Research in Autism Spectrum Disorders 6 (2012) 1004–1018

Contents lists available at SciVerse ScienceDirect

Research in Autism Spectrum Disorders

Journal homepage: http://ees.elsevier.com/RASD/default.asp

Review

Sensory integration therapy for autism spectrum disorders: A systematic review

Russell Lang [a,b,*], Mark O'Reilly [b], Olive Healy [c], Mandy Rispoli [d], Helena Lydon [c], William Streusand [e], Tonya Davis [f], Soyeon Kang [b], Jeff Sigafoos [g], Giulio Lancioni [h], Robert Didden [i], Sanne Giesbers [i]

[a] Texas State University-San Marcos, Clinic for Autism Research Evaluation and Support, San Marcos, TX, United States
[b] The Meadows Center for the Prevention of Educational Risk, University of Texas at Austin, Austin, TX, United States
[c] National University of Ireland, Galway, Ireland
[d] Texas A&M University, College Station, TX, United States
[e] University of Texas at Austin, Austin, TX, United States
[f] Baylor University, Waco, TX, United States
[g] Victoria University of Wellington, Wellington, New Zealand
[h] University of Bari, Bari, Italy
[i] Radboud University Nijmegen, Nijmegen, The Netherlands

ARTICLE INFO

Article history:
Received 27 January 2012
Accepted 30 January 2012

Keywords:
Sensory integration
Autism
Weighted vests
Proprioceptive
Vestibular
Sensory diet

ABSTRACT

Intervention studies involving the use of sensory integration therapy (SIT) were systematically identified and analyzed. Twenty-five studies were described in terms of: (a) participant characteristics, (b) assessments used to identify sensory deficits or behavioral functions, (c) dependent variables, (d) intervention procedures, (e) intervention outcomes, and (f) certainty of evidence. Overall, 3 of the reviewed studies suggested that SIT was effective, 8 studies found mixed results, and 14 studies reported no benefits related to SIT. Many of the reviewed studies, including the 3 studies reporting positive results, had serious methodological flaws. Therefore, the current evidence-base does not support the use of SIT in the education and treatment of children with autism spectrum disorders (ASD). Practitioners and agencies serving children with ASD that endeavor, or are mandated, to use research-based, or scientifically-based, interventions should not use SIT outside of carefully controlled research.

Contents

* Corresponding author at: Texas State University-San Marcos, Department of Curriculum and Instruction, 601 University Dr, San Marcos, TX 78666, United States.
E-mail address: russlang@txstate.edu (R. Lang).

1750-9467/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.rasd.2012.01.006

Autism spectrum disorders (ASD) are characterized by a combination of restrictive and repetitive behaviors and deficits in communication and social skills (American Psychological Association, 2000). Although not part of the diagnostic criteria, individuals with ASD may also appear to seek or avoid ordinary auditory, visual, tactile, and oral stimuli (Ben-Sasson et al., 2009). For example, individuals with ASD may perseverate on objects that have a specific texture or visual pattern, may cover their ears when they hear a specific noise (e.g., car horn), or may not respond to stimuli that should elicit their attention (e.g., someone calling their name). These unusual behaviors are sometimes described as "sensory behaviors" (Ben-Sasson et al., 2009; Kern et al., 2008; Rogers & Ozonoff, 2005; Lane, Young, Baker, & Angley, 2010).

A meta-analysis of 14 studies involving sensory processing symptoms in individuals with ASD suggested that sensory behaviors were common (Ben-Sasson et al., 2009). However, Rogers and Ozonoff (2005) reviewed 48 empirical papers and 27 theoretical or conceptual papers and found that the frequency, severity, and topography of these abnormal sensory behaviors varied greatly across samples of individuals with ASD. Further, Rogers and Ozonoff reported that there was insufficient evidence to suggest sensory behaviors could be used to differentiate ASD from other developmental disabilities.

Despite debate regarding the prevalence of these behaviors, researchers have sought to identify a biological cause for the abnormal behaviors observed in individuals with ASD. One hypothesis is that abnormal behaviors are caused by a defect in the nervous system in which sensory stimuli are processed and integrated abnormally (Ayres, 1972; Ayres & Tickle, 1980; Schaaf & Miller, 2005). Sensory integration therapy (SIT) is an extension of this hypothesis and further speculates that, given the nervous systems ability to change (neuroplasticity), providing specific forms of sensory stimulation in the appropriate dosage may improve the nervous system's ability to process sensory stimuli. Ultimately, the improved nervous system may then result in reductions in problem behaviors and more efficient learning (Baraneck, 2002; Lane et al., 2010; Schaaf & Miller, 2005). However, the exact nature of the nervous system's impairment and the influence of SIT on sensory processing is currently the subject of debate and ongoing research (Iarocci & McDonald, 2006; Lane & Schaaf, 2010; Smith, Mruzek, & Mozingo, 2005).

Implementation of SIT typically involves some combination of the child wearing a weighted vest, being brushed or rubbed with various instruments, riding a scooter board, swinging, sitting on a bouncy ball, being squeezed between exercise pads or pillows, and other similar activities. Ideally, the specific set of activities implemented is based upon an assessment of a child's sensory profile (e.g., Dunn, 1999) and adheres to the essential components of SIT described by Parham et al. (2011). Specifically, SIT should involve: (a) child safety, (b) opportunities to obtain tactile, vestibular, and/or proprioceptive sensory stimulation to support self-regulation, sensory awareness, or movement, (c) appropriate levels of participant alertness, (d) challenge to postural, ocular, oral, or bilateral motor control, (e) novel motor behaviors and efforts to organize movements in time and space, (f) preferences in the choice of activities and materials, (g) activities that are not too easy or too difficult, (h) activities in which the participant experiences success (i) support for intrinsic desire to play, and (j) a therapeutic reliance (Parham et al., 2011).

SIT is among the most common interventions delivered to children with ASD. Watling, Deitz, Kanny, and McLaughlin (1999) surveyed 72 occupational therapists (OT) working with children with autism and found that 99% regularly implemented SIT. Similarly, Case-Smith and Miller (1999) contacted 292 OTs and found SIT to be the most frequent intervention utilized by OTs with children with ASD. Finally, Green et al. (2006) surveyed 552 parents of children with autism and reported that 38.2% of parents said their child currently receives SIT and an additional 33.2% reported that their child has received SIT at some point in the past.

Previous reviews involving individuals with ASD and other diagnoses have arrived at varying conclusions regarding SIT's effectiveness (e.g., Hoehn & Baumeister, 1994; May-Benson & Koomar, 2010; Ottenbacher, 1982; Stephenson & Carter, 2005). Additionally, a recent review focusing only on individuals with ASD has not been conducted. Given discrepancies across previous reviews, the immense popularity and wide spread use of SIT within the ASD population (Green et al., 2006), and the increasing importance of implementing evidence-based practice (e.g., IDEIA, 2004) such a review is warranted.

The purpose of this current review was to systematically identify, analyze, and summarize research involving the use of SIT in the education and treatment of individuals with ASD. Herein we endeavor to determine if SIT can be classified as a research-based or scientifically-based intervention for individuals with ASD. A review of this type may provide useful information to practitioners and agencies interested in providing effective education/rehabilitation to individuals with ASD.

1. Methods

1.1. Search procedures

Systematic searches were conducted in four electronic databases: Medline, Education Resources Information Center (ERIC), Psychology and Behavioral Sciences Collection, and PsycINFO. Searches were limited to peer-reviewed studies

written in English. In all four databases, the terms "sensory" or "sensorimotor" or "weighted vest" or "brushing", or "swinging" or "deep pressure" or "vestibular stimulation" or "proprioceptive stimulation" plus "developmental disabil*" or "autis*" or "Asperger" were inserted as free text into the keywords field in pairs (e.g.; autism plus brushing). The abstracts of the resulting studies were reviewed to identify studies for inclusion (see Section 1.2). The reference lists for studies meeting these criteria were then reviewed to identify additional articles for possible inclusion. Hand searches; covering January to July 2011; were completed for the journals that had published studies included in the review. Searches of databases; journals; and reference lists occurred during June and July 2011.

1.2. Inclusion and exclusion criteria

In order to be included in this review, a study had to meet the following inclusion criteria. First, the study had to contain at least one participant diagnosed with an ASD (i.e., Autistic Disorder, Asperger's syndrome or Pervasive Developmental Disorder, Not Otherwise Specified). Second, the study had to implement some form of SIT in an effort to decrease the symptoms of ASD (e.g., decrease stereotypy, improve communication and/or social skills), improve quality of life, increase access to typical environments (e.g., school or community), and/or improve academics. In order to be considered SIT, interventions had to involve one or more of the following: weighted vests; swinging, brushing; joint compression; and/or alternative seating (e.g., sitting on therapy balls). Interventions described as providing "vestibular" or "proprioceptive' stimulation were only included if the authors described their intervention as "sensory integration". Interventions that claimed to manipulate participants' "sensory diet" were considered to be SIT, even if the exact procedures implemented as part of this "sensory diet" (i.e., a multicomponent SIT package), were not listed. Studies were excluded if SIT was involved, but the variable being evaluated was not SIT. For example, Jung et al. (2006) tested the effects of a virtual reality approach to implementing SIT, but this study was excluded because the experiment was designed to test the virtual reality approach to service delivery and not SIT.

1.3. Data extraction

Each study identified during the systematic search was first assessed for inclusion. Studies selected for inclusion in this review were then summarized in terms of the (a) participant characteristics, (b) assessments used to identify sensory deficits or behavioral functions, (c) dependent variables, (d) intervention procedures, (e) intervention outcomes, and (f) certainty of evidence. Various procedural aspects were also noted, including implementation setting, implementer, social validity, treatment fidelity, and inter-observer agreement (IOA).

Intervention outcomes of SIT were summarized as positive, negative, or mixed (e.g., Lancioni, O'Reilly, & Emerson, 1996; Machalicek et al., 2008). Results were classified as positive in single-case experimental designs if visual analysis of the graphed data suggested improvement in all of the dependent variable(s) for all participants in the study. Results were classified as positive for studies using between-group designs if statistically significant improvement was found for the SIT group on all dependent variables. Results were classified as negative in single-case experimental designs if visual analysis suggested no improvement for any participant on any dependent variable. Results were classified as negative for between-group designs if no statistically significant improvement was found in the SIT group on any dependent variable. Results were classified as mixed in single-case experimental designs if improvement was found in some, but not all, of the participants or dependent variables. Finally, results were classified as mixed for between-group designs if statistically significant improvement was found for some, but not all, dependent variables in the SIT group.

Certainty of evidence is a description of a study's methodological rigor. The ability of a study to provide certainty of evidence was rated as either "suggestive", "preponderant", or "conclusive" (Schlosser, 2009; Simeonsson & Bailey, 1991; Smith, 1981). The lowest level of certainty was suggestive evidence. Studies within this category might have used an AB or intervention-only design, but did not involve a true experimental design (e.g., group design with random assignment, multiple-baseline or an ABAB design). The second level of certainty was preponderance of evidence. Studies within this level had the following four attributes: (a) experimental design, (b) adequate inter-observer agreement (e.g., 20% of sessions with 80% or better agreement), (c) operationally-defined dependent variables, and (d) enough detail to enable replication of intervention procedures. The fifth quality of studies at the preponderant level was that they were in some way limited in their ability to control for alternative explanations for treatment effects. For example, if concurrent interventions (e.g., SIT and psychopharmacological) were targeting the same or related dependent variables and no design feature controlled for the influence of the non-SIT component, then the study was classified at the preponderant level. The final level of certainty was conclusive. Within this level, studies contained all of the attributes of the preponderance level, but also attempted to control for alternative explanations of intervention effects and contained a measure of treatment fidelity. In studies involving simple and obvious intervention procedures (e.g., sitting on a therapy ball or wearing a vest) a measure of treatment fidelity was not required for the study to be classified at the conclusive level of certainty (e.g., Reichow, Barton, Sewell, Good, & Wolery, 2010). However, studies involving more complex multi-component interventions did require a measure of treatment fidelity to be classified as conclusive (e.g., Case-Smith & Bryan, 1999).

This certainty of evidence classification system was applied in an effort to provide an overview of the quality of evidence across the corpus of reviewed studies and to inform the interpretation of an individual study's results (Schlosser & Sigafoos, 2007). For example, a study with positive findings and a conclusive level of certainty provides more evidence in support of

SIT than a study with positive findings classified at the suggestive level of certainty. Because level of certainty is independent of the results of intervention, it is possible for a study to have mixed results and a conclusive level of certainty. The interpretation of such a study should be that the experiment was rigorous but the findings were unclear or did not support the hypothesis that SIT is an effective intervention (e.g., Hodgetts, Magill-Evans, & Misiaszek, 2011a).

1.4. Reliability of search procedures and inter-coder agreement

The first, fourth, and fifth authors of this review independently conducted the database searches and screened the resulting articles for inclusion. These co-authors each produced a list of studies that should be further considered for inclusion. The reliability of the database search was then measured by calculating the percent of articles identified by all three co-authors out of the total number of articles across lists. A combined total of 53 articles were identified at this stage, of which 45 appeared on all three lists (87% initial agreement on the database searches). Using the combined list of 53 articles, a list of the journals that published at least two articles was created. The two most recent issues of those journals were then hand searched for additional studies to be considered for inclusion. Four studies were added to the list following the hand search. Co-authors then obtained complete copies of all 57 studies being considered for inclusion (53 from database plus four from hand searches), and the first, fourth, and fifth authors independently applied the inclusion and exclusion criteria to these 57 studies.

In order to ensure the accuracy of the application of the inclusion and exclusion criteria, the resulting lists of studies to include was compared across co-authors. Agreement as to whether a study should be included or excluded was 86% (i.e., agreement was obtained on 49 of the 57 studies). The disputed articles were then discussed by co-authors until 100% agreement was reached. The result was a list of 22 studies to be included. After this list of 22 included studies was agreed upon, the references of the included studies were searched for other studies that should be considered for inclusion. This ancestry search identified three more studies for inclusion. Agreement on the inclusion of those additional three studies was 100%. Ultimately, 25 studies were included in this review.

After the final list of 25 studies was agreed upon, information from each study was extracted by the first author to develop an initial summary of each study. In order to ensure the accuracy of these summaries, co-authors used a checklist designed to evaluate inter-coder agreement on the extraction of data. The checklist included six questions regarding various details of the study. Specifically: (a) is this an accurate description of the participants? (b) Is this an accurate description of the assessment procedures? (c) Is this an accurate description of the dependent variables? (d) Is this an accurate description of the intervention procedures? (e) Is this an accurate description of the outcomes? And, (f) is this an accurate description of the certainty of evidence? This approach was intended to ensure accuracy in the summary of studies and to provide a measure of inter-coder agreement on data extraction and analysis. There were 150 items on which there could be agreement or disagreement (i.e., 25 studies with six checklist items per study). Initial agreement was obtained on 142 items (95%). When summaries were considered inaccurate, co-authors discussed the study and the summary and then made corrections. This process was repeated until 100% agreement regarding the accuracy of the summaries was achieved. The resulting summaries were then used to create Table 1.

2. Results

The systematic search procedures and application of the predetermined inclusion criteria resulted in the inclusion of 25 studies in this review. Table 1 summarizes: (a) participant characteristics, (b) assessments used to identify sensory deficits or behavioral functions, (c) dependent variables, (d) intervention procedures, (e) intervention outcomes, and (f) certainty of evidence of the 25 included studies.

2.1. Participants

The 25 included studies provided SIT intervention to a total of 217 individuals with an ASD. Of these 217 participants, 176 (85%) were male, 31 (15%) were female and the sex of 10 participants (5%) was not reported. Participants ranged in age from 2 to 12 years ($M = 5.9$ years). The majority of participants ($n = 195$; 90%) were diagnosed with autism. For 140 (72%) of those participants with autism, not enough information was provided in the reviewed studies to determine the presence or absence of intellectual disability (ID). However, for the remaining participants with autism, standardized assessment scores and/or the authors' detailed descriptions of participants were used to classify three participants as having mild ID, 15 with moderate ID, and 37 with severe/profound ID. Twenty-one participants (10%) were diagnosed with PDD-NOS and one with Asperger's syndrome. In addition to ASD, three individuals also had a hearing impairment, one a visual impairment, one with epilepsy, and one with bipolar disorder. Overall, participants involved in these SIT interventions were mainly of elementary school age with autism and a large percentage also had moderate to profound ID.

2.2. Person implementing SIT and intervention settings

Occupational therapists (OTs) were the most common professionals involved in the SIT research. OTs either directly implemented SIT with participants (e.g., Case-Smith & Bryan, 1999; Linderman & Stewart, 1999; Watling & Dietz, 2007),

Table 1
Summary of included studies.

Citation	Participant characteristics	Assessment of behavior	Dependent variables	Intervention	Results and certainty of evidence
Ayres and Tickle (1980)	9 males and 1 female; 2 mild AU[a], 5 moderate AU, and 3 severe AU; 2 were also hearing impaired and one of those was also visually impaired; ages 3.5–13 years old (M = 7.4 years)	A sensory assessment designed to measure reaction to sensory stimuli was created. Reactions ranged from: 1 = "no reaction or definite under-reaction" to 5 = "definitive overreaction". Fourteen forms of sensory stimuli were then assessed	Dependent variables included language, awareness of the environment, engagement in purposeful activity, self-stimulation, and social-emotional behavior. These variables were "judged qualitatively and differently for each subject"	Participants received SIT from an expert twice per week for 1 year (1 participant received 11 months). SIT was individualized for each child and specific intervention procedures were NR[b]	Results: mixed. Based upon outcomes children were classified as "good responders" (n = 6) and "poor responders" (n = 4). Participants with normal or over-reaction responses in the areas of tactile defensiveness, reaction to movement, gravitational insecurity, and reaction to air puff stimuli were more likely to be "good responders". Certainty: Suggestive, due to insufficient operational definitions of target behaviors, insufficient detail to enable replication, and non-experimental design. IOA[c] and TF[d] were NR
Bagatell et al. (2010)	6 males; all with moderate to severe AU; ages NR but all children were in kindergarten to 1st grade	SPM: main classroom form categorizes children's sensory processing as "typical range", "some problems", and "definite dysfunction"	Total duration of time out of seat and/or not attending to teacher or task	Participants sat on therapy ball chairs during class circle time. The inflated therapy ball chairs individualized so that feet rested flat on ground and hips and knees were at a 90° angle. After 9 days, participants were given the choice of sitting in therapy ball chairs or regular chairs	Results: mixed. One student stayed in seat longer, one student was out of his seat more, and the out of seat behavior of 3 of the students did not change (data on out of seat behavior for 1 student was NR). Children spent less time attending to the teacher, less time on task, or showed no change from baseline when using the therapy ball. Certainty: suggestive, due to use of a non-experimental ABC design in which "A" represented baseline, "B" ball chairs, and "C" choice between seats. The ABC design was not embedded within a multiple baseline. TF was not NR
Bonggat and Hall (2010)	1 male with AU; 4 years old; 2 other participants without an ASD diagnosis were in the study but are omitted from this analysis	NR	Direct observation of attention to task and disruptive behavior	Sensory diet consisting of brushing, joint compression, therapy ball, hammock activities, and stretching was provided for 10 min in the morning	Results: negative. No improvement in attention or disruptive. behavior Certainty: suggestive, due to insufficient detail to enable replication of procedures. An alternating treatment design compared SIT and a control condition in which attention was given during non-SIT activities (e.g., reading and board games). TF was NR
Carter (2005)	1 male with profound AU and frequent severe sinus infections; 4 years old	Analogue functional analysis revealed an automatic reinforcement function	Direct observation of self injurious behavior	Wearing a weighted vest	Results: negative. The weighted vest had no effect on self-injury when the child did not have a sinus infection and caused increased levels of self-injury when the child did have a sinus infection. Certainty: conclusive, the effect of the weighted vest was evaluated in an ABABAB design. The presence and absence of a sinus infection was examined as an alternative explanation within an ABA in which A represented illness
Case-Smith and Bryan (1999)	5 males; all with AU, 1 with a bilateral hearing impairment and 1 with bipolar disorder; 4 to 5 years old	NR	Engagement was measured using Engagement check (Parson et al., 1989). The engagement check measures (a) mastery and non-mastery of play, (b) non-engaged behaviors, and (c) social interactions	10 weeks of SIT was provided 30 min daily by an OT[e]. Swings, brushing, bean bag chairs, rocking equipment, and water/sand tables were used	Results: mixed. Three out of 5 participants increased in mastery of play, 4 participants reduced non-engaged behaviors, 1 improved in interactions with adults, and none improved in interaction with peers. Certainty: suggestive, due to non-experimental AB designs. The authors report using a multiple baseline design. However, baseline duration was not staggered across participants. Insufficient details regarding how SIT was individualized prevent replication. Dependent variables were not operationally defined. TF was NR

Cox et al. (2009)	2 males and 1 female; all with AU; 5, 6, and 9 years old	Short sensory profile identified sensory processing deficits in all 3 participants	Direct observation of in-seat behavior	Wearing a weighted vest	Results: negative. The weighted vests had no effect on the amount of time any of the participants were in their seat. Certainty: conclusive. An alternating treatment design was used to compare weighted vests, vest without weights, and no vests. The vests without weight acted as a placebo control. TF and IOA were measured
Davis et al. (2011)	1 male with AU; 4 years old	Analogue functional analysis revealed stereotypy was maintained by automatic reinforcement	Direct observation of stereotypy	Arms, hands, back, legs, and feet were brushed with a soft surgical brush 7 times a day for 5 weeks. He was brushed with long strokes until entire exposed skin surfaced was brushed 3–10 times	Results: negative. Brushing did not result in a decrease in stereotypy. Certainty: suggestive, due to a non-experimental ABA design. A second "B" phase was planned, but the participant's parents would not allow the brushing procedure to be continued because they deemed it to be ineffective. TF was NR
Devlin et al. (2011)	4 males with AU; 1 also with 1 epilepsy; 6 to 11 years old ($M = 9$)	Questions About Behavioral Function, Functional Analysis Screening Tool, and/or analogue functional analysis revealed escape and access to tangible functions for 3 participants and an escape function for 1 participant	Direct observation of challenging behavior including self-injurious behavior	A sensory diet developed by an OT that involved swinging, jumping, rocking on a therapy ball, wrapping in blanket, crawling, joint compression, squeezing with bean bags, chewing a tube, and brushing was implemented for 15 min prior to desktop work 6 times per day or contingent upon challenging behavior	Results: negative. Sensory diet had no effect on challenging behavior and may have increased challenging behavior in 1 participant. Behavioral intervention reduced challenging behavior for all 4 participants. Certainty: preponderance, due to lack TF. Alternating treatment design compared function-based behavioral intervention to SIT
Devlin et al. (2009)	1 male with AU; 10 years old	Analogue functional analysis revealed an escape function for self-injurious behavior	Direct observation of self-injurious behavior	A sensory diet involving, swinging, beanbag compression, rocking, jumping, crawling, rolling in a blanket, chew tube, brushing, and joint compression was implemented 4 times per day for 30 min/min per session or contingent upon the occurrence of self-injury	Results: negative. The sensory diet had no effect on self-injury, but a subsequent function-based behavioral intervention reduced self-injury. Certainty: suggestive, due to insufficient detail to replicate sensory diet and absence of TF. An alternating treatment design was used to compare the effects of SIT and behavioral intervention
Fazlioglu and Baran (2008)	30 children with AU were randomly assigned to treatment or control. Treatment group contained 12 males and 3 females; all with severe AU; 7 to 11 years old	Sensory evaluation form for children	A checklist created by the researchers called the sensory evaluation form for children with autism was used to determine the severity of sensory processing abnormalities	A sensory diet consisting of brushing and joint compression followed by a set of activities designed to meet the child's sensory needs and integrated into the child's daily routine. Concurrently, prompting, reinforcement and extinction were used to teach specific target motor behaviors related to the sensory evaluation form for children	Results: positive. There was a statistically significant main effect for treatment group in total scores ($F = 5.84$, $p < .05$) as well as a main effect of test time (pre- and post-test) ($F = 98.38$, $p < .01$). The interaction of group and time was also significant ($F = 119.38$, $p < 01$). Certainty: suggestive, due to the simultaneous implementation of research-based behavioral intervention components that directly targeted skills related to the dependent variable, insufficient detail to enable replication of intervention procedures, and TF was NR

Table 1 (*Continued*)

Citation	Participant characteristics	Assessment of behavior	Dependent variables	Intervention	Results and certainty of evidence
Fertel-Daly et al. (2001)	3 males and 2 females; 4 with PDD-NOS and 1 with AU; 2 to 3 years old (M = 2.8 years old)	NR	Direct observation of time on task, number of distractions, and self stimulatory behaviors	Wearing a weighted vest	Results: mixed. All 5 children improved in time on task and number of distractions. Four of 5 of the children's self-stimulation decreased. However, self-stimulation did not return to baseline when intervention was removed for 2 children. One child's self-stimulation increased during intervention. Certainty: suggestive, due to a lack of experimental design and insufficient IOA (IOA was only collected during baseline). Additionally, all children were concurrently receiving behavioral intervention and many of the changes in dependent variables did not revert to baseline levels following removal of the vest, suggesting that the concurrent intervention may have contributed to some or all of the observed improvements
Hodgetts et al. (2011a)	5 males and 1 female; all with severe AU; 4–10 years old (M = 6.7 years)	Short sensory profile indicated that all participants were 2 standard deviations below mean for typical sensory processing	Stereotypy was measured via coding videotapes of sessions and heart rate was measured via a heart rate monitor	Wearing a vest styrofoam balls in place of weights and weighted vests with 5–10% of child's body weight	Results: negative. Weighted vest did not decrease motor stereotypy or heart rate for any participant. A small effect on verbal stereotypy was recorded in one child. Certainty: conclusive. An ABCBC design in which "A" represented no vest, "B" represented vests with Styrofoam to prevent rater bias and "C" represented vests with 5–10% of body weight. The sequence of conditions was counter balanced across participants. Data collectors were blind to condition. IOA and TF were measured
Hodgetts et al. (2011b)	8 males and 2 females; all with moderate to severe AU; 3 to 10 years old (M = 5.9 years)	Short sensory profile indicated that all participants were 2 standard deviations below mean for typical sensory processing	Duration of off-task behavior and time in seat was measured via direct observation. Teacher ratings of participant restlessness, impulsivity, and emotional liability were measured using the 10-item Conner's Global Index-Teacher (CGI-T; Conners, 1997)	Wearing a vest with styrofoam balls in place of weights and weighted vests with 5–10% of child's body weight	Results: mixed. Time in seat was measured for 3 participants and no effect was found. Vests decreased off-task behavior for 3 participants, had no effect for 5 participants, and 2 of the participants' data was uninterruptible due to illness and high levels of variability within phases. Results from CGI-T did not correspond with direct observation data. The CGI-T indicated improvements during 45% of the weighted vest conditions but did not correspond with the direct observation data. For one participant CGI-T scores corresponded with his actual behavior across all conditions. Certainty: conclusive. An ABCBC design in which "A" represented no vest, "B" represented vests with styrofoam to prevent rater bias and "C" represented vests with 5–10% of body weight. The sequence of conditions was counter balanced across participants. Data collectors were blind to condition. IOA and TF were measured
Kane et al. (2004)	2 males and 2 females; 3 with AU and 1 with PDD-NOS; 8 to 11 years old (M = 9 years)	NR	Direct observation of stereotypy and attention to task	Wearing a weighted vest	Results: negative. Weighted vest had no effect on stereotypy or attention. Certainty: Suggestive, due to a non-experimental ABC design in which baseline duration was not staggered across participants and "A" represented no west and "B" or "C" represented wearing a vest with and without weights. IOA was NR

Leew et al. (2010)	4 males; all with AU; 27 to 33 months (M = 30.5 months)	The infant/toddler sensory profile was given to assess the degree to which sensory processing issues affected infants daily life	Joint attention and behaviors that compete with joint attention were directly observed and The Parenting Morale Index (PMI) measured parenting morale	Wearing a weighted vest	Results: negative. Weighted vests did not increase joint attention or decrease competing behaviors. One mother's PMI increased but the other three did not. Certainty: Suggestive, although a multiple baseline across participants was used, there were insufficient operational definitions of target behaviors. TF was NR
Linderman and Stewart (1999)	2 males, 1 with mild AU and 1 with severe AU, both 3 years old	NR	A modified version of the Functional Behavior Assessment for Children with Integrative Dysfunction rated social interaction skills, approach to new activities, and response to hugging (participant 1) and social interaction skills, functional communication, and response to movement (participant 2) on a 10 point scale	Participant 1 received 1 h of SIT per week for 11 weeks and participant 2 received 1 h/week for 7 weeks. SIT included large pillows, jumping, a trapeze bar swing, "body socks", a bounce pad, and textured toys. Sessions were child-lead	Results: positive. Both participants made substantial gains on all dependent variables. Certainty: Suggestive, due to non-experimental AB designs, concurrent interventions, insufficient operational definitions, insufficient detail regarding intervention procedures, low IOA and TF was NR
Pfeiffer et al. (2011)	32 males and 5 females; 21 with AU and 16 with PDD-NOS; all with an additional diagnosis of sensory processing disorder; 6–12 years (M = 8.8 years); stratified random assignment to fine motor control group or SIT group	SPM was used to identify and described sensory processing deficits	(a) The Sensory Processing Measure (SPM; Parham & Ecker, 2007) is a 4-point Likert scale that assesses processing issues, praxis, and social participation. (b) The Social Responsiveness Scale (SRS; Constantino & Gruber, 2005) is a 65 item rating scale that measures social impairments, awareness, information processing, and communication, anxiety, and autism traits. The Goal Attainment Scale (GAS: Mailloux et al., 2007) was used to measure progress on individualize goals. The Quick Neurological Screening Test 2nd Edition (QNST-II: Mutti, Martin, Sterling, & Spalding, 1998) identifies possible neurological interference with learning	SIT involved 18 sessions 45 min each over a 6-week period. SIT included the 10 key therapeutic strategies identified by Parham et al. (2007)	Results: mixed. SIT group displayed significantly fewer autistic mannerisms than the fine motor group as measured by one subscale of the SRS. No differences between groups were found on the GAS, QNST-II, SPM, or other SRS subscales. Both groups made significant improvement on the GAS. Certainty: Suggestive. The research design was capable of providing a conclusive level of certainty. Specifically, random assignment to groups, blinding, and strong TF. However, there was insufficient detail regarding the SIT procedures to enable replication
Piravej et al. (2009)	60 children with autism were randomly assigned to a SIT group or a SIT + massage group. Across groups, 49 males and 11 females; 3–10 years old (M = 4.66 years old)	NR	Conners' Parent Rating Scale (CPRS) and Conners' Teacher Rating Scale (CTRS) (Conners, 1989) were used to measure conduct problems, learning problems, hyperactivity, anxiety, psychosomatic issues, and inattention. The parents also kept a sleep diary (SD)	Both groups received SIT by an OT twice per week, 1 h per session for 16 sessions. SIT involved "individualized therapeutic environments" and 10 key therapeutic strategies identified by Parham et al. (2007). In the SIT + massage group the child was instructed to lie down facing upward while the masseuse applied pressure to the soles of the feet for a few minutes and then rubbed the foot, leg, thigh, waist, hand, arm, shoulder. This process was repeated as the child changed positions	Results: mixed. Both groups showed significant improvement on the CTRS and sleep behavior. On the CPRS, only anxiety was reported to have improved. Significant improvements were not reported in conduct, learning, psychosomatic, impulsivity, and hyperactivity. Certainty: suggestive. Pre- and posttest data were analyzed using the Wilcoxon signed-rank test. Group comparisons of the pre- posttest difference scores were calculated by the Mann–Whitney U test. Parents were not blind to group assignment. TF was NR. Although substantial detail was provided regarding the massage component, insufficient detail was provided regarding SIT

Table 1 (*Continued*)

Citation	Participant characteristics	Assessment of behavior	Dependent variables	Intervention	Results and certainty of evidence
Quigley et al. (2011)	3 males; 1 with AS and 2 with AU; 4, 6, and 12 years old	Functional Behavior Assessment Interview and an analogue functional analysis revealed that the problem behavior of all three participants was maintained by escape from demands and, in one participant, also by access to tangibles	Direct observation of problem behavior and making a choice making between work and break	Wearing a weighted vest	Results: negative. The vests did not result in reductions in problem behavior for any of the children. Follow-up behavioral intervention (FCT[f]) did reduce problem behavior. Certainty: conclusive, a multielement treatment design embedded within a reversal design with three phases: (a) no vest, (b) unweight vest (control), and (c) vest with 5% to 10% of body weight. This was followed by a function-based intervention that demonstrated the alternative behavioral explanation for the behavior was potentially accurate
Ray et al. (1988)	1 male with severe AU; 9 years old	NR	Percentage of time spent producing vocalizations and verbalizations combined	Vestibular stimulation using a Southpaw Model PS-1800 platform swing with a bouncer attachment was provided daily for 17 days, 5 min per session. The child actively engaged with the swing by pushing his feet against the floor	Results: negative. Child made more noises while swinging, but post swinging noises were the same as pre-swing. The increase in noises made while swinging decreased across the four weeks. Certainty: suggestive, due to a lack of an experimental design and no IOA
Reichow et al. (2010)	2 males; both with AU and 5 years old. A third participant was included in the study but not in this analysis because he did not have an ASD	NR	Direct observation of engagement, stereotypy, and problem behavior	Wearing a weighted vest	Results: negative. The vest did not influence any of the dependent variables. Certainty: conclusive. An alternating treatment design compared (a) weighted vests, (b) vest with no weight (control for alternative explanations), and (c) no vest
Reilly et al. (1983)	15 males and 3 females; all with AU; 6.2 to 11.7 years old (M = 8.2 years old)	NR	The Autism Screening Instrument for Educational Planning (ASIEP; Krug et al., 1980) was used to measure variety, function, articulation, length, autistic speech, total language raw score, and rate of vocalizations	Two 30 min sessions of SIT compared to two 30 min sessions of table top fine motor activities. SIT involved activities that emphasized vestibular and proprioceptive input (e.g., straddling and swinging on the bolster swing, swinging in a net swing, swinging on an inner tube, and bouncing on an inner tube). Fine motor activities included non-SIT activities (e.g., puzzles and coloring)	Results: negative. The fine motor activities (not SIT) resulted in a more variety of speech, greater average length of utterances, and less autistic speech. No significant differences were found for function of speech, articulation, total language raw scores, or rate of vocalizations. Certainty: suggestive. Two-tailed t-tests for related measures were performed on each of the dependent variables. TF was NR. Insufficient detail to enable replication of either SIT or fine motor activity
Thompson (2011)	10 participants with AU were among the larger group of 50 participants with other disabilities. The results for the autism group were disaggregated	NR	Sustained focus was measured using an observation system created by the authors and evaluated in this study	A multi-sensory environment that included the 10 key therapeutic strategies identified by Parham et al. (2007)	Results: positive. The group with AU was found to have significantly higher levels of sustained focus during and after the multi-sensory room experience. Certainty: suggestive, due to a lack of information regarding statistical results. All participants were in one group and data was collected before, after, and during the multi-sensory environment. Data was analyzed using repeated measures ANOVA. Data collectors were not blind

Van Rie and Heflin (2009)	4 males; all with moderate to severe AU; 6–7 years old	NR	Direct observation of responses on academic tasks	Sensory activities included slow linear swings and bouncing on a therapy ball	Results: mixed. For one child there was no difference between conditions, for one child bouncing on the ball was associated with highest percent correct responses on academic task, and for two children the swing was associated with highest percent correct responses on academic tasks. Certainty: conclusive. An alternating treatment design compared SIT to a control condition matched for amount of attention given to the children (e.g., looking at a book with 1 to 1 attention). TF and IOA were measured
Watling and Dietz (2007)	4 males; all with AU; 3–4 years old	Sensory profile: infant/toddler or child version was used to identified and describe sensory processing deficits	Direct observation of undesired behaviors that interfered with task engagement and direct observation of engagement in play or purposeful activities	SIT based on results from sensory assessment. SIT involved "clinical reasoning" Therapist continual observed child and made modifications to SIT	Results: negative. No improvement in engagement or undesired behaviors for any participant. Certainty: suggestive, due to insufficient detail to enable replication of SIT procedures. TF and IOA were measured

[a] Autism.
[b] Not reported.
[c] Interobserver agreement.
[d] Treatment fidelity.
[e] Occupaional therapist.
[f] Functional communication training.

supervised the delivery of SIT (e.g., Pfeiffer, Koenig, Kinnealey, Sheppard, & Henderson, 2011), or recommended the specific SIT procedures (e.g., weighted vests and brushing) that were evaluated (e.g., Davis, Durand, & Chan, 2011; Devlin, Healy, Leader, & Hughes, 2011; Kane, Luiselli, Dearborn, & Young, 2004; Reichow et al., 2010). In the remainder of the studies SIT was implemented by specially trained therapists, classroom teachers or teacher assistants, or researchers who were not also OTs (e.g., Devlin, Leader, & Healy, 2009).

In 13 studies (52%) SIT was implemented in the participants' typical classroom. Of those 13 studies, four were self-contained special education classrooms (e.g., Hodgetts, Magill-Evans, & Misiaszek, 2011b). Three studies implemented SIT in clinical therapy rooms (e.g., Piravej, Tangtrongchitr, Chandarasiri, Paothong, & Sukprasong, 2009), two studies involved a room designed specifically for the delivery of SIT (Fazlioglu & Baran, 2008; Thompson, 2011), two studies were conducted in the children's homes (Davis et al., 2011; Linderman & Stewart, 1999), one in a summer camp (Pfeiffer et al., 2011), and one in an early childcare center (Reichow et al., 2010). The implementation setting was not reported in three studies.

2.3. Assessment of sensory processing

Seven studies attempted to confirm the presence of a sensory processing issue prior to implementing SIT. The most common assessment used for this purpose was the Short Sensory Profile (SSP; Dunn, 1999). The SSP is a standardized assessment intended for children 3–10 years of age. The scale consists of 38 items that are completed by a primary caregiver. Caregivers are asked to rate how their child responds to various sensory stimuli on a 5-point Likert scale. Scores are provided in seven categories intended to identify how a child's nervous system regulates and processes sensory input. The categories include: tactile sensitivity, taste/smell sensitivity, movement sensitivity, under responsive and seeks sensation, auditory filtering, low energy, and visual and/or auditory sensitivity. The SSP was used in three studies (Cox, Gast, Luscre, & Ayres, 2009; Hodgetts et al., 2011a, 2011b).

The Sensory Processing Measure (SPM: Glennon, Miller-Kuhaneck, Henry, Parham, & Ecker, 2007) was used to identify and describe sensory processing issues in two studies (Bagatell, Mirigliani, Patterson, Reyes, & Test, 2010; Pfeiffer et al., 2011). The SPM asks primary caregivers to complete 75 items (classroom form is 62-items completed by teachers) and generates eight standard scores that describe: social participation, vision, hearing, touch, body awareness, balance and motion, planning, and total sensory system. Ultimately, children's sensory processing is classified as "typical", "some problems" or "definite problems".

The Infant Toddler Sensory Profile (Dunn, 2002) was used to identify and describe sensory processing issues in two studies (Leew, Stein, & Gibbard, 2010; Watling & Dietz, 2007). This assessment is intended to measure sensory processing in children between birth and three years old. A primary caregiver rates a number of items on a 5-point Likert scale (number of items depends on age of child). Scores classify sensory issues as sensory seeking, low registration, sensory sensitivity, and sensory avoiding. Depending on the age of the child, sensory issues are summarized overall as "typical performance", "probable difference", or "definitive difference". Children birth to 6 months old can only be classified as typical or referred for evaluation later in life.

2.4. Assessment of behavioral functions

An analogue functional analysis (Hanley, Iwata, & McCord, 2003; Iwata, Dorsey, Slifer, Bauman, & Richman, 1994) was implemented in five studies to identify environmental factors that maintained participants' problem behavior (Carter, 2005; Davis et al., 2011; Devlin et al., 2009, 2011; Quigley, Peterson, Frieder, & Peterson, 2011). Carter (2005) and Davis et al. (2011) identified automatic reinforcement contingencies to be maintaining problem behavior. Devlin et al. (2009) identified an escape function for problem behavior, Devlin et al. identified an escape and tangible function for three participants and an escape function for one participant, and Quigley et al. (2011) identified an escape function for three children and the dual functions of escape and access to tangibles in one child. Devlin et al. (2011) also implemented the Questions About Behavioral Function (Matson & Vollmer, 1995) and the Functional Analysis Screening Tool Revised (Iwata & Deleon, 1996) to identify behavioral functions of challenging behavior.

2.5. Dependent variables

Across studies a variety of dependent variables were measured. Six studies evaluated the effects of SIT on behaviors thought to be self-stimulatory and/or stereotypic (Ayres & Tickle, 1980; Davis et al., 2011; Fertel-Daly, Bedell, & Hinojosa, 2001; Hodgetts et al., 2011a; Kane et al., 2004; Reichow et al., 2010). Specifically, Davis et al. evaluated the effects of a brushing procedure on the occurrence of hand flapping, finger flicking, and body rocking in one child with autism. Four studies examined the effects of SIT on communication and language skills (Linderman & Stewart, 1999; Pfeiffer et al., 2011; Ray, King, & Grandin, 1988; Reilly, Nelson, & Bundy, 1983). Specifically, Ray et al. measured the percentage of time a child with autism produced vocalizations and/or verbalizations, and Reilly et al. used the Autism Screening Instrument for Educational Planning (ASIEP: Krug, Arick, & Almond, 1980) to measure the various aspects of language use (e.g., articulation, length of utterance, rate of vocalization, and occurrence of autistic speech) in 18 individuals with autism. Finally, four studies evaluated SIT's potential benefit on social and/or emotional skills (Ayres & Tickle, 1980; Hodgetts et al., 2011b; Linderman & Stewart, 1999; Pfeiffer et al., 2011). For example, Linderman and Stewart modified the

Functional Behavior Assessment for Children with Integrative Dysfunction to rate social interaction skills of two children with autism.

A variety of additional skill deficits associated with ASD also served as dependent variables in these studies. For example, 13 studies measured engagement, focus, and/or attention (e.g., Bonggat & Hall, 2010; Case-Smith & Bryan, 1999). Eight studies measured problem behavior (e.g., Carter, 2005; Devlin et al., 2011; Quigley et al., 2011). Three studies measured variables related to sensory processing (Fazlioglu & Baran, 2008; Linderman & Stewart, 1999; Pfeiffer et al., 2011). Three studies measured how often the participants were out of their seat during classroom instruction (Bagatell et al., 2010; Cox et al., 2009; Hodgetts et al., 2011b). Three studies measured issues related to learning and/or academic behavior (Pfeiffer et al., 2011; Piravej et al., 2009; Van Rie & Heflin, 2009). Two studies measured the participants' awareness of their environment (Ayres & Tickle, 1980; Pfeiffer et al., 2011). And, one study each contained dependent variables related to joint attention (Leew et al., 2010), heart rate (Hodgetts et al., 2011a), stress (Devlin et al., 2011), sleep, hyperactivity and anxiety (Piravej et al., 2009).

2.6. Intervention procedures

Ten different activities designed to provide a variety of different types of sensory stimulation were investigated across studies. Specifically, in 10 studies the intervention involved the provision of weighted vests (e.g., Carter, 2005; Cox et al., 2009). Eight studies provided swinging or rocking stimulation (Bonggat & Hall, 2010; Case-Smith & Bryan, 1999; Devlin et al., 2009, 2011; Linderman & Stewart, 1999; Ray et al., 1988; Reilly et al., 1983; Van Rie & Heflin, 2009). Five studies involved brushing the child with a bristle or feather instrument (Bonggat & Hall, 2010; Davis et al., 2011; Devlin et al., 2009, 2011; Fazlioglu & Baran, 2008). Five studies involved activities designed to provide joint compression or stretching (Bonggat & Hall, 2010; Devlin et al., 2009, 2011; Fazlioglu & Baran, 2008; Linderman & Stewart, 1999). Three studies involved some form of alternative seating including use of bean bag chairs, therapy ball chairs, and hammocks (Bagatell et al., 2010; Bonggat & Hall, 2010; Case-Smith & Bryan, 1999). Four studies required the participant to jump or bounce (Devlin et al., 2009, 2011; Linderman & Stewart, 1999; Reilly et al., 1983). Three studies involved rolling the child in a blanket or putting them in a "body sock" (Devlin et al., 2009, 2011; Linderman & Stewart, 1999). Finally, playing with a water and sand sensory table (Case-Smith & Bryan, 1999), chewing on a rubber tube (Devlin et al., 2009, 2011), and playing with specially textured toys (Linderman & Stewart, 1999; Devlin et al., 2011) were evaluated in one to two studies each. Thirteen studies evaluated a combination of these intervention components simultaneously and 12 studies evaluated only a single SIT procedure in isolation (e.g., only brushing). In five of the studies there was insufficient information provided to identify the specific combination of procedures used (Ayres & Tickle, 1980; Pfeiffer et al., 2011; Piravej et al., 2009; Thompson, 2011; Watling & Dietz, 2007).

2.7. Outcomes and certainty of evidence

The results of 14 studies (56%) were classified as negative because no benefit to any participant on any dependent measure was found. Of those 14 studies, 4 suggested that SIT may have contributed to increases in stereotypy and problem behavior (Carter, 2005; Davis et al., 2011; Devlin et al., 2011; Kane et al., 2004). Across the studies reporting negative findings, eight were rated as providing a suggestive level of certainty (e.g., Watling & Dietz, 2007), one was rated at the preponderance level (Devlin et al., 2011) and five were rated as providing a conclusive level of certainty. All five studies with a conclusive level of certainty and negative findings involved wearing a weighted vest. The results of eight studies were classified as mixed because some but not all participants improved or some but not all dependent variables improved. For example, Ayres and Tickle (1980) classified six participants as "good responders" to SIT and four as "poor responders". Across the studies with mixed results, six were classified at the suggestive level of certainty and two were classified at the conclusive level of certainty (Hodgetts et al., 2011b; Van Rie & Heflin, 2009). The results of three studies were classified as positive all with a suggestive level of certainty (Fazlioglu & Baran, 2008; Linderman & Stewart, 1999; Thompson, 2011).

3. Discussion

The results from three of the 25 reviewed studies suggested that SIT was effective. In contrast, 8 studies reported mixed results, and 14 found no benefit following SIT. Chambless and Hollon (1998) offer criteria for identifying empirically supported interventions when some studies suggest an intervention is effective and other studies do not. Specifically, the relative methodological rigor of the conflicting research must be examined to determine if the more rigorous studies tend to suggest one conclusion over another. Of the three studies that reported positive findings (i.e., Fazlioglu & Baran, 2008; Linderman & Stewart, 1999; Thompson, 2011) all three were classified at the lowest level of certainty due to serious methodological limitations. The methodological problems of these three studies are outlined below.

Concurrently with SIT, Fazlioglu and Baran (2008) implemented a research-based behavioral intervention that targeted skills directly related to the dependent variables. Specifically, difficult tasks were broken down into smaller steps, tangible reinforcers were provided contingent upon successful participation, and a variety of prompts (i.e., verbal, model, physical, and gestural) were used and then gradually faded. These procedures are consistent with the principles of Applied Behavior Analysis (ABA) and thus the positive results in this study could be interpreted as providing support for ABA-based

procedures, rather than SIT. However, any interpretation of these results may be spurious given the methodological limitations.

In addition to using pre-experimental AB designs, Linderman and Stewart (1999) reported that both participants began receiving speech therapy after baseline. This additional therapy may have influenced the results given that one of their dependent variables was frequency of social initiations. Furthermore, the study is limited in that the specific procedures of the SIT intervention were not described in replicable terms. For example, the researchers stated that a wide array of materials and activities were selected and that the specific activity and duration was determined by the individual sensory needs of each participant, but the details of that assessment and how, specifically, the assessment results guided intervention was not reported.

Thompson (2011) provided what is perhaps the best evidence in support of SIT among these 25 reviewed studies. Unfortunately, the results from that study are also difficult to interpret because, while a repeated-measures ANOVA was conducted, an ANOVA results table was not provided. Instead, graphs of aggregate data and written descriptions regarding the results were provided. Without the tables, readers cannot determine if there was a statistically significant difference and must rely on the author's written descriptions. However, those written descriptions did not include sufficient information to interpret the results. For example, the author reported that mean self-injury behavior decreased by over 90%, but it remains unclear if that was a statistically significant difference because the corresponding *p*-value was not provided.

Out of the eight studies with mixed results, five were classified as mixed because some participants made gains and others did not. Therefore, in terms of those five studies, there is a failure to replicate results across participants even within the same study. For the three studies with mixed results in which all of the participants made improvement on at least one dependent variable, the variable(s) that improved were different across studies and/or participants in the same study. Therefore, failure to replicate results in terms of participants and/or dependent variables is found both within and across studies with mixed results. Hypotheses regarding why some participants and/or dependent variables improved and others did not were not stated a priori or directly tested. This lack of replication and post hoc explanations for discrepancies across variables and/or participants hinders the interpretation of this group of studies.

Across the 14 studies with negative results, five were classified at the conclusive level and one at the preponderance level of certainty. However, in the five studies that were judged to provide conclusive evidence, the SIT intervention consisted only of the use of a weighted vest, and in the remaining study (Devlin et al., 2011) the fidelity of implementation of the multicomponent sensory diet has been debated (Healy, Hughes, Leader, & Devlin, 2011; Schaaf & Blanche, 2011). Although the findings of this review do support Stephenson and Carter's (2005) review in which weighted vests were found to have no benefit for children with ASD, the immense differences in SIT procedures across studies and the lack of a treatment fidelity measure in the majority of studies prevents direct comparison of the studies with positive and negative findings.

When comparisons of certainty of evidence are insufficient to settle debates regarding discrepant findings across studies, Chambless and Hollon (1998) recommend in favor of the conservative conclusion; specifically, that the intervention in question should not be considered to be established as effective or even as possibly efficacious. However, several methods for evaluating a study's methodological rigor other than the certainty of evidence method used here exist. It is possible that the application of some other coding system may have yielded different conclusions. Similarly, criteria other than those provided by Chambless and Hollon (1998) are available for identifying evidence-based interventions. Further, it has been argued that all the existing criteria used to determine if an intervention is evidenced-based may be inappropriate when applied to ASD interventions (e.g., Mesibov & Shea, 2011; Reichow, Volkmar, & Cicchetti, 2008). Nevertheless, it is important to note that the summaries and analyses provided in this review indicate that SIT does not qualify as an evidence-based, or scientifically-based, intervention even when other common standards are used (e.g., Horner et al., 2005; Odom, Collet-Klingenberg, Rogers, & Hatton, 2010; Reichow et al., 2008). This review supports the omission of SIT from several recent peer-reviewed lists of evidenced-based practices for children with ASD (e.g., Mayton, Wheeler, Menendez, & Zhang, 2010; National Autism Center's National Standards Project, 2009; Rogers & Vismara, 2008).

Given the lack of scientific evidence, it would seem alarming how often SIT is reported delivered to individuals with ASD (Case-Smith & Miller, 1999; Green et al., 2006; Watling et al., 1999) by agencies that are mandated to use evidence-based interventions. For example, The Individuals with Disabilities Education Improvement Act (IDEIA, 2004) and No Child Left Behind Act (NCLB, 2001) require that schools implement evidenced-based interventions. Therefore, outside of a research context with approvals from relevant ethics committees and informed consent, our review suggests that SIT should not be commonly implemented within public schools that receive federal funding in the United States. However, 82% of the 292 OTs interviewed by Case-Smith and Miller reported they always use SIT with children with ASD and 66% of those OTs were employed by schools. Another 18% implemented SIT with children with ASD in early intervention programs that may also receive federal funding or reimbursements from insurance companies. This discrepancy between research findings, legal requirements, and actual practice is made more troublesome by the possibility that SIT may actually exacerbate behavior problems in some children (Carter, 2005; Davis et al., 2011; Devlin et al., 2011; Mason & Iwata, 1990).

Some researchers have argued that SIT may inadvertently cause an increase in problem behavior because SIT often provides access to enjoyable activities, attention from therapists, and breaks from work contingent upon the occurrence of problem behavior. This practice may inadvertently reinforce or strengthen abnormal behavior in the long term even when immediate reductions in problem behavior are observed. Even when SIT is delivered at set times during the child's day and contingent implementation is avoided, SIT may still undermine the effectiveness of concurrent research-based behavioral interventions by satiating the child on potential reinforcers or blurring the carefully programmed contingencies designed to

promote appropriate behavior. Unfortunately, only five of the 16 studies that addressed some form of problem behavior (e.g., stereotypy, self-injury, and off-task behavior) conducted any type of functional assessment prior to designing and implementing SIT.

The results of this systematic review were that SIT had no consistently positive effect as a treatment for children with ASD. These findings are in agreement with previous reviews of SIT involving individuals with ASD and/or other populations said to have "sensory integrative dysfunction" (e.g., Dawson & Watling, 2000; Hoehn & Baumeister, 1994; Ottenbacher, 1982; Polatajko, Kaplan, & Wilson, 1992; Smith et al., 2005). In conclusion, there is insufficient evidence to support the use of SIT as a therapy for children with ASD.

References[1]

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text revision). Washington, D.C. American Psychiatric Association.

Ayres, A. J. (1972). *Sensory integration and learning disorders*. Los Angles: Western Psychological Services.

*Ayres, A. J., & Tickle, L. S. (1980). Hyper-responsivity to touch and vestibular stimuli as a predictor of positive response to sensory integration procedures by autistic children. *American Journal of Occupational Therapy, 34*, 375–381.

*Bagatell, N., Mirigliani, G., Patterson, C., Reyes, Y., & Test, L. (2010). Effectiveness of therapy ball chairs on classroom participation in children with autism spectrum disorders. *American Journal of Occupational Therapy, 64*, 895–903.

Baraneck, G. T. (2002). Efficacy of sensory and motor interventions for children with autism. *Journal of Autism and Developmental Disorders, 32*, 397–422.

Ben-Sasson, A., Hen, L., Fluss, R., Cermak, S. A., Engel-Yeger, B., & Gal, E. (2009). A meta-analysis of sensory modulation symptoms in individuals with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39*, 1–11.

*Bonggat, P. W., & Hall, L. J. (2010). Evaluation of the effects of sensory integration-based intervention by a preschool special education teacher. *Education and Training in Autism and Developmental Disabilities, 45*, 294–302.

*Carter, S. L. (2005). An empirical analysis of the effects of a possible sinus infection and weighted vest on functional analysis outcomes of self-injury exhibited by a child with autism. *Journal of Early and Intensive Behavior Intervention, 2*, 252–258.

*Case-Smith, J., & Bryan, T. (1999). The effects of occupational therapy with sensory integration emphasis on preschool-age children with autism. *American Journal of Occupational Therapy, 53*, 489–497.

Case-Smith, J., & Miller, H. (1999). Occupational therapy with children with pervasive developmental disorders. *American Journal of Occupational Therapy, 53*, 506–513.

Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology, 66*, 7–18.

Conners, C. K. (1989). *Manual or Conner's rating scales*. North Tonawanda, NY: Multi-Health Systems.

Conners, C. K. (1997). *Conners' global index-teacher*. North Tonawanda, NY: Multihealth Systems Inc.

Constantino, J. N., & Gruber, C. P. (2005). *Social responsiveness scale manual*. Los Angeles: Western Psychological Services.

*Cox, A. L., Gast, D. L., Luscre, D., & Ayres, K. M. (2009). The effects of weighted vests on appropriate in-seat behaviors of elementary-age students with autism and severe to profound intellectual disabilities. *Focus on Autism and Other Developmental Disabilities, 24*, 17–26.

*Davis, T. N., Durand, S., & Chan, J. M. (2011). The effects of a brushing procedure on stereotypical behavior. *Research in Autism Spectrum Disorders, 5*, 1053–1058.

Dawson, G., & Watling, R. (2000). Interventions to facilitate auditory, visual, and motor integration in autism: A review of the evidence. *Journal of Autism and Developmental Disorders, 30*, 415–421.

*Devlin, S., Healy, O., Leader, G., & Hughes, B. M. (2011). Comparison of behavioral intervention and sensory-integration therapy in the treatment of challenging behavior. *Journal of Autism and Developmental Disorders, 41*, 1303–1320.

*Devlin, S., Leader, G., & Healy, O. (2009). Comparison of behavioral intervention and sensory-integration therapy in the treatment of self-injurious behavior. *Research in Autism Spectrum Disorders, 3*, 223–231.

Dunn, W. (1999). *Short sensory profile*. San Antonio, TX: Psychological Corporation.

Dunn, W. (2002). *Infant/toddler sensory profile*. San Antonio, TX: Therapy Skills Builders.

*Fazlioglu, Y., & Baran, G. (2008). A sensory integration therapy program on sensory problems for children with autism. *Perceptual and Motor Skills, 106*, 115–422.

*Fertel-Daly, D., Bedell, G., & Hinojosa, J. (2001). Effects of a weighted vest on attention to task and self-stimulatory behaviors in preschoolers with pervasive developmental disorders. *American Journal of Occupational Therapy, 55*, 629–640.

Glennon, T., Miller-Kuhaneck, H., Henry, D. A., Parham, L. D., & Ecker, C. (2007). *Sensory processing measure manual*. Los Angeles: Western Psychological Services.

Green, V. A., Pituch, K. A., Itchon, J., Choi, A., O'Reilly, M. F., & Sigafoos, J. (2006). Internet survey of treatments used by parents of children with autism. *Research in Developmental Disabilities, 27*, 70–84.

Hanley, G. P., Iwata, B. A., & McCord, B. E. (2003). Functional analysis of problem behavior: A review. *Journal of Applied Behavior Analysis, 36*, 147–185.

Healy, O., Hughes, B. M., Leader, G., & Devlin, S. (2011). Response to a letter to the editors re: Comparison of behavioral intervention and sensory-integration therapy in the treatment of challenging behavior. *Journal of Autism and Developmental Disorders, 41*, 139–1441.

*Hodgetts, S., Magill-Evans, J., & Misiaszek, J. E. (2011a). Weighted vests, stereotyped behaviors and arousal in children with autism. *Journal of Autism and Developmental Disorders, 41*, 805–814.

*Hodgetts, S., Magill-Evans, J., & Misiaszek, J. E. (2011b). Effects of weighted vests on classroom behavior for children with autism and cognitive impairments. *Research in Autism Spectrum Disorders, 5*, 495–505.

Hoehn, T. P., & Baumeister, A. A. (1994). A critique of the application of sensory integration therapy to children with learning disabilities. *Journal of Learning Disabilities, 27*, 338–350.

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children, 71*, 165–180.

Iarocci, G., & McDonald, J. (2006). Sensory integration and the perceptual experience of persons with autism. *Journal of Autism and Developmental Disorders, 36*, 77–90.

Individuals with Disabilities Education Improvement Act of 2004. 20 U.S.C. § 1400 et seq. (2004).

Iwata, B. A., & Deleon, I. G. (1996). *The functional analysis screening tool (FAST)*. Gainsville: The Florida Center on Self-Injury, University of Florida.

Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (1994). Toward a functional analysis of self-injury. *Journal of Applied Behavior Analysis, 27*, 197–209.

Jung, K., Lee, H., Lee, Y., Cheong, S., Choi, M., Suh, D., et al. (2006). The application of a sensory integration treatment based on virtual reality-tangible interaction for children with autistic spectrum disorder. *PsychNology Journal, 4*, 145–159.

*Kane, A., Luiselli, J. K., Dearborn, S., & Young, N. (2004). Wearing a weighted vest as intervention for children with autism/pervasive developmental disorder: Behavioral assessment of stereotypy and attention to task. *The Scientific Review of Mental Health Practice: Objective Investigations of Controversial and Unorthodox Claims in Clinical Psychology, Psychiatry, and Social Work, 3*, 19–24.

Kern, J. K., Garver, C. R., Carmody, T., Andrews, A. A., Mehta, J. A., & Trivedi, M. H. (2008). Examining sensory modulation in individuals with autism as compared to community controls. *Research in Autism Spectrum Disorders, 2*, 85–94.

[1] Asterisk symbol indicates the studies included in the review.

Krug, D. A., Arick, J. R., & Almond, P. J. (1980). *Autism screening instrument for educational planning*. Portland: ASIEP Ed Co.
Lancioni, G. E., O'Reilly, M., & Emerson, E. (1996). A review of choice research with people with severe and profound developmental disabilities. *Research in Developmental Disabilities, 17*, 391–411.
Lane, S. J., & Schaaf, R. C. (2010). Examining the neuroscience evidence for sensory-driven neuroplasticity: Implications for sensory-based occupational therapy for children and adolescents. *The American Journal of Occupational Therapy, 64*, 375–390.
Lane, A. E., Young, R. L., Baker, A. E. Z., & Angley, M. T. (2010). Sensory processing subtypes in autism: Association with adaptive behavior. *Journal of Autism and Developmental Disorders, 40*, 112–122.
*Leew, S. V., Stein, N. G., & Gibbard, W. B. (2010). Weighted vests' effect on social attention for toddlers with autism spectrum disorders. *Canadian Journal of Occupational Therapy/Revue Canadienne D'Ergotherapie, 77*, 113–124.
*Linderman, T. M., & Stewart, K. B. (1999). Sensory integrative-based occupational therapy and functional outcomes in young children with pervasive developmental disorders: A single-subject study. *American Journal of Occupational Therapy, 53*, 207–213.
Machalicek, W., O'Reilly, M. F., Beretvas, N., Sigafoos, J., Lancioni, G., Sorrells, A., et al. (2008). A review of school-based instructional interventions for students with autism spectrum disorders. *Research in Autism Spectrum Disorders, 2*, 395–416.
Mailloux, Z., May-Benson, T. A., Summers, C. A., Miller, L. J., Brett-Green, B., Burke, J. P., et al. (2007). Goal Attainment Scaling as a measure of meaningful outcomes for children with sensory integration disorders. *American Journal of Occupational Therapy, 61*, 254–259.
Mason, S. A., & Iwata, B. A. (1990). Artifactual effects of sensory integrative therapy on self-injurious behavior. *Journal of Applied Behavior Analysis, 23*, 361–370.
Matson, J. L., & Vollmer, T. (1995). *Questions about behavioral function (QABF)*. Baton Rouge, LA: Disability Consultants, LLC.
May-Benson, T. A., & Koomar, J. A. (2010). Systematic review of the research evidence examining the effectiveness of interventions using a sensory integrative approach for children. *The American Journal of Occupational Therapy, 64*, 403–414.
Mayton, M. R., Wheeler, J. J., Menendez, A. L., & Zhang, J. (2010). An analysis of evidence-based practices in the education and treatment of learners with autism spectrum disorders. *Education and Training in Autism and Developmental Disabilities, 45*, 539–551.
Mesibov, G. B., & Shea, V. (2011). Evidenced-based practices and autism. *Autism, 15*, 114–133.
Mutti, M. C., Martin, N. A., Sterling, H. M., & Spalding, N. V. (1998). *Quick neurological screening test manual* (2nd Ed.). Navato, CA: Academic Therapy.
National Autism Center's National Standards Project (NSP). (2009). *The national standards report*. http://www.nationalautismcenter.org/affiliates/model.php
No Child Left Behind Act of 2001. 20 U.S.C. 70 § 6301 et seq. (2002).
Odom, S. L., Collet-Klingenberg, L., Rogers, S. J., & Hatton, D. D. (2010). Evidence-based practices in interventions for children and youth with autism spectrum disorders. *Preventing School Failure, 54*, 275–282.
Ottenbacher, K. (1982). Sensory integration therapy: Affect or effect. *American Journal of Occupational Therapy, 36*, 571–578.
Parham, L. D., Cohn, E. S., Spitzer, S., Koomar, J. A., Miller, L. J., Burke, J. P., et al. (2007). Fidelity in sensory integration intervention research. *American Journal of Occupational Therapy, 61*, 216–227.
Parham, L. D., & Ecker, C. (2007). *Sensory processing measure manual*. Los Angeles: Western Psychological Services.
Parham, L. D., Roley, S. S., May-Benson, T. A., Koomar, J., Brett-Green, B., Burke, J. P., et al. (2011). Development of a fidelity measure for research on the effectiveness of the Ayres sensory integration® intervention. *The American Journal of Occupational Therapy, 65*, 133–142.
Parson, A. C., McWilliam, R. A., & Buysese, V. (1989). *A procedural manual for coding engagement in early childhood programs*. Chapel Hill, NC: Fran Porter Graham Child Development Center, University of North Carolina at Chapel Hill.
*Pfeiffer, B. A., Koenig, K., Kinnealey, M., Sheppard, M., & Henderson, L. (2011). Effectiveness of sensory integration interventions in children with autism spectrum disorders: A pilot study. *American Journal of Occupational Therapy, 65*, 76–85.
*Piravej, K., Tangtrongchitr, P., Chandarasiri, P., Paothong, L., & Sukprasong, S. (2009). Effects of Thai traditional massage on autistic children's behavior. *The Journal of Alternative and Complementary Medicine, 15*, 1355–1361.
Polatajko, H. J., Kaplan, B. J., & Wilson, B. N. (1992). Sensory integration treatment for children with learning disabilities: Its status 20 years later. *Occupational Therapy Journal of Research, 12*, 323–341.
*Quigley, S. P., Peterson, L., Frieder, J. E., & Peterson, S. (2011). Effects of a weighted vest on problem behaviors during functional analyses in children with pervasive developmental disorders. *Research in Autism Spectrum Disorders, 5*, 529–538.
*Ray, T. C., King, L. J., & Grandin, T. (1988). The effectiveness of self-initiated vestibular stimulation in producing speech sounds in an autistic child. *Occupational Therapy Journal of Research, 8*, 186–190.
*Reichow, B., Barton, E. E., Sewell, J. N., Good, L., & Wolery, M. (2010). Effects of weighted vests on the engagement of children with developmental delays and autism. *Focus on Autism and Other Developmental Disabilities, 25*, 3–11.
Reichow, B., Volkmar, F. R., & Cicchetti, D. V. (2008). Development of the evaluative method for evaluating and determining evidence-based practices in autism. *Journal of Autism and Developmental Disorders, 38*, 1311–1319.
*Reilly, C., Nelson, D. L., & Bundy, A. C. (1983). Sensorimotor versus fine motor activities in eliciting vocalizations in autistic children. *Occupational Therapy Journal of Research, 3*, 199–212.
Rogers, S. J., & Ozonoff, S. (2005). What do we know about sensory dysfunction in autism? A critical review of the empirical evidence. *Journal of Child Psychology and Psychiatry, 46*, 1255–1268.
Rogers, S. J., & Vismara, L. A. (2008). Evidenced-based comprehensive treatments for early autism. *Journal of Clinical Child & Adolescent Psychology, 37*, 8–38.
Schaaf, R., & Blanche, E. I. (2011). Comparison of behavioral intervention and sensory-integration therapy in the treatment of challenging behavior. *Journal of Autism and Developmental Disorders, 41*, 1436–1438.
Schaaf, R. C., & Miller, J. (2005). Occupational therapy using a sensory integrative approach for children with developmental disabilities. *Mental Retardation and Developmental Disabilities, 11*, 143–148.
Schlosser, R. (2009). The role of evidence-based journals as evidence-based information sources. *Applied behavior analysis international annual convention (conference)*. Phoenix: Az.
Schlosser, R., & Sigafoos, J. (2007). Editorial: Moving evidence-based practice forward. *Evidence-Based Communication Assessment and Intervention, 1*, 1–3.
Simeonsson, R., & Bailey, D. (1991). Evaluating programme impact: Levels of certainty. In D. Mitchell & R. Brown (Eds.), *Early intervention studies for young children with special needs*. New York: Chapman and Hall.
Smith, N. (1981). The certainty of evidence in health evaluations. *Evaluation and Program Planning, 4*, 273–278.
Smith, T., Mruzek, D. W., & Mozingo, D. (2005). Sensory integrative therapy. In J. W. Jacobson, R. M. Foxx, & J. A. Mulick (Eds.), *Controversial therapies for developmental disabilities: Fad, fashion and science in professional practice* (pp. 331–350). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
Stephenson, J., & Carter, M. (2005). The use of weighted vests with children with autism spectrum disorders and other disabilities. *Journal of Autism and Developmental Disorders, 39*, 105–114.
*Thompson, C. J. (2011). Multisensory intervention observational research. *International Journal of Special Education, 26*, 202–214.
*Van Rie, G. L., & Heflin, L. J. (2009). The effect of sensory activities on correct responding for children with autism spectrum disorders. *Research in Autism Spectrum Disorders, 3*, 783–796.
*Watling, R. L., & Dietz, J. (2007). Immediate effect of Ayres' sensory integration-based occupational therapy intervention on children with autism spectrum disorders. *American Journal of Occupational Therapy, 61*, 574–583.
Watling, R., Deitz, J., Kanny, E. M., & McLaughlin, J. F. (1999). Current practice of occupational therapy for children with autism. *American Journal of Occupational Therapy, 53*, 498–505.