Review Journal of Autism and Developmental Disorders (2018) 5:156–175
https://doi.org/10.1007/s40489-018-0130-z

REVIEW PAPER

Effects of Equine Therapy on Individuals with Autism Spectrum Disorder: a Systematic Review

Sudha M. Srinivasan[1] · David T. Cavagnino[2] · Anjana N. Bhat[2,3]

Received: 3 January 2018 / Accepted: 31 January 2018 / Published online: 20 February 2018

Abstract
Literature on effects of equine therapy in individuals with autism spectrum disorder (ASD) has grown in recent times. Equine therapy is an alternative multimodal intervention that involves utilizing a horse to remediate core impairments in ASD. Recent systematic reviews in this area have several limitations including inclusion of populations other than ASD, assessment of a variety of animal-assisted interventions other than equine therapy, and a failure to conduct quantitative analyses to provide accurate effect size estimates. We conducted a focused systematic review to address these limitations. Our review suggested that equine therapy has beneficial effects on behavioral and to some extent on social communication skills in ASD. The evidence for positive effects of equine therapy on perceptuo-motor, cognitive, and functional skills is currently limited.

Keywords Equine therapy · Hippotherapy · Therapeutic horseback riding · Social communication · Behavior · Autism

Autism spectrum disorder (ASD) is a neurodevelopmental disorder that affects multiple subsystems including social communication, behavioral, cognitive, and perceptuo-motor domains. Core impairments in ASD include social communication difficulties as well as the presence of stereotyped and repetitive behaviors and interests (American Psychiatric Association 2013). More specifically, social communication impairments include poor reciprocity during social interactions, reduced social gaze/eye contact, as well as delays in nonverbal and verbal communication (Mundy and Newell 2007; Tager-Flusberg 1999; Dawson et al. 2004). In addition, children with ASD also demonstrate several comorbidities such as significant difficulties with behavioral and emotional regulation (i.e., anxiety, aggression, depression, hyperactivity, temper tantrums, and self-injurious behaviors) (Bodfish et al. 2000; Lecavalier 2006; Loh et al. 2007; Mazefsky et al. 2012) and cognitive difficulties in attentional focus, executive functioning, and working memory (Ozonoff et al. 2006; Williams et al. 2006). There is also substantial evidence for sensorimotor impairments in children with ASD (Fournier et al. 2010; Bhat et al. 2011) including difficulty modulating sensory inputs (Tomchek and Dunn 2007; Baranek et al. 2005) and pervasive motor impairments during coordinated arm/leg movements (Fournier et al. 2010; Green et al. 2009; Vilensky et al. 1981; Hallett et al. 1993), balance tasks (Minshew et al. 2004), as well as imitation and praxis tasks (Mostofsky et al. 2006; Dewey et al. 2007). Furthermore, comorbidities in sensorimotor performance correlate with and influence social communication performance of individuals with ASD (Dziuk et al. 2007; Lord and Ulrich 2014). Limited movement exploration and motor clumsiness may lead to missed opportunities to develop social connections with peers and caregivers (Bhat et al. 2011; Leary and Hill 1996; Jansiewicz et al. 2006). Overall, individuals with ASD face difficulties in multiple developmental domains; this necessitates the use of multimodal interventions to effectively address both diagnostic impairments as well as comorbidities in this population.

Mainstream autism interventions based on principles of applied behavioral analysis such as Discrete Trial Training (Lovaas 1987), Pivotal Response Therapy (Koegel and Koegel 2006), Early Start Denver Model (Rogers and Dawson 2009), and Picture Exchange Communication System (Bondy and Frost 2003) as well as more recent

✉ Anjana N. Bhat
abhat@udel.edu

1 IDC School of Design, Indian Institute of Technology Bombay, Powai, Mumbai, Maharashtra 400076, India

2 Department of Physical Therapy, University of Delaware, 540 S College Ave., Newark, DE 19713, USA

3 Department of Psychological & Brain Sciences, University of Delaware, Newark, DE, 19713, USA

developmental approaches such as Floortime, Social Communication, Emotional Regulation and Transactional Support (SCERTS), Developmental Individual-Difference, Relationship-based model (DIR) (Greenspan and Wieder 1997; Wieder and Greenspan 2003; Prizant et al. 2006; Kasari et al. 2008; Landa et al. 2011), mainly focus on improving social communication, behavioral and academic skills of children with ASD. While ABA-based approaches use structured environments and incremental prompting to promote positive behaviors (Lovaas 1987; Bondy and Frost 2003; Mesibov et al. 2005), developmental approaches focus on facilitating specific early social communication skills such as joint attention and imitation (Landa et al. 2011; Kasari et al. 2008). Although these approaches have substantial evidence for improving social communication and functional/academic skills, they may not address the sensorimotor needs of the children with ASD. Through this systematic review (SR), we summarize the evidence for the use of *equine therapy as a multimodal intervention* to address diagnostic impairments in social communication skills as well as behavioral and sensorimotor comorbidities in individuals with ASD.

Equine therapy involves activities completed in the presence of a horse, including mounted activities such as hippotherapy and therapeutic horseback riding as well as non-mounted, equine-focused activities such as grooming and caring for the horse (Lentini and Knox 2015). Hippotherapy (HIP) involves purposeful manipulation of equine movement based on clinical reasoning of occupational therapists (OTs), physical therapists (PTs), and speech language pathologists (SLPs) to engage an individual's sensory, neuromotor, and cognitive systems and achieve certain functional outcomes (American Hippotherapy Association 2017). The focus in HIP is on using the horse and its movement as a tool by allied health professionals to achieve therapeutic goals such as improvement of balance, sensory processing skills, and arousal in clients. In contrast, in THR the focus is on teaching the student different types of riding skills (PATH International 2017). A typical THR session involves mounted and non-mounted equine-assisted activities to promote the physical, cognitive, emotional, and social well-being of individuals with special needs (Ward et al. 2013). Compared to HIP accreditation, THR instructor certification is accessible to a broader group of trainers including special educators, counselors, as well as equestrians, following completion of online coursework and 25 hours of training experience.

To date, few SRs have been conducted on the effects of equine therapy in individuals with ASD (O'Haire 2013, 2017; Lentini and Knox 2015; Weise et al. 2016; McDaniel and Wood 2017). O'Haire (2013, 2017) conducted a qualitative analysis of the effects of animal-assisted interventions as a whole (vs. equine therapy only) in ASD and reported positive effects in terms of social, communication/language, and stress/behavior, as well as a reduction in autism symptoms. Similarly, based on their broad review of 47 studies in individuals with diverse types of special needs, Lentini and Knox (2015) reported qualitative improvements in social emotional skills, self-regulation, self-esteem, and levels of anxiety and depression following equine therapy (Lentini and Knox 2015). Weise et al. (2016) reviewed 8 studies that assessed the effects of equine therapy on only social behavioral skills of individuals with ASD. Lastly, based on a comprehensive systematic review of equine-assisted interventions, McDaniel and Wood (2017) offered proof of concept evidence that equine therapies can benefit children and adolescents with ASD. Taken together, past SRs have focused on broader topics of animal-assisted interventions (vs. focusing on equine therapies) and diverse populations (special needs with various diagnoses or mental health issues), and used more qualitative as opposed to quantitative analytical methods (none of the reviews assessed the quality of research evidence in conjunction with the size of treatment effects obtained in each study). The current study addresses these limitations by conducting a comprehensive (qualitative and quantitative) review of the effects of equine therapy on impairments in multiple subsystems/domains in individuals with ASD. Our first aim was to examine the effects of equine therapy on specific domains including social, communication, behavioral, and sensorimotor skills as well as broader functional outcomes including overall adaptive functioning and quality of life. Our second aim was to systematically assess the methodological quality (using Sackett's level of evidence and PEDro scores) of the published research evidence to date and calculate the size of the treatment effects for all outcomes measured in individual studies included in the review.

Methods

Search Protocol

Studies were searched from five common electronic databases associated with the fields of health, psychology, allied health, and education, i.e., PubMed (1950–present), PsycINFO (1969–present), ERIC (1966–present), Scopus (1966–present), and CINAHL (1937–present). We searched for terms associated with "equine therapy" and "autism" (see details in Appendix 1). We also searched the reference sections of included articles and previously published reviews for additional relevant articles (see Fig. 1 for details).

Eligibility Criteria

Studies were screened based on the following *inclusion criteria*—(a) peer-reviewed articles and (b) studies reporting data on treatment effects of "equine" therapy using experimental or quasi-experimental study designs—and *exclusion*

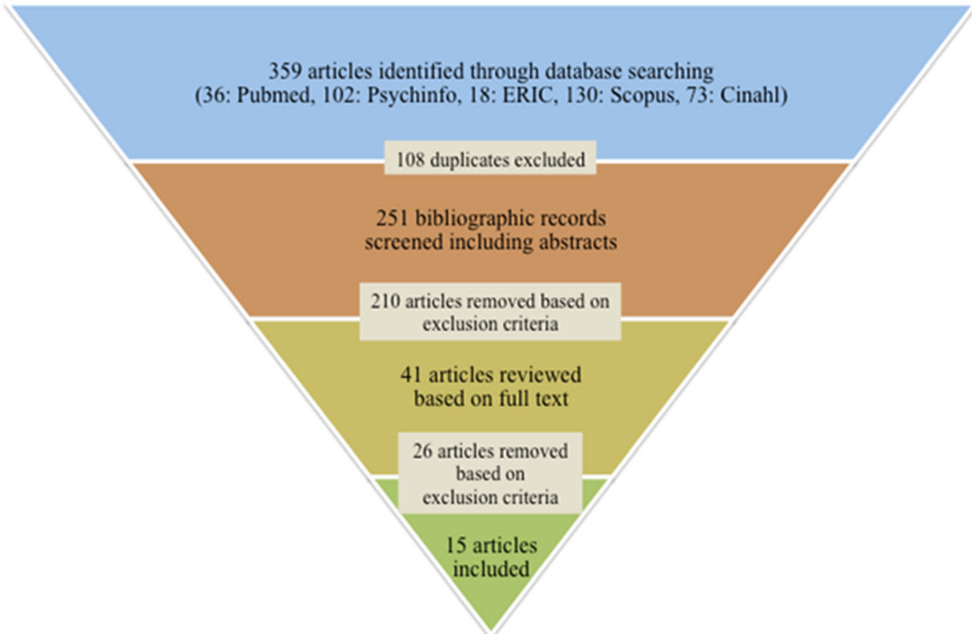

Fig. 1 A flowchart of the study selection process

criteria—(a) foreign language; (b) case reports, opinions, narrative reports, etc.; (c) animal-assisted or pet therapy not equine in nature, for example, using dolphins and dogs; and (d) unrelated topics, for example, PET imaging.

Data Extraction and Evaluation

After applying the eligibility criteria, we were left with 15 articles in this review. We used the levels of evidence described by Sackett et al. (1996) to classify the methodological quality of studies. These guidelines provide systematic criteria that can be used to classify individual studies based on study design into 5 levels of evidence from levels I–V. Level I studies provide the highest level of evidence and include randomized controlled trials, whereas level V studies include case reports and narrative statements. In addition, we used the Physiotherapy Evidence Database (PEDro) scale (Moseley et al. 2002) to evaluate the internal validity and interpretability of the randomized controlled trials (RCTs) and controlled clinical trials (CCTs) included in this review. The PEDro scale contains 11 items of which the last 10 items are scored to obtain the total PEDro score (Moseley et al. 2002). We coded each study for sample and study characteristics, methodological quality, assessment measures, dependent variables, and treatment effects (see Appendix 2 for the coding sheet). In addition to the qualitative coding of treatment effects based on the original study reports, we also report on quantitative effect size measures from each study. We used data reported in the original studies (wherever available) to calculate effect sizes using the standardized mean difference (*d*) index (Hedges 1981; Lipsey and Wilson 2001; Huedo-Medina and Johnson 2011). We used the excel code developed by Heudo-Medina and Johnson to calculate effect sizes and their 95% confidence intervals for all studies (Huedo-Medina and Johnson 2011). For studies that reported effect sizes within the original paper, we compared the reported (from original study) and calculated (based on our calculations) effect sizes. We also summarized the treatment effects reported per domain based on the reviewed studies, i.e., effects on social communication skills and motor skills. All three authors who are physical therapists by training were involved in the search process, extraction of data, and coding of studies. For the purpose of reliability, all three authors coded five out of the 15 articles using the coding form listed in Appendix A. We used intraclass correlation coefficients to calculate reliability on all the parameters coded for the individual studies. Each of the three authors achieved inter-rater reliability of over 87.5% after establishing consensus on scores they disagreed on. Each coder also established intra-rater reliability of over 99%. Following reliability, the remaining 10 articles were divided for review between the three authors.

Results

Description of Studies

Out of the 15 studies reviewed, 8 were conducted in USA, 2 in Spain (Tabares et al. 2012; Garcia-Gomez et al. 2014), 1 in Canada (Llambias et al. 2016), 1 in the UK (Anderson and Meints 2016), 1 in Taiwan (Wuang et al. 2010), 1 in Italy (Borgi et al. 2016), and 1 in Slovakia (Steiner and Kertesz 2015). All studies were published in peer-reviewed journals except one study, which was published as a peer-reviewed conference paper (Steiner and Kertesz 2015). All included studies were published between 2009 and 2016, although only one study mentioned the year of data collection in the published report (Wuang et al. 2010). Out of the 15 studies, 12 studies assessed the effects of THR whereas the remaining 3 studies assessed the effects of HIP in individuals with ASD.

Sample Characteristics

Our systematic review is based on a total of 428 subjects out of which 294 subjects received equine therapy and the remaining 134 subjects received some form of control intervention. The reviewed studies demonstrated wide variation in sample sizes ranging from 6 to 116 participants (Table 1). Eleven out of the 15 studies included only children with diagnoses of ASD, 2 studies included a mixed sample of children with ASD and Asperger Syndrome (Bass et al. 2009; Kern et al. 2011), 1 study included children with ASD and attention deficit hyperactivity disorder (ADHD) (Llambias et al. 2016), and 1 study included children with ASD who also had comorbid disorders such as ADHD and hypersensitivity and sensory integration disorder (HSID) (Anderson and Meints 2016).

Only four studies reported the use of standardized autism-specific measures such as the Autism Diagnostic Observation Schedule (ADOS), Social Communication Questionnaire (SCQ), or Childhood Autism Rating Scale (CARS) to establish the diagnosis of participants (Gabriels et al. 2012, 2015; Garcia-Gomez et al. 2014; Kern et al. 2011). Eleven other studies reported having confirmed diagnosis of subjects based on expert clinical opinion of a licensed health professional and four of the remaining studies did not report on the method of confirmation of subjects' diagnoses (Table 1). Seven out of the 15 studies also reported on some additional baseline variables such as intelligence quotient (IQ) and language use of subjects to characterize their adaptive skills and functional status.

All studies provided equine therapy to children or adolescents between 3 and 16 years (Table 1); interestingly, our literature search did not reveal any studies that provided interventions to adults with autism. In terms of gender distribution of subjects, two studies did not provide information on the male-to-female ratio of subjects in their final post-attrition sample (Lanning et al. 2014; Ajzenman et al. 2013). Out of the total 404 subjects included in the remaining 13 studies, 328 subjects were males and 76 were females. In terms of sample selection criteria, 13 out of the 15 studies excluded children with previous riding experience, 1 study included both children with and without prior riding experience (Tabares et al. 2012), and 1 study did not report on children's previous riding experiences (Steiner and Kertesz 2015, Table 1). Studies accounting for previous horseback riding experience either excluded subjects with previous riding experience or defined a time period prior to the intervention (ranging from 3 months to 3 years) in which subjects must have limited to no riding experience.

Study Characteristics

A variety of study designs were employed in the included studies (Table 2 for details). Out of the three randomized controlled trials (RCT) that qualified as level I evidence, 2 studies (Bass et al. 2009; Borgi et al. 2016) employed a waitlist control group, whereas in the third study (Gabriels et al. 2015), the control group engaged in horsemanship activities involving a stuffed horse, with no contact with live horses. The PEDro scores for these 3 studies ranged from 5 to 7/10 (see Tables 2 and 3).

Out of the five studies that were classified as level II evidence, one study (Wuang et al. 2010) used a crossover design, whereas the remaining 4 studies were controlled clinical trials (Gabriels et al. 2012; Steiner and Kertesz 2015; Garcia-Gomez et al. 2014; Lanning et al. 2014). All 5 studies employed non-randomized controls; however, only 2 of these studies conducted post hoc testing to ascertain baseline similarity of the experimental and control groups on select variables including age, gender, perceptuo-motor performance, nonverbal IQ, and presence of seizures (Wuang et al. 2010; Gabriels et al. 2012). Four of these Level II studies provided conventional therapies for control group subjects including occupational therapy (OT) (Wuang et al. 2010), special pedagogy exercises (Steiner and Kertesz 2015), medical and re-education treatment (Garcia-Gomez et al. 2014), and social circles (Lanning et al. 2014), while the last study (Gabriels et al. 2012) used a waitlist control group. In terms of PEDro scoring, all level II studies, scored in the range of 2–4/10 (see Tables 2 and 3 for details).

Our review included 5 Level III studies that employed a single group and compared outcomes within this group from pretest to post-test or at multiple time points in an interrupted fashion (Kern et al. 2011; Tabares et al. 2012; Ajzenman et al. 2013; Ward et al. 2013; Anderson and Meints 2016). Lastly, 2 studies in this review employed a single subject design and were qualified as level IV evidence (Jenkins & Reed 2013 (here on Jenkins & Reed); Llambias et al. 2016). Although the Jenkins & Reed (2013) study non-randomly assigned 4 of their 7 subjects to a treatment group and the remaining 3 to a waitlist control group, they did not conduct any between-group comparisons and only reported on individual data, therefore justifying their classification as level IV evidence. Since PEDro scores can only be calculated for RCTs or CCTs (Moseley et al. 2002), we could not report on these scores for the 7 level III and IV studies in this review.

Outcome Measures and Assessments

Given that ASD is a disorder that affects several subsystems, many studies in this review assessed the impact of equine therapies on more than one skill domain using a combination of subjective and objective measures including standardized

Table 1 Study characteristics

Participant characteristics						Intervention characteristics						
First author, year	Final sample size (E, C)	Age M (SD) (Range)	Diagnoses	Diagnostic measure	Other therapy	Duration in weeks (frequency sessions/week)	Session time in minutes	Type (HIP, THR, SHR)	Format (I or G)	Therapy provider accreditation	Past/recent riding experience	C group interventions
Bass et al. (2009)	34 (19, 15)	E: 6.95 (1.67) (5–10) C: 7.73 (1.65) (4–10)	ASD, AS	DSM IV	–	12 (1)	60	THR	I	–	None	Waitlist
Wuang et al. (2010)	60 (60, 0)	E only: F: 7.59 (0.31) M: 7.58 (0.34) Group range: (6.42–8.75)	ASD	–	OT	20 (2)	60	SHR, OT	I	n/a	None	OT
Kern et al. (2011)	24 (24, 0)	7.8 (2.9) (3–12)	ASD, AS	PC, CARS	CNAT	24 (1)	60	THR	I	–	None	n/a
Gabriels et al. (2012)	42 (26, 16)	8.7 (–) (6–16)	ASD	PC, ABC-C, SCQ, ADOS	CNAT	10 (1)	60	THR	G	PATH	<2 weeks in past 3 years	Waitlist
Steiner et al. (2012)	26 (13, 13)	(–) (–) (10–13)	ASD	–	–	5 (1)	30	THR	–	–	–	Special pedagogy exercises
Tabares et al. (2012)	8 (8, 0)	(–) (–) (8–16)	ASD	–	–	4 (1)	30	HIP	I	AZE	Yes	n/a
Ajzenman et al. (2013)	6 (6, 0)	8.4 (2.5) (5–12)	ASD	DSM IV	–	12 (1)	45	HIP	G	PATH	None	n/a
Jenkins and Reed (2013)	7 (4,3)	9.5 (–) (6–14)	ASD	PC	–	9 (1)	60	THR	G	PATH	None	Waitlist
Ward et al. (2013)	21 (21, 0)	8.1 (–) (5–12)	ASD	DSM IV	SLP (all students), OT (20 students), PT (1 student)	18 (1)	60	THR	G	PATH	None for 13 of the 21 students	n/a
Garcia-Gomez et al. (2014)	16 (8, 8)	(–) (–) (7–14)	ASD	CARS	CNAT	12 (2)	45	THR	–	PATH	None within 2 years	Medical and re-education treatment
Lanning et al. (2014)	18 (10, 8)	E: 7.5 (3.2) (4–15) C: 9.8 (3.2) (5–14)	ASD	–	–	12 (1)	60	THR	I or G	PATH	None within 6 months	Social circles: card games, board games, sensory activities
Gabriels et al. (2015)	116 (58, 58)	10.2 (3.0) (6–16)	ASD	SCQ, ADOS, ABC-C	–	10 (1)	45	THR	G	PATH	<2 h EAAT within 6 months	Barn activities
Anderson and Meints (2016)	15 (15, 0)	10 (3.8) (5–16)	ASD ADHD HSID	DSM (unspecified), PC	–	5 (1)	180	THR	G	BHS, RDA	None	n/a
	28	8.6 (1.7)	ASD	DSM IV,	CNAT	25 (1)	60–70	THR	G	FISE	None	Waitlist

Table 1 (continued)

Participant characteristics						Intervention characteristics						
First author, year	Final sample size (E, C)	Age M (SD) (Range)	Diagnoses	Diagnostic measure	Other therapy	Duration in weeks (frequency sessions/week)	Session time in minutes	Type (HIP, THR, SHR)	Format (I or G)	Therapy provider accreditation	Past/recent riding experience	C group interventions
Borgi et al. (2016)	(15, 13)	(6–12)		ICD-10								
Llambias et al. (2016)	7 (7, 0)	(–) (–) (4–12)	ASD (7) ADHD (2)	PC	SLP, PT, OT or BI	9–12 (1)	45–60	HIP	I	PATH	None within 3 months	n/a

– not reported, *ABC-C* Aberrant Behavior Checklist-Community, *ADHD* attention deficit hyperactivity disorder, *ADOS* Autism Diagnostic Observation Schedule, *AHA* American Hippotherapy Association, *AS* Asperger Syndrome, *ASD* autism spectrum disorder, *AZE* Association of Zootherapy of Extremadura, *BHS* British Horse Society, *BI* Behavioral Interventionist Services, *C* control group, *CARS* Childhood Autism Rating Scale, *CNAT* Continued Normal Activities and Therapies (subjects did not alter normal routine—specific therapies not specified), *E* experimental group, *EAAT* equine-assisted activities and therapies, FISE Federazione Italiana Sport Equestri, *G* Group, *HIP* hippotherapy (equine-assisted occupational or physical therapy), *HSID* hypersensitivity and sensory integration disorder, *ICD-10* International Statistical Classification of Disease and Related Health Problems, *I* individual, *M* mean, *n/a* not applicable, *OT* occupational therapy, *PATH* Professional Association of Therapeutic Horsemanship International, *PC* physician confirmation/diagnosis, *PIW* pre-intervention waiting period, *PT* physical therapy, *RDA* riding for the disabled, *SCQ* Social Communication Questionnaire, *SHR* simulated horseback riding, *SLP* Speech Language Pathology Services, *THR* therapeutic horseback riding

tests, quantitative measures, observational measures, parent/teacher-rated questionnaires, and video coding to assess treatment efficacy (Tables 4 and 6). Specifically, 12 out of the 15 studies used parent/teacher-rated questionnaires such as the Sensory Profile (SP), Vineland Adaptive Behavior Scale (VABS), and Ages & Stages Questionnaire (ASQ); 6 studies used standardized tests administered by expert clinicians, for instance, the Childhood Autism Rating Scale (CARS), Sensory Integration and Praxis Tests (SIPT), and Bruininks Oseretesky Test of Motor Proficiency (BOT); 3 studies used video coding or observational measures to assess changes in subjects' behaviors during training sessions; 2 studies used objective quantitative measures including kinematic analysis and force plate data; and 1 study used physiological measures including salivary hormonal levels to assess the effects of equine therapy (see Tables 4 and 7 for details). For the outcomes evaluated, only 7 studies reported reliability of the assessed variables. In terms of blinding assessors for clinical assessments or video coding, three studies blinded assessors to grouping/treatments (Borgi et al. 2016; Gabriels et al. 2015; Wuang et al. 2010), two other studies blinded assessors/coders to study goals but not treatment groups (Anderson and Meints 2016; Llambias et al. 2016), two studies admitted to not following blinding procedures (Ajzenman et al. 2013; Gabriels et al. 2012), and the remaining 8 studies did not specify details about blinding (see Table 7). In terms of the maintenance of treatment gains, only 4 of the 15 studies assessed the long-term effects of equine therapies through follow-up testing which was conducted at 1 month (Llambias et al. 2016), 1.5 months (Ward et al. 2013), 3 months (Steiner and Kertesz 2015), and 6 months post intervention (Wuang et al. 2010).

Treatment Effects

Out of the 15 studies, 6 reported only within-group effects, 2 reported only between-group effects, and 7 studies reported both between- and within-group effects (see Table 6). Although the critical first step in this literature is to assess if equine therapy is a feasible and useful treatment modality for individuals with ASD, equally if not more important is the accurate estimation of the size of this treatment effect, if present and its classification as small, medium, or large according to standard conventions (Cohen 1988). Effect sizes (ES) and their confidence intervals (CI) are required to understand the clinical meaningfulness of research findings (Page 2014; Sullivan and Feinn 2012). Interestingly, only 6 studies in our review reported ES for some of the outcomes (see Table 5). Even out of these 6 studies, Wuang et al. (2010) did not provide exact values of between-group ES. Similarly, although Bass et al. (2009) conducted a between-group study, they only reported within-group ES (see Table 5). Therefore, wherever possible, we calculated ES and their 95% CI using data reported in the original studies. Accordingly, we could calculate

Table 2 Methodological quality

First author, year	Study design	Control Group allocation	Exclusion criteria: past/recent riding experience?	Checks on treatment fidelity/integrity	PEDro score	Level of evidence
Bass et al. (2009)	RCT	Randomized	Yes	Not reported	5	I
Wuang et al. (2010)	Crossover design	Non-randomized but post-hoc testing of baseline similarity	Yes	Yes	4	II
Kern et al. (2011)	Pre-post design	n/a	Yes	Not reported	n/a	III
Gabriels et al. (2012)	CCT	Non-randomized but post-hoc testing of baseline similarity	Yes	Not reported	4	II
Steiner et al. (2012)	CCT	Not randomized	Not reported	Not reported	2	II
Tabares et al. (2012)	Pre-post design	n/a	Yes	Not reported	n/a	III
Ajzenman et al. (2013)	Pre-post design	n/a	Yes	Not reported	n/a	III
Jenkins et al. (2013)	Single case	Not randomized	Yes	Yes	n/a	IV
Ward et al. (2013)	Single-group interrupted time series	n/a	No	Not reported	n/a	III
Garcia-Gomez et al. (2014)	CCT	Not randomized	Yes	Not reported	4	II
Lanning et al. (2014)	CCT	Not randomized	Yes	Not reported	2	II
Gabriels et al. (2015)	RCT	Randomized	Yes	Yes	7	I
Anderson and Meints (2016)	Pre-post design	n/a	Yes	Not reported	n/a	III
Borgi et al. (2016)	RCT	Randomized	Yes	Not reported	5	I
Llambias et al. (2016)	Single-case reversal	n/a	Yes	Yes	n/a	IV

Sackett's levels of evidence: level 1: systematic reviews, meta-analyses, RCTs; level II: two-group, non-randomized studies; level III: one-group non-randomized studies; level IV: descriptive studies (single-subject designs); level V: case reports/expert opinions/consensus statements

CCT controlled clinical trial, *n/a* not applicable, *N* no, *RCT* randomized control trial, *Y* yes

ES for 6 more studies that did not provide these estimates in the original report (see Table 5). However, we could not calculate ES for the 3 remaining studies due to insufficient data provided in the published report (Llambias et al. 2016; Jenkins & Reed 2013; Steiner and Kertesz 2015). Comparisons between the calculated and reported ES are provided in Table 5. We have also provided 95% CI ranges for calculated ES and details of the number of calculated ES per study where the 95% CI does not include 0 (implying a truly significant non-zero treatment effect at 5% significance level) (see Table 5). Next, we summarize the salient findings of this review in a domain-wise manner.

Social Communication Skills This domain was most frequently assessed in equine therapy studies, with 9 out of the 11 studies that assessed social communication outcomes reporting improvements in skills following equine interventions (Table 6). Several studies employed multiple measures to evaluate therapy-related changes in skills—8 studies used generic standardized developmental tests or questionnaires (such as Childhood Autism Rating Scale (CARS), Vineland Adaptive Behavior Scales (VABS)), 3 studies used measures specific to social communication skills (such as Social Responsiveness Scale (SRS), Peabody Picture Vocabulary Test (PPVT)), and 2 studies used video coding to code for relevant behaviors during training sessions (see Tables 4 and 6). In terms of ES, there was considerable variability in the reported/calculated estimates across studies (see Table 6). However, the 5 methodologically strongest level I and II studies in this group as identified using Sackett's classification system and PEDro scores suggested that *equine therapies have small to medium size effects on social communication skills* (Gabriels et al. 2012, 2015; Bass et al. 2009; Borgi et al. 2016; Steiner and Kertesz 2015). For instance, Gabriels et al. (2015) reported that a 10-week intervention of THR in individuals with ASD led to improvements in the social cognition (ES = 0.41) and social communication (ES = 0.63) subscales of the SRS as well as in the total number of different and new words spoken (ES = 0.54) as assessed using the SALT test. Along the same lines, Bass et al. (2009) who also used the SRS scale, reported improvements in the overall scores (ES = 0.45) and the social motivation subscale (ES = 0.38) following a 12-week THR intervention. In contrast, Borgi et al. (2016) reported modest between-group improvements on the socialization subdomain (ES = 0.14) of the VABS following a 6-month THR intervention. Overall, out of the 21 ES estimates we calculated only 3 ES had CI that did not include 0 (Table 6). Taken together, there is limited evidence supporting the use of equine therapies for facilitating social communication skills in individuals with ASD.

Table 3 PEDro scoring for RCT/CCT

First author, year	Eligibility criteria specified	Random subject allocation	Allocation concealment	Baseline similarity of groups	Blinding of subjects	Blinding of therapists	Blinding of assessors	Measures of key outcomes	Intention to treat	Between-group comparisons	Point measures and variability measures	Total
Bass et al. (2009)	Y	Y	N	Y	N	N	N	Y	N	Y	Y	5
Wuang et al. (2010)	Y	N	N	Y	N	N	Y	N	N	Y	Y	4
Gabriels et al. (2012)	Y	N	N	Y	N	N	N	Y	N	Y	Y	4
Steiner et al. (2012)	Y	N	N	N	N	N	N	Y	N	N	Y	2
Garcia-Gomez et al. (2014)	Y	N	N	N	N	N	N	Y	Y	Y	Y	4
Lanning et al. (2014)	Y	N	N	N	N	N	N	N	N	Y	Y	2
Gabriels et al. (2015)	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	7
Borgi et al. (2016)	Y	Y	N	Y	N	N	N	Y	N	Y	Y	5

Behavioral Skills Out of the 7 studies that assessed the effects of equine therapies on behavioral skills including stereotyped/problem behaviors, affective responses, irritability and hyperactivity, and overall ability to regulate behaviors and moods, 5 studies found positive effects of equine-assisted activities (Table 6). In terms of outcome measures, 4 studies used assessments that specifically evaluated behavioral skills such as the Childhood Behavior Checklist (CBCL) and the Aberrant Behavior Checklist (ABC), whereas only 2 studies relied solely on generic assessments such as the CARS to evaluate effects of equine therapies on behavioral skills in individuals with ASD (see Tables 4 and 6). Although there was variability in ES estimates provided/calculated for these 7 studies, the 3 level I and II studies in this review reported *small to moderate effects of equine therapy on behaviors of children with ASD* (Gabriels et al. 2012, 2015; Garcia-Gomez et al. 2014). Gabriels et al. (2012) found medium to large between-group effect sizes for improvements in the irritability (ES = 0.87), lethargy/social withdrawal (ES = 0.81), stereotypy (ES = 0.69), and hyperactivity (ES = 0.80) subscales of the ABC in a pilot study involving a 10-week THR intervention in 42 children and adolescents with ASD compared to a waitlist control group. In a more recent, well-controlled RCT of identical duration by the same group, they found more moderate but significant between-group improvements in the irritability (ES = 0.5) and hyperactivity (ES = 0.53) subdomains of the ABC in 58 individuals with ASD (Gabriels et al. 2015). The third level II study by Garcia-Gomez et al. (2014) used the teacher-rated behavioral questionnaire, Behavior Assessment System for Children (BASC), as the outcome measure and found small within-group improvements (ES = 0.22) in only the aggressiveness domain of the questionnaire following 3 months of THR training in 8 children with ASD. Our own calculations suggested that out of the 19 ES computed for behavioral skills, the CI of 11 ES did not include 0 (Table 6). Overall, current literature provides modest evidence for the use of equine therapies to alleviate behavioral impairments in ASD.

Sensory Skills Sensorimotor skills are currently not considered a part of the cardinal diagnostic features of ASD; however, recently, there is growing evidence for the presence of perceptuo-motor impairments across the lifespan in ASD (Green et al. 2009; Bhat et al. 2011; Bedford et al. 2016; Ben-Sasson et al. 2009; Baranek et al. 2005; Srinivasan et al. 2015). We found that 3 of the 4 studies that assessed sensory skills reported positive effects following therapy (see Table 6). Three studies using parent/teacher-rated questionnaires (such as the Sensory Profile (SP) and the Sensory Profile School Companion (SPSC)), while the last study used a clinician-administered test called the Test of Sensory Integration Function (TSIF) (see Tables 4 and 6). Out of the 4 studies, 2 were classified as level I or II evidence

Table 4 Study-wise list of dependent variables and results

First author, year	Domains/variables tested	Type of effect (B/W)	Measures	Measures/variables showing improvement
Bass et al. (2009)	Social and sensory processing skills	B	SRS, SP	SRS overall and social motivation subscale scores; overall score and all subscales of SP except fine motor/perceptual subscale
Wuang et al. (2010)	Motor and sensory processing skills	B	BOT, TSIF	All subtests of BOT and TSIF
		W		All subtests of BOT and TSIF
Kern et al. (2011)	Social communication, behavioral, and sensory processing skills and QOL	W	CARS, TPCIS, SP, QOL ESS, TSS	CARS, QOL ESS
Gabriels et al. (2012)	Adaptive behavior, motor, sensory processing, and social communication skills	B	ABC- C, VABS, BOT, SIPT	All subscales of ABC-C except Inappropriate speech
		W		All subscales of ABC-C except Inappropriate speech; Adaptive behavior composite and social, communication, and daily living, subscales of VABS; BOT-2 short form scores; verbal and postural praxis subtests of SIPT
Steiner et al. (2012)	Social communication and motor skills	B	PAC, length of gait cycle	All subscales of PAC; length of gait cycle
		W		All subscales of PAC; length of gait cycle
Tabares et al. (2012)	Physiological variables as proxy for social skills	W	Hormone levels: cortisol, progesterone and cortisol/progesterone	Cortisol, progesterone and cortisol/progesterone hormone levels
Ajzenman et al. (2013)	Adaptive behavior, social communication skills, postural stability, functional participation	W	VABS, CACS, postural variables	Adaptive behavior composite and social and communication subscales of VABS; self-care, low demand leisure, and social interaction subscales of CACS; all postural stability variables
Jenkins et al. (2013)	Behavioral, affective, communication, and motor skills	B	CBCL-TRF, video coding of affect, problem behavior, and communication skills	No effect on any variables
		W	Behavioral video coding of posture	Improvements in posture
Ward et al. (2013)	Social communication, behavioral, and sensory processing skills	W	GARS, SPSC	Autism index and social interaction scales of GARS; registration, sensitivity, school factor 1, school factor 4, auditory, visual, and touch subscales on SPSC
Garcia-Gomez et al. (2014)	QOL and behavioral skills	B	QOLQ	Interpersonal relationship and social inclusion subscales of QOLQ
		W	BASC	Aggressiveness subscale of BASC
Lanning et al. (2014)	QOL	B	Pediatric QOL, CHQ	No improvement on any variables
		W		Physical health summary score, psychosocial health summary score, social functioning, emotional functioning, and physical functioning subscales of pediatric QOL
Gabriels et al. (2015)	Social communication, adaptive behavior, motor, and sensory processing skills	B	PPVT, VABS, ABC-C, SALT, SIPT, BOT, SRS	Irritability and hyperactivity subscales of ABC-C; social cognition and social communication scales of SRS; # of words and # of different words on SALT
Anderson and Meints (2016)	Adaptive behavior, social communication, problem solving, and affective skills	W	ASQ, VABS, EQ, SQ, EQ/SQ	ASQ, EQ, maladaptive behavior composite of VABS

Table 4 (continued)

First author, year	Domains/variables tested	Type of effect (B/W)	Measures	Measures/variables showing improvement
Borgi et al. (2016)	Adaptive behavior, social communication, motor, and executive functioning skills	B	VABS, TOLT	Social, communication, and daily living subscales of VABS; planning time, execution time, total time, # of correct solutions, and # of rule violations on the TOLT
		W		Social and motor subscales of VABS; planning time on TOLT
Llambias et al. (2016)	Social skills	W	Engagement measured through behavioral video coding	Engagement

B between-group effect, *W* within-group effect, *CI* confidence interval, *SRS* Social Responsiveness Scale, *SP* sensory profile, *BOT* Bruininks Oseretsky Test of Motor Proficiency, *TSIF* Test of Sensory Integration Function, *CARS* Childhood Autism Rating Scale, *TPCIS* Timberlawn Parent-Child Interaction Scale, *QOL ESS* quality of life enjoyment and satisfaction survey, *TSS* Treatment Satisfaction Survey, *SP* sensory profile, *ABC-C* Aberrant Behavior Checklist-Community, *VABS* Vineland Adaptive Behavior Scales, *SIPT* Sensory Integration and Praxis Tests, *PAC* Pedagogical Analysis and Curriculum, *CACS* Child Activity Card Sort, *CBCL-TRF* Childhood Behavior Checklist–Teacher Rating Form, *GARS* Gilliam Autism Rating Scale, *SPSC* Sensory Profile School Companion, *QOLQ* Quality of Life Questionnaire, *BASC* Behavior Assessment System for Children, *CHQ* Child Health Questionnaire, *PPVT* Peabody Picture Vocabulary Test, *SALT* Systematic Analysis of Language Transcripts, *ASQ* Ages & Stages Questionnaire, *EQ* Empathizing Quotient, *SQ* Systemizing Quotient, *TOLT* Tower of London Test of Executive Functioning, *MANOVA* multivariate analysis of variance

(Bass et al. 2009; Wuang et al. 2010), whereas the remaining 2 studies provided level III evidence (Kern et al. 2011; Ward et al. 2013). Although the same outcome measure was used by Bass et al. (2009) and Kern et al. (2011), the former study reported positive effects of 12 weeks of equine therapy whereas the latter found no changes in the SP following 24 weeks of equine therapy (see Tables 4 and 5). Furthermore, although the Bass et al. (2009) study was an RCT, the authors did not calculate between-group ES (see Table 5). Our own calculations suggested that ES varied from small to large across the different subscales of the SRS: overall score (ES = 0.46), sensory seeking (ES = 0.39), attention and distractibility (ES = 0.85), sensory sensitivity (ES = 0.48), and sedentary (ES = 0.59). In the Wuang et al. (2010) study, the authors employed a crossover design and found that a 20-week intervention involving simulated horseback riding led to large-sized improvements on all the subscales of the TSIF; our calculated between-group effect sizes (ES range = 1.78–5.29) corroborate the authors' claims. Out of the total 56 ES calculated from data presented in the 4 studies, 42 ES were significant (CI did not include 0, see Table 6). Overall, the review suggests promising *positive (small to large-sized) effects of equine therapies on sensory skills in ASD.*

Motor Skills Four out of the seven studies that assessed gross and fine motor skills suggested positive effects following equine therapy (see Table 6). To evaluate motor skills, 4 studies used standardized tests (such as the Bruininks Oseretsky Test of Motor Proficiency (BOT), the Sensory Integration and Praxis Tests (SIPT), and the motor subscale of the VABS), 1 study used force plates to assess postural sway during quiet stance (Ajzenman et al. 2013), 1 study used kinematic measures to assess gait of subjects (Steiner et al., 2012), and 1 study used video coding (Jenkins & Reed, 2013) to assess children's ability to maintain upright posture while seated on a horse (see Tables 4 and 6 for details). Five of these 7 studies qualified as level I or II evidence (see Tables 2 and 6). Amongst the methodologically strong studies, 2 studies reported improved motor skills following equine intervention (Wuang et al. 2010; Steiner et al., 2012), whereas the remaining 3 did not find robust between-group differences in motor skills following equine therapy sessions (Gabriels et al. 2012, 2015; Borgi et al. 2016). Out of the studies that supported the use of equine-assisted activities in ASD, we could corroborate the findings of large ES as reported by Wuang et al. (2010) on the BOT test. In contrast, although Steiner et al. (2012) reported positive effects of THR on gait of subjects, they did not report actual ES or include data that would allow independent calculation of ES. Although both the remaining level III and IV studies in this group suggested positive effects of equine therapy on motor skills, they either did not allow independent calculation of ES (Jenkins & Reed, 2013) or reported a wide range of ES across dependent variables (Ajzenman et al. 2013). Although 43 of the 44 calculated ES in this set of studies reported ES that were significantly different from 0 at a 95% CI, note that the Wuang study alone contributed to around 40 ES estimates (Table 6). Taken together, presently there is only *weak evidence for positive treatment effects (varying in magnitude from small to large) of equine therapy on motor skills.*

Functional Participation and Quality of Life Measures Four studies (2 level II and 2 level III evidence) evaluated changes in children's functional participation and quality of life (QOL)

Table 5 Study-wise list of reported and calculated effect sizes

First author, year	Reported ES (Y/N)	Type of effect (B/W)	Magnitude of reported ES	Magnitude of calculated ES	CI range for ES	# of ES per measure where CI does not include 0	Comments
Bass et al. (2009)	N	B	None	SRS (2): 0.38–0.45; SP (5): 0.39–0.85	SRS (2): − 1.13 to 0.3; SP (5): − 0.29 to 1.56	SRS: 0; SP: 1 (inattention/ distractibility subscale)	Although the study has a between-group design, they did not report between-group ESs. Instead they calculated within-group ESs for each group and tested the significance of these effects using dependent *t* tests within each group. ESs are also not reported for all measures and across both groups. Poor agreement between reported and calculated ESs
	Y	W	SRS (1): 0.66; SP (1): 0.059	SRS (1): 0.31; SP (1): 0.54	SRS (1): − 0.8 to 0.18; SP (1): 0.03 to 1.06	SRS: 0; SP: 1	
Wuang et al. (2010)	Y	B	BOT and TSIF: all effects were reported to be large using partial eta-squared	BOT (16): 0.56–7.98; TSIF (14): 1.78–5.29	BOT (16): 0.04 to 9.5; TSIF (14): 1.18 to 6.35	BOT: 16; TSIF: 14	This study did not report actual values of between-group effect sizes using Cohen's *d*. They assessed magnitude of between-group effects using partial eta-squared values obtained from their MANOVA run. Fairly good agreement between reported and calculated within-group effect sizes.
	Y	W	BOT (22): 4.50–13.31, retention of gains at follow-up; TSIF (14): 5.08–15.40, retention of gains at follow-up	BOT (16): 4.39–9.60, retention of gains at follow-up; TSIF: (14): 3.73–15.00, retention of gains at follow-up	BOT (16): 3.12 to 10.17; TSIF: (14): 2.64 to 19.15	BOT: 16; TSIF: 14	
Kern et al. (2011)	Y	W	CARS(2): 0.36–0.50; No calculations provided for QOL ESS	CARS (2): 0.36–0.50; Insufficient information to calculate ES for QOL ESS	CARS (2): − 0.81 to 0.098	CARS: 0	Excellent agreement between reported and calculated ESs
Gabriels et al. (2012)	N N	B W	None	ABC-C (4): 0.69–0.87; ABC-C (4): 0.63–0.81; VABS (4): 0.19–0.35; BOT (1): 0.50; SIPT (2): 0.38–0.45	ABC-C (4): 0.11 to 1.47; ABC-C (4): − 0.29 to − 1.17; VABS (4): − 0.05 to 0.68; BOT (1): 0.15 to 0.85; SIPT (2): 0.04 to 0.79	ABC-C: 4 ABC-C: 4; VABS: 1 (adaptive total score); BOT: 1; SIPT: 2	Good triangulation of ES estimates using means and SDs, *t* values, and *p* values
Steiner et al. (2012)	N N	B W	None	Data not available to make calculations	Not applicable	Not applicable	Study does not report any data. Only figures provided. Also the type of stats used were mentioned, no further details of statistical results provided
Tabares et al. (2012)	N	W	None	Hormonal changes (6): 0.33–2.09 2 ES could not be calculated due to insufficient data	Hormonal changes (6): − 2.09 to 3.81	Hormonal changes: 2 (cortisol and progesterone levels from pre 1st session to post 1st session)	It is not exactly clear from the report how certain variables were calculated for statistical analysis. Also there are some discrepancies between values reported in the text versus in one of the Figures and the abstract
Ajzenman et al. (2013)	Y	W	VABS: social (2): 0.35–0.36, communication (2): 0.47–1.19, ABC (1): 0.39; CACS (3): 0.62–0.91;	VABS: social (2): 0.32–0.34, communication (2): 0.41–1.01, ABC (1): 0.37; CACS (3): 0.44–0.78; postural variables: reliable calculation not possible	VABS: social (2): − 0.76 to 1.42, communication (2): − 0.7 to 2.46, ABC (1): − 0.73 to 1.46; CACS (3): − 0.5 to 2.08;	VABS: 0, CACS: 0;	Good agreement between reported and calculated ESs for VABS; Some discrepancies for CACS; ESs could not be reliably computed and corroborated for motor variables due to insufficient data

Table 5 (continued)

First author, year	Reported ES (Y/N)	Type of effect (B/W)	Magnitude of reported ES	Magnitude of calculated ES	CI range for ES	# of ES per measure where CI does not include 0	Comments
			postural variables (9): 0.2–1.91	due to lack of Means and SDs			
Jenkins et al. (2013)	N	B	None	Data not available to make calculations	Not applicable	Not applicable	–
	N	W	None				
Ward et al. (2013)	N	W	None	GARS (6): 0.35–0.51; SPSC (15): 0.24–0.67	GARS (6): − 0.99 to 0.12; SPSC (15): − 0.22 to 1.18	GARS: 1 (social interaction); SPSC: 7	–
Garcia-Gomez et al. (2014)	Y	B	QOLQ (2): 2.05–2.43	QOLQ (2): 1.95–2.31	QOLQ (2): 0.75 to 3.57	QOLQ: 2	Fairly good agreement between reported and calculated ESs
	Y	W	BASC (1): 0.22	BASC (1): 0.21	BASC (1): − 1.04 to 0.63	BASC: 0	
Lanning et al. (2014)	N	B	None	No significant effects	Not applicable	Not applicable	–
	N	W		Pediatric QOL (5): 0.62–0.68	Pediatric QOL (5): − 0.17 to 1.48	Pediatric QOL: 0	
Gabriels et al. (2015)	Y	B	ABC-C (2): 0.5–0.53; SRS (2): 0.41–0.63; SALT (2): 0.54	ABC-C (2): 0.40–0.43 SRS (2): 0.44–0.62; SALT (2): 0.21–0.30	ABC-C (2): − 0.8 to − 0.03 SRS (2): − 0.99 to − 0.07; SALT (2): − 0.15 to 0.66	ABC-C: 2; SRS: 2; SALT: 0	The study reported between-group ESs only. Good agreement between reported and calculated ES estimates for SRS, but some discrepancies for ABC-C and SALT, probably due to a different method of ES calculation used in the study
Anderson and Meints (2016)	N	W	None	ASQ (1): 0.06; VABS (1): 0.11; EQ (1): 0.12	ASQ (1): − 0.6 to 0.49; VABS (1): − 0.66 to 0.44 EQ (1): − 0.43 to 0.67	ASQ: 0; VABS: 0; EQ: 0	–
Borgi et al. (2016)	N	B	None	VABS (2): 0.14–0.4; TOLT (1): 0.76	VABS (2): − 0.7 to 1.5; TOLT (1): − 1.54 to 0.02	VABS: 0; TOLT: 0	Problems in triangulating calculated values of ES using means and SEs and F values. The paper reported that sample sizes were different for each of the outcomes measured. These differential sample sizes were used for our ES calculations. But probably the paper used some method of data imputation prior to statistical testing, which was not clearly specified in the paper. Hence, possibly the difficulties in triangulation
	N	W		Data not provided to calculate within-group ES for 3 variables on VABS and 5 variables on TOLT	Not applicable	Not applicable	
Llambias et al. (2016)	Y	W	Engagement: 100% IRD for all 7 subjects	No data provided to confirm results	Not applicable	Not applicable	–

Effect size (ESs) have been calculated and reported only for variables and measures where significant effects were reported in the original study. Effect sizes have been reported as absolute values in terms of magnitude only (for instance, for some variables a negative ES implies improvement). The numbers in parentheses reported next to the measure in columns 3 and 4 indicate the number of effect sizes calculated per measure. In case of multiple effect sizes calculated per measure, an ES range has been reported. Whenever possible, an attempt was made to triangulate ES values (for instance, if the study provide means and SDs as well as change values from pre- to post-intervention, ESs were calculated using both methods and checked for agreement). In addition, we have also reported confidence intervals (CI) for effect sizes. In cases where multiple effect sizes have been calculated per outcome measure, CI ranges have been reported. Lastly, we have also mentioned the number of statistically significant effect sizes (CI do not include 0) per outcome measure for each study

B between-group effect, *W* within-group effect, *CI* confidence interval, *SRS* Social Responsiveness Scale, *SP* sensory profile, *BOT* Bruininks Oseretsky Test of Motor Proficiency, *TSIF* Test of Sensory Integration Function, *CARS* Childhood Autism Rating Scale, *TPCIS* Timberlawn Parent-Child Interaction Scale, *QOL ESS* quality of life enjoyment and satisfaction survey, *TSS* Treatment Satisfaction Survey, *SP* sensory profile, *ABC-C* Aberrant Behavior Checklist-Community, *VABS* Vineland Adaptive Behavior Scales, *SIPT* Sensory Integration and Praxis Tests, *PAC* Pedagogical Analysis and Curriculum, *CACS* Child Activity Card Sort, *CBCL-TRF* Childhood Behavior Checklist–Teacher Rating Form, *GARS* Gilliam Autism Rating Scale, *SPSC* Sensory Profile School Companion, *QOLQ* Quality of Life Questionnaire, *BASC* Behavior Assessment System for Children, *CHQ* Child Health Questionnaire, *PPVT* Peabody Picture Vocabulary Test, *SALT* Systematic Analysis of Language Transcripts, *ASQ* Ages & Stages Questionnaire, *EQ* Empathizing Quotient, *SQ* Systemizing Quotient, *TOLT* Tower of London Test of Executive Functioning, *MANOVA* multivariate analysis of variance

Table 6 Domain-specific effects of hippotherapy

Domains/variables assessed	Measures used	Studies using the measure	# of studies with significant effects	# of studies with non-significant effects	ES magnitude (total # of ES calculated: B and W)	# of ES where CI does not include 0
Social communication	SRS, VABS, ABC-C, CARS, GARS, SALT, PPVT, ASQ, PAC, TPCIS, behavioral coding	Bass et al. 2009, Kern et al. 2011; Jenkins et al., 2013; Ward et al. 2013; Gabriels et al., 2012, 2015; Anderson and Meints 2016; Borgi et al. 2016; Llambias et al. 2016; Steiner et al., 2012; Ajzenman et al. 2013	9	2	Small to large (21: 6B and 15W)	3: 2B and 1W
Behavioral	BASC, ABC-C, VABS, EQ, SQ, CBCL-TRF, CARS, GARS, behavioral coding	Garcia-Gomez et al. 2014; Gabriels et al., 2012, 2015; Jenkins et al., 2013; Anderson and Meints 2016, Ward et al. 2013; Kern et al. 2011	5	2	Small to large (19: 6B and 13W)	11: 6B and 5W
Sensory	SP, TSIF, SPSC	Bass et al. 2009; Wuang et al. 2010; Kern et al. 2011; Ward et al. 2013	3	1	Small to large (56: 19B and 37W)	42: 14B and 28W
Motor	BOT, PAC, VABS, postural and gait variables, behavioral coding	Wuang et al. 2010; Gabriels et al., 2012, 2015; Ajzenman et al. 2013; Jenkins et al., 2013; Steiner et al., 2012; Borgi et al. 2016	4	3	Small to large (44: 16B and 28W)	43: 16B and 27W
Cognition/executive functioning	TOLT	Borgi et al. 2016	1	–	Medium (1: 1B)	0
Functional participation	CACS	Ajzenman et al., 2013;	1	–	Medium to large (3: 3W)	0
QOL	Pediatric QOL, CHQ, QOL ESS, custom-developed	Kern et al. 2011; Garcia-Gomez et al. 2014; Lanning et al. 2014;	2	1	Medium to large (7: 2B and 5W)	2: 2B
Physiological	Cortisol and progesterone levels	Tabares et al. 2012	1	–	Small to large (6: 6W)	2: 2W

In between-group studies, if for a specific variable, significant within-group effects were reported in the study in the absence of significant between-group effects, the study has been categorized as not demonstrating a significant effect for that variable/outcome (the effect was not robust enough to be detected at a between-group level). The qualitative effect size magnitude classification reported in the last column is based on the ESs reported in the original papers (wherever applicable); if unavailable in the original report, our calculated ESs have been used for the coding. ES magnitude coding: small: 0.2–0.49, medium: 0.50–0.79, large ≥ 0.8. We have also provided the # of effect sizes (between-group and within-group) where the CI does not include 0

B between-group effect, *W* within-group effect, *CI* confidence interval, *SRS* Social Responsiveness Scale, *SP* sensory profile, *BOT* Bruininks Oseretsky Test of Motor Proficiency, *TSIF* Test of Sensory Integration Function, *CARS* Childhood Autism Rating Scale, *TPCIS* Timberlawn Parent-Child Interaction Scale, *QOL ESS* quality of life enjoyment and satisfaction survey, *TSS* Treatment Satisfaction Survey, *SP* sensory profile, *ABC-C* Aberrant Behavior Checklist-Community, *VABS* Vineland Adaptive Behavior Scales, *SIPT* Sensory Integration and Praxis Tests, *PAC* Pedagogical Analysis and Curriculum, *CACS* Child Activity Card Sort, *CBCL-TRF* Childhood Behavior Checklist–Teacher Rating Form, *GARS* Gilliam Autism Rating Scale, *SPSC* Sensory Profile School Companion, *QOLQ* Quality of Life Questionnaire, *BASC* Behavior Assessment System for Children, *CHQ* Child Health Questionnaire, *PPVT* Peabody Picture Vocabulary Test, *SALT* Systematic Analysis of Language Transcripts, *ASQ* Ages & Stages Questionnaire, *EQ* Empathizing Quotient, *SQ* Systemizing Quotient, *TOLT* Tower of London Test of Executive Functioning, *MANOVA* multivariate analysis of variance

following short-term equine therapy interventions (see Table 6 for details). In terms of functional participation, following a 3-month HIP intervention, children increased their participation (ES range: 0.62–0.81) in age-appropriate leisure and self-care activities on the Child Activity Card Sort test (Ajzenman et al. 2013) (Note: our ES estimates based on data from the report are more conservative and CI of ES include 0, see Tables 5 and 6). In terms of QOL, Lanning et al. (2014) and Kern et al. (2011) used standardized measures such as the Pediatric Quality of Life Generic Core Scales (PedsQL), the Child Health Questionnaire (CHQ), and the General Activities Subscale of the Quality of Life Enjoyment and Satisfaction Questionnaire (QOL ESS), whereas Garcia-Gomez et al. (2014) developed a custom questionnaire to assess QOL. While large improvements (ES range = 2.05–2.43) were found on the custom-developed QOL questionnaire following a 3-month THR intervention, the other two studies reported lack of definitive improvements in QOL of subjects due to the equine interventions provided (Lanning et al. 2014; Kern et al. 2011). Moreover, out of the total 7 ES calculated from these studies, the

Table 7 Types of assessors/raters and methods for blinding

First author (year)	Type of assessment	Blinding of assessor (s)	Raters/informants	
			Authors/ research staff	Parent/ teacher
Bass et al. (2009)	Parent questionnaires	None specified		x
Wuang et al. (2010)	Clinical assessments	Blinded to groups	x	
Kern et al. (2011)	Observational assessment, questionnaires	None specified	x	x
Gabriels et al. (2012)	Questionnaires, parent interview, clinical assessments	Not blinded	x	x
Steiner et al. (2012)	Observational assessment, kinematic and kinetic gait analysis	None specified	x	x
Tabares et al. (2012)	Biochemical	None specified	x	
Ajzenman et al. (2013)	Parent questionnaires/interviews, motion/force analysis	Not blinded	x	x
Jenkins et al. (2013)	Teacher questionnaire, video recordings for behavioral coding	None specified	x	x
Ward et al. (2013)	Parent/expert questionnaire	None specified	x	x
Garcia-Gomez et al. (2014)	Parent questionnaires	None specified		x
Lanning et al. (2014)	Parent questionnaires	None specified		x
Gabriels et al. (2015)	Parent questionnaires, clinical assessments	Blinded to treatment	x	x
Anderson and Meints (2016)	Parent questionnaires, interviews	Blinded to study goals		x
Borgi et al. (2016)	Parent interview, clinical assessment	Testers blinded to group	x	x
Llambias et al. (2016)	Video coding of behaviors	Blinded to study goals	x	x

CI of only 2 ES did not include 0 (Table 6). It may well be the case that the relatively limited duration of interventions (12–24 weeks) provided by some of the studies in this group (Lanning et al. 2014; Kern et al. 2011) may have been insufficient to produce substantial changes in the QOL of the participants. To summarize, *the current state of literature in this field does not allow us to comment on the effect of equine therapy on QOL and functional participation of individuals with ASD.*

Other Skills Only 1 level I study (Borgi et al. 2016) assessed the effects of a 6-month equine therapy intervention on cognitive skills, specifically executive functioning and found that compared to a waitlist control group children with ASD reduced the latency of their first move (ES = 0.76, but CI includes 0) during a problem solving task following THR. Another study that assessed the effect of HIP on salivary cortisol and progesterone levels as a proxy for social skills found large reductions in cortisol levels (ES range = 0.96–2.09) and small to large increases in progesterone levels (ES range = 0.33–1.59) in 8 children with ASD (Tabares et al. 2012). However, we caution readers while interpreting the results of this study since it was methodologically poor (level III) and had inadequate reporting standards.

Long-Term Effects of Equine Therapies The utility of any therapy depends not just on its immediate effects but also more importantly on its long-term carryover effects and the ability to generalize learned skills to novel contexts. Interestingly, only 4 studies in this review assessed long-term effects of equine therapy and the generalization of treatment effects to settings not involving horses (Wuang et al. 2010; Llambias et al. 2016; Steiner and Kertesz 2015; Ward et al. 2013). Their results are mixed. Ward et al. (2013) and Steiner et al. (2012) found that although there were short-term benefits following equine therapy, children's behaviors returned to almost baseline levels following withdrawal of the intervention. On the other hand, Wuang et al. (2010) and Llambias et al. (2016) found sustained improvements in behaviors even after the completion of the equine therapy intervention. Overall, the conflicting evidence on long-term effects of equine-assisted therapies limits our ability to draw definitive conclusions regarding the sustenance of treatment effects following equine interventions in ASD.

Intervention Characteristics

The studies included in this review assessed the effects of relatively short-term equine-assisted interventions in children with ASD (Table 8). The mean duration in weeks of the delivered intervention was 12.67 (SD = 6.47 weeks, range = 4–25 weeks). Almost all studies provided sessions at a frequency of one session per week, with the exception of 2 studies that provided therapy twice a week (Garcia-Gomez et al. 2014; Wuang et al. 2010). The median session duration was 60 min ($N = 9$), with an average time of 61 min (SD = 34.7 min, range 30–180 min). The average number of sessions in the intervention was 14.8 (SD = 9.8 sessions, range 4–40 sessions).

Fourteen out of the 15 studies employed live horses for the purpose of intervention, whereas 1 study used innovative exercise equipment, Joba® that provided a simulated horse riding experience (Wuang et al. 2010). Out of the studies that used live horses, 11 provided therapeutic horseback riding (THR) treatment whereas the remaining 3 studies used hippotherapy (HIP) (see Table 2). Six out of the 11 THR studies and 2 out of the 3 HIP studies explicitly stated that the delivered intervention adhered to standard guidelines for equine-assisted activities as laid down by national-level governing organizations such as the Professional Association of Therapeutic Horsemanship (PATH) International, the Spanish Equestrian Federation, Italian Equestrian Federation, or the American Hippotherapy Association (see Table 1 for details). In terms of certification of instructors, 8 out of the 11 THR interventions and all 3 HIP studies stated that riding instructors were certified by national-level horsemanship organizations (see Table 1 for details). Out of the 14 studies that used live horses for therapy, 10 studies reported using assistants (trained personnel or parents) to guide/help children during therapy, 2 studies did not use assistants (Borgi et al. 2016; Llambias et al. 2016), and the remaining 2 studies provided no details on support staff (Tabares et al. 2012; Steiner and Kertesz 2015). However, only 2 studies that used assistants explicitly reported training them using an internally developed program specific to the riding facility (Anderson and Meints 2016; Gabriels et al. 2012).

Studies that provided THR typically focused on (1) horsemanship skills that aimed at developing the child's bond with their horse and (2) riding skills targeted towards enhancing children's motor, social communication, emotional, and cognitive skills (Table 8). Horsemanship skills involved grooming, feeding, taking care of, and leading horses. Riding skills include training children to mount and dismount from their horses, followed by more complex skills including walking, trotting, halting, steering, turning, and walking their horse around cones. In addition, a few studies (Bass et al. 2009; Ward et al. 2013; Borgi et al. 2016; Gabriels et al. 2012) also included horse-mounted games such as "Simon says," catch and throw, dropping a ring on a pole, and ball and cup games that focused on training desired skills within an enjoyable and interactive context. Typically, THR sessions started with some warm up exercises and activities to prepare the horse, followed by riding exercises/games on the horse, and finally leading the horse back to the stables and bidding farewell to the horse and the instructors. Amongst the HIP studies, Tabares et al. (2012) structured their therapy similar to a THR session. In contrast, Llambias et al. (2016) and Ajzenman et al. (2013) studies involved certified OTs using horses as an integral part of their therapy session to improve children's balance, postural control, motor planning and sequencing, fine motor, sensory regulation, and attentional skills. For instance, subjects were encouraged to maintain balance in different positions as the horse was led through different paths or participants practiced functional and social skills through games involving obstacle courses, ball catching,

Table 8 Recommendations for equine therapy treatment parameters

Characteristics	Recommendations for clinicians
Duration	30–60 min per session
Frequency	1–2 sessions per week
Time	3 to 6 months (1 month minimum)
Type	Therapeutic horseback riding or hippotherapy
Setting	Horse barn and nearby trails and/or therapy room for carryover activities
Environment	Outdoor and indoor environments
Providers	Certified riding instructors or OT/PT/SLP clinicians with hippotherapy certification
Assistants	2 side walkers, therapist/instructor could be horse leader
Components	- Warm up activities—whole body exercises or play activities
	- Greet horse and mount on horse using a mounting block
	- Travel while mounted—leader changes horse's pace/path or child's body position
	- Goal-oriented/functional activities on horse
	- Treat/groom/care for horse
	- Dismounting and carryover activities off-horse

carrying objects, and grabbing and dropping rings (Llambias et al. 2016 and Ajzenman et al. 2013). The only study that provided simulated riding experience using exercise equipment focused on training the child to ride on the horse simulator in seated and lying down positions and also engaging the child in games meant to facilitate their cognitive, affective, and social skills (Wuang et al. 2010).

Out of the 14 studies, in 6 studies, the intervention was tailored to the individual skills/needs of the child (Wuang et al. 2010; Gabriels et al. 2012; Jenkins & Reed 2013; Lanning et al. 2014; Anderson and Meints 2016; Llambias et al. 2016). For instance, the specific activities chosen or the format of the intervention (individual or group therapy) were tailored to the interests and abilities of the child. Finally, only 4 studies reported on procedures for ensuring treatment implementation fidelity across training sessions for study subjects (Gabriels et al. 2015; Jenkins & Reed 2013; Llambias et al. 2016; Wuang et al. 2010).

Discussion

Summary of Results

In the present review, we aimed to assess the quality and quantity of evidence in support of equine therapies as an adjunct therapy tool for individuals with ASD. While past reviews have provided a qualitative summary of the literature in this field, ours is the first review to report on the actual size of the treatment effects by calculating effect size estimates and their 95% CI wherever possible from data presented in individual study reports. We believe that a fair judgment of the merits of equine therapy as a treatment tool for ASD require consideration of the methodological quality of studies (PEDro rating and Sackett's level of evidence) in conjunction with the magnitude and direction of treatment effects on different outcomes.

Our review was based on 15 experimental and quasi-experimental studies that assessed the impact of THR and HIP interventions in 294 children and adolescents with ASD. In terms of methodological quality (Sackett et al. 1996), a little over half (8 out of 15) of these studies offered level I or II evidence, whereas the remaining qualified as level III or IV evidence. Eight of the studies designed their interventions by adhering to standard guidelines offered by national or international equine therapy certification agencies. Although most of the studies provided an adequate description of the study procedures including the setting of the study, personnel involved, and activities conducted, it was surprising that only about a quarter of the studies actually assessed treatment fidelity/integrity across training weeks. Similarly, only 4 of the studies assessed long-term benefits or generalization of training following short-term equine experiences in individuals with ASD.

Equine therapy by its very nature has the potential to impact multiple subsystems. Accordingly, we found that most studies assessed the effects of equine therapies on multiple skills in ASD. As is clear from Table 6, maximum work so far has been dedicated towards evaluating equine therapy effects on social communication, behavioral, and motor outcomes. Going by dual criteria of amount of positive evidence (# of studies demonstrating gains and # of significant non-zero ES) and strength of positive evidence (level I or II studies), our review suggests *consistent and reliable positive effects of short-term equine therapies on behavioral skills in ASD*. Although similar promising evidence is emerging with respect to the utility of equine therapies in enhancing social communication, sensory and perceptuo-motor skills in ASD, given the small number of studies (of high quality) in this area and somewhat conflicting results across studies using the same outcome measures (for example, Bass et al. 2009 and Kern et al. 2011 for the Sensory Profile), we see a clear need for rigorous research in these domains. Moreover, in case of sensory-motor skills, currently, the literature seems biased as a result of the few studies that calculated multiple ES (Wuang et al. 2010; Ward et al. 2013). Lastly, the current state of evidence on equine therapies does not allow us to make any claims with respect to the effects of these therapies on cognitive skills, physiological variables, functional skills or QOL of individuals with ASD.

Clinical Implications

In this section, based on our review, we provide recommendations pertaining to assessment and treatment, for clinicians working with children and adolescents with ASD. In terms of assessment measures, a majority of studies used caregiver-rated questionnaires followed by standardized tests to assess dependent variables (Table 7). In a related vein, some studies used broad assessment tools that evaluated treatment effects on multiple developmental domains, whereas others used tools specific to the domain assessed. The choice of instruments typically depends on multiple factors such as participants' functional skills including their ability to comprehend and comply with instructions, their ability to imitate others, and their motor and intellectual skill levels. Nevertheless, we recommend that wherever possible, clinicians use a combination of qualitative and quantitative tools (parent questionnaires, standardized tests, video coding) to assess the impact of equine-assisted activities in ASD. Similarly, by using both domain-general and domain-specific measures, clinicians will be able to gauge impact of the intervention on the overall development of the child as well as detect subtle and more-nuanced treatment effects within specific domains (see Table 6 for details).

In terms of treatment, our review suggests that in order to obtain appreciable behavioral changes in individuals with ASD, equine-based interventions should be provided for at least 1 month, although more robust improvements are seen only if the equine activities are continued for at least 3 to 6 months. A longer duration of intervention provides the individual sufficient time to get accustomed to the horse, overcome the anxiety associated with novel activities, and build a rapport with the horse. Interventions were mostly provided at the frequency of 1 session/week (with few studies providing 2 sessions/week) with each session ranging from 30 to 60 min in duration (see Table 8). Therapeutic activities practiced should focus on horsemanship skills, riding skills, as well as games and group activities that focus on addressing the core social communication and perceptuo-motor impairments in ASD. Given the novelty and dynamic nature of the activities involved, clinicians may need 1 to 3 trained volunteers to assist participants with mounting-dismounting, guiding, and steering the horse and also provide cues and prompts as needed during the session. Since ASD is a spectrum disorder and the challenges faced by individuals with this diagnosis varies significantly, we recommend that clinicians tailor the intervention to the specific needs of the individual, including the choice of activities, level of support provided, individual or group nature of the training, amount of practice provided, and the progression of activities during training.

Limitations and Future Directions for Research

Although we demonstrated preliminary support for the utility of equine therapy in ASD, our review highlights several lacunae in this literature that need to be addressed in future studies. There is a clear need for large sample size studies that employ narrow and clear inclusion criteria to study the impact of equine therapies on homogenous samples of individuals with ASD in terms of age range, diagnosis, adaptive behavior, and functional skills. Moreover, future studies need to ensure methodological rigor by using standardized tests for confirming participant diagnosis, employing well-matched control groups, random assignment of subjects, blinding of assessors, and strict controls on treatment fidelity. In order to strongly establish the utility of equine therapy in the treatment of ASD, more studies need to assess the short-term and long-term effects of equine therapy (THR and HIP) compared to conventional modes of therapy on multiple subsystems (social communication, sensory-motor, behavioral, and cognitive) and functional skills in individuals with ASD across the lifespan.

In the current review, we attempted to go beyond providing a qualitative summary of the current equine therapy literature by calculating quantitative ES estimates for each skill domain. We found that several studies did not report ES appropriate to their study design and also did not provide enough data to enable us to independently calculate ES. For some other studies that allowed independent ES estimation, there were conflicts between reported and calculated ES. Future studies should ensure adequate reporting of ES along with their 95 or 99% CI or at least provide necessary data to allow estimation of ES and their CI. Such reporting standards would enable a quantitative synthesis of the equine therapy literature in the form of a meta-analysis. A meta-analysis would be able to provide an aggregate estimate of the treatment effect in ASD and systematically account for the large heterogeneity in effect sizes (see Table 6) using moderator analyses.

Conclusions

Our literature review aimed at summarizing the evidence to date on the effects of equine-assisted activities and therapies including both THR and HIP in individuals with ASD. We searched several relevant health-related databases and finally selected and reviewed 15 articles assessing the impact of equine therapies on several different outcomes in children and adolescents with ASD. Although there was considerable variability between studies in terms of sample characteristics, intervention characteristics, outcomes and assessment tools used, as well as the magnitude and direction of treatment effects reported for different types of outcomes, our combined qualitative and quantitative analysis suggested promising immediate effects of short-term equine therapy interventions on behavioral skills in ASD. There is currently some limited evidence supporting the utility of equine therapies in enhancing social communication and perceptuo-motor skills in individuals with ASD. The effects of equine-assisted activities on cognitive and functional skills/QOL of individuals with ASD are at present poorly researched. Future studies should employ methodologically rigorous study designs with large sample sizes to evaluate the role of equine therapies in the standard-of-care treatment of individuals with ASD.

Funding DC and AB's work on this manuscript was supported by the Leadership Education in Neurodevelopmental and Related Disabilities (LEND) grant from the Maternal & Child Health Bureau (MCHB) of the Health Resources and Service Administration (HRSA) under grant number T73MC30116 (PI: Beth Mineo and Co-I: Brian Freedman). AB's work on this manuscript was also supported by an Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health under grant number U54-GM104941 (PI: Binder-Macleod).

Compliance with Ethical Standards

Conflict of Interest The author declare that they have no conflict of interest.

Appendix 1: Search Terms for Databases

("hippotherapy" OR "horse-riding" OR "therapeutic horseback riding" OR "horseback riding" OR "therapeutic riding"

OR "equine therapy" OR "pet therapy" OR "animal-assisted" OR "equine-assisted" OR "horse therapy" OR "equine movement" OR "equine facilitated therapy" OR "equine-assisted therapy"[MeSH Terms] OR "animal-assisted therapy"[MeSH Terms])

AND ("Autistic Disorder"[MeSH Terms] OR "autism" OR "autistic" OR "autism spectrum disorder" OR "ASD" OR "ASDs")

Appendix 2: Final Coding Sheet per Study

Coding Form: Effects of Equine Therapy on Individuals with Autism Spectrum Disorder: A Systematic Review.

References

Ajzenman, H. F., Standeven, J. W., & Shurtleff, T. L. (2013). Effect of hippotherapy on motor control, adaptive behaviors, and participation in children with autism spectrum disorder: a pilot study. *American Journal of Occupational Therapy, 67*(6), 653–663. https://doi.org/10.5014/ajot.2013.008383.

American Hippotherapy Association. *Hippotherapy as a treatment*. Available at: http://www.americanhippotherapy association.org/hippotherapy/hippotherapy-as-a-treatment-strategy/. Accessed June 16, 2017.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Association.

Anderson, S., & Meints, K. (2016). Brief report: the effects of equine-assisted activities on the social functioning in children and adolescents with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 46*(10), 3344–3352. https://doi.org/10.1007/s10803-016-2869-3.

Bass, M. M., Duchowny, C. A., & Llabre, M. M. (2009). The effect of therapeutic horseback riding on social functioning in children with autism. *Journal of Autism and Developmental Disorders, 39*(9), 1261–1267. https://doi.org/10.1007/s10803-009-0734-3.

Baranek, G., Parham, L., & Bodfish, J. (2005). Sensory and motor features in autism: assessment and intervention. In F. P. R. Volkmar, A. Klin, & D. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders* (pp. 831–857). Hoboken: Wiley.

Bedford, R., Pickles, A., & Lord, C. (2016). Early gross motor skills predict the subsequent development of language in children with autism spectrum disorder. *Autism Research, 9*(9), 993–1001. https://doi.org/10.1002/aur.1587.

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), 1–11. https://doi.org/10.1007/s10803-008-0593-3.

Bhat, A., Landa, R., & Galloway, J. C. (2011). Perspectives on motor problems in infants, children, and adults with autism spectrum disorders. *Physical Therapy, 91*(7), 1116–1129. https://doi.org/10.2522/ptj.20100294.

Bodfish, J. W., Symons, F. J., Parker, D. E., & Lewis, M. H. (2000). Varieties of repetitive behavior in autism: comparisons to mental retardation. *Journal of Autism and Developmental Disorders, 30*(3), 237–243. https://doi.org/10.1023/A:1005596502855.

Bondy, A., & Frost, A. (2003). Communication strategies for visual learners. In O. I. Lovaas (Ed.), *Teaching individuals with developmental delays: basic intervention techniques* (pp. 291–304). Austin: Pro-Ed.

Borgi, M., Loliva, D., Cerino, S., Chiarotti, F., Venerosi, A., Bramini, M., .. Cirulli, F. (2016). Effectiveness of a standardized equine-assisted therapy program for children with autism spectrum disorder. Journal of Autism and Developmental Disorders 46(1):1–9. doi: https://doi.org/10.1007/s10803-015-2530-6

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.).

Dawson, G., Toth, K., Abbott, R., Osterling, J., Munson, J., Estes, A., & Liaw, J. (2004). Early social attention impairments in autism: social orienting, joint attention, and attention to distress. *Developmental Psychology, 40*(2), 271–283. https://doi.org/10.1037/0012-1649.40.2.271.

Dewey, D., Cantell, M., & Crawford, S. G. (2007). Motor and gestural performance in children with autism spectrum disorders, developmental coordination disorder, and or attention deficit hyperactivity disorder. *Journal of the International Neuropsychological Society, 13*, 246–256. https://doi.org/10.1017/S1355617707070270.

Dziuk, M. A., Gidley Larson, J. C., Apostu, A., Mahone, E. M., Denckla, M., & Mostofsky, S. H. (2007). Dyspraxia in autism: association with motor, social, and communicative deficits. *Developmental Medicine and Child Neurology, 49*, 734–739. https://doi.org/10.1111/j.1469-8749.2007.00734.x.

Fournier, K. A., Hass, C. J., Naik, S. K., Lodha, N., & Cauraugh, J. H. (2010). Motor coordination in autism spectrum disorders: a synthesis and meta-analysis. *Journal of Autism and Developmental Disorders, 40*, 1227–1240. https://doi.org/10.1007/s10803-010-0981-3.

Gabriels, R. L., Agnew, J. A., Holt, K. D., Shoffner, A., Zhaoxing, P., Ruzzano, S., et al. (2012). Pilot study measuring the effects of therapeutic horseback riding on school-age children and adolescents with autism spectrum disorders. *Research in Autism Spectrum Disorders, 6*(2), 578–588.

Gabriels, R. L., Pan, Z., Dechant, B., Agnew, J. A., Brim, N., & Mesibov, G. (2015). Randomized controlled trial of therapeutic horseback riding in children and adolescents with autism spectrum disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 54*(7), 541–549. https://doi.org/10.1016/j.jaac.2015.04.007.

Garcia-Gomez, A., Risco, M. L., Rubio, J. C., & Garcia-Pena, I. M. (2014). Effects of a program of adaptive therapeutic horse-riding in a group of ASD children. *Electronic Journal of Research in Educational Psychology, 12*(1), 107–128.

Green, D., Charman, T., Pickles, A., Chandler, S., Loucas, T., Simonoff, E., & Baird, G. (2009). Impairment in movement skills of children with autistic spectrum disorders. *Developmental Medicine and Child Neurology, 51*(4), 311–316. https://doi.org/10.1111/j.1469-8749.2008.03242.x.

Greenspan, S. I., & Wieder, S. (1997). Developmental patterns and outcomes in infants and children with disorders in relating and communicating: A chart review of 200 cases of children with autistic spectrum diagnoses. *Journal of Developmental and Learning Disorders, 1*, 87–142.

Hallett, M., Lebiedowska, M. K., Thomas, S. L., Stanhope, S. J., Denckla, M. B., & Rumsey, J. (1993). Locomotion of autistic adults. *Archives of Neurology, 50*(12), 1304–1308.

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

Huedo-Medina, T. B., & Johnson, B. T. (2011). *Estimating the standardized mean difference effect size and its variance from different data sources: a spreadsheet*. Storrs: Authors.

Jansiewicz, E. M., Goldberg, M. C., Newschaffer, C. J., Denckla, M. B., Landa, R., & Mostofsky, S. H. (2006). Motor signs distinguish children with high functioning autism and Asperger's syndrome from controls. *Journal of Autism and Developmental Disorders, 36*(5), 613–621. https://doi.org/10.1007/s10803-006-0109-y.

Jenkins, S. R., & Reed, F. D. D. (2013). An experimental analysis of the effects of therapeutic horseback riding on the behavior of children with autism. *Research in Autism Spectrum Disorders, 7*(6), 721–740.

Kasari, C., Paparella, T., Freeman, S., & Jahromi, L. B. (2008). Language outcome in autism: randomized comparison of joint attention and play interventions. *Journal of Consulting and Clinical Psychology, 76*(1), 125–137. https://doi.org/10.1037/0022-006X.76.1.125.

Kern, J. K., Fletcher, C. R., Garver, C. R., Mehta, J. A., Grannemann, B. D., Knox, K. R., & Trivedi, M. H. (2011). Prospective trial of equine-assisted activities in autism spectrum disorder. *Alternative Therapies, 17*(3), 14–20 http://www.ncbi. nlm.nih.gov/pubmed/22164808.

Koegel, R. L., & Koegel, L. K. (2006). *Pivotal response treatments for autism: Communication, social, and academic development*. Baltimore, MD: Paul H Brookes Publishing.

Landa, R. J., Holman, K. C., O'neill, A. H., & Stuart, E. A. (2011). Intervention targeting development of socially synchronous engagement in toddlers with autism spectrum disorder: A randomized controlled trial. *Journal of Child Psychology and Psychiatry, 52*(1), 13–21.

Lanning, B. A., Baier, M. E., Ivey-Hatz, J., Krenek, N., & Tubbs, J. D. (2014). Effects of equine assisted activities on autism spectrum disorder. *Journal of Autism and Developmental Disorders, 44*(8), 1897–1907. https://doi.org/10.1007/s10803-014-2062-5.

Llambias, C., Magill-Evans, J., Smith, V., & Warren, S. (2016). Equine-assisted occupational therapy: increasing engagement for children with autism spectrum disorder. *American Journal of Occupational Therapy, 70*(6), 7006220040p1–7006220040p9. https://doi.org/10.5014/ajot.2016.020701.

Leary, M. R., & Hill, D. A. (1996). Moving on: autism and movement disturbance. *Mental Retardation, 34*, 39–53.

Lecavalier, L. (2006). Behavioral and emotional problems in young people with pervasive developmental disorders: relative prevalence, effects of subject characteristics, and empirical classification. *Journal of Autism and Developmental Disorders*, (8), 1101–1114. https://doi.org/10.1007/s10803-006-0147-5.

Lentini, J. A., & Knox, M. S. (2015). Equine-facilitated psychotherapy with children and adolescents: an update and literature review. *Journal of Creativity in Mental Health, 10*, 278–305. https://doi.org/10.1080/15401383.2015.1023916.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks: Sage Publications.

Loh, A., Soman, T., Brian, J., Bryson, S. E., Roberts, W., Szatmari, P., Smith, I. M., & Zwaigenbaum, L. (2007). Stereotyped motor behaviors associated with autism in high-risk infants: a pilot videotape analysis of a sibling sample. *Journal of Autism and Developmental Disorders, 37*(1), 25–36. https://doi.org/10.1007/s10803-006-0333-5.

Lovaas, O. I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology, 55*(1), 3–9.

MacDonald, M., Lord, C., & Ulrich, D. A. (2014). Motor skills and calibrated autism severity in young children with autism spectrum disorder. *Adapted Physical Activity Quarterly, 31*(2), 95–105. https://doi.org/10.1123/apaq.2013-0068.

McDaniel, P., & Wood, W. (2017). Autism and equine-assisted interventions: a systematic mapping review. *Journal of Autism and Developmental Disorders, 47*(10), 3220–3242. https://doi.org/10.1007/s10803-017-3219-9.

Moseley, A. M., Herbert, R. D., Sherrington, C., & Maher, C. G. (2002). Evidence for physiotherapy practice: a survey of the Physiotehrapy Evidence Database (PEDro). *The Australian Journal of Physiotherapy, 48*, 43–49.

Mazefsky, C. A., Pelphrey, K. A., & Dahl, R. E. (2012). The need for a broader approach to emotion regulation research in autism. *Child Development Perspectives, 6*, 92–97. https://doi.org/10.1111/j.1750-8606.2011.00229.x.

Mesibov, G. B., Shea, V., & Schopler, E. (2005). The TEACCH approach to autism spectrum disorders. Springer Science & Business Media.

Minshew, N. J., Sung, K. B., Jones, B. L., & Furman, J. M. (2004). Underdevelopment of the postural control system in autism. *Neurology, 63*(11), 2056–2061. https://doi.org/10.1212/01.WNL.0000145771.98657.62.

Mostofsky, S. H., Dubey, P., Jerath, V. K., Jansiewicz, E. M., Goldberg, M. C., & Denckla, M. B. (2006). Developmental dyspraxia is not limited to imitation in children with autism spectrum disorders. *Journal of the International Neuropsychological Society, 12*(3), 314–326. https://doi.org/10.1017/S1355617706060437.

Mundy, P., & Newell, L. (2007). Attention, joint attention, and social cognition. *Current Directions in Psychological Science, 16*, 269–274. https://doi.org/10.1111/j.1467-8721.2007.00518.x.

O'Haire, M. E. (2013). Animal-assisted intervention for autism spectrum disorder: a systematic literature review. *Journal of Autism and Developmental Disorders, 43*(7), 1606–1622. https://doi.org/10.1007/s10803-012-1707-5.

O'Haire, M. E. (2017). Research on animal-assisted intervention and autism spectrum disorder, 2012–2015. *Applied Developmental Science, 21*(3), 200–216.

Ozonoff, S., Pennington, B. F., & Rogers, S. J. (2006). Executive function deficits in high-functioning autistic individuals: relationship to theory of mind. *Journal of Child Psychology and Psychiatry, 32*(7), 1081–1105. https://doi.org/10.1111/j.1469-7610.1991.tb00351.x.

PATH International. PATH International home page. 2011. Available at: http://www.pathintl.org/. 2017. Accessed June 16, 2017.

Prizant, B. M. (2006). *The SCERTS model: A comprehensive educational approach for children with autism Spectrum disorders volume II Program Planning & Intervention*. Baltimore: Paul H. Brookes Publishing Company.

Rogers, S. J., & Dawson, G. (2009). *Play and engagement in early autism: The early start Denver model*. New York: Guilford.

Sackett, D., Rosenberg, J., Muir, G., Haynes, B., & Richardson, W. (1996). Evidence based medicine: what it is and what it isn't. *British Medical Journal, 312*, 71–72. https://doi.org/10.1136/bmj.312.7023.71.

Srinivasan, S., Kaur, M., Park, I., Gifford, T., Marsh, K., & Bhat, A. (2015). The effects of rhythm and robotic interventions on the imitation/praxis, interpersonal synchrony, and motor performance of children with autism spectrum disorder (ASD): a pilot randomized controlled trial. *Autism Research and Treatment, 2015*, 736516, 18 pages. https://doi.org/10.1155/2015/736516.

Steiner, H., & Kertesz, Z. (2015). Effects of therapeutic horse riding on gait cycle parameters and some aspects of behavior of children with autism. *Acta Physiologica Hungarica, 102*(3), 324–335. https://doi.org/10.1556/036.102.2015.3.10.

Sullivan, G. M., & Feinn, R. (2012). Using effect size—Or why the P value is not enough. *Journal of Graduate Medical Education, 4*(3), 279–282.

Tabares, C., Vicente, F., Sanchez, S., Aparicioc, S., Alejoc, S., & Cuberob, J. (2012). Quantification of hormonal changes by effects of hippotherapy in the autistic population. *Clinical Neurochemistry Neurochemical Journal, 6*(4), 311–316.

Tager-Flusberg, H. (1999). A psychological approach to understanding the social and language impairments in autism. *International Review of Psychiatry, 11*(4), 325–334.

Tomchek, S. D., & Dunn, W. (2007). Sensory processing in children with and without autism: a comparative study using the short sensory profile. *The American Journal of Occupational Therapy, 61*(2), 190–200.

Vilensky, J. A., Damasio, A. R., & Maurer, R. G. (1981). Gait disturbances in patients with autistic behavior: a preliminary study. *Archives of Neurology, 38*(10), 646–649.

Ward, S. C., Whalon, K., Rusnak, K., Wendell, K., & Paschall, N. (2013). The association between therapeutic horseback riding and the social communication and sensory reactions of children with autism. *Journal of Autism and Developmental Disorders, 43*(9), 2190–2198. https://doi.org/10.1007/s10803-013-1773-3.

Weise, C., Simpson, R., & Kumar, S. (2016). The effectiveness of equine-based therapy in the treatment of social and behavioural aspects of children with autism Spectrum disorder: a systematic review. *Internet Journal of Allied Health Sciences and Practice, 14*(3), 12.

Wieder, S., & Greenspan, S. I. (2003). Climbing the symbolic ladder in the DIR model through floor time/interactive play. *Autism, 7*(4), 425–435.

Williams, D. L., Goldstein, G., & Minshew, N. J. (2006). The profile of memory function in children with autism. *Neuropsychology, 20*(1), 21–29.

Wuang, Y. P., Wang, C. C., Huang, M. H., & Su, C. Y. (2010). The effectiveness of simulated developmental horse-riding program in children with autism. *Adapted Physical Activity Quarterly, 27*(2), 113–126.