31, no. 3, pp. 133–154.
54
55
- 56
57 8. Thomeer, ML, Lopata, C, Volker, MA, Toomey, JA., Lee, GK, Audrey M. and Smith,
58
59 RA. (2012). 'Randomized Clinical Trial Replication of a Psychosocial Treatment for
60

1
2
3 Children with High-Functioning Autism Spectrum Disorders'. *Psychology in the*
4
5 *Schools*, vol. 49, no. 10 ,pp. 153-178.
6
7

8
9 **The following studies were excluded because interventions were not based on ABA or**
10
11 **PBS principles (N=5):**
12

- 13
14 1. Bagner, DM and Eyberg, SM (2007) 'Parent–Child Interaction Therapy for Disruptive
15 Behavior in Children with Mental Retardation: A Randomized Controlled Trial',
16
17 *Journal of Clinical Child & Adolescent Psychology*, vol. 36, no. 3, pp. 418–429.
18
19
- 20
21 2. Roberts, C, Mazzucchelli, T, Studman, L, & Sanders, MR (2006) 'Behavioral family
22 intervention for children with developmental disabilities and behavioral problems',
23
24 *Journal of Clinical Child and Adolescent Psychology*, vol. 35, no. 2, pp. 180–193.
25
26
- 27
28 3. Ros, R, Hernandez, J, Graziano, PA and Bagner, DM (2016) 'Parent Training for
29 Children With or at Risk for Developmental Delay: The Role of Parental Homework
30 Completion', *Behavior Therapy*, vol. 47, no. 1, pp. 1–13.
31
32
- 33
34 4. Roux, G, Sofronoff, K and Sanders, M (2013) 'A randomized controlled trial of group
35 stepping stones triple P: A mixed-disability trial', *Family Process*, vol. 52, no. 3, pp.
36
37 411–424.
38
39
- 40
41 5. Whittingham, K, Sofronoff, K, Sheffield, J and Sanders, MR (2009) 'Do parental
42 attributions affect treatment outcome in a parenting program? An exploration of the
43 effects of parental attributions in an RCT of Stepping Stones Triple P for the ASD
44
45 population', *Research in Autism Spectrum Disorders*, vol. 3, no. 1, pp. 129–144.
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

The following studies were excluded because interventions did not aim to reduce challenging behaviour (N=6):

1. Mandell, DS, Stahmer, AC, Shin, S, Xie, M, Reisinger, E and Marcus, SC (2013) 'The role of treatment fidelity on outcomes during a randomized field trial of an autism intervention', *Autism*, vo. 17, no. 3, pp. 281–295.
2. Smith, R, Groen, A and Wynn, J (2000) 'Randomized trial of intensive early intervention for children with pervasive developmental disorder', *American Journal of Mental Retardation*, vol. 105, no. 4, pp. 269–285.
3. Sallows, GO and Graupner, TD (2005) 'Intensive behavioral treatment for children with autism: four-year outcome and predictors.', *American Journal of Mental Retardation*, vol. 110, no. 6, pp. 417–438.
4. Rickards, AL, Walstab, JE, Wright-Rossi, RA, Simpson, J and Reddihough, DS (2009) 'One-year follow-up of the outcome of a randomized controlled trial of a home-based intervention programme for children with autism and developmental delay and their families', *Child: Care, Health and Development*, vol. 35, no. 5, pp. 593–602.
5. Mohammadzaheri, F, Koegel, LK, Rezaee, M and Rafiee, SM (2014) 'A randomized clinical trial comparison between Pivotal Response Treatment (PRT) and structured Applied Behavior Analysis (ABA) intervention for children with autism', *Journal of Autism and Developmental Disorders*, vol. 44, no. 11, pp. 2769–2777.
6. Bryson, SE, Koegel, LK, Koegel, RL, Openden, D, Smith, IM and Nefdt, N (2007) 'Large Scale Dissemination and Community Implementation of Pivotal Response Treatment : Program Description and Preliminary Data', *Research And Practice For Persons With Severe Disabilities*, vol. 32, no. 2, pp. 142–153.

1
2
3 **The following study was excluded because results of fidelity measures have not been**
4 **reported (N=1):**
5
6
7

- 8
9 1. Eikeseth, S, Smith, T, Jahr, E and Eldevik, S (2002) 'Intensive Behavioral Treatment
10 at School for 4- to 7-Year-Old Children with Autism: A 1-Year Comparison
11 Controlled Study', *Behavior Modification*, vol. 26, no. 1, pp. 49–68.
12
13
14
15

16 **The following studies were excluded because more than 50% of sample had IQ above 70**
17 **(N=3):**
18
19
20
21

- 22 1. Bearss, K, Johnson, C, Handen, B, Smith, T and Scahill, L (2013) 'A pilot study of
23 parent training in young children with autism spectrum disorders and disruptive
24 behavior', *Journal of Autism and Developmental Disorders*, vol. 43, no. 4, pp. 829–
25 840.
26
27
28
29
30
31
32 2. Bearss, K, Johnson, C, Smith, T, Lecavalier, L, Swiezy, N, Aman, M, McAdam, DB,
33 Butter, E, Stillitano, C, Minshawi, N, Sukhodolsky, DG and Scahill, L (2015) 'Effect
34 of Parent Training vs Parent Education on Behavioral Problems in Children With
35 Autism Spectrum Disorder', *Jama*, vol. 313, no. 15, pp. 1525-1533.
36
37
38
39
40
41
42 3. Bearss, K, Lecavalier, L, Minshawi, N, Johnson, C, Smith, T, Handen, B,
43 Sukhodolsky, D, Aman, M, Swiezy, N, Butter, E and Scahill, L (2013) 'Toward an
44 exportable parent training program for disruptive behaviors in autism spectrum
45 disorders', *Neuropsychiatry*, vol. 3, no. 2, pp. 169–180.
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Supplementary material 3: Risk of bias in the included studies

Hassiotis et al, 2009

Elements of risk of bias	Reviewers' judgement	Description or Comments
Sequence generation (allocation bias)	Low risk of bias	Participants were randomised to the groups using a computer-driven randomization list.
Allocation concealment (selection bias)	Low risk of bias	Quote: "A set of sealed envelopes, each bearing only the name of the area and a number, were held by an independent administrator".
Blinding of personnel and participants (performance bias)	Unclear risk of bias	No information provided.
Blinding of outcome assessors (detection bias) for all outcomes	Unclear risk of bias	No information provided.
Incomplete outcome data (attrition bias)	Low risk of bias	Quote: "One participant in each arm died during the trial, and one participant in the intervention arm declined to participate in follow-up assessments".
Selective reporting (reporting bias)	High risk of bias	Stated that fidelity ratings were high but how it was measured and quantitative data were not provided.
Other potential threats to validity	High risk of bias	Moderate or larger treatment effects might be easier to achieve due to small sample.

Ingersoll et al, 2016

Elements of risk of bias	Reviewers' judgement	Description or Comments
Sequence generation (allocation bias)	Low risk of bias	Randomisation using a coin tossing procedure was reported in the paper.
Allocation concealment (selection bias)	High risk of bias	Using the coin tossing method the study personnel might be able to predict to which group the next participant will be assigned.
Blinding of personnel and participants (performance bias)	High risk of bias	No blinding of participants and personnel.
Blinding of outcome assessors (detection bias) for all outcomes	High risk of bias	Social skills and challenging behaviours were measured by a psychologist who was kept blind to group assignments. However, "the imitation measures for cohort 1 were administered by the teachers who worked with the adolescents and, for cohort 2, the third author who was a staff supervisor

		at the centre and was not blind to group assignment”.
Incomplete outcome data (attrition bias)	Low risk of bias	Small number of drop outs was reported and its reasons were provided in the paper.
Selective reporting (reporting bias)	Low risk of bias	We believe that all outcomes were reported.
Other potential threats to validity	High risk of bias	Moderate or larger treatment effects might be easier to achieve due to small sample. RIT sessions slightly longer than TAU sessions. RIT sessions administered in addition to PBS which also targets communication as a part of a plan. Unclear if results are due to specific RIT techniques or differences within the quality of PBS plans.

Johnson et al, 2013

Elements of risk of bias	Reviewers' judgement	Description or Comments
Sequence generation (allocation bias)	Unclear risk of bias	Participants were randomly assigned to the groups but specific methods employed to achieve randomisation were not specified.
Allocation concealment (selection bias)	Unclear risk of bias	No information provided.
Blinding of personnel and participants (performance bias)	Unclear risk of bias	No information provided.
Blinding of outcome assessors (detection bias) for all outcomes	High risk of bias	Sleep problems were rated by parents who participated in the intervention group.
Incomplete outcome data (attrition bias)	Low risk of bias	Small number of drop outs was reported and its reasons were provided in the paper.
Selective reporting (reporting bias)	Low risk of bias	All outcome data was reported in the paper.
Other potential threats to validity	High risk of bias	Moderate or larger treatment effects might be easier to achieve due to small sample.

Reitzel et al, 2013

Elements of risk of bias	Reviewers' judgement	Description or Comments
Sequence generation (allocation bias)	Low risk of bias	Participants were randomly assigned to treatment groups using a table of random numbers.
Allocation concealment (selection bias)	Low risk of bias	Quote: "Investigators and research staff were blind to participants' group allocation"
Blinding of personnel and participants (performance bias)	Low risk of bias	Quote: "...therapists involved in administering the FBST were blind to the study's objective and were not involved in assessing children's pre- and post-treatment performance"
Blinding of outcome assessors (detection bias) for all outcomes	Low risk of bias	Research team was blind to participants' group allocation.
Incomplete outcome data (attrition bias)	High risk of bias	Numbers of drop-outs were quite large (5 participants in the control group and 6 participants in the treatment group).
Selective reporting (reporting bias)	High risk of bias	Reported that fidelity checklists were administered to the intervention providers but quantitative data were not provided.
Other potential threats to validity	High risk of bias	In the treatment group children had higher levels of ASD symptom severity, cognitive abilities and adaptive functioning than children in the control condition. In the control condition 3 children were also involved in Intensive Behavioural Intervention and this might biased the results.

Singh et al, 2016

Elements of risk of bias	Reviewers' judgement	Description or Comments
Sequence generation (allocation bias)	Low risk of bias	The authors used random number generation to assign participants to the groups.
Allocation concealment (selection bias)	Unclear risk of bias	No information provided.
Blinding of personnel and participants (performance bias)	Unclear risk of bias	No information provided.

Blinding of outcome assessors (detection bias) for all outcomes	Unclear risk of bias	Reported that staff recorded outcome data but it is unclear if they were blinded to the treatment allocation.
Incomplete outcome data (attrition bias)	Low risk of bias	Small number of drop outs was reported and its reasons were provided in the paper.
Selective reporting (reporting bias)	Low risk of bias	We believe that all outcomes were reported.
Other potential threats to validity	High risk of bias	Several outcomes (aggression and physical restraints) were measured looking at records in the databases made by staff members. It is unclear whether staff members were blind to the treatment allocation and what steps were used to ensure that physical restraints were recorded accurately. Small sample size.