Fra: Appia klinikken <info@appia-klinikken.dk>
Sendt: 14. september 2020 09:13
Til: SST NKR Sekretariat <NKRSekretariat@SST.DK>
Emne: Høringssvar på NKR Behandling af autismespektrumforstyrrelser hos børn og unge

Kære Arbejdsgruppe
Denne retningslinje er dejlig overskuelig og anvendelig. Den forekommer dog som kliniker at være begrænset i de områder som den behandler – i daglig praksis ville det være rart, hvis behandlingsforslag (også selvom de ville være med svag evidens) til hvordan man behandler problemstillinger som

- Behandling der rækker ud over psykiatrisk center
- At disse børn kun må følges af psykiater i kortere forløb af 20 sessioners varighed
- Identitetsdannelse i teenageårene, især hvis de unge først bliver diagnosticeret i denne alder
- Spiseforstyrrelser, overtræning, håndtering af uddannelsesmål i ungdomsuddannelser (både faglige og boglige)
- Opbygning af støtteforanstaltninger der tager udgangspunkt i barnets eller den unges vanskeligheder og ressourcer
- Samarbejde med uddannelsesinstitutioner, kommuner, støttekorps og andre professionelle
- Deltagelse i ungdomskulturen med hvad der deraf følger af rusmidler, sex mmm.
- Skift – fordi det er det der står øverst på ”livets” huskeseddel når man er mellem 15 og 25 år (Og jeg er klar over at denne retningslinje kun vedrører unge op til det 18. år)
- Overgangen fra indskoling (med voksenstyrede rammer) til mellemtrinnet, til udskolingen, til uddannelse udenfor specialskole/almen folkeskole ….
- Hjælp til at håndtere skolevægring
- Behov for særlige tiltag omkring eksaminationer
- At man med denne diagnose ikke kan blive kendt uddannelsesparat
- Handicap SU
- Hvordan den viden man har som psykiater efter udredning af et barn/en ung bliver anvendt i det videre arbejde udenfor lægens kontor
- At antallet af tvangsfjernelser af børn og unge med denne diagnose er steget signifikant gennem de senere år, uden undersøgelser af om ikke netop disse børn og unge er særligt sårbare overfor denne type voldsomme skift i deres liv

I håbet om at der kommer flere retningslinjer der peger på muligheder for samarbejde mellem Sundhedsstyrelsen og kommunerne for dette livslange handicap for at vi kan få gavn af disse børn og unges ressourcer.

Med venlig hilsen
Mie Bonde

Dansk Selskab for Almen Medicin

Sundhedsstyrelsen
Evidens, Uddannelse og Beredskab

16. september 2020

DSAM's høringssvar vedrørende udkast til National klinisk retningslinje for behandling af autismespektrumforstyrrelser hos børn og unge

DSAM takker for muligheden for at kommentere på ovennævnte udkast.

Umiddelbart bemærker vi, at der ved udvælgelsen af "nedslagspunkter" er valgt relevante områder, men stadig kun en lille del af de spørgsmål, der kunne være relevante at få belyst, når man skal behandle autismespektrumforstyrrelser.

Vi har ikke kommentarer til de enkelte anbefalinger, som vi tager til efterretning.

Med venlig hilsen

Anders Beich
Formand, DSAM

Stockholmsgade 55
2100 København Ø

T: 7070 7431
dsam@dsam.dk
www.dsam.dk

Sundhedsstyrelsen
Sekretariatet for Nationale Kliniske Retningslinjer
Islands Brygge 67
2300 København S

Sendt pr. mail til **NKRsekretariat@sst.dk**

Dato: 15. september 2020

Sags ID:
Dok. ID:

E-mail: RIL@kl.dk
Direkte: 3370 3238

Weidekampsgade 10
Postboks 3370
2300 København S

www.kl.dk
Side 1 af 2

Høringssvar fra KL vedr. udkast til national klinisk retningslinje for behandling af autismespektrumforstyrrelser hos børn og unge

KL har modtaget udkast til national kliniske retningslinje for behandling af autismespektrumforstyrrelser hos børn og unge i høring.

Overordnede kommentarer

KL hæfter sig ved, at det beskrives i indledningen, at en national klinisk retningslinje alene indeholder konkrete handlingsanvisninger indenfor udvalgte, velafgrænsede kliniske problemstillinger (dvs. hvad der skal gøres og hvem er det relevant for), samt at den ikke har som primært formål at afklare visitation og organisering af indsatsen (hvem der skal tilbyde indsatsen) eller samfundsøkonomiske konsekvenser (hvad er den afledte effekt på ressourcerne og er disse til stede).

KL anbefaler, at dette afklares i forlængelse af udsendelsen af retningslinjen.

Der har gennem en årrække været en kraftig stigning i antallet at børn og unge, der diagnosticeres med autismespektrumforstyrrelser. I samme periode er der ikke afsat ekstra midler til, at kommunerne kan udvide deres aktuelle ansvar for at tilbyde pædagogiske, psykologiske eller socialfaglige indsatser for målgruppen.

Det vil således ikke være muligt for kommunerne indenfor de aktuelle rammer at tilbyde flere indsatser til målgruppen, uden at der følger midler med til at finansiere disse.

Forskning vedr. autisme

Det fremgår af udkastet, at en stor del af de gennemgåede studier har en lav grad af reliabilitet, bl.a. på grund af en stor risiko for bias mv. og at evidensen for/imod de undersøgte interventioner på mange områder er lav. KL mener, at det er problematisk, at der tilsyneladende ikke gennemføres mere solid og troværdig forskning på området.

Som nævnt ovenfor får flere og flere børn, unge (og voksne) i dag en autismediagnose og der er derfor brug for solid og troværdig forskning, som kan pege på, hvilke interventioner, der bedst gavner målgruppen.

Dato: 15. september 2020

Sags ID:
Dok. ID:

E-mail: RIL@kl.dk
Direkte: 3370 3238

Weidekampsgade 10
Postboks 3370
2300 København S

www.kl.dk
Side 2 af 2

Herunder anbefaler KL, at der også iværksættes forskning vedr. effekten af pædagogiske, psykologiske og sociale tiltag til målgruppen, så kommunerne får bedre grundlag for at prioritere, hvilke indsatser, de bør tilbyde børnene, de unge og deres familier.

Intervention målrettet børnenes sprogudvikling
Udkastet indeholder en svag anbefaling om at overveje at tilbyde sprogintervention til børn i alderen 18 måneder - 6 år med autisme med forsinket eller afvigende sprogudvikling.

KL bemærker til dette, at børn med autisme indgår som alle andre børn i målgruppen for PPR's indsats, samt at de er omfattet af § 11 i dagtilbudsloven. Det betyder, at der i hvert enkelt tilfælde tages stilling til den nødvendige intervention. Her indgår en helhedsvurdering af barnets behov, der ikke kun tager udgangspunkt i barnets diagnose. KL ser ingen faglige argumenter eller anden evidens der tyder på at der er behov for anden specifik indsats ift. børn med autisme.

Neurofeedback
Udkastet indeholder en stærk anbefaling imod at tilbyde neurofeedback til børn og unge med autisme i aldersgruppen 3-17 år.

KL finder det problematisk, at det i anbefalingen ikke fremgår, hvilke neurofeedback-metoder der ligger til grund for anbefalingen. Der er omkring 20 forskellige neurofeedback-metoder, som ikke kan sammenlignes 1:1. Det giver derfor ikke mening at fraråde brugen af neurofeedback, uden at det er mere specifikt beskrevet, hvilke metoder, som konkret frarådes.

KL anerkender, at neurofeedback ikke fjerner kernesymptomer for ASF. Hvis det diagnostiske arbejde er troværdigt, vil symptomerne altid være til stede, men der kan trænes udvikling. Det er her, at nogle kommuner oplever, at neurofeedback har sin berettigelse.

Kommunerne oplever, at neurofeedback har effekt på stress, angst og depression som ofte er komorbide lidelser til autismespektrumforstyrrelser. Dermed har metoden også effekt på livskvaliteten for børn og unge med autismespektrumforstyrrelser.

KL anerkender, at der er behov for yderligere forskning på området, og at evidensen som sådan er svag. KL mener dog, at mærkningen med "stærk anbefaling imod" står i betydelig modsætning til de erfaringer kommunerne indtil videre har med metoderne. Kommunerne ser manglende evidens som en invitation til, at der foretages yderligere forskning i brugen af metoderne. Kommunerne bidrager gerne til denne forskning.

Med venlig hilsen

Janet Samuel

Socialstyrelsen

Enhed:
Center for Handicap
og Psykisk Sårbarhed

Sagsnr.: 2020-11197

Dato:
16. september 2020

Att.: Sundhedsstyrelsen
Evidens, Uddannelse og Beredskab
Sekretariatet for Nationale Kliniske Retningslinjer

Høringssvar vedr. den nationale kliniske retningslinje for behandling af autismespektrumforstyrrelser hos børn og unge

Socialstyrelsen takker for høringsversionen af Behandling af autismespektrumforstyrrelser hos børn og unge – National klinisk retningslinje, som Sundhedsstyrelsen har sendt i høring den 3. september 2020.

Socialstyrelsen ser den nationale kliniske retningslinje som et væsentlig bidrag og inspiration til behandling af autismespektrumforstyrrelser hos børn og unge.

I det følgende præsenteres først Socialstyrelsens generelle bemærkninger til den nationale kliniske retningslinje. Efterfølgende præsenteres Socialstyrelsens specifikke bemærkninger til de enkelte afsnit i retningslinjen.

Generelle bemærkninger til retningslinjen

Den nationale kliniske retningslinje giver god inspiration til interventioner og indsatser i forhold til behandlingen af børn og unge med autismespektrumforstyrrelser. Socialstyrelsen finder det ligeledes positivt, at der i hvert enkelt afsnit om de forskellige interventioner og indsatser også er afsnit om patientpræferencer og andre overvejelser.

Socialstyrelsen finder det positivt, at der er medtaget non farmakologiske behandlingsindsatser. Der kan med fordel lægges yderligere vægt på den kontekst, barnet og den unge befinder sig i, således at også fx forældrenes evner og mestring tillægges den betydning, det har for barnets udvikling.

I forhold til anvendelsen af sundhedsfaglige termer er det Socialstyrelsens vurdering, at det sociale perspektiv kan skrives tydeligere frem. Eksempler herpå kan være at erstatte ”patientforløb” med ”behandlingsforløb”, ”patientgruppe” kan skiftes ud til ”målgruppe”, og ”patienter” kan erstattes med ”børn og unge med autisme”.

Socialstyrelsen finder målgruppen og symptomerne på autismespektrumforstyrrelser velbeskrevet. Det bemærkes dog at indadreagerende adfærd og selvskadende adfærd i mindre grad er beskrevet og eksempelvis med fordel kan indgå i afsnit 2 s. 11 – *Indledning.*

Konkrete bemærkninger til afsnit i retningslinjen

- Vedr. afsnit 12 s. 65 – *Baggrund*
  Socialstyrelsen vurderer, at der også bør beskrives behandlingsindsatser målrettet den kontekst børn og unge med autismespektrumforstyrrelser er en del af, eksempelvis indsatser der styrker mestring og mestringsevne hos børn og unge med autismespektrumforstyrrelser og særligt forældrene.

  Socialstyrelsen finder det relevant at medtage nye tal for antallet af børn og unge med autisme. Tallene forefindes via Landspatientregisteret, og det bør undersøges, om der også kan være nyere og mere præcise tal for antallet af børn i skolealderen med autismespektrumforstyrrelser og mental retardering.

- Vedr. afsnit 5 s. 25-31 – *Indsatser til træning i sociale færdigheder til børn i alderen 18 måneder-17 år med autisme*
  Socialstyrelsen bemærker, at der i indledningen står at børn og unge med autisme er forstyrrede i det sociale samspil og kommunikation. Det forslås at der i stedet eksempelvis skrives, at det sociale samspil og kommunikation udfordrer børnene og de unge i samspillet med andre, eller at børn og unge med autisme har udfordringer med det sociale samspil og kommunikation.

- Vedr. afsnit 7 s. 37-42 – *Intervention målrettet sprogudviklingen til børn i alderen 18 måneder – 6 år med autisme*
  I denne beskrivelse af en intervention er fokus på sprog, sprogforståelse og sprogudvikling. I forlængelse af dette afsnit kan det tilføjes at kommunikation og kommunikationsredskaber er særlig vigtige, da kommunikative udfordringer ofte er en væsentlig del af autismediagnosen, og særligt når det gælder børn og unge med autismespektrumforstyrrelser uden sprog.

Fra: Henning Damgaard Laugesen <henning@dadlnet.dk>
Sendt: 17. september 2020 07:06
Til: SST NKR Sekretariat <NKRSekretariat@SST.DK>
Emne: Høringssvar til den National kliniske retningslinje "behandling af Autismespektrumforstyrrelser hos børn og unge"

Høringssvar til den National kliniske retningslinje "behandling af Autismespektrumforstyrrelser hos børn og unge"

Jeg er speciallæge i voksenpsykiatri.

Jeg har de seneste 20 år særligt interesseret mig for udviklingsforstyrrelser som ADHD såvel som Autismespektrumforstyrrelser.

Jeg har siden 2013 uddannet mig i og arbejdet med Neurofeedback, ikke kun, men også i forhold til ovennævnte særlige interesseområde. Blandt mere end 20 forskellige apparaturer og software til Neurofeedback valgte jeg Othmer-metoden. Med dette apparatur kan man træne med flere forskellige typer af feedbackmetoder fra den oprindelige beta og SMR over alfa-theta, forskellige former for synkronisitets træning til det, der er specielt for Othmer-metoden, som er træning af frekvenser under 1 Hz, også kaldet Slow cortical potentials, som styrer default mode network. Metoden tager ikke udgangspunkt i diagnoser men almindelige symptomer, som kan identificeres uden medicinsk baggrund. Den vigtigste forudsætning er at være i stand til at indgå i relationer til den person, der skal trænes således, at man kan identificere og følge symptomerne i træningsforløbet. Metoden er nonverbal og kan i samarbejde med forældre anvendes til selv meget små børn. Metoden er ikke invasiv.

Selv om der er talrige varige bedringer i funktion efter træning, er der ikke varige bivirkninger. Der kan være beskedent ubehag i forbindelse med træningen, dette anvendes til justering af træningsprotokollen.

Othmer metoden er udviklet over 40 år. Jeg ved ikke om, det er den bedste, men de resultater, vi ser, giver rigtig god mening.

Neurofeedback historie:

Siden Sterman i 1968 i sit forsøg med katte ved NASA viste effekt i forhold til Epileptiske anfald ved optræning af skiftevis Beta og SMR-aktivitet (Bilag 1) har evidensen for effekt af Neurofeedback hos mennesker konsolideret sig vedlagt Metaanalyse fra 2009, der viser 61 % anfaldsreduktion ved behandlingsrefraktær Epilepsi med effekt vurderet op til 10 år efter Neurofeedbacktræningen. (Bilag 2)

Neurofeedback er brugt i forhold til ADHD på forskellige måder siden slut 70-erne.

Neurofeedback anbefales af tyske børne-ungdoms- og voksenpsykiatere som en behandlingsmulighed, der kan iværksættes på et hvilket som helst tidspunkt fra tidlig barndom til sen voksenalder, ikke som erstatning for andre gode tiltag, men som supplement. Anbefalingen er på højeste niveau, A++(bilag 3)

I forhold til Autismespektrum vil jeg henvise til vedlagte review af relevant litteratur om anbefalede indsatser for Autismespektrum specifikt.

I forhold til Neurofeedback refereres til adskillige undersøgelser med forskellige træningsprotokoller og forskellige apparaturer, dog ingen undersøgelser af automatiske apparatur til hjemme og selvbehandling.

Ingen undersøgelser er dobbeltblindede, fordi Neurofeedback indtil 2009 kræver et tæt samarbejde mellem klient og træner.

Neurofeedback er virksomt overfor talrige af de udfordringer, der kan opstå, når hjerner kommer ud af balance og den individuelle træningsprotokol skal tilrettelægges med hensyntagen til den enkelte klients udfordringer. Det er af væsentlig betydning, om man træner et menneske med Autisme der har haft en traumatisk opvækst eller en harmonisk opvækst. Det er vigtigt at vide om der er tendens til migræne, allergi eller angst. Det er vigtigt, at der i træningsforløbet justeres for alle de udfordringer, der måtte komme frem særligt i forhold til Autisme, hvilket stiller store krav til træner, da mennesker med Autisme ikke altid rapporterer problemer, som derfor først bliver tydelige i træningsforløbet. (bilag 4)

Jeg har vedlagt et pilotstudie fra 2008 som i sin enkelhed viser tydelig effekt af Neurofeedback. Når man læser metodeafsnittet, er det tydeligt at mennesker med Autismespektrumforstyrrelser har forskellige udfordringer og skal trænes med forskellige protokoller, justeret i forhold til symptomer under forløbet. (bilag 5)

Jeg har fuld forståelse for, at man forsøger at udvikle udstyr, som ikke kræver tilstedeværelse af træner i de enkelte træningssessioner, da dette jo selvsagt er bekosteligt. En anden fordel er, at det muliggør dobbelt blinding og dermed adgang til en gylden videnskabelig standard, men at undersøgelser af **ét** nyt apparatur på baggrund af ikke højsignifikante resultater og bivirkninger skal tilsidesætte årtiers forskning, finder jeg dybt kritisabelt.

Venlig hilsen

Henning Laugesen
Speciallæge i Psykiatri
Misbrugspsykiatrisk Klinik
Fuglebakkevej 54
8210 Århus V
http://misb.dk

SOCIALPÆDAGOGERNE

NOTAT

Kontakt:
Sofie Thal Fisker
son@sl.dk
Dok.nr.
5753301
Sagsnr.
2020-SLCPP-01420

17-09-2020

Høringssvar vedr. den nationale kliniske retningslinje for behandling af autismespektrumforstyrrelser hos børn og unge

Socialpædagogerne ønsker at afgive høringssvar til Sundhedsstyrelsens ”Nationale kliniske retningslinje for behandling af autismespektrumforstyrrelser hos børn og unge”.

Der ses i disse år en stor stigning i antallet af børn og unge, som bliver diagnosticeret med autismespektrumforstyrrelser. Derudover viste en undersøgelse foretaget af VIVE i 2019, at en af kommunernes største udfordringer på det sociale område er børn og unge med autismespektrumforstyrrelser. I VIVE’s undersøgelse oplever kommunerne især, at de mangler virkningsfulde indsatser til målgruppen, ligesom de oplever udfordringer med at etablere et relevant og effektivt samarbejde med børne- og ungdomspsykiatrien omkring det enkelte barn eller unge.

Socialpædagogerne anerkender behovet for specifikke kliniske retningslinjer for behandlingsindsatsen i sundhedsregi i forhold til børn og unge med autismespektrumforstyrrelser. Socialpædagogerne vil dog samtidig gerne opfordre til, at de forskellige systemer taler bedre sammen for at imødekomme kommunernes behov for at være bedre klædt på til opgave og for at sikre sammenhæng og koordination i de indsatser, der alle i sidste ende har samme mål: at sikre en bedre trivsel og udvikling hos børn og unge med autisme og deres familier.

Ligeledes opfordrer Socialpædagogerne til, at man fremadrettet ser mere på virkningsfulde praksisser forbundet med et helhedssyn på barnet og familien frem for et snævert ’sygdomsblik’, særligt hvis disse retningslinjer også skal benyttes i samarbejde med forældre til at lægge en behandlingsplan.

Med venlig hilsen

Marie Sonne
Forbundsnæstformand

d. 17. sep. 2020

ABAforeningen

Intensiv pædagogisk indsats for børn med autisme

Tomas Bjørn Pedersen, landsformand
Lektor, cand.mag, Master i mediefagsdidaktik
Æblehegnet 7
8320 Mårslet
22 15 41 32 / 86 19 08 26
formand@abaforeningen.dk

Til Sundhedsstyrelsen

Hermed høringssvar om "Anbefalinger om behandling af autisme hos børn og unge".

I ABA-foreningen sætter vi stor pris på, at der udarbejdes nationale retningslinjer på området for børn og unge med autisme og vil gerne i den anledning kvittere for det store arbejde, som høringsmaterialet er et udtryk for.

I ABA-foreningen vil vi gerne have lov at stille spørgsmåltegn ved den overordnede inddeling i retningslinjerne, hvor KAT udgør én klassificering, mens forældremedierede indsatser udgør en fællesbetegnelse – for på den måde kommer forældremedierede indsatser til at fremstå mere mudrede og uigennemskuelige i det samlede billede.

Vi tillader os således at stille spørgsmåltegn ved, at der ikke i Sundhedsstyrelses anbefalinger henvises til det massive opbud af state-of-art metanalyser, som fastslår effekten ved tidlig intensiv adfærdsbaseret indsats (EIBI). Disse studier er almindeligvis ikke RCT-studier, men effekten er veldokumenteret internationalt og indsatsen betragtes som veletableret i henhold til almindeligt etablerede kriterier for evidensbaseret praksis (se bilag). Derfor er det uhensigtsmæssigt, at effekten af sådanne studierne ender med at gå helt tabt i de nationale retningslinjer. Effekten af EIBI er både dokumenteret i den internationale forskning, når den er udført forældremedieret, og når den er udført i ”community settings” som for eksempel dagtilbud.

Sundhedsstyrelsen skal være opmærksom på, at retningslinjerne i det daglige ikke kun vil blive anvendt indenfor Sundhedsstyrelsens egne kerneområder (sygehuse, lægehuse, klinikker etc.), men også vil blive læst af de pædagogiske miljøer på hele børn- og ungeområdet i landets kommuner – og af forældre til børn og unge med autisme.

ABA-foreningen savner grundlæggende, at Sundhedsstyrelsens anbefalinger vil kunne rette sig imod tidlig, metodebaseret indsats i landets daginstitutioner og skoler – og dersom det ikke er muligt for Sundhedsstyrelsen også at favne den målgruppe – så burde det forhold eksplicit indskrives i den endelige tekst. Et forslag:

> "Det er vigtigt at understrege, at disse retningslinjer fortrinsvis gælder lægefaglige tilgange til autisme hos børn og unge – ikke nødvendigvis pædagogiske tilgange og indsatser."

At Sundhedsstyrelsens anbefalinger desværre ikke har en bredere målgruppe end Sundhedsstyrelsens egen kerneområder ses eksempelvis også afspejlet i, at anbefalingerne afviger betragteligt fra Socialstyrelsens anbefalinger.[1]

---

[1] https://socialstyrelsen.dk/udgivelser/mennesker-med-autisme

Tomas Bjørn Pedersen, landsformand
Lektor, cand.mag, Master i mediefagsdidaktik
Æblehegnet 7
8320 Mårslet
22 15 41 32 / 86 19 08 26
formand@abaforeningen.dk

I ABA-foreningens optik burde anbefalingerne fra Sundhedsstyrelsen blive samstemt med Socialstyrelsens – landets styrelser bør samstemme sine anbefalinger, ellers risikerer anbefalingerne at føre til unødige misforståelser. Disse misforstålser kan dog rettes ved at definere anbefalingernes målgrupper helt tydeligt i teksten.

I ABA-foreningen vil vi gerne takke for at få lov at bidrage i forhold til denne høring, og vi vil gerne i den anledning anmode om at blive inviteret med i sådan et arbejde fremover, i givet fald det måtte blive relevant.

Med venlig hilsen

Tomas Bjørn Pedersen, landsformand for ABA-foreningen

American Psychological Association

Psychological Bulletin

ISSN: 0033-2909

2020, Vol. 146, No. 1, 1–29
http://dx.doi.org/10.1037/bul0000215

Project AIM: Autism Intervention Meta-Analysis for Studies of Young Children

Micheal Sandbank
The University of Texas at Austin

Kristen Bottema-Beutel and Shannon Crowley
Boston College

Margaret Cassidy, Kacie Dunham, and Jacob I. Feldman
Vanderbilt University

Jenna Crank and Susanne A. Albarran
The University of Texas at Austin

Sweeya Raj
Vanderbilt University

Prachy Mahbub
Mount Holyoke College

Tiffany G. Woynaroski
Vanderbilt University Medical Center

In this comprehensive systematic review and meta-analysis of group design studies of nonpharmacological early interventions designed for young children with autism spectrum disorder (ASD), we report summary effects across 7 early intervention types (behavioral, developmental, naturalistic developmental behavioral intervention [NDBI], TEACCH, sensory-based, animal-assisted, and technology-based), and 15 outcome categories indexing core and related ASD symptoms. A total of 1,615 effect sizes were gathered from 130 independent participant samples. A total of 6,240 participants, who ranged in age from 0–8 years, are represented across the studies. We synthesized effects within intervention and outcome type using a robust variance estimation approach to account for the nesting of effect sizes within studies. We also tracked study quality indicators, and report an additional set of summary effect sizes that restrict included studies to those meeting prespecified quality indicators. Finally, we conducted moderator analyses to evaluate whether summary effects across intervention types were larger for proximal as compared with distal effects, and for context-bound as compared to generalized effects. We found that when study quality indicators were not taken into account, significant positive effects were found for behavioral, developmental, and NDBI intervention types. When effect size estimation was limited to studies with randomized controlled trial (RCT) designs, evidence of positive summary effects existed only for developmental and NDBI intervention types. This was also the case when outcomes measured by parent report were excluded. Finally, when effect estimation was limited to RCT designs and to outcomes for which there was no risk of detection bias, no intervention types showed significant effects on any outcome.

Micheal Sandbank, Department of Special Education, The University of Texas at Austin; Kristen Bottema-Beutel and Shannon Crowley, Lynch School of Education and Human Development, Boston College; Margaret Cassidy, College of Arts and Sciences, Vanderbilt University; Kacie Dunham, Vanderbilt Brain Institute, Vanderbilt University; Jacob I. Feldman, Department of Speech and Hearing Sciences, Vanderbilt University; Jenna Crank and Susanne A. Albarran, Department of Special Education, The University of Texas at Austin; Sweeya Raj, College of Arts and Sciences, Vanderbilt University; Prachy Mahbub, Department of Neuroscience and Behavior, Mount Holyoke College; Tiffany G. Woynaroski, Department of Hearing and Speech Sciences, Vanderbilt University Medical Center.

Jenna Crank is a paid employee of a private early intervention clinic that provides NDBI services. Susanne A. Albarran is a Board Certified Behavior Analyst who receives payment for providing ABA therapy. This work was funded in part by Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) of the National Institutes of Health (U54HD083211; PI: Neul), the Training Program in Fundamental Neuroscience of the National Institutes of Health (T32 MH064913; Winder) and the National Institute on Deafness and Other Communication Disorders (1R21DC016144; Tiffany G. Woynaroski). The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies. Data used to generate results reported in this article is included in the online supplementary materials. We would like to acknowledge Matt Sandbank for his assistance in typing.

Correspondence concerning this article should be addressed to Micheal Sandbank, Department of Special Education, The University of Texas at Austin, 1912 Speedway STOP D5300, Austin, TX 78712. E-mail: msandbank@austin.utexas.edu

Public Significance Statement

This comprehensive meta-analysis of interventions for young children with autism spectrum disorder (ASD) suggests that naturalistic developmental behavioral interventions and developmental intervention approaches have amassed enough quality evidence to be considered promising for supporting children with ASD in achieving a range of developmental outcomes. Behavioral intervention approaches also show evidence of effectiveness, but methodological rigor remains a pressing concern in this area of research. There is little evidence to support the effectiveness of TEACCH, sensory-based interventions, animal-assisted interventions, and interventions mediated solely through technology at this time.

Keywords: autism spectrum disorder, intervention, meta-analysis, robust variance estimation, study quality

Supplemental materials: http://dx.doi.org/10.1037/bul0000215.supp

Autism spectrum disorder (ASD) is a relatively common neurodevelopmental disorder with a varied impact. Current prevalence estimates suggest that one in 59 meet the criteria for ASD, though this prevalence varies by sex, with males having a higher (approximately four times greater) likelihood of being affected (Baio et al., 2018). The diagnosis is primarily associated with core challenges in social communication, as well as restricted interests and repetitive behaviors and differences in sensory function (American Psychiatric Association [APA], 2013). Individuals with ASD, however, may also exhibit difficulty in a number of related areas, such as language, adaptive behavior, and academic achievement.

A substantial portion of autistic[1] individuals report drawing a sense of identity and empowerment from the diagnosis, and advocate for a neurodiversity conceptualization of ASD as a natural form of human difference (den Houting, 2019). Researchers have recently articulated a view of early intervention that is consistent with a neurodiversity framework (e.g., Fletcher-Watson, 2018). Specifically, early intervention services provided throughout childhood may support children with ASD in developing competencies that will allow them to navigate into adulthood in ways they see fit. At present, long-term life outcomes of autistic individuals vary widely. Though a number of individuals that receive early diagnoses go on to develop adaptive and communicative skills within the average range, most require at least some support, and many require substantial support into adulthood (Renty & Roeyers, 2006). Importantly, quality of life among autistic adults also varies between individuals (Howlin & Magiati, 2017). Improving the quality of intervention provided in early childhood may be one way to increase the likelihood that long term life-satisfaction is attainable for all autistic people.

Research on Interventions in Early Childhood

Common Intervention Recommendations

Recommendations abound regarding the nature and amount of intervention that should be provided to support development in children with ASD. Scholars and professionals have routinely asserted that intervention should be provided as early as possible, beginning at or even before diagnosis in toddlerhood or infancy; that intervention should be intensive (i.e., provided for 25–40 hr per week for over a year or longer); and that it should be comprehensive (i.e., targeting broader development rather than specific skills; Boyd, Odom, Humphreys, & Sam, 2010; Lord et al., 2001; Lovaas, 1987; McEachin, Smith, & Lovaas, 1993; Odom, Boyd, Hall, & Hume, 2010). These recommendations are motivated by the theory that interventions provided in early childhood are likely to yield the most optimal effects by capitalizing on the neuroplasticity of the developing brain (Dawson & Zanolli, 2003; Kolb & Gibb, 2011), and are rooted in early influential studies which suggested that intensive intervention yielded substantial cognitive gains, and that such gains varied according to age at the onset of intervention (e.g., Lovaas, 1987; McEachin et al., 1993). However, it is notable that some subsequent studies exploring putative predictors of treatment response have reported that age at intake was not significantly associated with intervention outcomes (e.g., Eikeseth, Klintwall, Jahr, & Karlsson, 2012; Eikeseth, Smith, Jahr, & Eldevik, 2007).

Types of Intervention Approaches

Several approaches to intervention aim to address the core and related challenges associated with ASD. These approaches vary in their underlying theories on the nature of ASD and development, as well as in their procedures and instructional modalities.

Behavioral approaches. Behavioral interventions were among the first developed and clinically tested approaches for improving outcomes for children with autism (Ferster & Demyer, 1962). These approaches are derived from operant learning theory and are characterized by the discrete presentation of information (i.e., a stimulus), the prompted exhibition of target responses (i.e., desired academic, adaptive, and communicative behaviors), and the provision of extrinsic positive reinforcement (e.g., edible treats, toys, stickers, etc.) in the presence of those responses. Target skills are chosen based on functional areas of child need. Skills tend to be initially targeted in highly structured interactions within isolated

---

[1] Though researchers and clinicians often feel more comfortable with and advocate for using person-first language such as “individuals with autism,” some autistic individuals and their parents have endorsed identity-first language that incorporates autism as a component of their identity over person-first language (Gernsbacher, 2017; Kenny et al., 2016). In this article, we flexibly use identity-first and person-first language to acknowledge the diversity of opinions on this issue within the broader autism community (see Robison, 2019).

clinical contexts (e.g., in the course of one-on-one interactions at a clinic with a therapist), but more natural settings and interaction partners (e.g., mainstream classrooms and other children) are gradually integrated as a child demonstrates progress. Initial studies suggested that early intensive behavioral intervention (EIBI) could yield marked improvements in cognitive and academic placement outcomes for children with ASD, especially when provided before school age and with sufficient intensity (Lovaas, 1987; McEachin et al., 1993). In the wake of such research, a number of behavioral approaches were further developed and refined, and the Behavior Analyst Certification Board (BACB) was established to oversee the clinical certification associated with this approach. Other behavioral interventions include discrete trial training (DTT), picture exchange communication system (PECS), and positive behavioral supports (PBS). Together, these interventions are sometimes loosely described as applied behavior analysis (ABA) therapy and now constitute the primary approach used in clinical practice, according to parent and provider reports (Green et al., 2006; Stahmer, Collings, & Palinkas, 2005).

Developmental approaches. At times viewed in contrast to the aforementioned traditional behavioral interventions are those derived from developmental theories of learning (e.g., Ospina et al., 2008; Prizant & Wetherby, 1998). Developmental interventions are rooted in constructivist theory, which posits that development is the result of children's active exploration of their physical and social surroundings. This exploration is far from being a solitary endeavor, as children are supported in social and language development by their interactions with more competent interaction partners such as caregivers (Vygotsky, 1978). Foundational research on ASD within the developmental tradition has suggested that early deficits in social processes (joint attention being of particular importance) in children with ASD may in turn lead to difficulties in early caregiver-child social interactions. These early deficits are thus viewed as disrupting the primary context for subsequent language and social communication development. As such, developmental interventions focus on improving the synchrony, reciprocity, and duration of parent–child or child–child interactions as a pathway for ameliorating deficits in social communication and generating cascading improvements in developmentally related skills. These interventions are primarily delivered in the context of everyday routines such as play, and intervention goals are chosen based on the typical sequences of social communication and language development. Examples of classically developmental interventions include DIR/Floortime (Greenspan & Wieder, 2007) and Hanen models (Carter et al., 2011).

Naturalistic developmental behavioral interventions (NDBIs). In 2015, several interventions were categorized as belonging to a third type of intervention approach which has theoretical underpinnings in both behavioral and developmental theories of learning and development. NDBIs involve the use of behavioral principles of learning to teach skills chosen from a developmental sequence in naturalistic environments and using natural rewards (Schreibman et al., 2015). Skills selected as relevant for intervention are those that allow the child to participate more fully within reciprocal interactions with the adult. These interventions are delivered primarily in the context of play, but control of interactions within this context is shared by both the child and the adult, through balanced turn-taking. Interventions categorized as NDBIs include the Early Start Denver Model (Rogers & Dawson, 2010); Enhanced Milieu Teaching (EMT; Kaiser, 1993); Pivotal Response Treatment (Koegel, Koegel, & Carter, 1999); and Joint Attention, Symbolic Play, Engagement, and Regulation (JASPER; Kasari, Freeman, & Paparella, 2006).

TEACCH. The TEACCH (Treatment and Education of Autistic and related Communication-Handicapped Children) program was developed in 1972 by Eric Schopler and is based primarily in the state of North Carolina (Mesibov, Shea, & Schopler, 2004). We consider this specific intervention as distinct from other approaches because of the explicit focus on structured environmental design and self-monitoring, which is not the emphasis of any of the other interventions of interest to the present synthesis. The theoretical foundations of TEACCH are rooted neither in behavioral nor in developmental theories of learning. Rather, TEACCH procedures were designed according to Schopler's theorized profile of the learning strengths, preferences, and needs of individuals with ASD, which include relative visual strength and comfort with consistent routines. Thus, the TEACCH program is characterized by highly structured work routines and a heavy reliance on the visual presentation of information. TEACCH "work systems" organize individual student tasks to visually convey four pieces of information: (a) what activity the student will complete, (b) how many items need to be completed, (c) how to identify when the work is finished, and (d) what will happen after task completion. TEACCH classrooms tend to feature carefully planned and structured environmental arrangements, work areas with minimal distractions, consistent routines, and the extensive use of visual schedules and supports.

Sensory-based interventions. Sensory-based interventions are motivated by the theory that sensory function is foundational in nature, and that sensory disruptions, particularly early in life, may produce cascading effects on development across a number of domains, ultimately yielding the constellation of core and related characteristics associated with ASD (e.g., Bahrick & Todd, 2012). Within this framework, it is hypothesized that targeted treatments may thus have the potential not only to ameliorate reported sensory differences, but also to translate to effects on higher-order social, communication, and cognitive skills in children with ASD (Cascio, Woynaroski, Baranek, & Wallace, 2016). The most well-known of these sensory-based approaches to treatment is sensory integration therapy, in which children are presented with a series of individualized sensory-motor experiences intended to build foundational skills that will facilitate their engagement and participation in a range of activities of daily living (Ayres, 1979, 2005). Other sensory-based interventions, as broadly conceptualized, may include activities such as brushing, swinging, the use of weighted vests and blankets to improve sensory processing, and music therapy and auditory integration training approaches that aim to scaffold motor, social, and emotional development (e.g., Baranek, 2002; Case-Smith & Arbesman, 2008). Sensory based approaches are most often provided by occupational therapists in clinical contexts but may also be delivered by caregivers, educators, and/or other service providers across a broader range of home and community settings.

Animal-assisted interventions. Animal-assisted interventions are those that rely on interactions with animals as the primary context for facilitating developmental change (e.g., O'Haire, 2013, 2017; Trzmiel, Purandare, Michalak, Zasadzka, & Pawlaczyk, 2019). In the ASD intervention literature, the intervention most prominently represented in this category is equine-assisted activities and therapy (EAAT; see Gabriels et al., 2012 for a review of related terminology). Proponents of EAAT contend that the activities of horseback riding and horse care provide a multisensory experience that allows children the opportunity to practice skills across multiple domains. More broadly, animal-assisted interventions are theoretically motivated by the possibility that human–animal interactions are highly motivating and provide calming contexts which may support improved psychological wellbeing and social function.

Technology-based interventions. Technology-based interventions employ one or more of a variety of technologies (e.g., computers, videos, video games, robots) as the primary medium for delivery of instruction. These interventions attempt to capitalize on the reported special interest that many autistic individuals have in computer technology (Grynszpan, Weiss, Perez-Diaz, & Gal, 2014) and predictable formats of information delivery (Baron-Cohen, Golan, & Ashwin, 2012), which allow users to control the pace of the interaction (Knight, McKissick, & Saunders, 2013). Examples of technology-based interventions include computer-assisted instruction and The Transporters DVD series (e.g., Young & Posselt, 2012).

Previous Syntheses of Intervention Literature

The National Professional Development Center (NPDC) on autism spectrum disorders generated a list of 27 evidence-based practices for improving outcomes in individuals with ASD, based on prior reviews of single subject and group design research (Wong et al., 2015). Similarly, the National Standards Project (National Autism Center, 2015) described 14 intervention practices as established for children with ASD and an additional 18 as emerging, based on a review of single subject and group design literature. In 2011, Warren et al. (2011) systematically reviewed 34 group design studies examining interventions in children with ASD. Notably, only two of the studies included in the review by Warren et al. (2011) were RCTs, and only one of those was rated as high quality. Very recent systematic reviews suggest the publication of RCTs has precipitously increased in ASD since the publication of the aforementioned synthesis by Warren and collaborators. For example, French and Kennedy (2018) systematically reviewed RCTs of interventions targeting any outcome in children with ASD below age 6, and found a total of 48 RCTs, 40 of which had been published since 2010.

Previous efforts to synthesize this literature have a number of shortcomings. First, NPDC and NSP review procedures attempted to synthesize evidence from RCTs, quasi-experimental studies, and single subject design studies (SSDs), when there is currently no agreed upon way of doing so. Though multiple methodologies can contribute to knowledge about effective practices, studies employing group designs, in particular high-quality RCTs, are the best equipped to control for alternative explanations and threats to internal validity. Syntheses that attempt to combine RCTs, quasi-experimental studies, and SSDs may overestimate the effectiveness of a given intervention approach. Inclusion of SSDs also limits the extent to which summary effects of intervention can be quantified with meta-analytic approaches. Though effect sizes that quantify change observed in SSDs have been proposed, many of these approaches fail to account for first order autocorrelation of data, ignore the logic of within study replication that is critical to interpretation of SSD data, and yield highly inflated and positively biased effect sizes which are not comparable with mean group differences that index treatment effects in group design (Wolery, Busick, Reichow, & Barton, 2010; Zimmerman et al., 2018).

Second, in previous reviews, limited consideration was given to the nature of outcomes measured. That is, prior syntheses of intervention literature have predominantly sought to ascertain whether various approaches to interventions are "evidence based," but they have largely failed to summarize the extent to which interventions effected meaningful change. Interventions that were shown to effect change that was overly specific to intervention targets were generally not distinguished from those that impacted scores on broader standardized assessments of developmentally advanced skills as administered by independent assessors. A synthesis is needed which asks not only "what works and for whom," but also, "for what?"

Third, none of the prior reviews seeking to synthesize effects for the broad range of interventions geared toward young children with ASD attempted to identify the summary effects of varied interventions on any outcomes using meta-analytic tools. Although a narrative synthesis approach allows for tallying the number of studies that have shown an effect for a given outcome, they do not allow for deriving an estimate of the combined magnitude of the effect, or determining whether or not the combined effect is significantly different from zero. Additionally, narrative synthesis methods are unable to offer information about variables that may moderate effect sizes. Moderator effects offer vital information for understanding for whom interventions are effective, and for identifying study design features that result in potentially inflated effect sizes.

Crucial Quality Considerations

Although systematic reviews and meta-analyses are purported to provide the most reliable summary of evidence of intervention effects, their conclusions are limited by the quality of evidence which they summarize (Higgins et al., 2011; Murad, Asi, Alsawas, & Alahdab, 2016). Several aspects of study design pose risk of biasing outcomes. Thus, examination of any set of intervention literature must include an assessment of several study-level quality indicators. We outline those that are particularly important in studies of nonpharmacological interventions of children with ASD.

Random Assignment

Though some have questioned the feasibility of conducting randomized controlled trials to test the effects of "real-world" interventions with individuals with disabilities (Oliver et al., 2002), random assignment remains the most rigorous control for rival explanations of findings. Though random assignment does not ensure pretreatment statistical equivalence between groups on all variables, it is the best procedural guard against systematic differences between groups that would limit confidence in conclusions about causal associations between the intervention and dependent variables (Kasari, 2002). Historically, randomized tests of interventions have been exceptionally rare in ASD research (Warren et al., 2011). However, the recent proliferation of RCTs in this field suggests that random assignment is feasible and employed frequently enough to permit an evaluation of evidence from randomized trials versus quasi-experimental studies.

Independence of Assessors

Detection bias refers to the risk of bias that arises when assessors are aware of the group assignment of individual participants. This type of bias manifests in different ways in studies of autism intervention, and the degree of risk may vary depending on the extent to which nonindependent assessors are involved in outcome assessment. It is likely that detection bias poses the greatest threat when caregivers participate in outcome assessment, either as reporters or interaction partners, though the threat is still substantial in situations wherein outcomes are assessed or coded by professionals that are aware of group assignment.

Caregiver/teacher report. It is common for researchers to rely on parents or teachers to assess outcomes via standardized interview and/or report forms in pediatric psychology and adjacent fields. Because caregivers observe and engage with children for extended periods of time across a variety of contexts, they can draw on their cross-context knowledge of a child's abilities when reporting on an outcome, and may therefore produce scores that are more representative of a child's generalized abilities, compared to scores derived from brief assessments administered by unfamiliar examiners. However, parents and teachers are virtually always aware of the extent and nature of a child's participation in an intervention study. Moreover, they are likely to be personally invested in the outcome of intervention. This combination of awareness of group assignment and strong investment in positive outcomes can yield a "placebo by proxy" effect, which can positively bias results in favor of the treatment group (Grelotti & Kaptchuk, 2011). Prior placebo-controlled studies of pharmacological interventions such as secretin have demonstrated that these effects can be rather large (Williams, Wray, & Wheeler, 2012), and present even in simulated clinical trials where no intervention was provided (Jones, Carberry, Hamo, & Lord, 2017). Thus, outcomes from caregiver report are highly subject to systematic measurement error and may positively bias summary estimates of intervention effects.

Outcomes assessed in interactions with caregivers. Even in situations that do not involve standardized report, caregivers can exert undue influence on outcome measurement. This occurs when caregivers participate as interaction partners in observational measures of outcomes of interest. Autism researchers frequently use observational measurement to capture social communication and related skills in the natural contexts in which they arise. For example, scores of language and communication are often derived from free play sessions with parents, or from interactions with teachers in the classroom. These scores are fundamentally dyadic; though they are often assumed to solely represent the skills or behavior of the child, they actually index the child's response to the interaction partner. When interaction partners are aware of the administration of a treatment, they may subconsciously or consciously shift their behavior to better elicit skill demonstration from the child. Though this threat arises often in studies of interventions on language and communication outcomes, it is not limited to measures of those domains. Therefore, outcomes measured in the context of natural interaction with caregivers are also subject to bias and may influence intervention effect sizes.

Outcomes assessed or coded by professionals aware of group assignment. Even unfamiliar professionals can influence outcomes when administering standardized assessments or coding observational measures of behaviors. A recent systematic review of medical literature that contained assessment of binary outcomes from both independent and nonindependent assessors found that assessors that were aware of group assignment exaggerated odds ratios as much as 36% (Hróbjartsson et al., 2012).

Influential Outcome Characteristics

The Cochrane Collaboration has delineated a set of quality indicators that are applicable to intervention literature in most fields, but additional field-specific sources of bias also exist for autism early intervention literature. Further, when it comes to studies of intervention for children with ASD, we contend that various aspects of outcome measurement can also serve as sources of bias and should therefore be considered (Yoder, Bottema-Beutel, Woynaroski, Chandrasekhar, & Sandbank, 2013). We summarize two particularly important dimensions of outcome variables below (boundedness and proximity), and we review one additional source of bias related to study design that we hypothesize has the potential to influence effect sizes observed across studies of treatment effects on outcomes of young children with ASD (correlated measurement error [CME] that arises when parents or teachers are trained in the intervention and then participate in the data collection).

Boundedness of Outcomes to Intervention Context

Whether or not an intervention effects change that generalizes beyond the context of an intervention is a question of great importance. While the context of intervention is generally contrived and temporary, changes effected by intervention are often assumed to (or at least intended to) extend to natural environments and the routines of daily life. However, dependent variables vary in the extent to which they index generalized change. Those that are measured within the context of intervention, or in a context that is similar to intervention across several dimensions (i.e., materials, setting, interaction partners, interaction style), may reflect changes that are potentially bound to the intervention context. In contrast, dependent variables that are measured in a context that differs from the intervention on several dimensions should reflect highly generalized changes. For example, in the hypothetical study of an intervention that is administered during play with a therapist, outcomes measured in a play-based interaction with a familiar therapist and similar toys may index change that is bound to that context. The outcome measure does not afford any degree of confidence that the treatment has induced changes in child behavior that would generalize to other contexts. In contrast, outcomes measured using standardized assessment procedures (i.e., different interaction style and materials) administered by an unfamiliar examiner (i.e., different interaction partner) would likely reflect change that reaches across a wide range of contexts. Similarly, outcomes measured in the home environment in an interaction with a parent (i.e., different setting, interaction partner, and interaction style, assuming the parent has not been trained in the intervention), would serve as a naturalistic assessment of highly generalized change in this hypothetical study. In theory, generalized change is more difficult to effect than context-bound change, so effect sizes for generalized outcomes are likely to be smaller relative to effect sizes of outcomes that are potentially context-bound.

Proximity of Outcomes to Intervention Targets

Outcomes may also vary by their proximity to the targets or goals of the intervention. Ideally, interventions would be able to demonstrate change not only on outcomes that are directly taught or addressed by the intervention (i.e., *proximal* outcomes), but also on outcomes that are developmentally downstream from what is directly taught or addressed (i.e., *distal* outcomes). When interventions are able to demonstrate growth on distal outcomes, they are essentially providing evidence that the intervention is influencing children's development, which may mean that the intervention will continue to have effects long after the intervention has stopped. However, prior best evidence syntheses have shown that early interventions for children with ASD show much larger effects for proximal as compared with distal outcomes (Yoder et al., 2013).

Correlated Measurement Error in Parent/Teacher Mediated Interventions

In addition to other commonly cited sources of bias, studies of autism-specific interventions are frequently threatened by CME that occurs when parents or teachers are the interventionists and also participate in assessment procedures. Parents and teachers are primary figures in the lives of children with ASD, and this makes them ideal mediators of intervention. For this reason, researchers have developed a number of interventions that target parents and teachers as interventionists, and tested their effectiveness in parent- or teacher-training studies. Trouble arises when natural interaction partners are trained as interventionists over the course of a study while simultaneously participating as assessors, either by rating child outcomes via a standardized report, or by serving as the interaction partner in an observational assessment context. The risk of bias posed by this specific study design flaw extends beyond that posed by detection bias related to the nonindependence of assessors. This is because, in addition to being aware of group assignment, the assessors and assessment context has also changed from pre- to postintervention in a manner that favors the intervention group. For example, a study might test the effects of parent-training for improving communication in children with ASD by examining the frequency of child communication during free play with parents. Prior to intervention, the assessment context in both groups would feature a parent naive to strategies for eliciting communication. However, after intervention, the assessment context in the treatment group would feature a parent who is more adept at eliciting communication while the assessment context in the control group remained the same. Though these two assessment contexts seem identical, they are fundamentally different. Though studies of parent- and teacher-led interventions are not unique to this population, they are well-represented in autism intervention literature. As such, any assessment of study quality should include an evaluation of the potential influence of this field-specific source of bias.

Study Purpose and Research Questions

The purpose of this study is to gather and synthesize all available studies of nonpharmacological interventions targeting any outcome in children with ASD below the age of 8 years. Our specific research questions were:

1. Across all eligible quasi-experimental and experimental studies, are summary effects positive and significant for targeted outcomes for each of seven intervention types (behavioral, developmental, NDBI, TEACCH, sensory-based, animal-assisted, and technology-based)?

2. Are summary effects positive and significant for targeted outcomes for each of the aforementioned seven intervention types when only outcomes from studies with basic quality controls (i.e., random assignment, independent assessors) are included?

3. Across intervention and outcome types, are summary effects for proximal outcomes larger than summary effects for distal outcomes?

4. Across intervention and outcome types, are summary effects for outcomes that measure context-bound behaviors larger than summary effects for outcomes that measure more highly generalized characteristics?

Method

Search

Search terms and databases. To gather the peer-reviewed literature included in the current meta-analysis, the following nine online databases were searched: Academic Search Complete, CINAHL Plus with Full Text, Education Source, Educational Administration Abstracts, ERIC, MEDLINE, PsycINFO, Psychology and Behavioral Sciences Collection, and SocINDEX with Full Text. Search terms were used in various combinations to capture the diagnostic criteria and intervention designs included within the search. The individual databases were searched using the following terms: autis*, ASD, PDD, Aspergers, intervention, therapy, teach*, treat*, program, package, assign*, control group, BAU, "wait list," RCT, random*, quasi, "treatment group," "intervention group," "group design," and trial. This initial search yielded 12,933 results from academic journals, dissertations, books, reports, conference materials, and reviews.

To gather gray literature, or studies not published in peer-reviewed journals, investigators who received federal grants to study autism were identified through a search of the National Database for Autism Research (NDAR), the National Institutes of Health (NIH) Matchmaker, and Institute of Education Sciences (IES) websites. A list of researchers ($n = 106$) was generated, and 90 of these investigators were emailed with a request for eligible data. The contact information for the remaining investigators could not be found.

Screening process. A preliminary screen of abstracts was first completed using abstrackr (Wallace, Small, Brodley, Lau, & Trikalinos, 2012). Studies were screened at the full-text level if they met the following inclusion criteria: (a) published in English,

(b) published from 1970–present, (c) group design that included both an intervention and control group, (d) a simple majority of participants were reported to have a diagnosis of ASD, and (e) the average age of included participants was between 0 and 8 years. In many instances, though studies met inclusion criteria, insufficient information was provided to enable the extraction of unadjusted effect sizes. In these cases, authors were identified and emailed with a request to provide unadjusted postintervention means and standard deviations. The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) diagram in Figure 1 summarizes the search process and provides justifications for exclusion of articles.

Coding Procedure

Included studies were coded for participant characteristics, intervention characteristics, study characteristics (including quality indicators), outcome characteristics, and effect size information. The coding manual is available upon request from Micheal Sandbank.

Participant characteristics. Participant characteristics coded from studies included average age of participant samples in months, percentage of sample that was male, and average language age in months (either receptive, expressive, or total) whenever it was reported.

Intervention characteristics. Intervention approaches were categorized based on the specific techniques used and the underlying philosophies that motivated the approach. A set of candidate categories (behavioral, developmental, NDBI, sensory-based, technology-based, cognitive behavior therapy, other) were drafted in the first instantiation of the coding manual for this synthesis based on authors' knowledge of intervention literature. Based on the results of our literature search and screening process, as well as the range of intervention approaches encountered across our team's initial training on coding precision and reliability, intervention categories were further refined to include "animal-assisted therapy." This intervention approach was found to be motivated by a distinct theoretical framework and to have amassed a sufficient number of group design studies to permit prior systematic review and meta-analysis (O'Haire, 2013, 2017; Trzmiel et al., 2019). Thus, interventions were initially coded as animal-assisted therapy, behavioral, developmental, NDBI, cognitive behavior therapy, sensory-based, technology-based, or other. After completion of coding, the set of interventions coded as "other" were reexamined to determine whether there existed a sufficient set of similar studies (e.g., five or more) that could be meaningfully combined to comprise an additional category. This was the case for studies of the TEACCH intervention. Studies of TEACCH

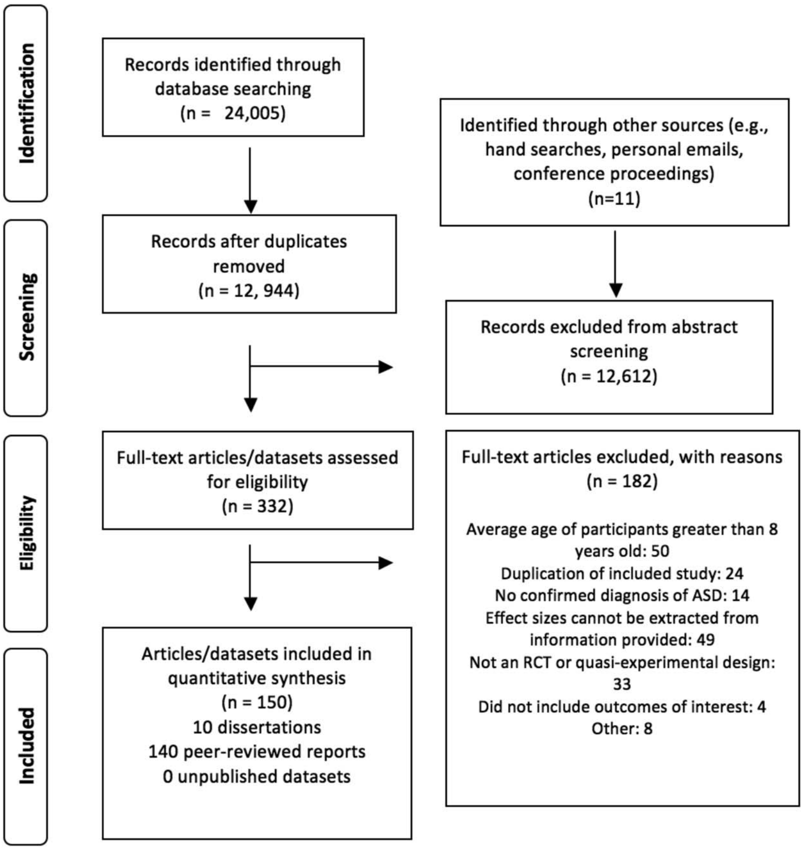


Figure 1. PRISMA (preferred reporting items for systematic reviews and meta-analyses) flow. RCT = randomized controlled trial; ASD = autism spectrum disorder.

that were initially coded as "other" were, therefore, recoded as "TEACCH."

Animal-assisted therapy. Interventions coded as animal-assisted therapy were those mediated through the presence of an animal. Equine assisted therapy was an example listed in the coding manual.

Behavioral. Interventions were coded as behavioral if authors described the intervention as being heavily situated in operant theories of learning, or if they relied heavily on behavior analytic techniques, such as didactic instruction, prompting, shaping, and extrinsic reinforcement. Examples of behavioral interventions listed in the coding manual included EIBI, PECS, discrete trial training, verbal behavior, autism partnership, and the Lovaas model.

Developmental. Interventions were coded as developmental if they were described as being motivated by constructivist theories of learning, or if they were heavily child-led and implemented according to a typical developmental sequence, with the goal of facilitating the development of foundational skills that would translate to gains in developmentally downstream domains. Examples of developmental interventions featured in the coding manual included DIR/Floortime, Hanen models, and responsive teaching.

NDBIs. Interventions were coded as NDBIs if they were one of any of the named interventions in the consensus article on this intervention approach (Schreibman et al., 2015), or if they combined adult-led, behavioral teaching methods with child-led routines and taught to a natural developmental progression within naturalistic settings. These included incidental teaching, pivotal response treatment, the early start denver model (ESDM), enhanced milieu teaching (EMT), reciprocal imitation training (RIT), project ImPACT (Improving Parents As Communication Teachers), joint attention symbolic play engagement and regulation (JASPER), social communication/emotional regulation/transactional supports (SCERTS), early achievements, and prelinguistic milieu teaching (PMT). Although PMT is not explicitly listed as an example of an NDBI in the consensus article by Schreibman et al. (2015), we contacted Paul Yoder, a leading researcher of this intervention, while drafting the coding manual for this meta-analysis to verify that this would be the appropriate category for this intervention approach (Paul Yoder, personal communication, March 29, 2018).

Cognitive behavior therapy. Interventions were coded as cognitive behavior therapy if a study explicitly named the intervention as such.

Sensory-based interventions. Interventions were coded as sensory-based if they incorporated targeted exposure to sensory or multisensory (e.g., auditory, visual, tactile, olfactory) stimuli. Examples listed in the coding manual included sensory integration, music therapy, massage, acupuncture, auditory integration, and weighted blankets. This category was drafted based on precedent across prior reviews of sensory-based interventions (Baranek, 2002; Case-Smith & Arbesman, 2008; Weitlauf, Sathe, McPheeters, & Warren, 2017).

Technology-based interventions. Interventions were coded as technology-based if the intervention was primarily delivered on a computer or electronic device (i.e., iPad, DVD).

TEACCH. Interventions were recoded as TEACCH if a study explicitly identified using this method.

Other. Interventions that did not fit into the previously defined categories were coded as other.

Study characteristics. Study-level characteristics that were coded include design type (i.e., RCT or quasi-experimental), publication status (i.e., indexed, nonindexed, unpublished), and several features of study quality. Studies were coded as a randomized controlled trial if the text indicated that participants were randomly assigned to an intervention group and a control group or contrasting treatment, or if the authors referred to the study as "randomized." Studies were coded as quasi-experimental when authors made no indication that the process of group allocation was random. If a contrasting treatment model was used, the group receiving the treatment that was hypothesized by the authors to effect greater change was considered the treatment group. In the case of studies testing multiple active treatment groups compared with a passive control, treatment characteristics and effects were coded separately in contrast to control.

Publications were coded for whether they were published or unpublished. Published studies included indexed and nonindexed journals, and unpublished studies included dissertations and theses. Despite our extensive attempts to locate, obtain, and include unpublished data apart from dissertations and theses, no researchers provided us with data or effect sizes from unpublished reports.

Studies were coded for several indicators of study quality. These indicators included those specified by the Cochrane Collaboration's tool for assessing risk of bias (e.g., selection bias, performance bias, detection bias, reporting bias; Higgins et al., 2011), as well as additional indicators which we proposed in prior work (e.g., potential presence of CME related to parent/teacher training, sufficient number of participants to justify statistical analysis, reliance on parent or teacher report; Yoder et al., 2013). Selection bias related to insufficient randomization procedures and allocation concealment was coded as "high," "low," or "unclear" for studies coded as randomized controlled trials, and as "not applicable" for quasi-experimental studies. For subsequent Cochrane quality indicators, risk of bias was coded as "high" or "low" if studies explicitly indicated or provided sufficient information to ascertain the presence or absence of such risk, and as "unclear" if information related to risk potential was not detailed. Risk of selective reporting bias was coded as high if outcomes were reported to have been collected at post but were not reported in results, or if an entire assessment was administered but only selective subscores were reported without sufficient justification. Performance bias risk was assessed in consideration of the participants' and families' awareness of their group assignment. Detection bias accounted for the independence of assessors and coders. We elected to include interaction partners in naturalistic observational measures as "assessors," given that they may transiently influence child behavior during interactions. Attrition bias was coded with respect to the number of participants recruited and the number of participants included in analysis. Specifically, attrition bias was considered low if attrition was lower than 20% or if intent-to-treat analysis was utilized.

Outcome characteristics. In addition to the above quality indicators, we coded for outcome-level quality indicators that are especially important for research on intervention in young children with ASD. These quality indicators captured the boundedness and distality of outcomes, as well as the potential presence of CME related to parent or teacher training. Outcomes were coded as context-bound if they were measured in or very near the context of the intervention, and as generalized if they were measured in a context that differed from the context of intervention on multiple dimensions (e.g., interaction partners, materials, setting, interaction style). Outcomes taken from standardized parent/teacher reports were coded as potentially context-bound if reporters were also the primary mediators of intervention, based on the rationale that their report could be based on their observance of the outcome as it occurred within the context of the intervention they provided. Outcomes were coded as proximal if they indexed skills that were directly taught, modeled, or prompted during the intervention, and otherwise as distal. Outcomes indexed by developmentally scaled assessments were automatically coded as distal, based on the reasoning that these assessments are meant to tap generalized development rather than specific skills. We recognize that it is possible for an intervention to directly target specific items of a developmentally scaled assessment, but reasoned that in the absence of an extremely detailed description of intervention procedures, we should assume these assessments captured constructs beyond what was directly taught in intervention. Decision trees used to judge distality and boundedness are presented in Figures 2 and 3, respectively. Correlated measurement error related to parent/teacher training was coded as potentially present when parents or teachers operated as both the mediators of intervention as well as the outcome assessors.

Outcome categorization. Each dependent variable was categorized as either a core feature of ASD (i.e., social communication; restricted/repetitive patterns of behaviors, interests, or activities; sensory) or a related outcome (i.e., language, motor, adaptive, cognitive, academic, play, sleep, brain imaging, social emotional/challenging behavior). If outcomes were reported at multiple time points, immediate and follow-up outcomes were coded separately.

Effect size information. Unadjusted means, *SD*s, and *n*s were extracted from all eligible studies that reported a group difference between participants receiving the specified intervention and those not receiving the specified intervention. Group difference effect sizes were calculated for each outcome using the standardized mean difference (*d*), as derived via the Campbell Collaboration Practical Meta-Analysis Effect Size Calculator (Lipsey & Wilson, 2001) and then converted to the effect size metric used for analyses, Hedge's g (*g*). Effect sizes were reported in such a way that higher *g* scores indicated superior performance in the treatment group.

We were unable to extract effect sizes from some eligible studies due to insufficient information (e.g., authors did not report means and *SD*s, reported only mean change scores, or reported means and *SD*s that were adjusted for baseline covariates and therefore could not be meta-analyzed with unadjusted means and *SD*s). When this occurred for articles published within the last 10 years, we contacted the corresponding author(s) in an attempt to obtain either the unadjusted post means and *SD*s, or any other statistical information that would allow us to calculate the standardized mean difference between treatment and control/contrast groups after intervention. Fifty-five studies did not have sufficient information to allow effect size extraction for all outcomes. In the case of nine of these studies, effect size extraction was possible for some but not all outcomes, so eligible outcomes were included from those studies. Authors responded and supplied effect size information for 14 additional studies.

Reliability

A primary coder (Micheal Sandbank) read and coded all studies. All studies were also independently coded for reliability by one coder from a team of nine. Both coding sheets were then sent to a separate coding auditor who examined codesheets for discrepancies and reported any disagreements between coders. Original primary and reliability codes were then saved for reliability analyses in a separate folder, and all disagreements were addressed in

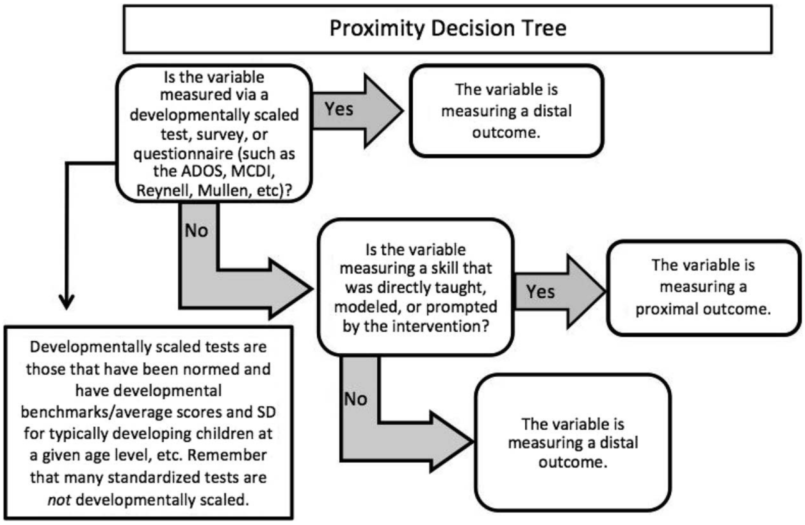


Figure 2. Decision tree used to code whether a study outcome was proximal or distal to treatment targets adapted from "Social communication intervention effects vary by dependent variable type in preschoolers with autism spectrum disorders," by P. Yoder, K. Bottema-Beutel, T. Woynaroski, R. Chandrasekhar, and M. Sandbank, 2013, Evidence-based Communication Assessment and Intervention, 170. Copyright 2013 by Taylor and Francis.

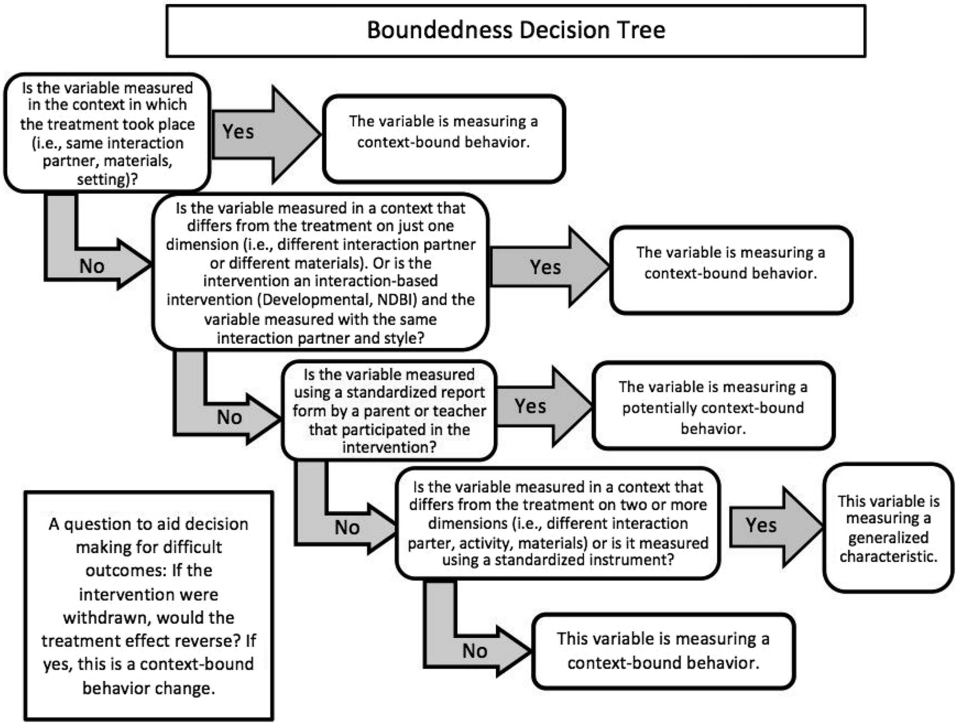


Figure 3. Decision tree used to code whether a study outcome measured a potentially context-bound or more highly generalized characteristic adapted from "Social communication intervention effects vary by dependent variable type in preschoolers with autism spectrum disorders," by P. Yoder, K. Bottema-Beutel, T. Woynaroski, R. Chandrasekhar, and M. Sandbank, 2013, Evidence-based Communication Assessment and Intervention, 171. Copyright 2013 by Taylor and Francis.

discrepancy discussions between the primary and reliability coders. Discrepancies were considered resolved once both coders agreed to a final consensus code, which was then added to the dataset used for the final analyses. Therefore, we are able to report reliability data from the original coding and also confirm that *all* disagreements were resolved prior to statistical analysis.

All reliability calculations were completed in R studio (R Core Team, 2017) using the *irr* package (Gamer, Lemon, Fellows, & Singh, 2012). Reliability was indexed using unweighted kappa for all categorical variables (Cohen, 1960) and one-way random intraclass correlation coefficients for all continuous variables (ICC; Shrout & Fleiss, 1979). Kappas ranged from 0.602–0.923, and average kappa across all categorical variables was 0.751. ICCs ranged from 0.676–0.999, and average ICC across all continuous variables was 0.916.

Analysis

All analyses were conducted in R (R Core Team, 2017). To account for the nesting of multiple effect sizes within overlapping participant samples, we used robust variance estimation (RVE) with small sample adjustments when synthesizing effect sizes and conducting metaregressions (Hedges, Tipton, & Johnson, 2010; Tanner-Smith, Tipton, & Polanin, 2016). These procedures account for the nonindependence of effect size statistics drawn from overlapping samples, and provide test statistics and confidence intervals that are adjusted based on how the effect sizes are clustered.

Effect sizes were aggregated based on type of outcome (see Outcome Characteristics section) within each type of intervention (see Intervention Characteristics section). Aggregating the results in this manner provided a summary statistic for the effect of each intervention type on each outcome type. Metaregression analyses were conducted on the coded variables of distality and boundedness (see Outcome Characteristics section) to determine whether the magnitude of the effects across intervention and outcome types were moderated by these categorical characteristics related to measurement. The threshold level of significance for these tests was set at $p < .10$, given that we had clear directional hypotheses for each potential moderator, meriting one-tailed tests of significance. To examine the potential presence of publication bias, we examined funnel plots of effect size estimates against their standard errors, and corresponding Egger's tests of funnel plot asymmetry, for each summary effect estimate. Due to the large number of significance tests this demanded, we applied the Benjimini-Yekutieli false discovery rate correction to the significance values from the Egger's tests to correct for spurious findings using the Hmisc package in R (Harrell, 2018). The Robumeta package in R (Fisher, Tipton, & Zhipeng, 2017) was used to conduct these analyses while the metafor package (Viechtbauer, 2010) was used to graph the forest plots and funnel plots.

Results

Descriptives of Included Study Samples and Outcomes

The search and screening process yielded 1,615 effect sizes gathered from 130 independent study samples (from a total of 150 reports) representing 6,240 participants. Across all studies, the average age of participants was 54.21 months (SD = 18.98), the average proportion of male participants per sample was 0.84 (SD = 0.07), and the average language age of participants in studies for which it was reported was 22.68 months (SD = 11.91). An average of 12.4 outcomes were reported for a single study sample (MIN = 1, MAX = 100, MDN = 8). Participant characteristics according to intervention type are reported in Table 1. There were 27 studies of behavioral interventions, 14 studies of developmental interventions, 26 studies of NDBIs, seven studies of sensory-based interventions, 10 studies of technology-based interventions, and six studies of TEACCH included in the synthesis. The RVE approach requires that at least five studies contribute to the generation of effect sizes, so the studies representing animal-assisted intervention (n = 4), cognitive–behavioral therapy (n = 2), and other varied approaches that could not be meaningfully combined into intervention types (n = 29) were excluded from summary effect estimation.

Study Quality

Figures 4 and 5 illustrate the proportion of studies or outcomes that received each quality rating (i.e., low risk of bias, high risk of bias, unable to determine) for seven key quality indicators, according to intervention type. These figures include only studies that contributed to summary effect estimation. Because it is almost always impossible for participants to be naive to intervention delivery in studies of nonpharmacological interventions for ASD, performance bias was rated as high for all but one study included in summary effect estimation and, thus, is not reported separately for each intervention type (see Corbett, Shickman, & Ferrer, 2008 for the lone exception).

Behavioral intervention studies. Figure 4 reflects information regarding quality indicator ratings for studies of behavioral interventions. Notably, only 29.63% of studies of behavioral interventions were RCTs. Detection bias was rated as high for 77.05% of outcomes in behavioral studies. High detection bias in this set of studies was largely driven by an overreliance on reports completed by individuals who were aware of intervention assignment—60.33% of outcomes were based on parent or teacher report. Correlated measurement error related to parent/teacher training threatened 53.77% of outcomes reported in behavioral studies. Because many of the studies relied on standardized report forms, and because most of these studies only loosely described intervention targets, 86.23% of outcomes tracked in behavioral intervention studies were categorized as distal to the intervention targets. Half (50.49%) of outcomes were categorized as generalized, and 10.49% were classified as context bound. The remaining 39.02% of outcomes were categorized as potentially context bound, because they were derived from caregiver reports in studies where caregivers participated as interventionists (meaning that it is unclear if the outcome could be demonstrated in interactions with individuals who were not trained as interventionists). Bias related to substantial attrition (i.e., >20% of the study sample) was rated as high for 15.41% of all outcomes.

Developmental intervention studies. Figure 4 reflects quality indicator ratings for studies of developmental interventions. A large majority (78.57%) of included developmental studies were RCTs. Detection bias was rated as high for 53.97% of outcomes, but this was not due entirely to overreliance on caregiver report. Nearly a third (29%) of outcomes were taken from parent/teacher report. The remainder of outcomes flagged for high detection bias (approximately half of the outcomes tracked in these studies) reflects the common practice of measuring language and communication outcomes in the context of interactions with natural communication partners (primarily parents, who were aware of group assignment). CME related to parent/teacher training threatened three quarters (75%) of all outcomes in developmental studies. Because many of the developmental interventions were explicitly described as targeting language and social communication, and many of the outcomes were observational measures of language and social communicative behaviors, approximately half (53.57%) of outcomes were categorized as proximal to intervention targets. Approximately a quarter (27.84%) of outcomes were categorized as generalized, a quarter (25%) were categorized as potentially context-bound, and approximately one half (47.16%) were categorized as context-bound. Over a third (34.66%) of all outcomes were subject to high bias from substantial attrition.

Naturalistic developmental behavioral intervention studies. Figure 4 illustrates quality indicator ratings for included studies of NDBIs. A large majority (76.92%) of included studies of NDBIs were RCTs. Detection bias was rated as high for 59.42% of outcomes. This was due, in part, to the common use of observational measures of skills coded from natural interactions with interaction partners who were aware of group assignment. Only 17% (the lowest of any intervention type) of outcomes were collected from parent/teacher report. However, CME related to parent/teacher training threatened 47.09% of outcomes, due to a prevalence of parent-training studies which included outcomes derived from parent–child interactions. Because many NDBIs were described as specifically targeting symbolic play, early social communication, and language, researcher-created measures of these skills were coded as proximal to intervention targets. Thus, nearly half (47.59%) of outcomes in NDBI studies were categorized as proximal. Nearly a quarter (22.22%) were cate-

Table 1
Participant Sample Characteristics Averaged by Intervention Type

Intervention	Age[a]	Percent male	Language age[a]
Behavioral	49.65	84.93	16.62
Developmental	42.36	84.96	21.42
NDBI	38.71	81.03	24.92
Sensory	65.45	84.27	31
TEACCH	48.22	83.85	18.6
Technology based	66.36	87.36	26.28

Note. NDBI = naturalistic developmental behavioral intervention; TEACCH = Treatment and Education of Autistic and related Communication-Handicapped Children. Participant language age was infrequently reported. Figures may not be representative of the full sample.
[a] In months.

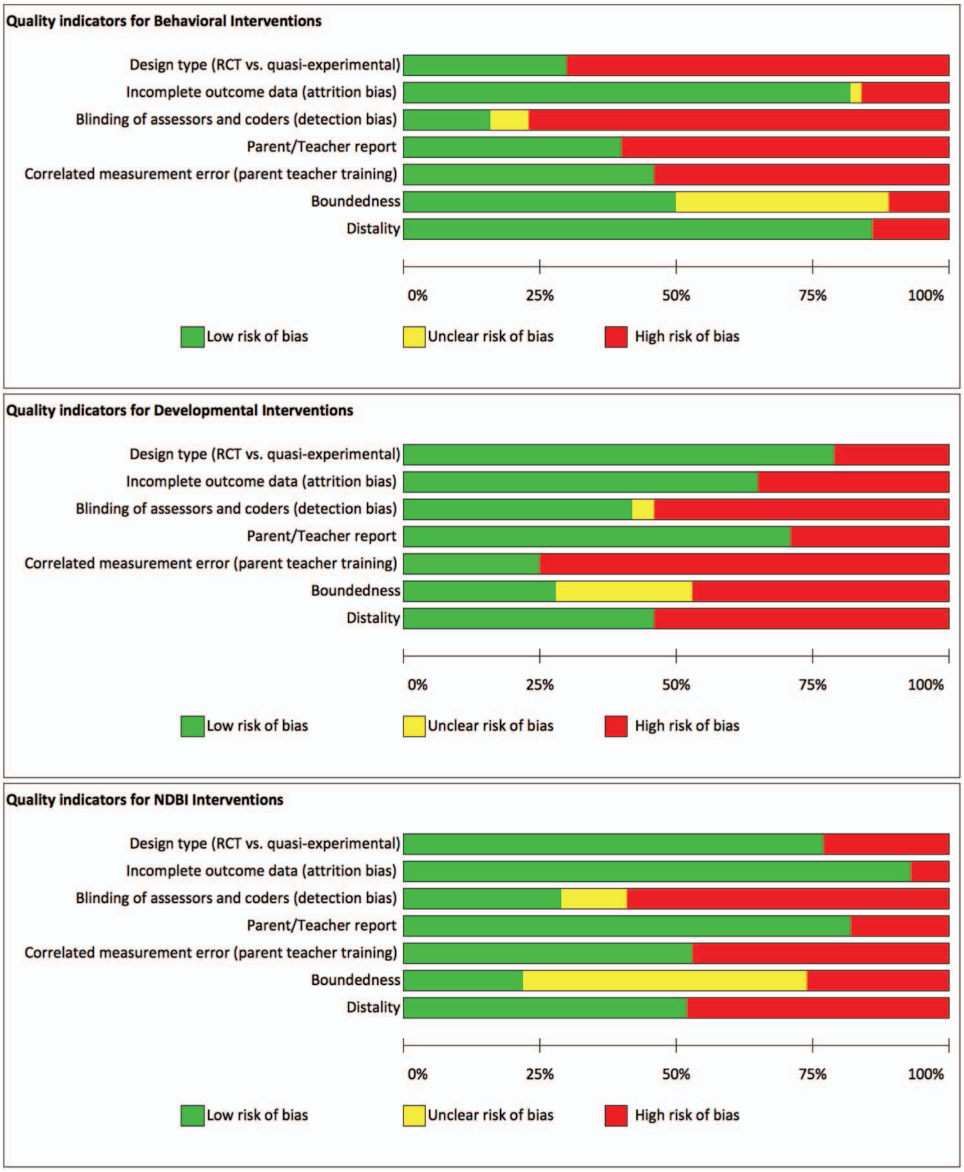


Figure 4. Summary of quality indicator ratings for studies of behavioral, developmental, and naturalistic developmental behavioral intervention (NDBI) types. RCT = randomized controlled trial. See the online article for the color version of this figure.

gorized as generalized, 52.41% were categorized as potentially context-bound, and another quarter (26.36%) were categorized as context-bound. Only 7.25% of outcomes were subject to bias from high attrition.

Sensory-based intervention studies. Figure 5 reflects quality indicator ratings for sensory-based intervention studies that were included in summary effect size estimation. All of the seven studies included in effect size estimation were RCTs. Because language was the only outcome category for which there were a sufficient number of sensory-based intervention studies to permit summary effect size estimation, the following outcome-level quality indicator ratings apply only to the language outcomes (n = 13) tracked in these studies. Detection bias was rated as high for nearly half (46.15%) of all language outcomes. Nearly a third (30.77%) of all outcomes were based on parent/teacher report, and these same outcomes were also subject to CME related to parent training. The overwhelming majority (92%) of outcomes were categorized as distal, because few sensory-based interventions were described as directly targeting language. Nearly a third (30.77%) were categorized as generalized, 53.86% were categorized as potentially context-bound, and 15.38% of outcomes were categorized as context-bound. Attrition bias was rated as high for 15.38% of outcomes.

TEACCH studies. Figure 5 illustrates quality indicator ratings for studies of TEACCH that were included in summary effect size estimation (n = 6). Only two (33%) of these studies were

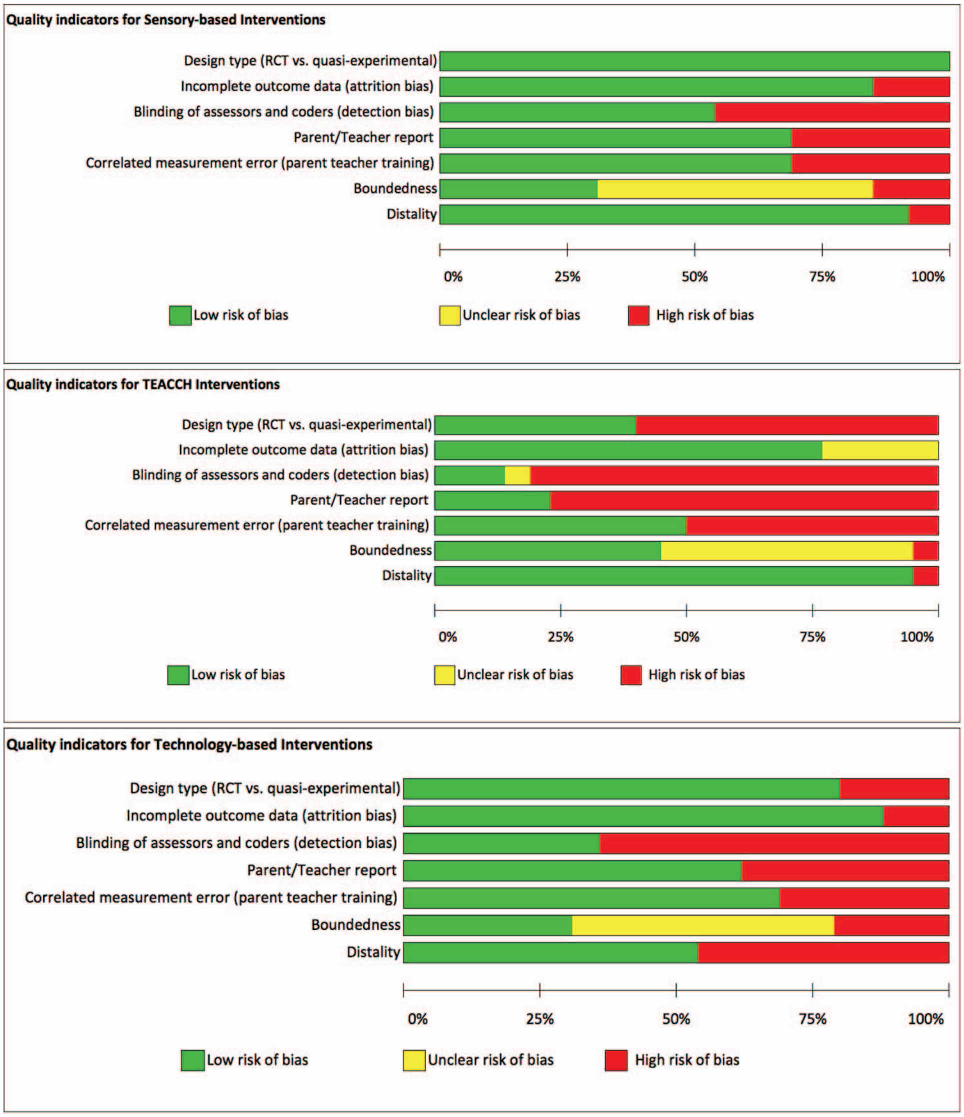


Figure 5. Summary of quality indicator ratings for sensory-based, Treatment and Education of Autistic and related Communication-Handicapped Children (TEACCH), and technology-based intervention types. RCT = randomized controlled trial. See the online article for the color version of this figure.

RCTs. Detection bias was rated as high for the majority (81.81%) of outcomes, and this was largely driven by an overreliance on parent/teacher report, from which 77.27% of outcomes were derived. CME related to parent/teacher training threatened half (50%) of all outcomes. Given that the explicit individual intervention targets of TEACCH were not thoroughly described, and that the majority of outcomes were taken from standardized parent/teacher reports, almost all (95.45%) outcomes were assumed to be distal. Nearly half (45.45%) of outcomes were categorized as generalized, half (50%) were categorized as potentially context-bound, and the remaining 4.54% were categorized as context-bound. None (0%) of the studies reported substantial attrition.

Technology-based intervention studies. Figure 5 illustrates quality indicator ratings for studies of technology-based interventions. Of 10 technology-based intervention studies included in summary effect estimation, eight (80%) were RCTs. Detection bias was rated as high for 64.28% of all outcomes. Over a third (38.1%) of outcomes were taken from parent/teacher report. CME related to parent/teacher training threatened 30.95% of outcomes. Over half (53.57%) of outcomes were categorized as distal. Nearly a third (30.95%) of outcomes were categorized as generalized, nearly half (47.62%) were categorized as potentially context-bound, and 21.43% were categorized as context-bound. Bias related to substantial attrition was rated as high for 15.38% of outcomes.

Summary Effects by Intervention and Outcome Type

Summary effects across all studies without consideration of quality indicators. Figure 6 reflects summary effect size estimates within interventions and outcome types. These estimates were derived

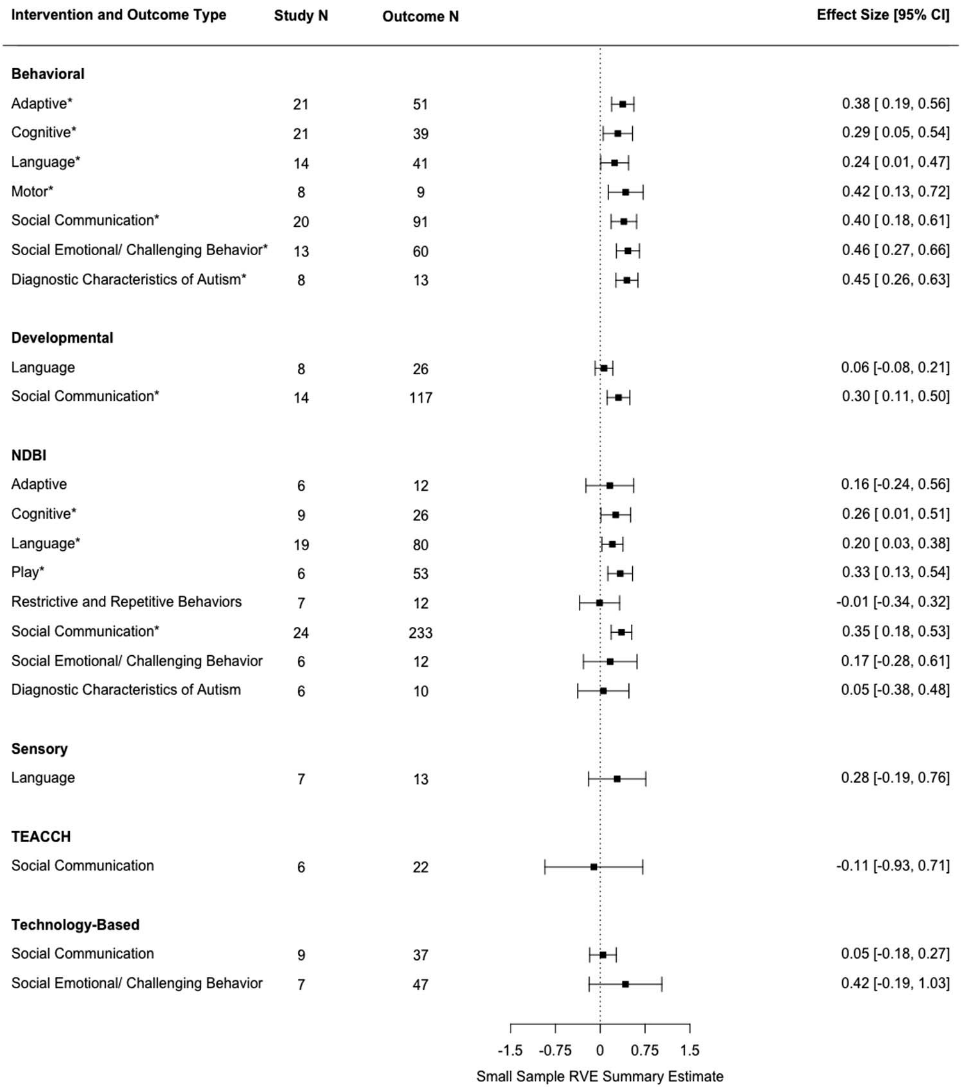


Figure 6. Forest plot of robust variance estimation (RVE) summary estimates with small sample bias correction for each outcome by intervention type, when all outcomes from quasi-experimental and RCT group design studies are included. NDBI = naturalistic developmental behavioral interventio; TEACCH = Treatment and Education of Autistic and related Communication-Handicapped Children. * Denotes summary effect size estimates with confidence intervals that do not overlap with zero.

using all available effect sizes, both from quasi-experimental studies and RCTs. Summary effects were computed when effect sizes associated with a given outcome and intervention type were available from at least five independent participant samples. Thus, we were able to estimate the summary effects of behavioral interventions on adaptive outcomes, cognitive outcomes, language outcomes, motor outcomes, social communication outcomes, social emotional/challenging behavior outcomes, and outcomes quantifying broader autism symptomatology. Summary effects for behavioral interventions across outcome types ranged from 0.24 to 0.46 and were all statistically significant. For developmental interventions, only language and social communication outcomes were measured in a sufficient number of studies to permit the estimation of summary effects. The summary effects of developmental interventions on these outcomes were 0.06 and 0.30, respectively, and only the estimate for social communication was statistically significant. The summary effects of NDBIs were separately estimated for adaptive outcomes, cognitive outcomes, language outcomes, play outcomes, restrictive and repetitive behaviors, social communication outcomes, social emotional/challenging behavior outcomes, and outcomes that quantified broader autism symptomatology. These summary effects ranged from −0.01 to 0.35. The summary effect estimates of NDBIs on cognition, language, play, and social communication outcomes were statistically significant. For sensory-based interventions, only language outcomes were measured in a sufficient number of studies to permit the estimation of summary effects. This summary effect estimate was 0.28, and was not significant. For TEACCH, summary effects could be generated only for social communication outcomes. This summary effect estimate was −0.11 and was not significant. For technology-based interventions, the most frequently tracked outcomes were social communication and social emotional/challenging behavior. Summary effect estimates for these outcomes were 0.05 and 0.42, respectively, and neither were significant.

Summary effects from RCTs. Figure 7 reflects summary effect size estimates derived exclusively from outcomes extracted from RCTs, according to intervention and outcome type. There were not enough RCTs of behavioral interventions to permit summary effect estimation for any outcome type. For developmental interventions, the summary effect across social communication outcomes from RCTs was 0.27 and significant. For NDBIs, a sufficient number of RCTs permitted the estimation of summary effects on cognition, language, play, and social communication. These estimates ranged from 0.18 to 0.42, and were significant for language, play, and social communication. All of the studies tracking the effect of sensory-based interventions on language outcomes were RCTs. Therefore, this summary effect estimate remains identical to that of the initial model. For technology-based interventions, there were only enough RCTs to permit estimation of a summary effect for social communication. This was 0.06 and was not significant. There were no RCTs examining the effects of the TEACCH intervention on any outcome.

Summary effects from RCTs excluding outcomes from caregiver reports. Figure 8 reflects summary effects estimated exclusively from outcomes that were extracted from RCTs and that were not based on caregiver report. For developmental interventions, a sufficient number of studies and outcomes permitted the estimation of a summary effect for social communication, which was 0.31 and statistically significant. For NDBIs, summary effect estimation was possible for cognition, language, play, and social communication. These effects ranged from 0.18 to 0.47, and were significant in the cases of play and social communication outcomes. For sensory-based interventions, summary effect estimation was possible for language only. This estimate was 0.28 and was not significant.

Summary effects from RCTs excluding all outcomes subject to a high threat of detection bias. Figure 9 reflects summary effects estimated exclusively from outcomes that were extracted from RCTs where assessors were unaware of group assignment. There were enough studies/effect sizes of this nature to permit estimation of the summary effects of NDBIs on language and social communication only. These estimates were 0.17 and 0.17, respectively, and were not significant.

Publication Bias Analyses

Funnel plots and Egger's test results are included in the online supplementary materials accompanying this report. Corrected *p* values for Egger's tests for funnel plot asymmetry were significant for adaptive and social communication outcomes from studies of NDBIs, suggesting that publication bias may have threatened these summary estimates.

Moderator Analyses

Metaregression analyses across the entire dataset suggested that summary effects were significantly larger for outcomes that were proximal compared with those that were distal ($\beta = 0.171$, $p = .024$). Boundedness was also a significant source of effect size variance; effect sizes coded as generalized ($\beta = -0.170$, $p = .076$) were smaller than those coded as potentially context-bound or context bound ($\beta = -0.115$, $p = .22$).

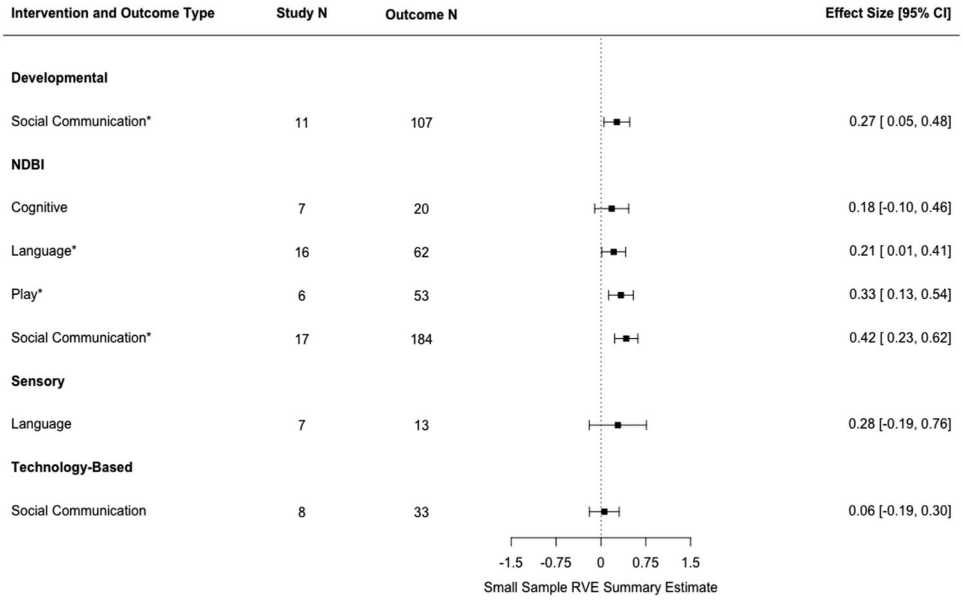


Figure 7. Forest plot of robust variance estimation (RVE) summary estimates with small sample bias correction for each outcome by intervention type, when all outcomes from RCTs are included. NDBI = naturalistic developmental behavioral intervention. * Denotes summary effect size estimates with confidence intervals that do not overlap with zero.

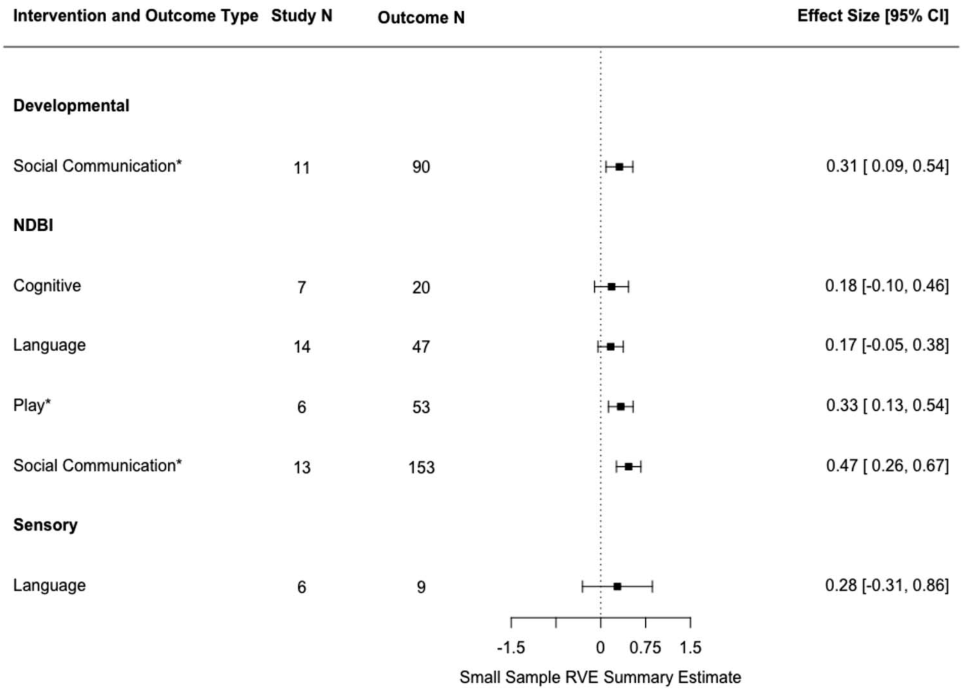


Figure 8. Forest plot of robust variance estimation (RVE) summary estimates with small sample bias correction for each outcome by intervention type, when only noncaregiver report outcomes from RCTs are included. NDBI = naturalistic developmental behavioral intervention. * Denotes summary effect size estimates with confidence intervals that do not overlap with zero.

Discussion

The purpose of this study was to locate, evaluate, and synthesize all available quasi-experimental and RCT investigations of nonpharmacological interventions for children with ASD in terms of methodological quality and summary effect. Results suggest that some intervention approaches show promise for improving a range of outcomes, while others have amassed relatively limited evidence of effectiveness to date. The number of RCT investigations in this area have increased precipitously, but low methodological rigor remains a concern.

Promising Intervention Types

We consider intervention types for which significant summary effects were shown for at least one outcome, when two important quality indicators were taken into account (randomization and abstention from using caregiver reports) to be “promising.” NDBIs and developmental interventions meet these criteria.

NDBI. This is the first article to report summary effects of NDBIs since the 2015 consensus article that established this new category of intervention as a blend of traditional behavioral and

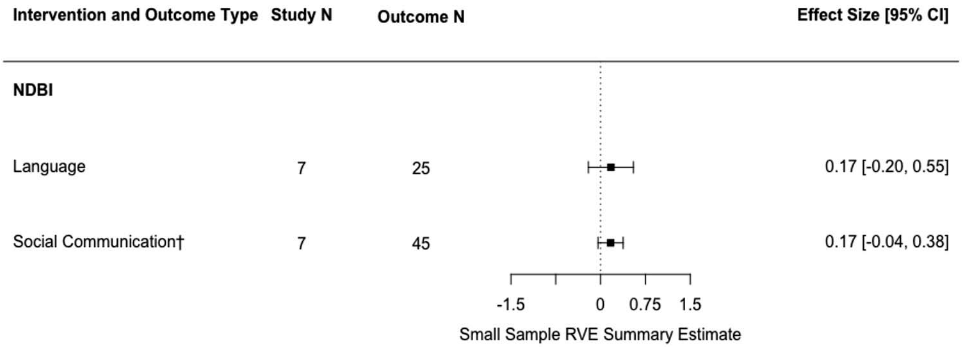


Figure 9. Forest plot of robust variance estimation (RVE) summary estimates with small sample bias correction for each outcome by intervention type, when only outcomes from RCTs that are not threatened by detection bias are included. NDBI = naturalistic developmental behavioral intervention. † Denotes summary effect size estimates that have p values $<.10$.

developmental approaches. By far, NDBIs have emerged as the intervention type most supported by evidence from RCTs. These studies suggest NDBIs may be particularly useful for supporting development of social communication, language, and play skills. Studies of NDBIs were also the least likely to rely on caregiver report as a primary index of intervention effectiveness. However, we note that when outcomes subject to all forms of detection bias were excluded from summary effect estimation, there was no category of outcomes for this intervention type that reached significance. In addition, our results suggest that publication bias may have threatened overall summary estimates for adaptive and social communication outcome types. However, asymmetry in these funnel plots may also be due to other methodological design flaws, such as the presence of detection bias.

Developmental. Evidence suggests that developmental interventions may be particularly effective for supporting the acquisition of social communication skills, which represents a core challenge for young children with ASD. This conclusion is supported even when outcomes from quasi-experimental studies and caregiver report are excluded. However, a substantial portion of outcomes were subject to high detection bias due to interaction partners or assessors that were aware of group assignment. When these outcomes are excluded, the remaining studies are too few in number to permit summary effect estimation for any outcome type. A key assumption of developmental interventions is that targeted gains in social communication will facilitate cascading developments in the domain of language. This assumption was not supported by our meta-analysis, as the summary effect of developmental interventions on language outcomes was not significant. However, we did locate compelling evidence suggesting that early targeted improvements in the synchrony of parent–child interactions can yield longitudinal improvements in the core challenges associated with ASD, which are detectable with standardized, independently administered assessments (Green et al., 2010; Pickles et al., 2016). Green et al. (2010) study of the Preschool Autism Communication Trial (PACT) supports the notion that proximal changes effected by intervention can facilitate long-term change in developmentally distal outcomes, even in the absence of continued intervention. It also provides an example of methodological rigor to which the field should aspire, as it employed random assignment, preregistered analyses, independent evaluators, and clearly defined proximal and distal outcomes.

Intervention Types With Some Evidence of Effectiveness: Behavioral Interventions

Behavioral intervention, specifically EIBI and related variants, is the most commonly recommended intervention approach for children with ASD, with many states specifying behavioral interventions explicitly in insurance coverage mandates ("Autism and Insurance Coverage," 2018). Indeed, the large number of behavioral intervention studies ($n = 27$) that met our search criteria also suggests this is the most studied intervention approach for this population. Considered as a whole, without regard to quality of evidence, these studies support the effectiveness of behavioral interventions for improving a wide range of outcomes for children with ASD. However, only a fraction of past studies exploring the effects of traditional behavioral interventions were RCTs, and the majority of outcomes contributing to summary effect sizes were taken from caregiver report. Thus, the relatively low quality of this set of intervention literature limits our confidence in the accuracy of the summary effect sizes estimated in the initial model. A notable exception is the sole RCT which examined the effects of EIBI on standardized measures of cognition and language administered by independent evaluators (Smith, Groen, & Wynn, 2000). Though the positive results of this study are encouraging, they have persisted without replication for nearly 20 years. The dramatic increase in published RCTs in the intervening years since this study's publication stand as proof that high-quality group experimental investigations of autism-specific interventions are both possible and necessary in order to unquestionably establish the effectiveness of interventions that are so routinely recommended. In the meantime, clinicians are encouraged to expand their knowledge and skills to include naturalistic approaches that center the principles of early childhood development. States with insurance mandates that explicitly cover traditional behavioral interventions should furthermore revise their policies to also include NDBI and developmental approaches, given that these approaches have now accrued substantial evidence for effects in young children on the autism spectrum from recently published RCTs.

Intervention Types With Little Evidence of Effectiveness

Sensory. Several previous systematic reviews have concluded that sensory-based interventions have amassed little evidence supporting their effectiveness to date (e.g., Barton, Reichow, Schnitz, Smith, & Sherlock, 2015; Case-Smith, Weaver, & Fristad, 2015). Our results are consistent with these conclusions. Relatively few group design studies of sensory-based interventions specifically focused on young children with ASD (i.e., with a mean age <8 years) were located. Furthermore, there were not a sufficient number of studies measuring and reporting sensory outcomes in a manner that permitted extraction of effect size information and estimation of the summary effect of this intervention approach on what would presumably be the most proximal outcome (i.e., improvements in sensory function). This is particularly concerning in light of the fact that sensory differences are highly prevalent in this population (e.g., Ausderau et al., 2014; Ben-Sasson et al., 2009; Leekam, Nieto, Libby, Wing, & Gould, 2007) and have been found to be associated with some aspects of child stress (Corbett, Schupp, Levine, & Mendoza, 2009). Unfortunately, across all included studies, we found no evidence that *any* intervention type had the potential to influence sensory outcomes in children with ASD. When we were able to estimate summary effects of sensory-based interventions, as was the case for language outcomes, the relative paucity of studies limited the precision of our estimates. Though the summary effect estimate for sensory-based interventions on language outcomes is similar in magnitude to those of behavioral and NDBI approaches, this estimate is surrounded by a much wider confidence band, which overlaps with zero (i.e., the effect is not significant).

It should be noted that our category of sensory-based interventions was broad and included intervention approaches as distinct as sensory integration therapy, Tomatis Sound Therapy™, and music therapy. The heterogeneity of these intervention approaches may limit the conclusions that can be drawn from this summary effect size estimate, as the theoretical underpinnings and clinical procedures do vary across approaches. It may be useful to consider the evidence for each of these intervention approaches separately, though the limited number of studies for each prevented us from computing subgroup effect sizes here. However, we did not come across any noteworthy high-quality studies that suggested that any of the aforementioned intervention approaches had markedly positive effects on outcomes (though see Schaaf et al., 2014 which unfortunately did not report outcome data in a manner that would permit derivation of effect size information for synthesis). We did locate two exceptionally high-quality studies demonstrating null effects of two sensory-based interventions, music therapy (Bieleninik et al., 2017) and auditory stimulation (Corbett et al., 2008). Therefore, our conclusion that there is limited high-quality evidence to date to support sensory-based interventions for young children with ASD is based on our quantitative findings as well as our more fine-grained qualitative observations about this set of literature. Given that sensory features are now a core diagnostic criteria of ASD (APA, 2013), and given the already widespread implementation of sensory-based interventions for this population (e.g., Goin-Kochel, Mackintosh, & Myers, 2009; Schaaf & Case-Smith, 2014), we suggest that more rigorous research of these interventions be conducted to precisely determine their effects for children with ASD.

TEACCH. Though TEACCH was among the first interventions designed specifically for individuals with ASD, it also remains relatively understudied compared with several other intervention approaches geared toward this population. Few eligible studies of TEACCH were located, and most were quasi-experimental. This may be because TEACCH is often conceptualized as a classroom wide intervention, necessitating large, cluster-randomized trials that are substantially more expensive to implement than clinically based RCTs. The negative summary effect estimated across these studies suggests that there is limited evidence to support the effectiveness of TEACCH for improvement of social communication skills, and almost no evidence to support the effectiveness of TEACCH for the improvement of other core and related symptoms of ASD.

Technology-based interventions. Although assistive technology is an important support that must be accessible to autistic individuals, early interventions mediated *entirely* through technology have little evidence to support their effectiveness for improving social communication or social emotional outcomes in children with ASD. Both of the summary effect sizes for these outcome types had confidence intervals which included zero. The majority of technology-based interventions represented in this meta-analysis were DVDs or video games that targeted social emotional learning and social communication skills. The limited effectiveness of these interventions may be attributable to the near or total absence of a human interaction partner in these intervention contexts. Though technological supports have characteristics that might make them particularly useful to autistic people (e.g., predictable formats of information delivery, self-paced usage, highly motivating), these supports likely need to be integrated into interpersonal interactions, which could include computer-mediated interpersonal interactions, rather than replacing interaction partners entirely in learning situations. This may be particularly true when the targeted developmental achievements are social in nature. In fact, the integration of technological supports into other interaction-based interventions is an approach that is supported by high-quality studies. For example, Kasari et al. (2014) integrated speech generating devices (SGD) into their JASP-EMT early intervention approach, and found gains on a variety of communication outcomes for preschoolers who were initially minimally verbal, compared with those that received the same intervention without use of the SGD. In this study, technology was integrated into an already well-developed intervention, that had amassed some degree of empirical support.

This may be a sensible path forward for conceptualizing the utility of new technologies for early intervention. That is, technology may be most useful when it is integrated into previously developed and validated approaches as a means to expand the populations of children with ASD for whom the intervention is accessible, rather than as an intervention in its own right. In this regard, it is important to consider that the ultimate use of technology is usually separable from the means by which children are taught to use it, so even the most intuitively designed technologies will still need to be paired with a validated teaching approach to ensure that children are able to learn to use the technology in a meaningful way.

Animal-assisted interventions. Although we did locate studies of animal-assisted interventions, there were too few to permit estimation of summary effect sizes for any outcomes. The two interventions represented in these studies were EAAT and canine assistance. Several of these studies relied on caregiver report to index change, and two were flagged for possible unreported conflicts of interest, as the authors currently provide the interventions in question for profit (Bass, Duchowny, & Llabre, 2009; Page, 2012). Therefore, there is little quality evidence to support the effectiveness of animal-assisted interventions for any outcomes for children with ASD at this time.

Issues Related to Quality Indicators

The results of this study indicate that study quality remains an issue plaguing intervention research in young children with ASD. Three issues appear especially important to point out, including the preponderance of quasi-experimental group designs, reliance on caregiver/teacher report, and correlated measurement error due to interaction partners or assessors who participated in the intervention.

Although it is well established that randomized controlled trials offer the best protection against alternative explanations for intervention effects, quasi-experimental studies continue to be relied upon in autism intervention research. There are some circumstances wherein quasi-experimental methods may be appropriate, such as studies aiming to move established interventions into community settings where groups are already intact and randomizing participants would be prohibitively costly (e.g., Vivanti et al., 2014). However, our results suggest that we do not yet have intervention types that can be considered “established” to an extent that would warrant this strategy. Because there were too few studies to permit the estimation of summary effects once study design and performance bias were taken into account, we suggest that researcher and funding resources should continue to focus on establishing study efficacy using the highest quality designs.

Another area of particular concern is continued reliance on parent/teacher report. These measures are nearly impossible to administer in such a way that the respondent is unaware of the child's participation in an intervention. Indeed, research has shown that when caregivers complete such measures, an intervention "effect" will be demonstrated if they believe their child is receiving an intervention even when no intervention has actually occurred (Jones et al., 2017). We therefore suggest that early intervention researchers should not rely on such measures, and instead seek alternative measurement systems that can be administered and scored by assessors who are unaware of group assignment.

Finally, correlated measurement error that occurs when parents or teachers are trained in an intervention and also participate as assessors is a common threat to validity that has received little attention from the field. Continued use of observational measures taken from interactions with trained caregivers may be fruitful for mediation analyses, in order to verify that posttreatment group differences in developmentally distal and generalized outcomes are explained at least in part by changes in reciprocal interactions with caregivers within the context of intervention. However, researchers should recognize that these measures are biased in favor of the intervention group, and should therefore not rely on them as a primary index of intervention effects. Researchers should also employ valid, standardized, independently administered assessments as primary outcomes whenever possible. While changes in interactions between a trained caregiver and child may be important to measure if those interactions are expected to be the "mechanism" through which the child achieves later developmental milestones, these interactions may not themselves index improvements in the child's interactional repertoire. If researchers consider interactions with a familiar person as the most valid context for outcome assessment, they can avoid this threat to validity by relying on observational measures taken from interactions with familiar but untrained interaction partners (e.g., untrained teachers, untrained parents, untrained siblings, or untrained peers). Use of untrained interaction partners that are also naive to group assignment will further help researchers address the added threat of detection bias.

Understanding Intervention Outcomes—Boundedness and Proximity

Replicating previous research syntheses (Fuller & Kaiser, 2019; Yoder et al., 2013) and confirming our hypotheses, effect sizes were larger for indices of context-bound behaviors as compared with generalized child characteristics. This finding confirms that interventions (broadly considered) produce larger effects on behaviors that are potentially bound to the treatment context, which are likely easier to change, than on more highly generalized characteristics of young children with ASD. In certain circumstances, context-bound behavior change may be considered important. For example, if a study aims to improve children's classroom engagement, many would consider it acceptable if these effects did not generalize beyond the classroom, as the effects are likely only relevant in classroom contexts.

However, many stakeholders may expect interventions aiming to improve child characteristics associated with longer-term development (e.g., social communication) to produce gains that generalize to contexts beyond intervention settings. If developmentally important effects cannot be demonstrated outside intervention settings, it is unlikely that they will continue to be a part of the child's behavioral repertoire, in any context, once the intervention has stopped. Unfortunately, researchers do not always indicate whether their measurement system was restricted to detecting context-bound behaviors, or if it was able to detect gains in generalized child characteristics. We encourage researchers to make this distinction clear when presenting their study design, and when describing potential limitations in the case of studies that exclusively examine context-bound behavior change.

Our hypothesis was also confirmed in regards to proximity; effect sizes for proximal outcomes were larger than effect sizes for distal outcomes. Parallel to our findings on boundedness, this indicates that interventions are more effective at achieving gains on outcomes that reflect what was directly addressed in the intervention than gains on outcomes that are broader or beyond what was directly taught. Evidence of distal effects provide some evidence that the intervention is tapping into a developmental pathway, which can give researchers confidence that the intervention will continue to influence children's development after the intervention period is over.

There are some caveats to our approach in categorizing outcome proximity. One is that this concept is likely more accurately described as continual rather than binary. There are degrees of proximity and distality that we were not able to capture by restricting our coding to only two categories. A second caveat is that we were limited to the information about the intervention provided by study authors, which was often quite sparse. When delineating the focus of the intervention, authors did not always clarify if they were describing the immediate targets of the intervention, or a developmentally downstream target. Similarly, many studies did not offer a detailed description of the intervention, which hampered our ability to determine which outcomes were directly addressed by intervention procedures. Finally, proximity and distality are conflated with type of measurement system. Norm-referenced, standardized measures generally assess broad contexts which by definition cannot be directly targeted by intervention procedures and are therefore categorized as distal. On the other hand, observational measures of particular behaviors are often designed by researchers specifically to detect the most immediate effects of intervention (e.g., observational measures of joint engagement for interventions that seek to increase the amount of time children spend jointly engaged), which would be categorized as proximal. Thus, proximal measures may be more sensitive to change than distal measures, while distal measures are likely more construct valid than researcher-created proximal measures.

Interpreting Findings in Light of the Exclusion of Evidence From SSDs

It should be reiterated that we exclusively synthesized findings from randomized and nonrandomized group design studies of interventions for children with ASD. By excluding studies with single group pretest–posttest designs and SSDs, we have omitted a substantial body of research that has been used to draw conclusions about evidence-based practice, particularly in regards to the effectiveness of behavior analytic approaches. In fact, as of 2015, the majority of the available studies of intervention techniques for children with autism employed SSD (Wong et al., 2015), though our review and other reviews published since attest to the recent precipitous increase in group design literature published in this field (French & Kennedy, 2018).

Our decision to exclude SSDs from this meta-analysis was rooted primarily in the lack of adequate and agreed upon effect size metrics for synthesizing effects (Kratochwill et al., 2013). However, we believe there are additional insights to be gained from limiting our conclusions specifically to evidence offered by group design studies. Though SSDs are well-equipped to identify effective techniques for teaching specific targeted skills, group design studies are particularly useful for determining whether interventions can facilitate gains in generalized development. The repeated measurement that is a hallmark of SSDs may allow investigators to understand variability in specific behaviors associated with careful and controlled changes in the independent variable, but it limits reliance on validated standardized assessments as outcome measures. Such assessments, though often time consuming to administer, are likely better equipped to tap improvements in generalized development than researcher-created operationalizations of specific behaviors. Thus, if we wish to evaluate whether intervention facilitates developmental progress in young children with autism on average, an evaluation of group design studies may, arguably, be more methodologically suited for this purpose. However, even though group design studies may be preferable in this regard, ours and other recent work has shown that a substantial portion of the outcome measures used in clinical trials were overly specific to the intervention context and targets (Provenzani et al., 2019). Thus, fragmented measurement approaches continue to limit the conclusions that can be drawn regarding the effectiveness of autism interventions, both in SSDs and group design studies. This remains a limitation, both for the body of evidence as a whole, and our conclusions here.

Recommendations for Primary Intervention Research

Given the results of this series of meta-analyses, we propose several recommendations. While our confidence in summary effect estimates for *any* intervention type is hindered by a lack of high-quality studies, we do have single examples of studies that meet the majority of quality indicators (e.g., Green et al., 2010). This suggests that designing a high-quality study is not an unreachable challenge for early intervention researchers. It would perhaps incentivize future high-quality research if funding agencies held investigators to a higher standard and required basic quality features such as randomized trials and measurement systems that can be administered in such a way that assessors remain naive to treatment status. At the very least, caregiver and teacher reports should likely be discarded altogether, as it is already clear that they introduce bias and render findings largely uninterpretable (Jones et al., 2017). For some domains, this may mean that new measures will need to be developed and validated that are low-cost to administer and adequately sensitive to change.

A second recommendation, also related to measurement systems, is that researchers should provide detailed descriptions of each measure (especially if they are researcher-created), and the assessment process in which each measure is used. This will allow for an adequate assessment of the kinds of bias introduced or avoided by particular approaches to measurement, and will allow for a determination of whether measures are capturing context-bound behavior change or generalized characteristics. To make this latter determination, aspects such as the measurement context, who administered the measure, and the materials and activities used during measurement should be made clear.

Third, we were quite struck by how little information many studies contained in regards to the intervention that was tested. Though it is not necessary that every study on a given intervention provide minute detail of the procedures, it would be helpful if there were at least one manualized protocol available for each intervention that describes the full set of strategies and activities involved in implementing the intervention. This would encourage independent replication of intervention studies, and would allow for a determination of whether the outcomes measured were proximal or distal to the intervention procedures. To make this distinction, researchers need to go beyond describing the aims of the intervention—they need to specifically describe the protocol in such a way that the immediate outcomes of implementing the intervention are readily discernible.

Fourth, 50 studies were excluded because relevant effect size information was not published or extractable. In many cases, this was due to exclusive reporting of change scores or postintervention means adjusted for various baseline covariates, which should not be meta-analyzed alongside standardized mean differences extracted from unadjusted means (Deeks, Higgins, & Altman, 2019). Though we contacted authors in every case wherein studies were less than 10-years-old, many failed to respond. Therefore, we recommend that authors reporting results that control for covariates include unadjusted means and *SD*s in online supplementary materials, to facilitate future attempts at meta-analysis.

Finally, we suggest that an “optimal” intervention design would include paired proximal and distal measures (or perhaps even include a third, far distal measure) that are expected to be developmentally connected and malleable to change. The proximal measure should be selected to capture the immediate effects of the intervention, while the distal measure should be selected to measure effects hypothesized to be developmentally downstream from proximal effects. Mediation analyses, in which the proximal measure is the mediating variable and the distal measure is the outcome variable, could then confirm whether the proposed developmental pathway between proximal and distal effects was activated by participation in the intervention. This would allow for a better understanding of the mechanisms or “active ingredients” through which interventions achieve cascading developmental gains.

Limitations and Future Meta-Analytic Research

There are at least three limitations to consider when interpreting the results of this study. First, despite our best efforts, we were unable to collect any unpublished effect sizes or data sets apart from dissertations and theses. This could mean that the effect size estimates presented here are larger than the “true” effects (an interpretation supported by inspection of funnel plots). Our attempts to gather unpublished effect sizes included searching NIH, NDAR, and IES databases, and requesting data directly from investigators who were reported to have received funding for group design intervention research in children with ASD. However, we did not receive any unpublished data from any researchers, suggesting there may be reticence among researchers to share their unpublished data. This is unfortunate, as access to unpublished data is critical for accurately estimating effect sizes, and accurately assessing the "state of the science." Further, data sharing practices are critical to ensure replicability of findings (Nuijten, 2018).

A second limitation to consider is that there were too few studies to adequately synthesize effect sizes for all outcomes and intervention types. This was especially true when quality indicators were taken into consideration. Researchers will need to commit to conducting high-quality intervention research in order for future syntheses to accurately draw conclusions about intervention effectiveness for outcomes of interest in children with autism.

Finally, the heterogeneity of variables within each "outcome type" and treatments represented within each "intervention type" may limit the interpretability of our summary effect estimates. While we note that variables and intervention approaches were similar enough to be categorized with high reliability—kappa coefficients for outcome type and intervention type coding were 0.862 and 0.907, respectively—categorization of items that differ on a continuum will always result in the loss of information, and this information may be important for understanding key components that drive intervention effects. For example, the same intervention provided with different intensities (i.e., number of hours per week) may yield different effects. Similarly, intervention effects may differ for variables that share a domain but are distinct (e.g., social communication variables such as responding to joint attention and initiating joint attention). More fine-grained analyses within each outcome type could allow us to answer questions about putative moderators, as well as to calculate subgroup effect sizes for identical outcome types across studies (e.g., Vineland scores), or identical interventions (e.g., PECS) as the literature base on treatment effects in children with ASD continues to grow.

Conclusions

The current study differs from existing reviews on intervention in children with ASD in two important ways. First, this study is one of few attempts to consider all intervention types and intervention outcomes as broadly as possible. This allows us to report the state of the science in regards to which interventions have accrued the most convincing evidence of effectiveness for young children with ASD, and to report on the full range of outcomes that these interventions are able to influence. Second, this study accounts for rigorous quality criteria that are common considerations in other areas of psychology, but that are applied less often to evaluations of autism research (e.g., Reichow, Volkmar, & Cicchetti, 2008). Finally, several syntheses that are similar to ours in scope consider some of the design factors of included studies in order to classify intervention types according to levels of evidence (e.g., Wong et al., 2015). However, these syntheses have not provided an examination of intervention effects according to characteristics of the outcome variable, which prevents researchers from drawing conclusions in regards to whether interventions are able to influence generalized characteristics that extend beyond the skills directly targeted by the interventions. Our findings echo recent sentiments from intervention researchers who are heartened by the relative increase in RCTs over the past 15 years, but also raise concerns in regards to the availability of high-quality study designs that reliably and consistently link established interventions with meaningful child outcomes (Charman, 2019).

Even given these concerns, the evidence base regarding intervention for children with ASD has been rapidly transforming. The last decade has seen the publication of over 100 group design studies of intervention, including at least 50 RCTs. These studies attest to the fact that access to intervention in early childhood can yield a range of positive outcomes for the children receiving it. NDBIs have emerged as a new intervention category with significant summary effects even when several quality indicators are taken into account. High-quality studies also suggest that developmental intervention can improve some core challenges associated with ASD, particularly difficulties in social communication. Traditional behavioral intervention approaches show some evidence of effectiveness, but methodological rigor remains a pressing concern for this body of research. There is little evidence to date, however, to support the effectiveness of several other interventions that are geared toward young children with autism, including TEACCH, sensory-based interventions, animal-assisted interventions, and interventions mediated solely through technology (though approaches that integrate technology, such as high-tech augmentative and alternative communication devices, into more established interventions appear promising). More high-quality randomized-controlled trials that feature independently administered assessments are needed to unquestionably establish the efficacy of any intervention type. Finally, researchers should consider the characteristics (i.e., distality and boundedness) of outcomes being tracked in intervention studies and interpret findings accordingly to permit a more ready assessment of the extent to which any particular treatment approach is likely to yield desired effects on developmental trajectories of young children affected by autism.

References

References marked with an asterisk indicate studies included in the meta-analysis.

*Adkins, K. W., Molloy, C., Weiss, S. K., Reynolds, A., Goldman, S. E., Burnette, C., . . . Malow, B. A. (2012). Effects of a standardized pamphlet on insomnia in children with autism spectrum disorders. *Pediatrics, 130,* S139–S144. http://dx.doi.org/10.1542/peds.2012-0900K

*Aldred, C., Green, J., & Adams, C. (2004). A new social communication intervention for children with autism: Pilot randomised controlled treatment study suggesting effectiveness. *Journal of Child Psychology and Psychiatry, 45,* 1420–1430. http://dx.doi.org/10.1111/j.1469-7610.2004.00338.x

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

Ausderau, K., Sideris, J., Furlong, M., Little, L. M., Bulluck, J., & Baranek, G. T. (2014). National Survey of Sensory Features in Children With ASD: Factor structure of the Sensory Experience Questionnaire (3.0). *Journal of Autism and Developmental Disorders, 44,* 915–925.

Autism and Insurance Coverage | State Laws. (2018). Retrieved from http://www.ncsl.org/research/health/autism-and-insurance-coverage-state-laws.aspx

Ayres, A. J. (1979). *Sensory integration and the child*. Los Angeles, CA: Western Psychological Services.

Ayres, A. J. (2005). *Sensory integration and the child, 25th anniversary edition*. Los Angeles, CA: Western Psychological Services.

*Azarbehi, A. C. (2009). *The effectiveness of early intervention programs for children with autism: A one-year follow-up study of intense behavioural intervention versus preschool integration* (Unpublished doctoral dissertation). Academic Unit of Psychology of The University of New Brunswick, Fredericton, New Brunswick, Canada.

Bahrick, L. E., & Todd, J. T. (2012). Multisensory processing in autism spectrum disorders: Intersensory processing disturbance as a basis for atypical development. In B. E. Stein & M. T. Wallace (Eds.), *The new handbook of multisensory processes* (pp. 1453–1508). Cambridge, MA: MIT Press.

*Bailey, B., Arciuli, J., & Stancliffe, R. J. (2017). Effects of ABRACADABRA literacy instruction on children with autism spectrum disorder. *Journal of Educational Psychology, 109,* 257–268. http://dx.doi.org/10.1037/edu0000138

Baio, J., Wiggins, L., Christensen, D. L., Maenner, M. J., Daniels, J., Warren, Z., . . . Dowling, N. F. (2018). Prevalence of autism spectrum disorder among children aged 8 years—Autism and developmental disabilities monitoring network, 11 sites, United States, 2014. *Morbidity and Mortality Weekly Report, 67,* 1–23. http://dx.doi.org/10.15585/mmwr.ss6706a1

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

Baron-Cohen, S., Golan, O., & Ashwin, E. (2012). Educational cognitive neuroscience: Designing autism-friendly methods to teach emotion recognition. In S. Della Salla & M. Anderson (Eds.), *Neuroscience in Education: The good, the bad and the ugly* (pp. 299–311). Oxford, UK: Oxford University Press.

Barton, E. E., Reichow, B., Schnitz, A., Smith, I. C., & Sherlock, D. (2015). A systematic review of sensory-based treatments for children with disabilities. *Research in Developmental Disabilities, 37,* 64–80. http://dx.doi.org/10.1016/j.ridd.2014.11.006

*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,* 1261–1267. http://dx.doi.org/10.1007/s10803-009-0734-3

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.

*Bieleninik, L., Geretsegger, M., Mössler, K., Assmus, J., Thompson, G., Gattino, G., . . . Gold, C. (2017). Effects of improvisational music therapy vs enhanced standard care on symptom severity among children with autism spectrum disorder: The TIME-A randomized clinical trial. *Journal of the American Medical Association, 318,* 525–535. http://dx.doi.org/10.1001/jama.2017.9478

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

Boyd, B. A., Odom, S. L., Humphreys, B. P., & Sam, A. M. (2010). Infants and toddlers with autism spectrum disorder: Early identification and early intervention. *Journal of Early Intervention, 32,* 75–98. http://dx.doi.org/10.1177/1053815110362690

*Boyd, B. A., Watson, L. R., Reszka, S. S., Sideris, J., Alessandri, M., Baranek, G. T., . . . Belardi, K. (2018). Efficacy of the ASAP intervention for preschoolers with ASD: A cluster randomized controlled trial. *Journal of Autism and Developmental Disorders, 48,* 3144–3162. http://dx.doi.org/10.1007/s10803-018-3584-z

*Bremer, E., Balogh, R., & Lloyd, M. (2015). Effectiveness of a fundamental motor skill intervention for 4-year-old children with autism spectrum disorder: A pilot study. *Autism, 19,* 980–991. http://dx.doi.org/10.1177/1362361314557548

*Brian, J. A., Smith, I. M., Zwaigenbaum, L., & Bryson, S. E. (2017). Cross-site randomized control trial of the Social ABCs caregiver-mediated intervention for toddlers with autism spectrum disorder. *Autism Research, 10,* 1700–1711. http://dx.doi.org/10.1002/aur.1818

Case-Smith, J., & Arbesman, M. (2008). Evidence-based review of interventions for autism used in or of relevance to occupational therapy. *The American Journal of Occupational Therapy, 62,* 416–429.

*Carter, A. S., Messinger, D. S., Stone, W. L., Celimli, S., Nahmias, A. S., & Yoder, P. (2011). A randomized controlled trial of Hanen's 'More Than Words' in toddlers with early autism symptoms. *Journal of Child Psychology and Psychiatry, 52,* 741–752. http://dx.doi.org/10.1111/j.1469-7610.2011.02395.x

Cascio, C. J., Woynaroski, T., Baranek, G. T., & Wallace, M. T. (2016). Toward an interdisciplinary approach to understanding sensory function in autism spectrum disorder. *Autism Research, 9,* 920–925. http://dx.doi.org/10.1002/aur.1612

*Casenhiser, D. M., Binns, A., McGill, F., Morderer, O., & Shanker, S. G. (2015). Measuring and supporting language function for children with autism: Evidence from a randomized control trial of a social-interaction-based therapy. *Journal of Autism and Developmental Disorders, 45,* 846–857. http://dx.doi.org/10.1007/s10803-014-2242-3

*Casenhiser, D. M., Shanker, S. G., & Stieben, J. (2013). Learning through interaction in children with autism: Preliminary data from asocial-communication-based intervention. *Autism, 17,* 220–241. http://dx.doi.org/10.1177/1362361311422052

Case-Smith, J., Weaver, L. L., & Fristad, M. A. (2015). A systematic review of sensory processing interventions for children with autism spectrum disorders. *Autism, 19,* 133–148. http://dx.doi.org/10.1177/1362361313517762

*Chang, Y. C., Shire, S. Y., Shih, W., Gelfand, C., & Kasari, C. (2016). Preschool deployment of evidence-based social communication intervention: JASPER in the classroom. *Journal of Autism and Developmental Disorders, 46,* 2211–2223. http://dx.doi.org/10.1007/s10803-016-2752-2

Charman, T. (2019). Editorial: Trials and tribulations in early autism intervention research. *Journal of the American Academy of Child & Adolescent Psychiatry, 58,* 846–848. http://dx.doi.org/10.1016/j.jaac.2019.03.004

*Cheldavi, H., Shakerian, S., Boshehri, S. N. S., & Zarghami, M. (2014). The effects of balance training intervention on postural control of children with autism spectrum disorder: Role of sensory information. *Research in Autism Spectrum Disorders, 8,* 8–14. http://dx.doi.org/10.1016/j.rasd.2013.09.016

*Chiang, C. H., Chu, C. L., & Lee, T. C. (2016). Efficacy of caregiver-mediated joint engagement intervention for young children with autism spectrum disorders. *Autism, 20,* 172–182. http://dx.doi.org/10.1177/1362361315575725

*Clionsky, L. N. (2012). *Child directed interaction training for young children with autism spectrum disorders: The impact on child language, social skills, adaptive skills, and problematic behaviors* (Unpublished doctoral dissertation). University of Florida, Gainesville, FL.

*Cohen, H., Amerine-Dickens, M., & Smith, T. (2006). Early intensive behavioral treatment: Replication of the UCLA model in a community setting. *Journal of Developmental and Behavioral Pediatrics, 27,* S145–S155. http://dx.doi.org/10.1097/00004703-200604002-00013

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20,* 37–46. http://dx.doi.org/10.1177/001316446002000104

Corbett, B. A., Schupp, C. W., Levine, S., & Mendoza, S. (2009). Comparing cortisol, stress, and sensory sensitivity in children with autism. *Autism Research, 2,* 39–49. http://dx.doi.org/10.1002/aur.64

*Corbett, B. A., Shickman, K., & Ferrer, E. (2008). Brief report: The effects of Tomatis sound therapy on language in children with autism. *Journal of Autism and Developmental Disorders, 38,* 562–566. http://dx.doi.org/10.1007/s10803-007-0413-1

*Dawson, G., Jones, E. J., Merkle, K., Venema, K., Lowy, R., Faja, S., . . . Webb, S. J. (2012). Early behavioral intervention is associated with normalized brain activity in young children with autism. *Journal of the American Academy of Child & Adolescent Psychiatry, 51,* 1150–1159. http://dx.doi.org/10.1016/j.jaac.2012.08.018

*Dawson, G., Rogers, S., Munson, J., Smith, M., Winter, J., Greenson, J., . . . Varley, J. (2010). Randomized, controlled trial of an intervention for toddlers with autism: The Early Start Denver Model. *Pediatrics, 125,* e17–e23. http://dx.doi.org/10.1542/peds.2009-0958

Dawson, G., & Zanolli, K. (2003). Early intervention and brain plasticity in autism. In M. Rutter (Ed.), *Autism: Neural bases and treatment possibilities* (pp. 266–280). London, UK: Novartis.

Deeks, J. J., Higgins, J. P., & Altman, D. G. (2019). Analysing data and undertaking meta-analyses. In J. P. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page, & V. A. Welch (Eds.), *Cochrane handbook for systematic reviews of interventions* (pp. 243–295). http://dx.doi.org/10.1002/9781119536604.ch10

*D'Elia, L., Valeri, G., Sonnino, F., Fontana, I., Mammone, A., & Vicari, S. (2014). A longitudinal study of the TEACCH program in different settings: The potential benefits of low intensity intervention in preschool children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 44,* 615–626. http://dx.doi.org/10.1007/s10803-013-1911-y

den Houting, J. (2019). Neurodiversity: An insider's perspective. *Autism, 23,* 271–273. http://dx.doi.org/10.1177/1362361318820762

*Drew, A., Baird, G., Baron-Cohen, S., Cox, A., Slonims, V., Wheelwright, S., . . . Charman, T. (2002). A pilot randomised control trial of a parent training intervention for pre-school children with autism. Preliminary findings and methodological challenges. *European Child & Adolescent Psychiatry, 11,* 266–272. http://dx.doi.org/10.1007/s00787-002-0299-6

*Duifhuis, E. A., den Boer, J. C., Doornbos, A., Buitelaar, J. K., Oosterling, I. J., & Klip, H. (2017). The effect of pivotal response treatment in children with autism spectrum disorders: A non-randomized study with a blinded outcome measure. *Journal of Autism and Developmental Disorders, 47,* 231–242. http://dx.doi.org/10.1007/s10803-016-2916-0

*Durand, V. M., Hieneman, M., Clarke, S., Wang, M., & Rinaldi, M. L. (2013). Positive family intervention for severe challenging behavior I: A multisite randomized clinical trial. *Journal of Positive Behavior Interventions, 15,* 133–143. http://dx.doi.org/10.1177/1098300712458324

*Eikeseth, S., Klintwall, L., Jahr, E., & Karlsson, P. (2012). Outcome for children with autism receiving early and intensive behavioral intervention in mainstream preschool and kindergarten settings. *Research in Autism Spectrum Disorders, 6,* 829–835. http://dx.doi.org/10.1016/j.rasd.2011.09.002

*Eikeseth, S., Smith, T., Jahr, E., & Eldevik, S. (2002). Intensive behavioral treatment at school for 4- to 7-year-old children with autism. A 1-year comparison controlled study. *Behavior Modification, 26,* 49–68. http://dx.doi.org/10.1177/0145445502026001004

*Eikeseth, S., Smith, T., Jahr, E., & Eldevik, S. (2007). Outcome for children with autism who began intensive behavioral treatment between ages 4 and 7: A comparison controlled study. *Behavior Modification, 31,* 264–278. http://dx.doi.org/10.1177/0145445506291396

*Elder, L. M. (2011). *The relationship between parent training and parent-child interaction in autism* (Unpublished doctoral dissertation). University of Washington, Seattle, WA.

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

*ElGarhy, S., & Liu, T. (2016). Effects of psychomotor intervention program on students with autism spectrum disorder. *School Psychology Quarterly, 31,* 491–506. http://dx.doi.org/10.1037/spq0000164

*Espanola, E. (2015). *A randomized comparison of two instructional sequences for imitation intervention for children with autism spectrum disorders* (Unpublished doctoral dissertation). Florida International University, Miami, FL.

*Estes, A., Munson, J., Rogers, S. J., Greenson, J., Winter, J., & Dawson, G. (2015). Long-term outcomes of early intervention in 6-year-old children with autism spectrum disorder. *Journal of the American Academy of Child & Adolescent Psychiatry, 54,* 580–587. http://dx.doi.org/10.1016/j.jaac.2015.04.005

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

*Fernell, E., Hedvall, Å., Westerlund, J., Höglund Carlsson, L., Eriksson, M., Barnevik Olsson, M., . . . Gillberg, C. (2011). Early intervention in 208 Swedish preschoolers with autism spectrum disorder. A prospective naturalistic study. *Research in Developmental Disabilities, 32,* 2092–2101. http://dx.doi.org/10.1016/j.ridd.2011.08.002

Ferster, C. B., & Demyer, M. K. (1962). A method for the experimental analysis of the behavior of autistic children. *American Journal of Orthopsychiatry, 32,* 89–98. http://dx.doi.org/10.1111/j.1939-0025.1962.tb00267.x

Fisher, Z., Tipton, E., & Zhipeng, H. (2017). *robumeta: Robust variance meta-regression* (Version 2.0). Retrieved from https://CRAN.R-project.org/package=robumeta

*Flanagan, H. E., Perry, A., & Freeman, N. L. (2012). Effectiveness of large-scale community-based Intensive Behavioral Intervention: A wait-list comparison study exploring outcomes and predictors. *Research in Autism Spectrum Disorders, 6,* 673–682.

Fletcher-Watson, S. (2018, June). Is early intervention compatible with neurodiversity? [Blog post]. Retrieved from https://dart.ed.ac.uk/intervention-neurodiversity/

*Fletcher-Watson, S., Petrou, A., Scott-Barrett, J., Dicks, P., Graham, C., O'Hare, A., . . . McConachie, H. (2016). A trial of an iPad™ intervention targeting social communication skills in children with autism. *Autism, 20,* 771–782. http://dx.doi.org/10.1177/1362361315605624

French, L., & Kennedy, E. M. M. (2018). Annual research review: Early intervention for infants and young children with, or at-risk of, autism spectrum disorder: A systematic review. *Journal of Child Psychology and Psychiatry, 59,* 444–456. http://dx.doi.org/10.1111/jcpp.12828

*Fridenson-Hayo, S., Berggren, S., Lassalle, A., Tal, S., Pigat, D., Meir-Goren, N., . . . Golan, O. (2017). 'Emotiplay': a serious game for learning about emotions in children with autism: Results of a cross-cultural evaluation. *European Child & Adolescent Psychiatry, 26,* 979–992. http://dx.doi.org/10.1007/s00787-017-0968-0

Fuller, E. A., & Kaiser, A. P. (2019). The effects of early intervention on social communication outcomes for children with autism spectrum disorder: A meta-analysis. *Journal of Autism and Developmental Disorders*. Advance online publication. http://dx.doi.org/10.1007/s10803-019-03927-z

Gabriels, R. L., Agnew, J. A., Holt, K. D., Shoffner, A., Zhaoxing, P., Ruzzano, S., . . . Mesibov, G. (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,* 578–588. http://dx.doi.org/10.1016/j.rasd.2011.09.007

Gamer, M., Lemon, J., Fellows, I., & Singh, P. (2012). *irr: Various coefficients of interrater reliability and agreement. R package v. 0.84*. Retrieved from http://cran.r-project.org/web/packages/irr/index.html

Gernsbacher, M. A. (2017). Editorial perspective: The use of person-first language in scholarly writing may accentuate stigma. *Journal of Child Psychology and Psychiatry, 58,* 859–861. http://dx.doi.org/10.1111/jcpp.12706

Goin-Kochel, R. P., Mackintosh, V. H., & Myers, B. J. (2009). Parental reports on the efficacy of treatments and therapies for their children with autism spectrum disorders. *Research in Autism Spectrum Disorders, 3,* 528–537.

*Golan, O., Ashwin, E., Granader, Y., McClintock, S., Day, K., Leggett, V., & Baron-Cohen, S. (2010). Enhancing emotion recognition in children with autism spectrum conditions: An intervention using animated vehicles with real emotional faces. *Journal of Autism and Developmental Disorders, 40,* 269–279. http://dx.doi.org/10.1007/s10803-009-0862-9

*Grahame, V., Brett, D., Dixon, L., McConachie, H., Lowry, J., Rodgers, J., . . . Le Couteur, A. (2015). Managing repetitive behaviours in young children with autism spectrum disorder (ASD): Pilot randomised controlled trial of a new parent group intervention. *Journal of Autism and Developmental Disorders, 45,* 3168–3182. http://dx.doi.org/10.1007/s10803-015-2474-x

*Green, J., Charman, T., McConachie, H., Aldred, C., Slonims, V., Howlin, P., . . . Pickles, A. (2010). Parent-mediated communication-focused treatment in children with autism (PACT): A randomised controlled trial. *Lancet, 375,* 2152–2160. http://dx.doi.org/10.1016/S0140-6736(10)60587-9

Green, V. A., Pituch, K. A., Itchon, J., Choi, A., O'Reilly, M., & Sigafoos, J. (2006). Internet survey of treatments used by parents of children with autism. *Research in Developmental Disabilities, 27,* 70–84. http://dx.doi.org/10.1016/j.ridd.2004.12.002

Greenspan, S. I., & Wieder, S. (2007). The developmental individual-difference, relationship-based (DIR/Floortime) model approach to autism spectrum disorders. In E. Hollander & E. Anagnostou (Eds.), *Clinical manual for the treatment of autism* (pp. 179–209). Arlington, VA: American Psychiatric Publishing, Inc.

Grelotti, D. J., & Kaptchuk, T. J. (2011). Placebo by proxy. *British Medical Journal, 343,* d4345. http://dx.doi.org/10.1136/bmj.d4345

Grynszpan, O., Weiss, P. L., Perez-Diaz, F., & Gal, E. (2014). Innovative technology-based interventions for autism spectrum disorders: A meta-analysis. *Autism: An International Journal of Research and Practise, 18,* 346–361. http://dx.doi.org/10.1177/1362361313476767

*Gulsrud, A. C., Hellemann, G. S., Freeman, S. F., & Kasari, C. (2014). Two to ten years: Developmental trajectories of joint attention in children with ASD who received targeted social communication interventions. *Autism Research, 7,* 207–215. http://dx.doi.org/10.1002/aur.1360

*Gulsrud, A. C., Kasari, C., Freeman, S., & Paparella, T. (2007). Children with autism's response to novel stimuli while participating in interventions targeting joint attention or symbolic play skills. *Autism, 11,* 535–546. http://dx.doi.org/10.1177/1362361307083255

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

Harrell, F. E. (2018). Hmisc: Harrell Miscellaneous (Version 4.1–1). Retrieved from https://CRAN.R-project.org/package=Hmisc

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods, 1,* 39–65. http://dx.doi.org/10.1002/jrsm.5

*Hemdi, A., & Daley, D. (2017). The effectiveness of a psychoeducation intervention delivered via WhatsApp for mothers of children with autism spectrum disorder (ASD) in the Kingdom of Saudi Arabia: A randomized controlled trial. *Child: Care, Health and Development, 43,* 933–941. http://dx.doi.org/10.1111/cch.12520

Higgins, J. P., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., . . . Sterne, J. A. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *British Medical Journal, 343,* d5928. http://dx.doi.org/10.1136/bmj.d5928

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

*Howlin, P., & Rutter, M. (1989). Mothers' speech to autistic children: A preliminary causal analysis. *Journal of Child Psychology and Psychiatry, 30,* 819–843. http://dx.doi.org/10.1111/j.1469-7610.1989.tb00285.x

Hróbjartsson, A., Thomsen, A. S. S., Emanuelsson, F., Tendal, B., Hilden, J., Boutron, I., . . . Brorson, S. (2012). Observer bias in randomised clinical trials with binary outcomes: Systematic review of trials with both blinded and non-blinded outcome assessors. *British Medical Journal, 344,* e1119. http://dx.doi.org/10.1136/bmj.e1119

*Hudson, R. F., Sanders, E. A., Greenway, R., Xie, S., Smith, M., Gasamis, C., . . . Hackett, J. (2017). Effects of emergent literacy interventions for preschoolers with autism spectrum disorder. *Exceptional Children, 84,* 55–75. http://dx.doi.org/10.1177/0014402917705855

*Iadarola, S., Shih, W., Dean, M., Blanch, E., Harwood, R., Hetherington, S., . . . Smith, T. (2018). Implementing a manualized, classroom transition intervention for students with ASD in underresourced schools. *Behavior Modification, 42,* 126–147. http://dx.doi.org/10.1177/0145445517711437

*Ichikawa, K., Takahashi, Y., Ando, M., Anme, T., Ishizaki, T., Yamaguchi, H., & Nakayama, T. (2013). TEACCH-based group social skills training for children with high-functioning autism: A pilot randomized controlled trial. *BioPsychoSocial Medicine, 7,* 14. http://dx.doi.org/10.1186/1751-0759-7-14

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

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

*Ingersoll, B., Wainer, A. L., Berger, N. I., Pickard, K. E., & Bonter, N. (2016). Comparison of a self-directed and therapist-assisted telehealth parent-mediated intervention for children with ASD: A pilot RCT. *Journal of Autism and Developmental Disorders, 46,* 2275–2284. http://dx.doi.org/10.1007/s10803-016-2755-z

*Jalili, M., Jahangiri, N., Yazdi, A. A., & Ashrafzadeh, F. (2014). The effects of imitative vs. Cognitive methods on the speech development of children with autism. *Iranian Journal of Child Neurology, 8,* 37–46.

*Jocelyn, L. J., Casiro, O. G., Beattie, D., Bow, J., & Kneisz, J. (1998). Treatment of children with autism: A randomized controlled trial to evaluate a caregiver-based intervention program in community day-care centers. *Journal of Developmental and Behavioral Pediatrics, 19,* 326–334. http://dx.doi.org/10.1097/00004703-199810000-00002

*Johnson, C. R., Turner, K. S., Foldes, E., Brooks, M. M., Kronk, R., & Wiggs, L. (2013). Behavioral parent training to address sleep disturbances in young children with autism spectrum disorder: A pilot trial. *Sleep Medicine, 14,* 995–1004. http://dx.doi.org/10.1016/j.sleep.2013.05.013

Jones, R. M., Carberry, C., Hamo, A., & Lord, C. (2017). Placebo-like response in absence of treatment in children with Autism. *Autism Research, 10,* 1567–1572. http://dx.doi.org/10.1002/aur.1798

*Jouen, A. L., Narzisi, A., Xavier, J., Tilmont, E., Bodeau, N., Bono, V., . . . Cohen, D. (2017). GOLIAH (Gaming Open Library for Intervention in Autism at Home): A 6-month single blind matched controlled exploratory study. *Child and Adolescent Psychiatry and Mental Health, 11,* 17. http://dx.doi.org/10.1186/s13034-017-0154-7

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

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

Kaiser, A. P. (1993). Parent-implemented language intervention: An environmental system perspective. In A. P. Kaiser & D. B. Gray (Eds.), *Enhancing children's communication: Research foundations for intervention* (Vol. 2, pp. 63–84). Baltimore, MD: Brookes.

*Kalyva, E., & Avramidis, E. (2005). Improving communication between children with autism and their peers through the 'circle of friends': A small-scale intervention study. *Journal of Applied Research in Intellectual Disabilities, 18,* 253–261. http://dx.doi.org/10.1111/j.1468-3148.2005.00232.x

*Kamps, D., Heitzman-Powell, L., Rosenberg, N., Mason, R., Schwartz, I., & Romine, R. S. (2016). Effects of Reading Mastery as a small group intervention for young children with ASD. *Journal of Developmental and Physical Disabilities, 28,* 703–722. http://dx.doi.org/10.1007/s10882-016-9503-3

Kasari, C. (2002). Assessing change in early intervention programs for children with autism. *Journal of Autism and Developmental Disorders, 32,* 447–461. http://dx.doi.org/10.1023/A:1020546006971

*Kasari, C., Freeman, S., & Paparella, T. (2006). Joint attention and symbolic play in young children with autism: A randomized controlled intervention study. *Journal of Child Psychology and Psychiatry, 47,* 611–620. http://dx.doi.org/10.1111/j.1469-7610.2005.01567.x

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

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

Kasari, C., Kaiser, A., Goods, K., Nietfeld, J., Mathy, P., Landa, R., . . . Almirall, D. (2014). Communication interventions for minimally verbal children with autism: A sequential multiple assignment randomized trial. *Journal of the American Academy of Child & Adolescent Psychiatry, 53,* 635–646. http://dx.doi.org/10.1016/j.jaac.2014.01.019

*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,* 125–137. http://dx.doi.org/10.1037/0022-006X.76.1.125

*Keen, D., Couzens, D., Muspratt, S., & Rodger, S. (2010). The effects of a parent-focused intervention for children with a recent diagnosis of autism spectrum disorder on parenting stress and competence. *Research in Autism Spectrum Disorders, 4,* 229–241. http://dx.doi.org/10.1016/j.rasd.2009.09.009

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

Knight, V., McKissick, B. R., & Saunders, A. (2013). A review of technology-based interventions to teach academic skills to students with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 43,* 2628–2648. http://dx.doi.org/10.1007/s10803-013-1814-y

Koegel, R. L., Koegel, L. K., & Carter, C. M. (1999). Pivotal teaching interactions for children with autism. *School Psychology Review, 28,* 576–594.

Kolb, B., & Gibb, R. (2011). Brain plasticity and behaviour in the developing brain. *Journal of the Canadian Academy of Child & Adolescent, 20,* 265–276.

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

Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2013). Single-case intervention research design standards. *Remedial and Special Education, 34,* 26–38. http://dx.doi.org/10.1177/0741932512452794

*LaGasse, A. B. (2014). Effects of a music therapy group intervention on enhancing social skills in children with autism. *Journal of Music Therapy, 51,* 250–275. http://dx.doi.org/10.1093/jmt/thu012

*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,* 13–21. http://dx.doi.org/10.1111/j.1469-7610.2010.02288.x

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

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

*Leaf, J. B., Leaf, J. A., Milne, C., Taubman, M., Oppenheim-Leaf, M., Torres, N., . . . Yoder, P. (2017). An evaluation of a behaviorally based social skills group for individuals diagnosed with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 47,* 243–259. http://dx.doi.org/10.1007/s10803-016-2949-4

Leekam, S. R., Nieto, C., Libby, S. J., Wing, L., & Gould, J. (2007). Describing the sensory abnormalities of children and adults with autism. *Journal of Autism and Developmental Disorders, 37,* 894–910.

*Lerna, A., Esposito, D., Conson, M., Russo, L., & Massagli, A. (2012). Social-communicative effects of the Picture Exchange Communication System (PECS) in autism spectrum disorders. *International Journal of Language & Communication Disorders, 47,* 609–617. http://dx.doi.org/10.1111/j.1460-6984.2012.00172.x

*Lim, H. A. (2010). Effect of "developmental speech and language training through music" on speech production in children with autism spectrum disorders. *Journal of Music Therapy, 47,* 2–26. http://dx.doi.org/10.1093/jmt/47.1.2

Lipsey, M. W., & Wilson, D. B. (2001). *Applied social research methods series: Vol. 49. Practical meta-analysis*. Thousand Oaks, CA: Sage.

Lord, C., Bristol-Power, M., Cafiero, J. M., Filipek, P. A., Gallagher, J. J., & Harris, S. L. (2001). *Educating children with autism*. Washington, DC: National Academy Press.

Lovaas, O. I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology, 55,* 3–9. http://dx.doi.org/10.1037/0022-006X.55.1.3

*Magiati, I., Charman, T., & Howlin, P. (2007). A two-year prospective follow-up study of community-based early intensive behavioural intervention and specialist nursery provision for children with autism spectrum disorders. *Journal of Child Psychology and Psychiatry, 48,* 803–812. http://dx.doi.org/10.1111/j.1469-7610.2007.01756.x

*Mahoney, G., & Solomon, R. (2016). Mechanism of developmental change in the PLAY project home consultation program: Evidence from a randomized control trial. *Journal of Autism and Developmental Disorders, 46,* 1860–1871. http://dx.doi.org/10.1007/s10803-016-2720-x

*Malow, B. A., Adkins, K. W., Reynolds, A., Weiss, S. K., Loh, A., Fawkes, D., . . . Clemons, T. (2014). Parent-based sleep education for children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 44,* 216–228. http://dx.doi.org/10.1007/s10803-013-1866-z

*Mandell, D. S., Stahmer, A. C., Shin, S., Xie, M., Reisinger, E., & Marcus, S. C. (2013). The role of treatment fidelity on outcomes during a randomized field trial of an autism intervention. *Autism, 17,* 281–295. http://dx.doi.org/10.1177/1362361312473666

*McCalla, M. K. (2014). *Enhancing the social interactions of preschool children with autism: The effectiveness of a peer-mediated* (Unpublished doctoral dissertation). The University of Alabama at Birmingham, Birmingham, AL.

*McEachin, J. J., Smith, T., & Lovaas, O. I. (1993). Long-term outcome for children with autism who received early intensive behavioral treatment. *American Journal on Mental Retardation, 97,* 359–372.

Mesibov, G. B., Shea, V., & Schopler, E. (2004). *The TEACCH approach to autism spectrum disorders*. New York, NY: Springer Science & Business Media. http://dx.doi.org/10.1007/978-0-306-48647-0

Murad, M. H., Asi, N., Alsawas, M., & Alahdab, F. (2016). New evidence pyramid. *British Medical Journal Evidence-Based Medicine, 21,* 125–127. http://dx.doi.org/10.1136/ebmed-2016-110401

*Murdock, L. C., Dantzler, J. A., Walker, A. N., & Wood, L. B. (2014). The effect of a platform swing on the independent work behaviors of children with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities, 29,* 50–61. http://dx.doi.org/10.1177/1088357613509838

National Autism Center. (2015). *Findings and conclusions: National Standards Project, Phase 2*. Randolph, MA: Author.

*Nefdt, N., Koegel, R., Singer, G., & Gerber, M. (2010). The use of a self-directed learning program to provide introductory training in pivotal response treatment to parents of children with autism. *Journal of Positive Behavior Interventions, 12,* 23–32. http://dx.doi.org/10.1177/1098300709334796

Nuijten, M. B. (2018). Practical tools and strategies for researchers to increase replicability. *Developmental Medicine and Child Neurology*. Advance online publication. http://dx.doi.org/10.31234/osf.io/emyux

Odom, S. L., Boyd, B. A., Hall, L. J., & Hume, K. (2010). Evaluation of comprehensive treatment models for individuals with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 40,* 425–436. http://dx.doi.org/10.1007/s10803-009-0825-1

O'Haire, M. E. (2013). Animal-assisted intervention for autism spectrum disorder: A systematic literature review. *Journal of Autism and Developmental Disorders, 43,* 1606–1622.

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

Oliver, P. C., Piachaud, J., Done, J., Regan, A., Cooray, S., & Tyrer, P. (2002). Difficulties in conducting a randomized controlled trial of health service interventions in intellectual disability: Implications for evidence-based practice. *Journal of Intellectual Disability Research, 46,* 340–345. http://dx.doi.org/10.1046/j.1365-2788.2002.00408.x

Ospina, M. B., Seida, J. K., Clark, B., Karkhaneh, M., Hartling, L., Tjosvold, L., . . . Smith, V. (2008). Behavioural and developmental interventions for autism spectrum disorder: A clinical systematic review. *PLoS ONE, 3,* e3755.

*Ozonoff, S., & Cathcart, K. (1998). Effectiveness of a home program intervention for young children with autism. *Journal of Autism and Developmental Disorders, 28,* 25–32. http://dx.doi.org/10.1023/A:1026006818310

*Page, C. E. (2012). *The social and emotional benefits of therapeutic riding on children with autism spectrum disorder* (Unpublished doctoral dissertation). Institute of Transpersonal Psychology, Palo Alto, CA.

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

*Pan, C. Y. (2010). Effects of water exercise swimming program on aquatic skills and social behaviors in children with autism spectrum disorders. *Autism, 14,* 9–28. http://dx.doi.org/10.1177/1362361309339496

*Paul, R., Campbell, D., Gilbert, K., & Tsiouri, I. (2013). Comparing spoken language treatments for minimally verbal preschoolers with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 43,* 418–431. http://dx.doi.org/10.1007/s10803-012-1583-z

*Paynter, J., & Peterson, C. C. (2013). Further evidence of benefits of thought-bubble training for theory of mind development in children with autism spectrum disorders. *Research in Autism Spectrum Disorders, 7,* 344–348. http://dx.doi.org/10.1016/j.rasd.2012.10.001

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

*Pickles, A., Le Couteur, A., Leadbitter, K., Salomone, E., Cole-Fletcher, R., Tobin, H., . . . Green, J. (2016). Parent-mediated social communication therapy for young children with autism (PACT): Long-term follow-up of a randomised controlled trial. *Lancet, 388,* 2501–2509. http://dx.doi.org/10.1016/S0140-6736(16)31229-6

*Piravej, K., Tangtrongchitr, P., Chandarasiri, P., Paothong, L., & Sukprasong, S. (2009). Effects of Thai traditional massage on autistic children's behavior. *Journal of Alternative and Complementary Medicine, 15,* 1355–1361. http://dx.doi.org/10.1089/acm.2009.0258

*Poslawsky, I. E., Naber, F. B., Bakermans-Kranenburg, M. J., van Daalen, E., van Engeland, H., & van IJzendoorn, M. H. (2015). Video-feedback Intervention to promote Positive Parenting adapted to Autism (VIPP-AUTI): A randomized controlled trial. *Autism, 19,* 588–603. http://dx.doi.org/10.1177/1362361314537124

*Powell, G., Wass, S. V., Erichsen, J. T., & Leekam, S. R. (2016). First evidence of the feasibility of gaze-contingent attention training for school children with autism. *Autism: An International Journal of Research and Practise, 20,* 927–937. http://dx.doi.org/10.1177/1362361315617880

Prizant, B. M., & Wetherby, A. M. (1998). Understanding the continuum of discrete-trial traditional behavioral to social-pragmatic developmental approaches in communication enhancement for young children with autism/PDD. *Seminars in Speech and Language, 19,* 329–353.

Provenzani, U., Fusar-Poli, L., Brondino, N., Damiani, S., Vercesi, M., Meyer, N., . . . Politi, P. (2019). What are we targeting when we treat autism spectrum disorder? A systematic review of 406 clinical trials. *Autism: An International Journal of Research and Practise*. Advance online publication. http://dx.doi.org/10.1177/1362361319854641

R Core Team. (2017). R: A language and environment for statistical computing (Version 3.4.1). Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Reichow, B., Volkmar, F. R., & Cicchetti, D. V. (2008). Development of the evaluative method for evaluating and determining evidence-based practices in autism. *Journal of Autism and Developmental Disorders, 38,* 1311–1319. http://dx.doi.org/10.1007/s10803-007-0517-7

*Reitzel, J., Summers, J., Lorv, B., Szatmari, P., Zwaigenbaum, L., Georgiades, S., & Duku, E. (2013). Pilot randomized controlled trial of a Functional Behavior Skills Training program for young children with autism spectrum disorder who have significant early learning skill impairments and their families. *Research in Autism Spectrum Disorders, 7,* 1418–1432. http://dx.doi.org/10.1016/j.rasd.2013.07.025

*Remington, B., Hastings, R. P., Kovshoff, H., degli Espinosa, F., Jahr, E., Brown, T., . . . Ward, N. (2007). Early intensive behavioral intervention: Outcomes for children with autism and their parents after two years. *American Journal on Mental Retardation, 112,* 418–438. http://dx.doi.org/10.1352/0895-8017(2007)112[418:EIBIOF]2.0.CO;2

Renty, J. O., & Roeyers, H. (2006). Quality of life in high-functioning adults with autism spectrum disorder: The predictive value of disability and support characteristics. *Autism, 10,* 511–524. http://dx.doi.org/10.1177/1362361306066604

*Rentz, E. A. (2015). *Autism spectrum disorders: Developmental, individual-difference, relationship-based (DIR)/Floortime outcomes* (Unpublished doctoral dissertation). Fordham University, Bronx, NY.

*Rice, L. M., Wall, C. A., Fogel, A., & Shic, F. (2015). Computer-assisted face processing instruction improves emotion recognition, mentalizing, and social skills in students with ASD. *Journal of Autism and Developmental Disorders, 45,* 2176–2186. http://dx.doi.org/10.1007/s10803-015-2380-2

*Rickards, A. L., Walstab, J. E., Wright-Rossi, R. A., Simpson, J., & Reddihough, D. S. (2007). A randomized, controlled trial of a home-based intervention program for children with autism and developmental delay. *Journal of Developmental and Behavioral Pediatrics, 28,* 308–316. http://dx.doi.org/10.1097/DBP.0b013e318032792e

*Roberts, J., Williams, K., Carter, M., Evans, D., Parmenter, T., Silove, N., . . . Warren, A. (2011). A randomised controlled trial of two early intervention programs for young children with autism: Centre-based with parent program and home-based. *Research in Autism Spectrum Disorders, 5,* 1553–1566. http://dx.doi.org/10.1016/j.rasd.2011.03.001

Robison, J. E. (2019). Talking about autism-thoughts for researchers. *Autism Research, 12,* 1004–1006. http://dx.doi.org/10.1002/aur.2119

Rogers, S. J., & Dawson, G. (2010). *Early Start Denver Model for young children with autism: Promoting language, learning, and engagement*. New York, NY: Guilford Press.

*Rogers, S. J., Estes, A., Lord, C., Vismara, L., Winter, J., Fitzpatrick, A., . . . Dawson, G. (2012). Effects of a brief Early Start Denver Model (ESDM)-based parent intervention on toddlers at risk for autism spectrum disorders: A randomized controlled trial. *Journal of the American Academy of Child & Adolescent Psychiatry, 51,* 1052–1065. http://dx.doi.org/10.1016/j.jaac.2012.08.003

*Rogers, S. J., Hayden, D., Hepburn, S., Charlifue-Smith, R., Hall, T., & Hayes, A. (2006). Teaching young nonverbal children with autism useful speech: A pilot study of the Denver Model and PROMPT interventions. *Journal of Autism and Developmental Disorders, 36,* 1007–1024. http://dx.doi.org/10.1007/s10803-006-0142-x

*Ruble, L. A., Dalrymple, N. J., & McGrew, J. H. (2010). The effects of consultation on individualized education program outcomes for young children with autism: The collaborative model for promoting competence and success. *Journal of Early Intervention, 32,* 286–301. http://dx.doi.org/10.1177/1053815110382973

*Sainato, D. M., Morrison, R. S., Jung, S., Axe, J., & Nixon, P. A. (2015). A comprehensive inclusion program for kindergarten children with autism spectrum disorder. *Journal of Early Intervention, 37,* 208–225. http://dx.doi.org/10.1177/1053815115613836

*Sallows, G. O., & Graupner, T. D. (2005). Intensive behavioral treatment for children with autism: Four-year outcome and predictors. *American Journal on Mental Retardation, 110,* 417–438. http://dx.doi.org/10.1352/0895-8017(2005)110[417:IBTFCW]2.0.CO;2

*Salt, J., Shemilt, J., Sellars, V., Boyd, S., Coulson, T., & McCool, S. (2002). The Scottish Centre for autism preschool treatment programme. II: The results of a controlled treatment outcome study. *Autism, 6,* 33–46. http://dx.doi.org/10.1177/1362361302006001004

*Sampanthavivat, M., Singkhwa, W., Chaiyakul, T., Karoonyawanich, S., & Ajpru, H. (2012). Hyperbaric oxygen in the treatment of childhood autism: A randomised controlled trial. *Diving and Hyperbaric Medicine, 42,* 128–133.

*Scarpa, A., & Reyes, N. M. (2011). Improving emotion regulation with CBT in young children with high functioning autism spectrum disorders: A pilot study. *Behavioural and Cognitive Psychotherapy, 39,* 495–500. http://dx.doi.org/10.1017/S1352465811000063

Schaaf, R. C., Benevides, T., Mailloux, Z., Faller, P., Hunt, J., van Hooydonk, E., . . . Kelly, D. (2014). An intervention for sensory difficulties in children with autism: A randomized trial. *Journal of Autism and Developmental Disorders, 44,* 1493–1506. http://dx.doi.org/10.1007/s10803-014-2111-0

Schaaf, R. C., & Case-Smith, J. (2014). Sensory interventions for children with autism. *Journal of Comparative Effectiveness Research, 3,* 225–227.

*Schertz, H. H., Odom, S. L., Baggett, K. M., & Sideris, J. H. (2013). Effects of joint attention mediated learning for toddlers with autism spectrum disorders: An initial randomized controlled study. *Early Childhood Research Quarterly, 28,* 249–258. http://dx.doi.org/10.1016/j.ecresq.2012.06.006

*Schertz, H. H., Odom, S. L., Baggett, K. M., & Sideris, J. H. (2018). Mediating parent learning to promote social communication for toddlers with autism: Effects from a randomized controlled trial. *Journal of Autism and Developmental Disorders, 48,* 853–867. http://dx.doi.org/10.1007/s10803-017-3386-8

Schreibman, L., Dawson, G., Stahmer, A. C., Landa, R., Rogers, S. J., McGee, G. G., . . . Halladay, A. (2015). Naturalistic developmental behavioral interventions: Empirically validated treatments for autism spectrum disorder. *Journal of Autism and Developmental Disorders, 45,* 2411–2428. http://dx.doi.org/10.1007/s10803-015-2407-8

*Schreibman, L., & Stahmer, A. C. (2014). A randomized trial comparison of the effects of verbal and pictorial naturalistic communication strategies on spoken language for young children with autism. *Journal of Autism and Developmental Disorders, 44,* 1244–1251. http://dx.doi.org/10.1007/s10803-013-1972-y

*Sheinkopf, S. J., & Siegel, B. (1998). Home-based behavioral treatment of young children with autism. *Journal of Autism and Developmental Disorders, 28,* 15–23. http://dx.doi.org/10.1023/A:1026054701472

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

*Shire, S. Y., Gulsrud, A., & Kasari, C. (2016). Increasing responsive parent–child interactions and joint engagement: Comparing the influence of parent-mediated intervention and parent psychoeducation. *Journal of Autism and Developmental Disorders, 46,* 1737–1747. http://dx.doi.org/10.1007/s10803-016-2702-z

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86,* 420–428. http://dx.doi.org/10.1037/0033-2909.86.2.420

*Silva, L., & Schalock, M. (2013). Treatment of tactile impairment in young children with autism: Results with qigong massage. *International Journal of Therapeutic Massage & Bodywork, 6,* 12–20.

*Silva, L. M., Schalock, M., Gabrielsen, K. R., Budden, S. S., Buenrostro, M., & Horton, G. (2015). Early intervention with a parent-delivered massage protocol directed at tactile abnormalities decreases severity of autism and improves child-to-parent interactions: A replication study. *Autism Research and Treatment*. Advance online publication. http://dx.doi.org/10.1155/2015/904585

*Simpson, K., Keen, D., & Lamb, J. (2013). The use of music to engage children with autism in a receptive labelling task. *Research in Autism Spectrum Disorders, 7,* 1489–1496. http://dx.doi.org/10.1016/j.rasd.2013.08.013

*Simpson, K., Keen, D., & Lamb, J. (2015). Teaching receptive labelling to children with autism spectrum disorder: A comparative study using infant-directed song and infant-directed speech. *Journal of Intellectual & Developmental Disability, 40,* 126–136. http://dx.doi.org/10.3109/13668250.2015.1014026

*Smith, T., Eikeseth, S., Klevstrand, M., & Lovaas, O. I. (1997). Intensive behavioral treatment for preschoolers with severe mental retardation and pervasive developmental disorder. *American Journal on Mental Retardation, 102,* 238–249. http://dx.doi.org/10.1352/0895-8017(1997)102<0238:IBTFPW>2.0.CO;2

*Smith, T., Groen, A. D., & Wynn, J. W. (2000). Randomized trial of intensive early intervention for children with pervasive developmental disorder. *American Journal on Mental Retardation, 105,* 269–285. http://dx.doi.org/10.1352/0895-8017(2000)105<0269:RTOIEI>2.0.CO;2

*Sokmun, S., Singh, S. J., & Vandort, S. (2018). The impact of Hanen More Than Words programme on parents of children with ASD in Malaysia. *Malaysian Journal of Health Sciences, 15,* 43–51.

*Solomon, R., Van Egeren, L. A., Mahoney, G., Quon Huber, M. S., & Zimmerman, P. (2014). PLAY Project Home Consultation intervention program for young children with autism spectrum disorders: A randomized controlled trial. *Journal of Developmental and Behavioral Pediatrics, 35,* 475–485. http://dx.doi.org/10.1097/DBP.0000000000000096

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

*Srinivasan, S. M., Eigsti, I. M., Gifford, T., & Bhat, A. N. (2016). The effects of embodied rhythm and robotic interventions on the spontaneous and responsive verbal communication skills of children with Autism Spectrum Disorder (ASD): A further outcome of a pilot randomized controlled trial. *Research in Autism Spectrum Disorders, 27,* 73–87. http://dx.doi.org/10.1016/j.rasd.2016.04.001

*Srinivasan, S. M., Kaur, M., Park, I. K., Gifford, T. D., Marsh, K. L., & Bhat, A. N. (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*. Advance online publication. http://dx.doi.org/10.1155/2015/736516

*Srinivasan, S. M., Park, I. K., Neelly, L. B., & Bhat, A. N. (2015). A comparison of the effects of rhythm and robotic interventions on repetitive behaviors and affective states of children with autism spectrum disorder (ASD). *Research in Autism Spectrum Disorders, 18,* 51–63. http://dx.doi.org/10.1016/j.rasd.2015.07.004

*Stadnick, N. A., Stahmer, A., & Brookman-Frazee, L. (2015). Preliminary effectiveness of Project ImPACT: A parent-mediated intervention for children with autism spectrum disorder delivered in a community program. *Journal of Autism and Developmental Disorders, 45,* 2092–2104. http://dx.doi.org/10.1007/s10803-015-2376-y

Stahmer, A. C., Collings, N. M., & Palinkas, L. A. (2005). Early intervention practices for children with autism: Descriptions from community providers. *Focus on Autism and Other Developmental Disabilities, 20,* 66–79. http://dx.doi.org/10.1177/10883576050200020301

*Strain, P. S., & Bovey, E. H., II. (2011). Randomized, controlled trial of the LEAP model of early intervention for young children with autism spectrum disorders. *Topics in Early Childhood Special Education, 31,* 133–154. http://dx.doi.org/10.1177/0271121411408740

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

*Sullivan, K. A. (2013). *The Early Start Denver Model: Outcomes and moderators of an intervention for toddlers with Autism* (Unpublished doctoral dissertation). University of Washington, Seattle, WA.

*Suzuki, M., Yamada, A., Watanabe, N., Akechi, T., Katsuki, F., Nishiyama, T., . . . Furukawa, T. A. (2014). A failure to confirm the effectiveness of a brief group psychoeducational program for mothers of children with high-functioning pervasive developmental disorders: A randomized controlled pilot trial. *Neuropsychiatric Disease and Treatment, 10,* 1141–1153. http://dx.doi.org/10.2147/NDT.S60058

Tanner-Smith, E. E., Tipton, E., & Polanin, J. R. (2016). Handling complex meta-analytic data structures using robust variance estimates: A tutorial in R. *Journal of Developmental and Life-Course Criminology, 2,* 85–112. http://dx.doi.org/10.1007/s40865-016-0026-5

*Tellegen, C. L., & Sanders, M. R. (2014). A randomized controlled trial evaluating a brief parenting program with children with autism spectrum disorders. *Journal of Consulting and Clinical Psychology, 82,* 1193–1200. http://dx.doi.org/10.1037/a0037246

*Thompson, G. A., McFerran, K. S., & Gold, C. (2014). Family-centred music therapy to promote social engagement in young children with severe autism spectrum disorder: A randomized controlled study. *Child: Care, Health and Development, 40,* 840–852. http://dx.doi.org/10.1111/cch.12121

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

Trzmiel, T., Purandare, B., Michalak, M., Zasadzka, E., & Pawlaczyk, M. (2019). Equine assisted activities and therapies in children with autism spectrum disorder: A systematic review and a meta-analysis. *Complementary Therapies in Medicine, 42,* 104–113. http://dx.doi.org/10.1016/j.ctim.2018.11.004

*Tsang, S. K., Shek, D. T., Lam, L. L., Tang, F. L., & Cheung, P. M. (2007). Brief report: Application of the TEACCH program on Chinese pre-school children with autism—Does culture make a difference? *Journal of Autism and Developmental Disorders, 37,* 390–396. http://dx.doi.org/10.1007/s10803-006-0199-6

*Venker, C. E., McDuffie, A., Ellis Weismer, S., & Abbeduto, L. (2012). Increasing verbal responsiveness in parents of children with autism: A pilot study. *Autism, 16,* 568–585. http://dx.doi.org/10.1177/1362361311413396

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36,* 1–48. http://dx.doi.org/10.18637/jss.v036.i03

*Vivanti, G., Paynter, J., Duncan, E., Fothergill, H., Dissanayake, C., & Rogers, S. J. (2014). Effectiveness and feasibility of the Early Start Denver Model implemented in a group-based community childcare setting. *Journal of Autism and Developmental Disorders, 44,* 3140–3153. http://dx.doi.org/10.1007/s10803-014-2168-9

Vygotsky, L. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.

Wallace, B., Small, K., Brodley, C. E., Lau, J., & Trikalinos, T. A. (2012). Deploying an interactive machine learning system in an evidence-based practice center: Abstrackr. *Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium* (pp. 819–824). New York, NY: ACM.

Warren, Z., McPheeters, M. L., Sathe, N., Foss-Feig, J. H., Glasser, A., & Veenstra-Vanderweele, J. (2011). A systematic review of early intensive intervention for autism spectrum disorders. *Pediatrics, 127,* e1303–e1311. http://dx.doi.org/10.1542/peds.2011-0426

*Weiner, R. H., & Greene, R. L. (2014). Intention-based therapy for autism spectrum disorder: Promising results of a wait-list control study in children. *EXPLORE: The Journal of Science and Healing, 10,* 13–23. http://dx.doi.org/10.1016/j.explore.2013.10.005

Weitlauf, A. S., Sathe, N., McPheeters, M. L., & Warren, Z. E. (2017). Interventions targeting sensory challenges in autism spectrum disorder: A systematic review. *Pediatrics, 139,* e20170347.

*Welsh, K. (2009). *The use of dogs to impact joint attention in children with autism spectrum disorders* (Unpublished doctoral dissertation). Walden University, Minneapolis, MN.

*Welterlin, A., Turner-Brown, L. M., Harris, S., Mesibov, G., & Delmolino, L. (2012). The home TEACCHing program for toddlers with autism. *Journal of Autism and Developmental Disorders, 42,* 1827–1835. http://dx.doi.org/10.1007/s10803-011-1419-2

*Wetherby, A. M., & Woods, J. J. (2006). Early social interaction project for children with autism spectrum disorders beginning in the second year of life: A preliminary study. *Topics in Early Childhood Special Education, 26,* 67–82. http://dx.doi.org/10.1177/02711214060260020201

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

*Whittingham, K., Sofronoff, K., Sheffield, J., & Sanders, M. R. (2009). Stepping Stones Triple P: An RCT of a parenting program with parents of a child diagnosed with an autism spectrum disorder. *Journal of Abnormal Child Psychology, 37,* 469–480. http://dx.doi.org/10.1007/s10802-008-9285-x

*Wild, D. L. (2012). *The impact of canine assistance for children with autism and the family unit* (Unpublished doctoral dissertation). Walden University, Minneapolis, MN.

*Williams, B. T., Gray, K. M., & Tonge, B. J. (2012). Teaching emotion recognition skills to young children with autism: A randomised controlled trial of an emotion training programme. *Journal of Child Psychology and Psychiatry, 53,* 1268–1276. http://dx.doi.org/10.1111/j.1469-7610.2012.02593.x

Williams, K., Wray, J. A., & Wheeler, D. M. (2012). Intravenous secretin for autism spectrum disorders (ASD). *Cochrane Database of Systematic Reviews, 18,* CD003495. http://dx.doi.org/10.1002/14651858.CD003495.pub3

Wolery, M., Busick, M., Reichow, B., & Barton, E. E. (2010). Comparison of overlap methods for quantitatively synthesizing single-subject data. *The Journal of Special Education, 44,* 18–28. http://dx.doi.org/10.1177/0022466908328009

Wong, C., Odom, S. L., Hume, K. A., Cox, A. W., Fettig, A., Kucharczyk, S., . . . Schultz, T. R. (2015). Evidence-based practices for children, youth, and young adults with autism spectrum disorder: A comprehensive review. *Journal of Autism and Developmental Disorders, 45,* 1951–1966. http://dx.doi.org/10.1007/s10803-014-2351-z

*Wong, V. C., & Kwan, Q. K. (2010). Randomized controlled trial for early intervention for autism: A pilot study of the Autism 1-2-3 Project. *Journal of Autism and Developmental Disorders, 40,* 677–688. http://dx.doi.org/10.1007/s10803-009-0916-z

*Woo, C. C., Donnelly, J. H., Steinberg-Epstein, R., & Leon, M. (2015). Environmental enrichment as a therapy for autism: A clinical trial replication and extension. *Behavioral Neuroscience, 129,* 412–422. http://dx.doi.org/10.1037/bne0000068

*Woo, C. C., & Leon, M. (2013). Environmental enrichment as an effective treatment for autism: A randomized controlled trial. *Behavioral Neuroscience, 127,* 487–497. http://dx.doi.org/10.1037/a0033010

Yoder, P. J., Bottema-Beutel, K., Woynaroski, T., Chandrasekhar, R., & Sandbank, M. (2013). Social communication intervention effects vary by dependent variable type in preschoolers with autism spectrum disorders. *Evidence-Based Communication Assessment and Intervention, 7,* 150–174. http://dx.doi.org/10.1080/17489539.2014.917780

*Young, R. L., & Posselt, M. (2012). Using the transporters DVD as a learning tool for children with autism spectrum disorders (ASD). *Journal of Autism and Developmental Disorders, 42,* 984–991. http://dx.doi.org/10.1007/s10803-011-1328-4

*Yun, S. S., Choi, J., Park, S. K., Bong, G. Y., & Yoo, H. (2017). Social skills training for children with autism spectrum disorder using a robotic behavioral intervention system. *Autism Research, 10,* 1306–1323. http://dx.doi.org/10.1002/aur.1778

*Zachor, D. A., Vardi, S., Baron-Eitan, S., Brodai-Meir, I., Ginossar, N., & Ben-Itzchak, E. (2017). The effectiveness of an outdoor adventure programme for young children with autism spectrum disorder: A controlled study. *Developmental Medicine and Child Neurology, 59,* 550–556. http://dx.doi.org/10.1111/dmcn.13337

Zimmerman, K. N., Pustejovsky, J. E., Ledford, J. R., Barton, E. E., Severini, K. E., & Lloyd, B. P. (2018). Single-case synthesis tools II: Comparing quantitative outcome measures. *Research in Developmental Disabilities, 79,* 65–76. http://dx.doi.org/10.1016/j.ridd.2018.02.001

Received April 18, 2019
Revision received August 16, 2019
Accepted September 29, 2019 ■

Journal of Clinical Child & Adolescent Psychology
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t775648094

Meta-Analysis of Early Intensive Behavioral Intervention for Children With Autism

Sigmund Eldevik [abc]; Richard P. Hastings [a]; J. Carl Hughes [a]; Erik Jahr [d]; Svein Eikeseth [b]; Scott Cross [e]
[a] School of Psychology, Bangor University, [b] Faculty of Behavioral Science, Akershus University College, [c] Highfield Centre, [d] Akershus University Hospital, [e] Lovaas Institute for Early Intervention,

To cite this Article Eldevik, Sigmund, Hastings, Richard P., Hughes, J. Carl, Jahr, Erik, Eikeseth, Svein and Cross, Scott(2009)'Meta-Analysis of Early Intensive Behavioral Intervention for Children With Autism',Journal of Clinical Child & Adolescent Psychology,38:3,439 — 450

To link to this Article: DOI: 10.1080/15374410902851739

URL: http://dx.doi.org/10.1080/15374410902851739

Journal of Clinical Child & Adolescent Psychology, 38(3), 439–450, 2009
Copyright © Taylor & Francis Group, LLC
ISSN: 1537-4416 print/1537-4424 online
DOI: 10.1080/15374410902851739

Meta-Analysis of Early Intensive Behavioral Intervention for Children With Autism

Sigmund Eldevik
School of Psychology, Bangor University; Faculty of Behavioral Science, Akershus University College; and Highfield Centre

Richard P. Hastings and J. Carl Hughes
School of Psychology, Bangor University

Erik Jahr
Akershus University Hospital

Svein Eikeseth
Faculty of Behavioral Science, Akershus University College

Scott Cross
Lovaas Institute for Early Intervention

A systematic literature search for studies reporting effects of Early Intensive Behavioral Intervention identified 34 studies, 9 of which were controlled designs having either a comparison or a control group. We completed a meta-analysis yielding a standardized mean difference effect size for two available outcome measures: change in full-scale intelligence and/or adaptive behavior composite. Effect sizes were computed using Hedges's *g*. The average effect size was 1.10 for change in full-scale intelligence (95% confidence interval = .87, 1.34) and .66 (95% confidence interval = .41, .90) for change in adaptive behavior composite. These effect sizes are generally considered to be large and moderate, respectively. Our results support the clinical implication that at present, and in the absence of other interventions with established efficacy, Early Intensive Behavioral Intervention should be an intervention of choice for children with autism.

There is a developing evidence base for the positive effects of comprehensive interventions for children with autism spectrum disorders (ASD). Two recent narrative reviews have focused on a range of comprehensive interventions for children with autism (Eikeseth, 2009; Rogers & Vismara, 2008). The conclusion from both of these reviews is that Early Intensive Behavioral Intervention (EIBI) is an effective intervention when compared against no intervention controls or eclectic/autism-specific special education interventions. When applying more formal criteria (Chambless et al., 1998; Chambless & Hollon, 1998; Chambless et al., 1996), Rogers and Vismara found that EIBI (or what they call the "Lovaas treatment approach") should be considered "well established" and that no other intervention presently qualifies for this status.

We thank Dr. Michael Borenstein for helpful advice in carrying out the present analysis. We also thank Alison Sharland and Iwona Ondire for coding papers. Furthermore we express our gratitude to the Sharland Foundation for a research grant supporting this work. This paper was submitted as part of the first author's Ph.D. thesis at Bangor University.

Correspondence should be addressed to Sigmund Eldevik, Akershus University College, P. O. Box 423, N-2001 Lillestrøm, Norway. E-mail: eldevik@online.no

EIBI programs (including the Lovaas treatment approach) have been described by Green, Brennan, and Fein (2002, p. 70, see also Eikeseth, 2009, for a similar definition) as having the following common elements: (a) intervention is individualized and comprehensive, addressing all skill domains; (b) many behavior analytic procedures are used to build new repertoires and reduce interfering behavior (e.g., differential reinforcement, prompting, discrete-trial instruction, incidental teaching, activity-embedded trials, task analysis, and others); (c) one or more individuals with advanced training in applied behavior analysis and experience with young children with autism directs the intervention; (d) normal developmental sequences guides the selection of intervention goals and short-term objectives; (e) parents serve as active co-therapists for their children; (f) intervention is delivered in one-to-one fashion initially, with gradual transitions to small-group and large-group formats when warranted; (g) intervention typically begins in the home and is carried over into other environments (e.g., community settings), with gradual, systematic transitions to preschool, kindergarten, and elementary school classrooms when children develop the skills required to learn in those settings; (h) programming is intensive, is year round, and includes 20 to 30 hr of structured sessions per week plus informal instruction and practice throughout most of the children's other waking hours; (i) in most cases, the duration of intervention is 2 or more years; and (j) most children start intervention in the preschool years, when they are 3 to 4 years of age.

In addition to narrative reviews, there have been two recent systematic reviews of outcome research on EIBI. The first review presented a systematic description of the research published to date and pointed to challenges for future research (Howlin, Magiati, & Charman, in press). Eleven studies were identified using the following inclusion criteria: The study had to have a control or comparison group with a minimum of 10 participants in each group, participant at intake had to be younger than 6 years of age, and intervention had to be provided for at least 12 hr a week for 12 months. Howlin et al. discussed a number of problems associated with drawing conclusions about the efficacy of EIBI. First, although not accurately reported in some of the studies, they estimated that the EIBI groups on average received significantly more hours of intervention than did control groups. Second, a variety of assessment instruments were used across children and studies that made it difficult to compare results across studies and may have led to results being spuriously positive. Third, in some studies it was unclear at what points in time the assessments were conducted, particularly at posttreatment when in some cases assessments were undertaken years after treatment had ended. Fourth, the studies reported test scores in different ways that including standard scores, age equivalents, and raw scores.

Howlin et al. (in press) concluded that in general the average effects of EIBI were favorable compared to controls but that the variability across individual children in the EIBI studies was substantial. Howlin et al. could not identify any reliable predictors of outcome. Intake IQ was found by some researchers to be related to better outcomes but others found no such relationship. Furthermore, age at intake was not found to be related to outcome in any of the studies. However, age range was limited with all children being younger than 7 years of age. Initial language ability was identified as a possible predictor only in some of the studies that explored this, and autism symptomatology was found to be related both with better and with worse outcomes, in two different studies. Given these problems they concluded that conducting a meta-analysis of the evidence was not appropriate.

The second recent systematic review conducted by Reichow and Wolery (2009) addressed similar questions to Howlin et al. (in press) and drew similar conclusions. However, unlike Howlin et al., they included a meta-analysis. The authors argued that a meta-analysis of EIBI is feasible but that it had to be limited to change in intelligence scores and that to have enough studies, they would have to include studies that were not controlled. Thus, the meta-analysis used standardized mean change effect sizes and not the more methodologically rigorous standardized mean difference effect size. The mean change effect size is computed without comparison or control group data and, as the authors point out, any conclusions are limited by threats to validity such as maturation. In addition, the standardized mean change effect size may inflate effect size estimates (Morris, 2000). Based on 12 studies, Reichow and Wolery reported a weighted mean effect-size for change in intellectual functioning following EIBI of .69.

A second aspect of the Reichow and Wolery (2009) analysis that may affect the validity of conclusions was that studies using a variety of outcome measures for intelligence were included. For example, studies which primarily relied on performance based nonverbal measures of intelligence such as the Merrill-Palmer (Stutsman, 1948) and the Leiter-R (Roid & Miller, 1997) were treated as equivalent to studies which reported full-scale IQ measures. Because the performance-based tests measure areas where children with autism often are relatively strong (e.g., visual-spatial tasks), scores tend to be higher than on full-scale IQ tests (Lord et al., 2006). Not separating these tests in an analysis on the effects of EIBI may affect the conclusions drawn, especially in cases where these different measures are used interchangeably pre- and post-treatment. A minor methodological problem with the Reichow and Wolery analysis was that they reported reliability estimates only on the coding of information from the selected studies (research methods, participants, and intervention characteristics); no such data were provided for the initial procedure for selecting studies to be included in the review.

The purpose of our study was to provide a replication and extension of the Reichow and Wolery (2009) meta-analysis, with a focus on methodological improvements. First, we selected studies with comparison/control groups only, while employing a more precise definition of EIBI (Green et al., 2002) and the control/comparison groups. This makes it possible to apply more methodologically rigorous mean difference effect size measures. Second, we required more uniformity in outcome measurement and included only full-scale measures of intelligence. Third, we were able to add a meta-analysis of changes in adaptive behavior. Fourth, because our literature search was conducted later in time and employed a somewhat different definition of EIBI, some additional recent published studies were included. Fifth, we included interrater reliability for our literature search and initial selection procedure for studies in the review. Sixth, we based our analysis on individual raw data gathered from authors rather than group average data reported in the original papers. This made it possible to prevent children from being represented more than once if they were included in more than one published outcome study and to ensure the selection of evaluation periods as similar as possible across studies.

METHODS

Search Strategy and Data Collection

We searched the PsycINFO, PubMed, and ERIC databases (up to March 2008) using a combination of the following terms: *behavior analytic, behavioral, early, intervention*, and *autism and/or PDD*. The first author read the headings and abstracts of all the papers from this initial search to decide whether the study warranted a more detailed coding. If it was possible that the study reported outcome data on the effects of behavioral treatment for children with ASD, the study was obtained for more detailed coding. In addition, the reference sections of obtained papers were browsed in an attempt to locate studies that might have been missed in the electronic search.

A coding scheme was used for coding all the selected studies (available from the first author). First, it was coded whether the children had received behavioral intervention that generally adhered to common elements described by Green et al. (2002, p. 70). Second, a series of true/false scores were given for the following: (a) the participants were on average between 2 and 7 years of age when intervention started; (b) the children were independently diagnosed with autism or PDD-NOS; (c) a full-scale measure of intelligence and/or a standardized measure of adaptive behavior such as the Vineland Adaptive Behavior Scales (VABS; Sparrow, Balla, & Cicchetti, 1984) was conducted at intake and after intervention (primarily administering a nonverbal intelligence measure such as the Leiter–R [Roid & Miller, 1997] or the Merrill–Palmer Scale of Mental Tests [Stutsman, 1948] led to the study's exclusion); (d) the duration of intervention was between 12 and 36 months; (e) the study was not a case study (or series of case studies); (f) the results had been published in a peer-reviewed journal; and (g) the study included either a control or comparison group. The studies were classified as either a comparison or a control study (or both). If it was specified that the children in the study had received intervention(s) other than EIBI of similar duration and intensity in terms of 1:1 hours, it was classed as a comparison study. Although it would probably be impossible to determine whether the children in the comparison groups had similar eclectic or specialist autism provision (even within a single study), classifying the studies in this way could still yield useful information. For example, it may facilitate the exploration of whether it is the number of 1:1 hours itself (i.e., "intensity") that makes a difference. Where no intervention (or a considerably less intensive one) was provided, or a poorly specified intervention was described, the study was classified as a control study.

The electronic and manual searches resulted in 2,150 potential hits. Through the first screening process, we selected 34 papers for detailed coding. One of the database searches, resulting in 607 hits, was chosen for a reliability check. The screening results from the first author were compared to that of a second screener (another author) using the same decision criteria. Agreement was high overall in terms of whether the paper should be subject to further analysis (Cohen's $\kappa = .85$). Disagreements occurred only because the second screener included fewer studies than the first screener. Thus, there were no instances of the second screener including a study for further analysis that was not already included by the first screener.

The 34 studies that remained after this initial screening were coded by the first author and two independent scorers (master's students in behavior analysis). Agreement was calculated between the first author and each of the independent scorers separately by dividing the total number of agreements by the total number of agreements plus disagreements and multiplying by 100. Initial agreement was high in both cases (91% and 94%, respectively) and the few disagreements that occurred were resolved after brief discussions. We excluded 25 of the 34 studies for one or more of the following reasons: (a) seven had inadequate intake and/or outcome data, such as primarily reporting performance IQ instead of full-scale IQ (Bibby, Eikeseth, Martin, Mudford, & Reeves, 2002; Drew et al., 2002; Fenske, Zalenski, Krantz, & McClannahan, 1985; Luiselli, Cannon, Ellis, & Sisson, 2000; Magiati, Charman, & Howlin, 2007; Sheinkopf & Siegel, 1998; Solomon, Necheles, Ferch, & Bruckman, 2007); (b) five had an intervention duration that was too short to meet inclusion criteria (Harris, Handleman, Gordon, Kristoff, & Fuentes, 1991; Ingersoll, Schreibman, & Stahmer, 2001; Reed, Osborne, & Corness, 2007a, 2007b; Stahmer & Ingersoll, 2004); (c) two reported data from case studies only (Butter, Mulick, & Metz, 2006; Green et al., 2002); (d) three reported data that were already included in other studies (Beglinger & Smith, 2005; Eikeseth, Smith, Jahr, & Eldevik, 2007; McEachin, Smith, & Lovaas, 1993); (e) upon closer inspection, one of the studies provided intervention that did not meet the definition of behavioral treatment (Gabriels, Hill, Pierce, Rogers, & Wehner, 2001); and (f) seven did not have a control or comparison group (Anderson, Avery, DiPietro, Edwards, & Christian, 1987; Ben-Itzchak & Zachor, 2007; Harris & Handleman, 2000; Hayward, Eikeseth, Gale, & Morgan, in press; Sallows & Graupner, 2005; Smith, Buch, & Gamby, 2000; Weiss, 1999).

Individual data needed to calculate effect sizes from the nine remaining studies were obtained by contacting the authors of each study. We asked them to provide the age, IQ, and adaptive behavior scores at intake and after 2 years in intervention (or as close as possible). Also, we asked if any of the children either in the EIBI or comparison/control groups was represented in other published studies. Thus, all computations in our study were conducted by recalculating pre- and postgroup means and standard deviation on outcome measures rather than data reported in the original papers or extrapolated from these reports. Individual data from the second control group ($n = 21$) in the Lovaas (1987) study were not available, and 4 children in the comparison group from one study (Eldevik, Eikeseth, Jahr, & Smith, 2006) were excluded because they were already in the comparison group of an earlier study also included in the analysis (Eikeseth, Smith, Jahr, & Eldevik, 2002). Figure 1 summarizes the study search and selection process.

The total number of children in the nine intervention studies was 297–153 in the EIBI groups, 105 in control groups and 39 in comparison groups. Table 1 summarizes the main characteristics of the children included in this analysis, including mean age at intake, IQ, and Vineland Adaptive Behavior Composite (ABC) scores at intake and posttreatment. Not all authors reported both IQ and ABC data, or were able to give the exact duration of intervention for each individual child. The average intensity in terms of weekly hours and duration is provided in Table 1. The research design and assignment procedures employed are briefly described along with any inclusion criteria described in the original paper. If a study reported outcome data at more than one point in time, we chose the point that was closest to 2-year duration of treatment. All of the aforementioned calculations were conducted in SPSS (version 16.0) using raw data provided from the authors. Hence, the pre- and postgroup means and standard deviations may differ from those reported in the original published papers.

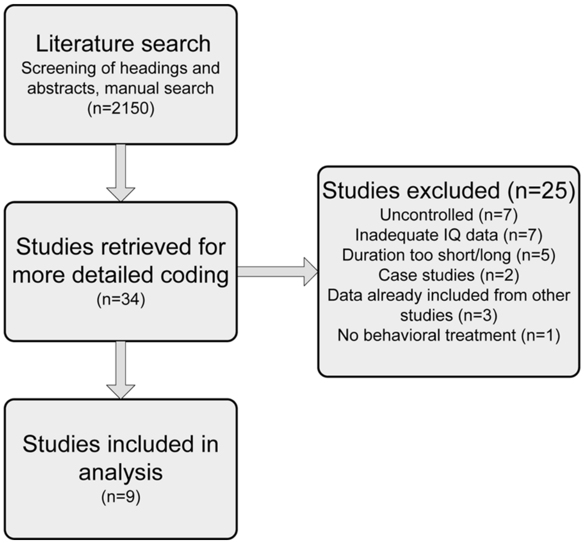


FIGURE 1 The search and selection procedure.

Child Measures

Intellectual functioning. The Bayley Scales of Infant Development, either the first or second edition (Bayley, 1969, 1993) were used for the youngest children or the children that scored below the basal on intelligence tests standardized for their chronological age. The Bayley Scales of Infant Development is a measure of mental development for children up to 42 months. It will yield a mental developmental index, which is considered broadly equivalent to an IQ score. For the older and higher functioning children the most frequently used measures of intelligence were the Stanford–Binet Intelligence Scale, Fourth Edition (Thorndike, Hagen, & Sattler, 1986), the Wechsler Preschool and Primary Scale Intelligence–Revised (Wechsler, 1989), the Wechsler Intelligence Scale for Children–Revised (Wechsler, 1974), or the Wechsler Intelligence Scales for Children–Third Edition (Wechsler, 1993). If the child scored below the norms on a test, researchers had generally computed a ratio IQ score by dividing the obtained mental age with chronological age and multiplying by 100. All of the tests have been used extensively and validated for children

TABLE 1
Main Characteristics of the Studies Included in the Analysis of Outcome. Studies Included in Other Recent Reviews are Shown

			Pre		*Post*		*Intensity*				
Study	*Group*	*Age*	*IQ*	*ABC*	*IQ*	*ABC*	*hrs*	*mths*	*n*	*Design/Assignment/ Inclusion*	*Comments*
USA											
(Lovaas, 1987) 1, 2, 3, 4	EIBI	34.6 (8.9)	62.9 (13.7)		83.3 (28.6)		40	24–36	19	QCT/Staff availability and archives. Included if CA < 40 months if mute or CA < 46 months if echolalic and prorated mental age of >11 months at CA 30 months.	Five subjects deemed untestable at intake, 3 in experimental group and 2 in control group 1. Intelligence scores based on mental age scores from the Vineland Social Maturity Scale (Doll, 1953) were used in these cases.
	Control	40.9 (10.3)	57.1 (14.5)		50.1 (22.4)		<10	24	19		
Smith, Eikeseth, Klevstrand, & Lovaas, 1997) 1, 3, 4	EIBI	36.0 (6.9)	27.8 (4.9)	50.3 (9.1)	35.8 (14.3)	51.7 (17.9)	30	24	11	QCT/Archival data. Included if CA < 46 months and IQ < 35	Post measures conducted 3–4 years after treatment in some cases. VABS data only available for 6 of the 11 children in the ABA group. Control group received minimal treatment.
	Control	38.0 (5.4)	27.3 (5.4)		24.0 (8.2)		<10	24	10		
(Howard, Sparkman, Cohen, Green, & Stansislaw, 2005) 2, 3	EIBI	30.9 (5.2)	58.5 (18.2)	70.5 (11.9)	89.9 (20.9)	81.3 (11.1)	25–40	14	29	QCT/Parental preference and IEP teams. 16 Included if CA < 48 months.	EIBI: Multiple settings (home, school and community. 25–30 h per week under 3 years of age, 35–40 h per week over 3 years of age. AP (autism educational programming): Public classroom for children with autism. 1:1 or 1:2 staff: child ratio. 25–30 hrs per week of intervention, supervision by special education teacher. Intervention eclectic (PECS, SIT; TEACCH, DTT). 7 children received 1–2 session per week of speech therapy. GP (generic educational programming): Local community special education classrooms. Average of 15 hrs per week intervention, 1:6 staff: child ratio. 13 children received speech and language therapy for 1–2 times per week.

(*Continued*)

TABLE 1
Continued

			Pre		*Post*		*Intensity*				
Study	*Group*	*Age*	*IQ*	*ABC*	*IQ*	*ABC*	*hrs*	*mths*	*n*	*Design/Assignment/ Inclusion*	*Comments*
	Comparison (AP)	37.4 (5.7)	53.7 (13.5)	69.8 (10.5)	62.1 (19.6)	69.1 (12.9)	25	13	16		
	Control (GP)	34.6 (6.5)	59.9 (14.8)	71.6 (10.5)	68.8 (15.3)	68.3 (9.9)	15	15	16		
(Smith, Groen, & Wynn, 2000) 1, 2, 3, 4	EIBI	36.1 (6.0)	50.5 (11.2)	63.7 (9.6)	66.5 (24.1)	61.3 (28.7)	24.5	24	15	RCT/Matched-pair random. Included if CA < 42 months and ratio IQ between 35 and 75.	Clinic directed group: Number of hours for ABA group are for first year in treatment. Gradual reductions in year two. Treatment phased out after 18 months for children responding slowly. Average duration 33 months. Parent managed group: 5 hrs a week of parent training for first 3–9 months, parents asked to do 5 hours a week in between sessions: Total <10 hours per week of ABA +12.5 hours of special education classes per week. ABA treatment hours second year presumed to be gradually decreasing, school hours presumed to be the same. Follow-up testing at CA 7–8 years. Duration between testing on average 54 months. Autism and PDD-NOS lumped together in the present analysis.
	Comparison	35.7 (5.4)	50.7 (13.9)	65.2 (9.0)	50.5 (20.4)	59.9 (16.7)	<10	~24	13		
(Cohen, Amerine-Dickens, & Smith, 2006) 1, 2, 3, 4	EIBI	34.4 (5.4)	62.0 (16.4)	64.0 (8.4)	81.1 (21.8)	79.5 (13.4)	35–40	24	21	QCT/Parental preference. Included if CA < 48 months and ratio IQ > 35.	Community-non-university setting. Community services selected by family. In control group 1 child had a Early Start Autism Intervention Program 9 hrs a week, 2 children home-based development program 1–4 hrs a week, 17 SDC (special day class) eclectic, ratio 1:1 to 3:1, 3–5

											days a week for up to 5 hrs. Speech, behavioral and occupational therapies 0–5 hrs per week. 3 where mainstreamed for up to 45 minutes a day.
	Control	33.2 (3.7)	59.4 (14.7)	71.9 (11.5)	65.9 (16.5)	70.7 (13.3)	–	24	21		
Australia (Birnbrauer & Leach, 1993) 1, 2, 4	EIBI	38.1 (7.1)	45.3 (17.9)	47.5	57.6 (18.7)	41.0 (5.1)	19	22	9	QCT/Parent willingness and geographical. 5 Included if CA between 24 and 48 months.	Untestable subjects set to IQ of 30. Ratio scores computed for the rest of subjects. Scores post treatment are deviation IQ used where available. Ratio VABS scores calculated both pre and post.
	Control	33.2 (10.3)	45.0 (9.4)	51.5	43.2 (15.0)	42.5 (4.9)	–	22	5		
Norway (Eikeseth, Smith, Jahr, & Eldevik, 2002) 1, 2, 3, 4	EIBI	66.3 (11.3)	61.9 (11.3)	55.8 (9.0)	79.1 (18.1)	67.0 (16.3)	28	12	13	QCT/Staff availability. Included if CA between 48 and 84 months and IQ ≥ 50	Comparison received eclectic treatment of similar intensity, but with less supervsision.
	Comparison	64.8 (9.9)	65.2 (15.0)	60.0 (13.2)	68.9 (18.8)	60.2 (11.7)	29	14	12		
(Eldevik, Eikeseth, Jahr, & Smith, 2006) 1, 2, 3, 4	EIBI	53.1 (9.5)	41.0 (15.2)	52.5 (3.9)	49.2 (16.6)	52.4 (9.2)	13	20	13	QCT/Archival data. Included if CA < 72 months.	Comparison received eclectic treatment of similar intensity. Four subjects from comparison group taken out here, because included in Eikeseth et al. (2002).
	Comparison	45.1 (16.5)	42.8 (13.0)	50.1 (9.2)	38.5 (15.5)	44.6 (7.5)	12	23	11		
United Kingdom (Remington et al., 2007) 2, 3	EIBI	35.7 (4.0)	61.4 (16.7)	60.2 (5.8)	73.5 (27.3)	61.5 (15.4)	26	24	23	QCT/Parent preference. Included if CA between 30 and 42 months.	Control group received TAU, special school, mainstream or mix, but little or no 1:1, speech therapy, TEACCH etc.
	Control	38.4 (4.4)	62.3 (16.6)	57.0 (6.8)	60.1 (27.8)	54.6 (13.1)	16	24	11		

QCT = Quasi-experimental controlled clinical trial, RCT = Randomized controlled clinical trial.

1. Included in Reichow & Wolery (2009); 2. Included in Eikeseth (2009); 3. Included in Magiati, Howlin & Charman (2009); 4. Included in Rogers & Vismara (2008).

with pervasive developmental disorders and intellectual disabilities (Newsom & Hovanitz, 1997).

Adaptive behavior. The VABS (Sparrow et al., 1984) was the only measure for adaptive behavior used in the studies included in our analysis. The VABS yields standard scores on four domains: communication; daily living skills; socialization; and, for children younger than 6 years old, motor skills. Based on these scores it will also yield a standardized ABC. In our study we only used this composite score as we did not have access to the domain scores for most of the children. The VABS is widely regarded as the best available instrument for assessing adaptive behavior in children with autism (Newsom & Hovanitz, 1997).

Tests of Homogeneity and Publication Bias

Data were entered into the Comprehensive Meta-Analysis Software (Borenstein, Hedges, Higgins, & Rothstein, 2005). To determine whether all studies were drawn from a population of studies with a common mean effect size, we performed a test of homogeneity using the *Q*-statistic and I^2, utilizing these options in the Comprehensive Meta-Analysis Software. These tests were conducted for the whole group of nine studies together. The I^2 gives the proportion of the variance that be explained by between-study variance. Using the software, we also assessed potential publication bias by a funnel plot of the standard error and effect size for each study (Egger, Smith, Schneider, & Minder, 1997) and the trim and fill method (Duval & Tweedie, 2000) for both IQ and ABC outcomes.

Effect Size Measures

The standardized mean difference effects size for EIBI were computed for IQ and ABC using the same software. Effect sizes were computed for each study separately, but we also computed an overall effect size against the comparison and control groups. We used the Hedges's *g* effect size measure (Hedges & Olkin, 1985) to adjust for the relatively small sample sizes in the studies, typically less than 20 in each group. When computing an overall (meta-analytic) effect size the individual studies were weighted using the inverse of the variance, as is widely considered to be the best practice (Borenstein, Hedges, Higgins, & Rothstein, 2009).

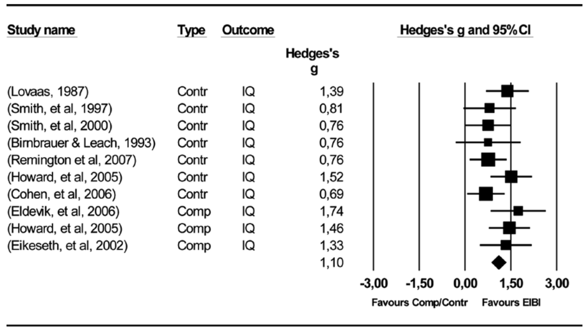
RESULTS

Homogeneity

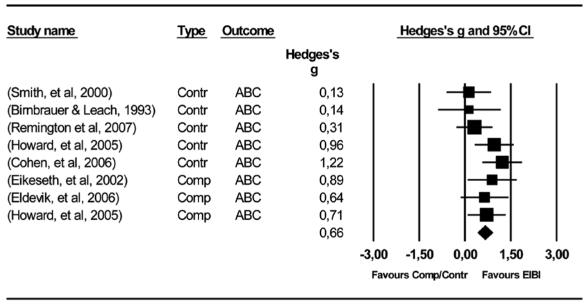
The *Q*-statistic was not statistically significant for either IQ, $Q(9) = 10.07$, $p = .345$, or ABC, $Q(7) = 8.50$, $p = .291$, scores across the nine identified studies. This indicated that all of the studies could be combined for one common effect size. We also calculated the between-study variance for IQ ($I^2 = 10.66$) and for ABC scores ($I^2 = 17.65$), and these data supported the homogeneity conclusion in that relatively small proportions of variance were explained by between-study variance. Given these findings, we used a fixed effects model for computing all effect sizes.

Effect Size Measures

The standardized mean difference effect size was calculated for IQ and ABC. The Howard, Sparkman, Cohen, Green, and Stanislaw (2005) study contributed both a control and a comparison group. We decided to calculate separate effect sizes for these. Hence, the total number of effect sizes for IQ was 10 from the nine studies included. Four studies had a comparison group and six studies had a control group. For ABC, four studies had a comparison group and four studies had a control group, the Howard et al. study again contributing one to each group, making the total number of effect sizes eight from the seven studies included. A forest plot of the effect sizes for each study and an overall effect size for IQ and ABC are shown in Figure 2.



Full-scale intelligence



Adaptive Behavior Composite

FIGURE 2 Forrest plots of standardized mean difference effect sizes (Hedges's *g*) and 95% confidence intervals (CIs). In the first plot effect sizes for full-scale IQ are shown and in the second ditto for adaptive behavior composites. Studies are grouped as either comparison or control. The fixed model effect size is computed against both the comparison and control studies and also an overall effect size is computed. EIBI = Early Intensive Behavioral Intervention.

The overall effect size for IQ change was 1.103 (95% confidence interval [CI] = .871, 1.335). The overall effect size for change in adaptive behavior composite scores was .660 (95% CI = .41, .90).

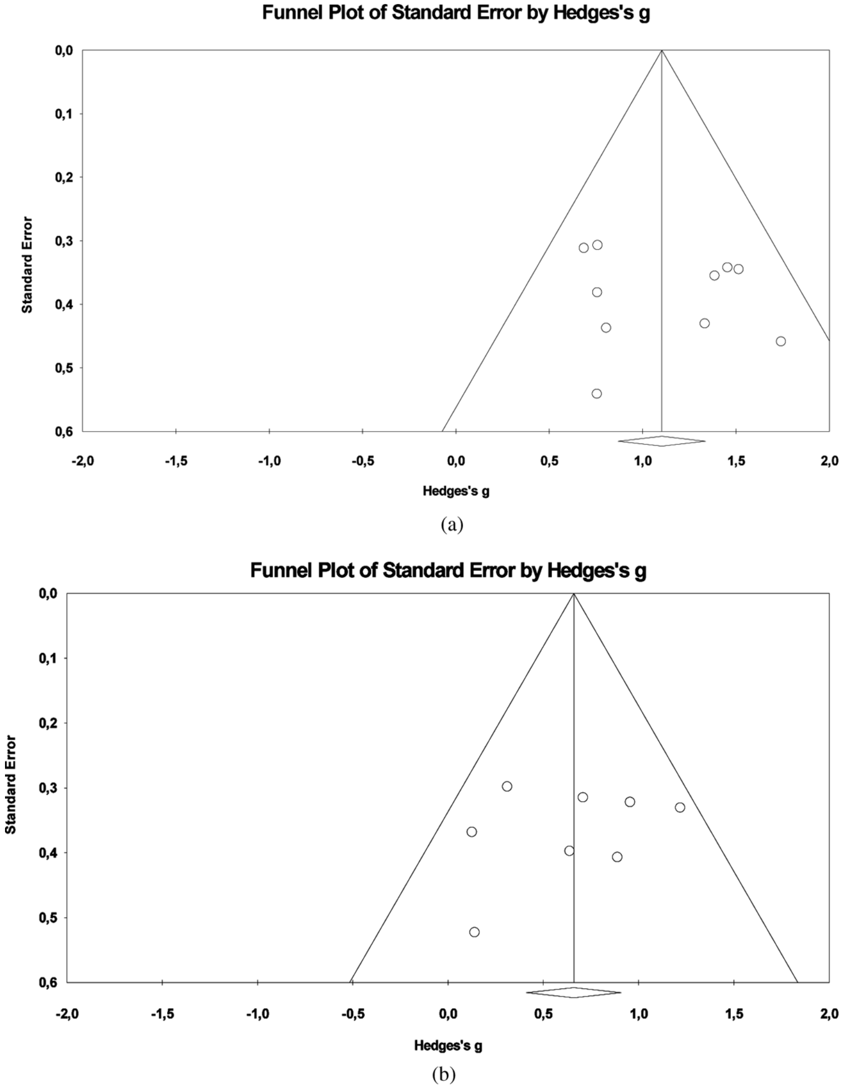
Publication Bias

We found no statistical or visual evidence of publication bias. Funnel plots of the standard error against effect sizes for IQ and ABC changes are shown in Figure 3. The Duval and Tweedie (2000) trim and fill method did not suggest the potential absence of any studies. However, the limitations of these techniques, particularly when there are few studies, means we cannot exclude publication bias.

DISCUSSION

Following EIBI treatment, our meta-analysis found an average large effect size for IQ change (based on 10 comparisons) and an average medium effect size for ABC change (based on 8 comparisons; Cohen, 1992). These estimates of effect size were also statistically significant from zero (the 95% CIs did not contain zero). We also found that the sample of studies was reasonably homogenous and that there was no evidence of publication bias.

The only other published meta-analysis we have found (Reichow & Wolery, 2009) reported an effect size of .69 for IQ change. The effect size for IQ change is thus somewhat higher in our analysis. There may be several explanations for the difference between our results



(a)

(b)

FIGURE 3 Funnel plots of effect sizes against the standard errors. In the first plot the full-scale IQ data are shown, and in the second plot the adaptive behavior composites are shown. The circles represent the studies included in the analysis, and the diamond represents the average effect size with a 95% confidence interval.

and those of Reichow and Wolery that relate to the steps we took to expand and improve on their analysis. First, to be able to use the more methodologically rigorous standardized mean difference effect size, we only included studies that had a control or a comparison group. In contrast, Reichow and Wolery computed their mean effect size based on the change within the EIBI group only. Second, we applied a more precise yet inclusive definition of EIBI, introduced by Green et al., (2002), that seems to us to be more in keeping with how other EIBI professionals define their field (e.g., Eikeseth, 2009). Reichow and Wolery employed more restricted criteria for including studies in their review, in that it had to be a replication of Lovaas's UCLA/YAP model and/or based on their treatment manuals. This may be the reason why the Howard et al., (2005) and the Remington et al., (2007) studies were not included in their analysis. However, in our opinion, the failure to include these two studies represents an inconsistency in the selection process. To us, the treatment provided in these studies is as much EIBI as the treatment provided in the studies that were included by Reichow and Wolery. Along the same lines, it seems inappropriate to us to include in the calculation of effect size a study that compared two models of EIBI service delivery (Sallows & Graupner, 2005). It is quite clear from the description of the provisions in this study that, although the groups differ in some respects, they are both examples of EIBI.

A third difference in our analysis is that we only included studies that reported full-scale intelligence scores. Reichow and Wolery did not make a distinction between performance based and full-scale intelligence measures. As we have noted earlier, this may skew results in either direction, especially when tests are used interchangeably over time. Fourth, our analysis was based on individual raw data from each study rather than the data reported in the published papers. This meant that we had a slightly different sample of children, even from the studies in common to both analyses (see Table 1).

As an extension to Reichow and Wolery's meta-analysis, we were able to include an analysis of another important outcome measure, namely, the adaptive behavior composite. This measure adds substantial validity to the outcomes, because it tells us more about the children's skills in daily life. Of interest, effects sizes were lower than for IQ. We also tested if intensity of treatment in itself may account for differences in outcome. This was possible by employing stricter criteria for what should constitute a comparison group. In our study, the comparison groups had to be given a provision of similar intensity (measured as weekly hours of 1:1 provision) as the EIBI groups. In the studies included here, this meant an "eclectic" provision. Although we agree with Reichow and Wolery's (2009) point that it is still hard to determine whether this means a specific common provision, we think it is valuable to treat them as a group, especially as eclectic provision is probably similar to a treatment as usual for many children with autism. In eclectic programs, the particular composition of treatments is to be adjusted to the individual child's needs and may thus vary a great deal across children and across time for a given child. Attempts to measure this have been made (e.g., Eikeseth et al., 2002), but it proved difficult for teachers in the eclectic groups (and thus for the researchers) to say what specific treatment they were using because they tended to blend and apply them depending on the child's behavior and needs through the day. Although difficult to specify, the eclectic approach seems to be the most common provision offered to children with ASD in service settings currently, even among those clinicians with behavior analysis training (Schreck & Mazur, 2008). Our results add to the serious concerns raised by Rogers and Vismara (2008) about eclectic treatment models.

Although we were able to refine Reichow and Wolery's meta-analysis, there are some serious limitations that remain, such that any conclusions need to be drawn with caution and to be considered tentative. First, the number of studies included in our analysis may be considered small, although it is above the median for reviews listed in the the Cochrane Database of Systematic Reviews. This database currently includes more than 3,000 reviews, and the median number of studies in a review is six (Borenstein et al., 2009). Second, a more serious limitation is the quality of the studies on effects of EIBI. Because of the lack of random assignment, only one study included in the present analysis met Type 1/highest level criteria of methodological rigor (Nathan & Gorman, 2002). Furthermore, the literature lacks comparisons between EIBI and other approaches, perhaps other than the eclectic one. Third, although there is a clear difference in outcome between EIBI and the comparison intervention, it should be noted that this may be due to differences in the amount and frequency of supervision and training. We did not have enough data to control for this in the present study. However, based on the information in the studies included, it is clear that the EIBI group in general received more frequent and more total hours of supervision and training. This remains a threat for the validity of conclusions about the superiority of EIBI in relation to comparison intervention. Fourth, we decided to include two effect sizes from the Howard et al. (2005) study, one for EIBI against the comparison group and one for EIBI against the control group. This is problematic because they are not independent of each other as both involve contrasts with a single EIBI group. We did all calculation only including the effect size from the comparison group and this did not alter the overall results in any significant manner; Hedges's $g = 1.048$ (95% CI = .80, 1.30) for IQ, and Hedges's $g = .607$ (95% CI = .34, .87) for ABC. Fifth, because of the limited number of studies and available variables we decided not to conduct an analysis of moderator variables that may explain variation in intervention outcome. However, this is certainly a priority for the future when more studies are published and more potential moderator variables can be analyzed.

Implications for Research, Policy, and Practice

With these limitations in mind, our general conclusions are very similar to those of other recent reviews: EIBI produces large to moderate effect sizes for changes in IQ and ABC scores for children with ASD when compared with no intervention controls and eclectic provision. These results support the clinical implication that EIBI at present should be an intervention of choice for children with ASD. However, randomized controlled trials comparing EIBI to other interventions are still needed. In particular, studies are needed where the comparison intervention is of similar intensity and where staff receive similar training and supervision.

REFERENCES

Anderson, S. R., Avery, D. L., DiPietro, E. K., Edwards, G. L., & Christian, W. P. (1987). Intensive home-based early intervention with autistic children. *Education & Treatment of Children*, *10*, 352–366.

Bayley, N. (1969). *Bayley Scales of Infant Development*. New York: Psychological Corporation.

Bayley, N. (1993). *Bayley Scales of Infant Development* (2nd ed.). San Antonio, TX: Psychological Corporation.

Beglinger, L., & Smith, T. (2005). Concurrent validity of social subtype and IQ after Early Intensive Behavioral Intervention in children with autism: A preliminary investigation. *Journal of Autism and Developmental Disorders*, *35*, 295–303.

Ben-Itzchak, E., & Zachor, D. A. (2007). The effects of intellectual functioning and autism severity on outcome of early behavioral intervention for children with autism. *Research in Developmental Disabilities*, *28*, 287–303.

Bibby, P., Eikeseth, S., Martin, N. T., Mudford, O. C., & Reeves, D. (2002). "Progress and outcomes for children with autism receiving parent-managed intensive interventions": Erratum. *Research in Developmental Disabilities*, *23*, 79–104.

Birnbrauer, J. S., & Leach, D. J. (1993). The Murdoch Early Intervention Program after 2 years. *Behaviour Change*, *10*, 63–74.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2005). *Comprehensive meta analysis Version 2*. Englewood, NJ: Biostat.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. (2009). *Introduction to meta-analysis*. Hoboken, NJ: Wiley.

Butter, E. M., Mulick, J. A., & Metz, B. (2006). Eight case reports of learning recovery in children with pervasive developmental disorders after early intervention. *Behavioral Interventions*, *21*, 227–243.

Chambless, D. L., Baker, M. J., Baucom, D. H., Beutler, L. E., Calhoun, K. S., & Crits-Christoph, P. (1998). Update on empirically validated therapies, II. *The Clinical Psychologist*, *51*, 3–16.

Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology*, *6*, 7–18.

Chambless, D. L., Sanderson, W. C., Shoham, V., Bennett Johnson, S., Pope, K. S., & Crits-Christoph, P. (1996). An update on empirically validated therapies. *The Clinical Psychologist*, *49*, 5–18.

Cohen, H., Amerine-Dickens, M., & Smith, T. (2006). Early Intensive Behavioral Treatment: Replication of the UCLA Model in a community setting. *Developmental and Behavioral Pediatrics*, *27*, 145–155.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155–159.

Drew, A., Baird, G., Baron-Cohen, S., Cox, A., Slonims, V., Wheelwright, S., et al. (2002). A pilot randomised control trial of a parent training intervention for pre-school children with autism: Preliminary findings and methodological challenges. *European Child & Adolescent Psychiatry*, *11*, 266–272.

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*, 455–463.

Egger, M., Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple graphical test. *British Medical Journal*, *315*, 629–634.

Eikeseth, S. (2009). Outcome of comprehensive psycho-educational interventions for young children with autism. *Research in Developmental Disabilities*, *30*, 158–178.

Eikeseth, S., Smith, T., Jahr, E., & Eldevik, S. (2002). Intensive behavioral treatment at school for 4- to 7-year-old children with autism: A 1-year comparison controlled study. *Behavior Modification*, *26*, 49–68.

Eikeseth, S., Smith, T., Jahr, E., & Eldevik, S. (2007). Outcome for children with autism who began intensive behavioral treatment between ages 4 and 7: A comparison controlled study. *Behavior Modification*, *31*, 264–278.

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

Fenske, E. C., Zalenski, S., Krantz, P. J., & McClannahan, L. E. (1985). Age at intervention and treatment outcome for autistic children in a comprehensive intervention program. *Analysis & Intervention in Developmental Disabilities*, *5*, 49–58.

Gabriels, R. L., Hill, D. E., Pierce, R. A., Rogers, S. J., & Wehner, B. (2001). Predictors of treatment outcome in young children with autism. *Autism*, *5*, 407–429.

Green, G., Brennan, L. C., & Fein, D. (2002). Intensive behavioral treatment for a toddler at high risk for autism. *Behavior Modification*, *26*, 69–102.

Harris, S. L., & Handleman, J. S. (2000). Age and IQ at intake as predictors of placement for young children with autism: A four- to six-year follow-up. *Journal of Autism & Developmental Disorders*, *30*, 137–142.

Harris, S. L., Handleman, J. S., Gordon, R., Kristoff, B., & Fuentes, F. (1991). Changes in cognitive and language functioning of preschool children with autism. *Journal of Autism & Developmental Disorders*, *21*, 281–290.

Hayward, D. W., Eikeseth, S., Gale, C., & Morgan, S. (in press). Assessing progress during treatment for young children with autism receiving intensive behavioural interventions. *Autism*.

Hedges, L., & Olkin, I. (1985). *Statistical models for meta-analysis*. New York: Academic Press.

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

Howlin, P., Magiati, I., & Charman, T. (2009). A systematic review of early intensive behavioural interventions (EIBI) for children with autism. *American Journal on Intellectual and Developmental Disabilities*, *114*, 23–41.

Ingersoll, B., Schreibman, L., & Stahmer, A. (2001). Brief report: Differential treatment outcomes for children with Autistic Spectrum Disorder based on level of peer social avoidance. *Journal of Autism and Developmental Disorders*, *31*, 343–349.

Lord, C., Risi, S., DiLavore, P. S., Shulman, C., Thurm, A., & Pickles, A. (2006). Autism from 2 to 9 years of age. *Archives of General Psychiatry*, *63*, 694–701.

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

Luiselli, J. K., Cannon, B. O. M., Ellis, J. T., & Sisson, R. W. (2000). Home-based behavioral intervention for young children with autism/pervasive developmental disorder: A preliminary evaluation of outcome in relation to child age and intensity of service delivery. *Autism: The International Journal of Research and Practice*, *4*, 426–438.

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

McEachin, J. J., Smith, T., & Lovaas, O. I. (1993). Long-term outcome for children with autism who received early intensive behavioral treatment. *American Journal on Mental Retardation*, *97*, 359–372.

Morris, S. B. (2000). Distribution of the standardized mean change effect size for meta-analysis on repeated measures. *The British Journal of Mathematical and Statistical Psychology*, *53*, 17–29.

Nathan, P., & Gorman, J. M. (2002). *A guide to treatments that work*. New York: Oxford University Press.

Newsom, C., & Hovanitz, C. A. (1997). Autistic disorder. In *Assessment of childhood disorders* (3rd ed., pp. 408–452). New York: Guilford.

Reed, P., Osborne, L. A., & Corness, M. (2007a). Brief report: Relative effectiveness of different home-based behavioral approaches to early teaching intervention. *Journal of Autism and Developmental Disorders*, *37*, 1815–1821.

Reed, P., Osborne, L. A., & Corness, M. (2007b). The real-world effectiveness of early teaching interventions for children with autism spectrum disorder. *Exceptional Children*, *73*, 417–433.

Reichow, B., & Wolery, M. (in press). Comprehensive synthesis of early intensive behavioral interventions for young children with autism based on the UCLA Young Autism Project model. *Journal of Autism and Developmental Disorders*.

Remington, B., Hastings, R. P., Kovshoff, H., Espinosa, F., Jahr, E., Brown, T., et al. (2007). Early intensive behavioral intervention: Outcomes for children with autism and their parents after two years. *American Journal on Mental Retardation*, *112*, 418–438.

Rogers, S. J., & Vismara, L. A. (2008). Evidence-based comprehensive treatments for early autism. *Journal of Clinical Child and Adolescent Psychology*, *37*, 8–38.

Roid, G. M., & Miller, L. M. (1997). *Leiter International Performance Scale–Revised: Examiners manual*. Wooddale, IL: Stoelting.

Sallows, G. O., & Graupner, T. D. (2005). Intensive behavioral treatment for children with autism: Four-year outcome and predictors. *American Journal on Mental Retardation*, *110*, 417–438.

Schreck, K. A., & Mazur, A. (2008). Behavior analyst use of and beliefs in treatments for people with autism. *Behavioral Interventions*, *23*, 201–212.

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

Smith, T., Buch, G. A., & Gamby, T. E. (2000). Parent-directed, intensive early intervention for children with pervasive developmental disorder. *Research in Developmental Disabilities*, *21*, 297–309.

Smith, T., Eikeseth, S., Klevstrand, M., & Lovaas, O. I. (1997). Intensive behavioral treatment for preschoolers with severe mental retardation and pervasive developmental disorder. *American Journal on Mental Retardation*, *102*, 238–249.

Smith, T., Groen, A. D., & Wynn, J. W. (2000). Randomized trial of intensive early intervention for children with pervasive developmental disorder. *American Journal on Mental Retardation*, *105*, 269–285.

Solomon, R., Necheles, J., Ferch, C., & Bruckman, D. (2007). Pilot study of a parent training program for young children with autism: The PLAY Project Home Consultation program. *Autism*, *11*, 205–224.

Sparrow, S. S., Balla, D. A., & Cicchetti, D. V. (1984). *Vineland Adaptive Behavior Scales*. Circle Pines, MN: American Guidance Service.

Stahmer, A. C., & Ingersoll, B. (2004). Inclusive programming for toddlers with autism spectrum disorders: Outcome from the Children's Toddler School. *Journal of Positive Behavior Interventions*, *6*, 67–82.

Stutsman, R. (1948). *Guide for administering the Merrill-Palmer Scale of Mental Tests*. New York: Harcourt, Brace & World.

Thorndike, R. L., Hagen, E. R., & Sattler, J. M. (1986). *The Stanford-Binet intelligence scale* (4th ed.). Chicago: Riverside.

Wechsler, D. (1974). *The Wechsler Intelligence Test for Children–Revised*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (1989). *Wechsler Preschool and Primary Scale of Intelligence–Revised*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (1993). *Wechsler Intelligence Scale for Children* (3rd ed.). San Antonio, TX: Psychological Corporation.

Weiss, M. J. (1999). Differential rates of skill acquisition and outcomes of early intensive behavioral intervention for autism. *Behavioral Interventions*, *14*, 3–22.

Research in Autism Spectrum Disorders 4 (2010) 577–593

Contents lists available at ScienceDirect

Research in Autism Spectrum Disorders

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

Review

A meta-analytic review of the effectiveness of behavioural early intervention programs for children with Autistic Spectrum Disorders

Maria K. Makrygianni [*], Phil Reed

Swansea University, Swansea, UK

ARTICLE INFO

Article history:
Received 23 December 2009
Accepted 13 January 2010

Keywords:
Meta-analysis
Autism
Effectiveness
Behavioural early intervention program

ABSTRACT

The effectiveness of behavioural intervention programs for children with Autistic Spectrum Disorders was addressed by a meta-analysis, which reviewed 14 studies. The findings suggest that the behavioural programs are effective in improving several developmental aspects in the children, in terms of their treatment gains, and also relative to eclectic-control programs in the same studies. Factors that were found to be correlated with the effectiveness of the behavioural programs were the intensity and the duration of the programs, the parental training, as well as the age and the adaptive behaviour abilities of the children at intake. The review showed that the high intensity of the programs is correlated with better gains in some developmental domains. Moreover, the high adaptive behaviour abilities and the young age of the children are also related with better progress the children have in some specific domains.

Contents

---

[*] Corresponding author at: Department of Psychology, Swansea University, Singleton Park, Swansea, SA2 8PP, UK. Tel.: +44 0 1792 295714; fax: +44 0 1792 295679.
E-mail addresses: mmakrigianni@yahoo.gr (M.K. Makrygianni), p.reed@swansea.ac.uk (P. Reed).

1750-9467/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.rasd.2010.01.014

1. Introduction

Autistic Spectrum Disorders (ASD) were historically viewed as lifelong developmental disabilities, characterised by intellectual, communication, social, emotional, and adaptive behavioural deficits (APA, 1994; WHO, 1992). Many intervention approaches have been suggested to help children with ASD (e.g., behavioural therapy, Son-Rise Program, TEACCH, PECS, holding therapy, diets, secretin, sensory integration therapy, music therapy, speech therapy, homeopathy, etc.). Despite the considerable number of intervention programs, the majority of the literature relating to these programs remains at the level of description, and there is little evidence provided for the effectiveness of many of these approaches. One exception to this generalization concerns treatments based on the applied behaviour analysis (ABA) approach, which are the most thoroughly evaluated, and among the best known interventions for ASD.

Research on behavioural treatments for children with ASD started around 40 years ago (see Larsson, 2007). However, the research paper that initiated most recent debates in this area was that reported by Lovaas (1987). The outcomes reported by Lovaas (1987) were remarkable, but at the same time there were significant methodological flaws with the study. Those methodological flaws provoked a strong debate, and a further set of studies concerning the effectiveness of behavioural EIPs, which attempted to address some of the issues arising from the perceived flaws in the original Lovaas (1987) study (e.g., Anderson, Avery, DiPietro, Edwards, & Christian, 1987; Bibby, Eikeseth, Martin, Mudford, & Reeves, 2002; Birnbrauer & Leach, 1993; Smith, Eikeseth, Klevstrand, & Lovaas, 1997). The findings obtained from these studies, on the whole, were promising, and most noted that behavioural EIPs were effective as a treatment approach for ASD, in that they improved the overall functioning of many children with ASD. Nevertheless, these findings have been challenged in terms of the magnitude of gains, and criticized for a further range of methodological weaknesses. For example, some of these studies (e.g., Howard, Sparkman, Cohen, Green, & Stanislaw, 2005; Sallows & Graupner, 2005) showed significant improvements in the developmental trajectory of children with ASD, while some other studies found less improvement (e.g., Bibby et al., 2002; Smith et al., 1997). Some of these studies found significant differences between behavioural EIPs and control groups (e.g., Howard et al., 2005; Sheinkopf & Siegel, 1998), while some did not (e.g., Farrell, Trigonaki, & Webster, 2005). The majority of the studies contain some methodological weaknesses (see Eikeseth, 2001), such as small sample sizes (e.g., Hoyson, Jamieson, & Strain, 1984; Smith, Buch, & Gamby, 2000), no comparison group (e.g., Bibby et al., 2002; Harris & Handleman, 2000), no random assignment of the children in the groups (e.g., Farrell et al., 2005; Smith et al., 1997), no matched groups (e.g., Eldevik, Eikeseth, Jahr, & Smith, 2006; Fenske, Zalenski, Krantz, & McClannahan, 1985), and different measures used for assessment between and within children (e.g., Howard et al., 2005; Sheinkopf & Siegel, 1998).

Despite the range of individual studies, a definitive answer regarding the effectiveness of behavioural EIPs in general has not yet been given. Further analysis of this question, beyond the discussion of individual reports, remains a necessity for various reasons. The demonstration of general effectiveness that cannot come from an individual study could have implications in cases, such as that of the Republic of Ireland, where the funding of ABA schools is in danger, because of a perceived lack of evidence regarding the general, as opposed to specific, effectiveness of this approach. Additionally, for many commissioning authorities, the behavioural EIPs have a relatively high operational cost per person, at least initially (although there is a debate relative to the long term costs of not providing treatment; Jacobson & Mulick, 2000; Jacobson, Mulick, & Green, 1998; Marcus, Rubin, & Rubin, 2000; Papps & Dyson, 2004). Notwithstanding this financial issue, the use of this kind of EIP for children with ASD has increased (Bibby et al., 2002; Jacobson et al., 1998). Consequently, both parents, and funding bodies, need reliable answers about the effectiveness of those programs and some guidance about the conditions under which such programs might be effective.

One way to study the literature on the field is meta-analysis. Meta-analysis is only one of many ways to summarise, integrate and interpret selected sets of studies. It is a research tool, developed at the end of the 1970s (Lipsey & Wilson, 2001), which translates the results from different studies to a common metric, and statistically explores the relations between the study characteristics, and their findings (Mullen, 1989; O'Mara, Marsh, & Craven, 2005).

When the present meta-analysis was initiated, there were no meta-analytic reviews published in the field of behavioural early intervention programs for children with ASD. However, recently, three meta-analyses (Eldevik et al., 2009; Reichow & Wolery, 2009; Spreckley & Boyd, 2009) have appeared in peer-reviewed journals, which, however, are characterised by significant differences in their design, inclusion criteria, and included studies and findings. Spreckley and Boyd (2009) carried out a review to the effectiveness of applied behaviour intervention programs for preschool children with ASD, in their cognitive, language, and adaptive behaviour abilities. This meta-analysis involved only four studies, in which ABA programs were compared with standard care. The analysis suggested that there is not adequate evidence for the superiority of ABA programs' outcomes to those of standard care, but the study had several limitations such as high variability in the included studies, and poor homogeneity (Spreckley & Boyd, 2009).

In a more inclusive meta-analysis with different inclusion criteria, Reichow and Wolery (2009) carried out a systematic review, which used meta-analytic techniques and was aiming mainly to provide a quantitative analysis of the studies on EIBI. That meta-analysis included 13 studies, which varied in their experimental design and quality. It provided a descriptive analysis of the studies, an effect sizes (ESs) analysis (both comparing baseline follow-up and experimental-control group data), while meta-analytic techniques were only used for the within-group changes in IQ and not for the other developmental aspects or the between group differences. Regarding the ESs analysis, there is a concern about the use of the same formula for both the calculation of developmental changes (baseline follow-up) and the differences between experimental and control programs. The findings of this study, contrary to the Spreckley's and Boyd's (2009), indicated that EIBI programs are effective, on average for children with ASD.

Lastly, Eldevik et al. (2009), in a replication and extension of the Reichow and Wolery (2009) meta-analysis, with focus on methodological improvement, employed a new meta-analysis, with stricter inclusion criteria (e.g., more precise definition of EIBA, more uniformity in outcomes, use of the raw data and inclusion of interrater reliability). That meta-analysis included nine studies and focused mainly on effect sizes based on differences between experimental and control groups, for IQ and adaptive behaviour abilities. The results from this analysis also supported the effectiveness of EIBI. Thus, aforementioned three meta-analyses constituted an important step towards the study of the effectiveness of the EIBI. However, many questions remain to be answered, such as the effectiveness of EIBI in improving other developmental aspects of children with ASD, as well as about the factors which are related to the effectiveness.

2. Purpose of the meta-analysis limitation

The present meta-analysis could be considered an extension to the previous meta-analyses about the effectiveness of the behavioural EIPs for children with ASD, which also tried to address some constraints of them. It included more studies than Spreckley's and Boyd's (2009) study, more developmental dimensions and measures of the effectiveness of the intervention programs than Eldevik's et al. (2009), and used the suggested formula by Lipsey and Wilson (2001) for standardised mean change ESs, which is not the same as the formula for mean difference ESs (see Reichow & Wolery, 2009). Additionally, it focused on slightly younger children (nursery school children), it included only moderate and high quality studies, which were analyzed separately, as well as some more recent published ones. It followed only the Lipsey and Wilson (2001) manual of meta-analysis, for both the calculation of effect sizes, and meta-analysis, and not a combination of the Hedges and Olkin (1985) and Lipsey and Wilson (2001) procedures. Lastly, more questions were addressed than in the previous meta-analyses (Eldevik et al., 2009; Reichow & Wolery, 2009; Spreckley & Boyd, 2009), such as more expressions of the effectiveness of behavioural EIPs, and the relationship between effect sizes and some factors such as the intensity the duration of the program, the parent training and the developmental characteristics (age, intellectual, language and adaptive abilities) of the child at baseline. However, due to the changes in methodology, and in the included studies, parts of the Reichow and Wolery's (2009) meta-analysis were repeated, such as the descriptive analysis of the studies and the analysis of ESs.

Consequently, the purpose of the present study was to provide a comprehensive synthesis of the research literature on the outcomes and effectiveness of the behavioural EIPs for children with ASD. Specifically, the first goal of this review was to evaluate the effectiveness of the behavioural EIPs, approaching this issue from two directions: firstly, comparing the baseline and follow-up assessment with one another; and secondly, comparing the performances of children in the behavioural EIPs with those in any eclectic-control programs (the most commonly employed control condition, and probably the most commonly employed educational intervention; see Howard et al., 2005; Reed, Osborne, & Corness, 2007b). The second goal was to identify the impact on the effectiveness of the EIPs of the characteristics of the children, such as: their age, intellectual abilities, language skills, and adaptive behaviour, and of the programs, such as: the intensity and the duration of the program, the staff number and training and the parental training.

Of course, as with any research tool, meta-analysis has some disadvantages (e.g., the quality of the included studies, the heterogeneity of the studies, the neglect of some effects and some limitations). The present meta-analysis tried to accommodate the above issues in a number of ways. The quality of the studies was assessed, and this factor was taken into account for their inclusion into the meta-analysis, as well during for the actual analysis. With few precise inclusion criteria, a homogeneous group of studies as was possible was created for this analysis. All the factors that are usually studied in relation to the effectiveness of behavioural EIPs, and that could be included in the meta-analytic study, were included in the analyses. Finally, the limitations of this meta-analysis are presented in detail throughout this paper.

3. Methodology

3.1. Selection of studies

A major threat to the validity of meta-analytic studies is the selection of studies, which is often characterised as 'biased', if it is based only on published studies (Lipsey & Wilson, 2001; Torgerson, 2006). This problem is colloquially known as the 'file drawer' problem, and refers to those unpublished studies that might exist, and whose results may fail to support the pattern established by published findings (Mullen, 1989). A very common technique that is used in meta-analysis in order to detect potential publication bias is the creation of the 'funnel-plots'. The 'funnel-plot' is a scatter-plot of the ESs with the sample size (Lipsey & Wilson, 2001). In the present meta-analysis, there is no evidence for publication bias since the plots revealed a wide dispersion of results of small sample size (similarly to previous meta-analysis). Moreover, most of the studies have a small sample size anyway, while the few unpublished studies that they were located (Beadle-Brown, Dorey, & Murphy, 2004; Mulick, 2003; Yamamoto & Nakano, 2002) had similar positive results to the published articles.

3.2. Inclusion and exclusion criteria

From the numerous of studies that have been carried in this field, the present meta-analysis focused on the published in peer-reviewed journals, longitudinal studies. Consequently, unpublished studies (e.g., Beadle-Brown et al., 2004; Mulick, 2003; Yamamoto & Nakano, 2002), reviews (e.g. Bassett, Green, & Kazanjian, 2000; Shea, 2004; Smith, 1999), retrospective studies, or studies with no pre-test assessments (e.g., Boyd & Corley, 2001; Harris, Handleman, Gordon, Kristoff, & Fuentes, 1991), as well as studies with test instruments with dubious validity or insufficient report of the outcomes (e.g., Fenske et al., 1985; Hoyson et al., 1984; Smith, Buch, et al., 2000), were excluded from the meta-analysis. Finally, single-case research and case studies were excluded, because they describe solitary cases, and do not constitute a description of the general population. Inclusion of such cases would have a disproportionate impact on the estimation of the effect sizes, and average means, giving a distorted picture of the population. Moreover, inclusion of this kind of study would have impact on the homogeneity of the included studies, and they would have required special statistical manipulation.

The rest of the studies had to meet the following seven criteria in order to be included.

3.2.1. Evaluate comprehensive treatment using the principles, methods and research findings of behaviour analysis

The included studies had to being evaluating a 'behaviour analytic treatment', or an 'early behavioural treatment', or a replication of the UCLA young project of satisfactory quality. This first inclusion criterion allowed in every study which was claiming that it was assessing an ABA program or a program based on Young Autism Project or it was a replication of Lovaas study. A further assessment of the quality of the programs and reduction of them was conducted by the sixth inclusion criterion.

3.2.2. Include children with ASD

The participants in the included studies had to have a diagnosis of autism, Autistic Spectrum Disorders (ASD), Autistic Disorder (AD), Pervasive Developmental Disorders-Not-Otherwise-Specified (PDD-NOS), and/or Pervasive Developmental Disorders (PDD).

3.2.3. Include children averaged 54 month-old, or younger, at treatment onset

This study focused on young children undergoing behavioural programs. There are several suggestions that the age of children at intake is a key factor in the effectiveness of EIPs (e.g., Bibby et al., 2002; Fenske et al., 1985). This meta-analysis studied very young children (nursery school, or first classes of primary school), where the optimum age of the child for the best outcomes may be found (based on the premise that early intervention will be more effective than remedial treatment later).

3.2.4. Provide assessments of the children's intellectual, language, and/or adaptive behaviour

These aspects of development constitute the main indices of the effectiveness of the EIP for this meta-analysis and they were chosen because they are the most commonly reported developmental aspects. Thus, studies that did not report at least one of these aspects of the children's development (e.g., intellectual, language, or adaptive behaviour) were excluded from this meta-analysis (e.g. Zachor, BenItzchak, Rabinovich, & Lahat, 2007).

3.2.5. Provide comparable results for at least one developmental aspect

Studies with inexact outcomes for the children's progress, or qualitative descriptions of the outcomes, were excluded (e.g., Mulick, 2003). These studies were excluded because such descriptions of outcomes could not be coded, and compared with those from the other studies that were included.

3.2.6. Have moderate to high methodological quality

The quality of the results of a meta-analysis depends on the quality of the primary studies. In order to control any potential bias because of the methodological quality of studies included in this meta-analysis, all the studies were assessed with an eleven-item scale. This scale was based on the research-quality indicators recommended by Reichow, Volkmar, and Cicchetti (2008). The 11 criteria on which each study was assessed were: random assignment, inter-observer agreement over 0.80, precise description of independent variable (e.g., treatment) and dependent variable, comparison group, fidelity, independent raters, reported effect sizes, participant characteristics, link between research question and data analysis, and appropriate statistical analysis with adequate power ($n > 10$). Studies which met less than half of these criteria can be considered as very weak (Reichow et al., 2008), and they were excluded from these analyses (e.g., Bibby et al., 2002; Birnbrauer & Leach, 1993; Farrell et al., 2005; Fenske et al., 1985; Sheinkopf & Siegel, 1998). The remaining studies were organized into two groups: studies with high methodological quality, which met over 9 criteria; and studies with low methodological quality, which met 6–8 criteria, and were analyzed separately in order to avoid bias because of the quality of the study.

Finally, it should be noted that studies with small samples, or which did not employ a random assignment of children to the programs, or which did not employ matched groups, were not excluded from the analysis. This was because at least one of these factors constituted a methodological weakness of almost all the studies in the field.

3.3. Search strategy

A thorough search of the literature was conducted from March 2006 until December, 2007, following the example of Smith (1999), in order to retrieve studies that met the above inclusion criteria. Several sources of potentially relevant studies were consulted, and an extensive search was carried out using search engines, and computerized bibliographic databases.

To cross-check whether all of the relevant published studies had been located, the respective citations in reviews by Bassett et al. (2000), Eldevik et al. (2009), Kasari (2002), Matson, Benavidez, Compton, Paclawskyj, and Baglio (1996), Reichow and Wolery (2009), Shea (2004), and Smith (1999), and Spreckley and Boyd (2009), as well as citations in the included studies, and relevant other studies, were examined. Finally, recommendations from experts in the field were taken into account. However, no more papers than the papers that we had already found on the basis of the above search were located.

As a result of these searches, a total of 15 eligible studies were identified. From those studies, Lovaas' (1987), and McEachin, Smith, and Lovaas (1993), used the same participants as one another, therefore, one of the two studies had to be excluded from the meta-analysis. The study that was chosen to be included was Lovaas' (1987), because its design was closer to the other included studies. Additionally, in McEachin et al. (1993) study the follow-up assessments carried out long after the baselines and during this period the children were not any longer in the studied experimental or control programs. Thus, 14 studies were included (see Appendix A), all of which were published between 1984 and 2007.

3.4. Coding

Experimental group in this study was any behavioural group, while control group was any other intervention approaches (usually 'eclectic' in nature) used to be compared with the behavioural one. Thus, the 14 studies coded for this meta-analysis were sometimes characterised by the lack of a control group (e.g., Weiss, 1999), while some studies included multiple behavioural groups of interest (e.g., Lovaas, 1987), or had one or more control groups (e.g., Reed et al., 2007b). For this reason, the unit of analysis in this meta-analysis was defined as the behavioural treatment group paired with every eclectic-control group of the study, where they existed. The collapse of multiple behavioural groups, or multiple control groups, into a single comparison from each study would have resulted in the loss of some information. At the same time, the use of each individual group from each study would give all of the possible correlations, and allows the best possible use of the data offered by each study. Thus, the 14 included studies, gave 21 units of analysis (see Appendix A). At this point, it should be clarified that each behavioural group was counted once for the descriptive factors, and for the calculation of pre–post-treatment differences (mean ES_1). Only for the differences between experimental and control groups (mean ES_2) were some experimental groups counted twice in order to be compared with all the control group in those studies (e.g., Reed, Osborne, & Corness, 2007b; Howard et al., 2005); and some control groups had to be counted twice in order to be compared with all the experimental groups of the study (e.g., Lovaas, 1987).

Due to the important role of the quality of the included studies in the meta-analysis, the 14 included studies were organized according to their methodological quality in to two groups; studies with high, and low, methodological design (see above), and these groups were analyzed separately. At this point, a possible argument could be that the ABA groups that had been used as control groups are liable to be subject to research bias, an argument which cannot be ignored. For this reason, all the behavioural-control groups were placed on the group of studies with low methodological quality.

The behavioural EIP groups were initially coded on 30 factors, categorized into three broad groups: (1) the developmental characteristics of children (baselines, follow-ups, ES_1 and ES_2) developmental characteristics of children, (2) the characteristics of the EIPs, and (3) the contextual factors related to methodological issues (e.g., the sample, the design). From all those factors only 18 were found to be statistically important and they were included in the meta-analysis. Those factors were the baseline assessments and the effect sizes (ES_1s and ES_2s) for intellectual, language and adaptive behaviour abilities (communication, daily living skills, and socialization), the number of participants, the age of the children at intake, the parent training, the intensity and the duration of the EIP and the quality of the study (see Appendices A and B). 36% of the studies (five studies) were coded by two independent researchers, and the inter-coder agreement was 95.5%.

3.5. *Statistical analysis*

The first stage of the analysis was the transformation of the provided outcomes into a common metric; in order for the same formulas to be used at the following steps of the statistical analysis. The common metric that was chosen was the standard scores, because most of the studies provided such a score in their data. If a study was not providing the outcomes in standard scores, the outcomes were not included in the next step of the analysis (e.g., Anderson et al., 1987).

The next critical step in the meta-analysis was the calculation of the effect sizes (ESs). ES is the name given to a family of indices that measure the magnitude of a treatment effect (Hedges & Olkin, 1985; Lipsey & Wilson, 2001; O'Mara et al., 2005). There are various useful ES statistics available (e.g., see Lipsey & Wilson, 2001; O'Mara et al., 2005), from which the most appropriate ones for the present study were carefully selected, according to the statistical forms in which the results are reported (usually mean standard scores and standard deviations), and the expression of the treatment effectiveness needed (comparison of pre–post-treatment performances, and/or behavioural-control groups performances). Specifically, the current study attempted to provide two expressions of the effectiveness of the behavioural EIPs, comparing the pre- and post-treatment performances of the children in the behavioural EIP (ES_1), and the performances of the children in the behavioural and control groups (ES_2) (Lipsey & Wilson, 2001).

3.5.1. *Calculation of ES_1*

For the calculation of ES_1 the index of ES for standardised mean gain was used. The formula for this index was developed for meta-analysis applications by Becker (1988) (see also Lipsey & Wilson, 2001, p. 44) and is provided in Table 1.

The standard deviation of the gain scores (SD_d) sometimes was provided by the studies, sometimes the raw data were given, and the SD_d could be estimated using SPSS, and some times these values were not provided at all. In these latter cases, an approximation of the SD_d value was attempted using regression equations, which were calculated based on the given pairs of gains (D) and SD_ds. These regression equations were: (a) for intellectual abilities $SD_d = 12.944 + 0.143D$ ($r = 0.166$); (b) for language, $SD_d = 12.301 + 0.351D$ ($r = 0.973$); and (c) for adaptive behaviour (with the exclusion of experimental group in Smith, Groen and Wynn (2000b), as an outlier), $SD_d = 5.877 + 0.756D$ ($r = 0.812$).

For the interpretation of ES_1s, the suggestions made by Cohen (1988) were used, according to which an ES ≤ 0.20 is small, an ES = 0.50 is of medium size, and an ES ≥ 0.80 is a large effect size.

Since the sample size of the included studies was variable, it was considered advisable to use the calculation of the weighted mean ES_1, instead of the simple unweighted mean ES_1 (Lipsey & Wilson, 2001). The weighted mean ES_1 was computed by weighting each ES_1 with the inverse variance weight (w) (see Table 1).

Table 1
Formulas used for the statistical analysis (Lipsey & Wilson, 2001).

ES_1	
$ES_1 = \frac{\bar{x}_{post} - \bar{x}_{pre}}{SD_d/\sqrt{2(1-r)}}$ (1)	where $\bar{x}_{post}$ and $\bar{x}_{pre}$ are the means of the post-treatment and pre-treatment assessments, respectively; r is the correlation between the pre-treatment and post-treatment scores; and SD_d is the standard deviation of the gain scores
$r = \sqrt{\frac{t^2}{t^2 + df}}$ (2)	where df is the degrees of freedom; and the t test values were calculated by formula (3)
$t = \frac{\bar{x}_{post} - \bar{x}_{pre}}{SD_d/\sqrt{n}}$ (3)	where n is the number of children in the behavioural group, $\bar{x}_{post}$ and $\bar{x}_{pre}$ are the means of the post- and pre-treatment assessments, respectively and SD_d is the standard deviation of the gain scores
Weighted mean ES_1	
$\overline{ES_1} = \frac{\sum(wES_1)}{\sum w}$ (4)	where $w = \frac{1}{SE^2}$
Standard error for ES_1	
$SE = \sqrt{\frac{2(1-r)}{n} + \frac{ES_1^2}{2n}}$ (5)	where n is the number of children in the behavioural group
Cochran's Q statistic for homogeneity	
$Q = (\sum wES_1^2) - \frac{(\sum wES_1)^2}{\sum w}$ (6)	where w is the inverse variance weight, and ES_1 is the ES_1 for each behavioural group
ES_2	
$ES_2 = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{((n_1-1)SD_1^2 + (n_2-1)SD_2^2)/((n_1-1)+(n_2-1))}}\left[1 - \frac{3}{4(n_1+n_2)-9}\right]$ (7)	where $\bar{x}_1$ and $\bar{x}_2$ are the mean standard scores for behavioural and control groups respectively; SD_1 and SD_2 are their standard deviations, respectively; and n_1 and n_2 are the numbers of children in the groups, respectively
Standard error for ES_2	
$SE = \sqrt{\frac{n_1+n_2}{n_1 n_2} + \frac{ES_2^2}{2(n_1+n_2)}}$ (8)	where n_1 and n_2 are the numbers of children in the behavioural and control groups, respectively

The next essential step for a meta-analysis was to check the data for errors, and outliers, before running any computations (Guilliksen, 1986). The purpose of every meta-analysis is to arrive at a reasonable summary of the quantitative findings of a body of research studies. This purpose is not usually served well by the inclusion of extreme ES values that are notably discrepant from those found in the research of interest, and which are possibly unrepresentative of the results of the research. In addition, extreme ES values have disproportionate influence on the values of the mean variance and other statistics used in meta-analysis, and may distort them in misleading ways (Lipsey & Wilson, 2001). Thus, the ESs were examined, and the extreme values were excluded (those with values discrepant from the rest of the values, and $<5\%$ or $>95\%$). The studies which were excluded were: Smith et al. (1997), and Magiati, Charman, and Howlin (2007), for ES_1 for IQ of the low methodological quality group; and the Magiati et al. (2007) for ES_1 for VABS of the low methodological quality.

The meta-analysis next continued with a check for the homogeneity of ESs, which determines the subsequent statistical process to be used in order to establish any relationships between the dependent and independent variables. In the meta-analysis, various ES_1s from each of the studies were to be averaged into a mean value. The key issue related to the homogeneity of the ESs is whether all of these separate ES_1s provide an estimate from the same population ES_1. If this were the case, then each ES_1 should not differ from the population mean ES_1 more than the sampling error (SE) would indicate. When the variability of the ES_1s is larger than expected on the basis of the SE, then the separate ES_1s do not describe the same population, and null hypothesis of homogeneity has to be rejected (Lipsey & Wilson, 2001). The homogeneity statistic used in this study was the Cochran's Q statistic, which is given by the formula (6) (see Table 1).

3.5.2. Calculation of ES_2

The second expression of outcome effectiveness (ES_2) comes from the comparison of the behavioural EIP with the control groups. The procedure is almost the same as for ES_1s, and only some of the formulas differ, because, in one case (ES_1), the values are related to one another, and, in the other case (ES_2), they are independent from one another (see formulas (7) and (8) in Table 1).

All of the rest of the procedure, and of the statistical analysis, and all of the rest of the formulas, were the same as described above, but where ES_1, is replaced with ES_2.

4. Results and discussion

The results from the meta-analysis are reported in two sections. The first section has some descriptive statistical elements which regard the variables which were found to be correlated with the effectiveness of the behavioural EIPs, and were included in the meta-analysis. Apart from those variables that will be presented, several other variables were assessed, but either no correlation was found with the progress of the children, or their frequency was small, and that is why they are not be presented. Some of those variables were: staff training, staff qualifications, staff number, supplementary treatments, the random assignment of the children in the study, the presence of matched groups, the type of instructions, and the location of the assessment. The second section consists of the meta-analysis, and the relationships between the dependent variables (e.g., the mean ESs of the developmental achievements of children with ASD), and the independent variables (e.g., the characteristics of the behavioural EIP, and the baseline characteristics of the children) controlled for the methodological quality of the studies.

4.1. Descriptive characteristics of studies

4.1.1. Baseline child characteristics

The mean age of the children at intake was approximately 38 months in both groups of methodological quality (see Table 2). This is a young age, as this study specifically focused on very young children with ASD. The mean IQ of the children at intake was 53 in both groups, which is substantially below the population with normal development (average = 100), and describes a group of children with mild mental retardation. The other developmental dimensions of interest were linguistic, and adaptive behavioural functioning. The mean standard score for the children's language abilities at intake was 42.71 (SD = 11.18) for the behavioural EIP on the high methodological quality studies and 47.48 (SD = 5.48) for the programs in the low methodological quality studies ($t = 1.060$, $p = 0.33$). Similarly, the mean composite standard score of adaptive behaviour at intake for these children was 57.11 (SD = 7.53), and 58.97 (SD = 6.45), for the programs in the high and low methodological quality studies, respectively, and the difference was not statistically significant ($t < 1$). Thus, the scores for both the above developmental abilities describe a population that lies well below the average in the normal distribution of developmental characteristics.

4.2. Characteristics of the behavioural EIPs

4.2.1. Qualitative characteristics – parent training

The 40% of the behavioural groups of high methodological quality and 53.8% of the behavioural groups of low methodological quality gave information about the parents training in basic behaviour analytic strategies. Although it was recommended in the majority of the studies that the parents should be actively involved in their children's treatment, in order to assist in the maintenance and generalization of the newly acquired skills across persons and settings, and to implement the program with the children outside of regularly scheduled intervention hours, only few of the parents actually received training.

Table 2
Sample characteristics, baseline assessments and quantitative characteristics of the behavioural EIPs.

	High methodological quality group		Low methodological quality group			
	N[a]	Mean (SD)	*N*[b]	Mean (SD)	*t*	Sig.
Sample characteristics						
n of subjects in behaviour groups	5	17.00 (4.69)	14	17.21 (6.99)	0.063	0.95
Age at intake	5	37.99 (8.72)	14	37.56 (4.63)	−0.143	0.89
IQ at intake	5	53.08 (8.66)	13	52.95 (14.87)	−.018	0.99
Language at intake	4	42.71 (11.18)	4	47.48 (5.48)	1.060	0.33
Adaptive behaviour at intake	4	57.11 (7.53)	9	58.97 (6.45)	−.587	0.57
	N[c]	Mean (SD)	*N*[a]	Mean (SD)	*t*	Sig.
EIPs characteristics						
Intensity (h/w)	5	27.54 (10.47)	14	25.89 (10.27)	−0.307	0.76
Duration	5	27.51 (14.83)	16	37.26 (15.89)	−1.919	0.07
n of staff per child	4	3.63 (1.12)	11	4.48 (1.94)	0.821	0.43

[a] The number of studies out of 5 which provide the required data.
[b] The number of studies out of 14 which provide the required data.
[c] The number of studies which provide information about each variable.

4.2.2. *Quantitative characteristics*

These aspects of the EIPs are shown in Table 2. The mean intensity of the studied behavioural EIPs was 27.54 (SD = 10.47) h per week (h/w) for the group of studies with high methodological quality, and 25.89 (SD = 10.27) h per week for the group of studies with low methodological quality. The difference between the two groups was not statistically significant, $t < 1$. As regards the duration that the behavioural programs (see Table 2), this varied between the two groups, but this apparent difference was not statistically significant ($t = 1.92$, $p = 0.07$). Specifically, the mean duration of the low methodological quality studies (37.26 months) was *longer* than that of the high quality studies (27.51 months).

4.3. *Effect sizes*

As the studies differed in their sample size, it was considered preferable for the combination of ESs to be based on the weighted mean ES, instead of the simple unweighted mean ES (Lipsey & Wilson, 2001; see Section 3.5 above). Thus, all of the ESs were weighted, and are presented in Table 3, along with the standard errors, and the *Q* statistics of homogeneity, separately for the two quality groups. On the same table the *t*-tests from the comparison of the mean weighted ESs are also provided.

These results are reported after the exclusion of any extreme values making the sample heterogeneous (Smith et al., 1997, and Magiati et al., 2007, for ES$_1$ for IQ of the low methodological quality group; and Magiati et al., 2007, for ES$_1$ for VABS of the low methodological quality group). In few cases (for ES$_2$ for adaptive behaviour for both groups, and ES$_2$ for IQ of low methodological quality group), where the *Q* was a bit higher than the tabulated value, a Bonferonni correction was applied. The adjusted *p* values (0.05/*n*), according to the Bonferroni correction were smaller than the tabulated values, which shows that the ESs were homogeneous.

4.3.1. *Effect size 1*

ES$_1$ describes the effectiveness of the behavioural EIPs in terms of the difference between the pre-treatment and post-treatment performances. This effect size was large for the improvement noted in the children's intellectual abilities in both

Table 3
The weighted ESs for the developmental aspects.

	High methodological quality group				Low methodological quality group			
	N[a]	Weighted Mean ES[b]	Mean SE[c]	Q	*N*[a]	Weighted Mean ES[b]	Mean SE[c]	Q
ES$_1$								
IQ	5	0.950	0.132	0.535	11	0.909	0.079	17.73
Language (total)	4	0.990	0.134	1.672	4	0.897	0.148	3.298
Adaptive behaviour (total)	4	0.421	0.154	7.990	7	0.474	0.108	8.032
VABS communication	4	0.967	0.115	2.100				
VABS daily living skills	4	−0.055	0.150	6.068				
VABS socialization	4	0.736	0.141	7.800				
ES$_2$								
IQ	3	0.568	0.192	5.076	8	0.730	0.123	19.431
Language (total)	2	0.534	0.244	0.404	4	0.910	0.177	1.996
Adaptive behaviour	2	0.971	0.256	4.310	5	0.656	0.153	11.523

[a] The number of studies for which the mean weighted ESs was calculated.
[b] Weighted mean effect sizes.
[c] Mean standard error.

groups (high methodological quality group: $w\mathrm{ES}_1 = 0.950$, SE = 0.132; low methodological quality group: $w\mathrm{ES}_1 = 0.909$, SE = 0.079), and also for the improvement in the children's language abilities (high methodological quality group: $w\mathrm{ES}_1 = 0.990$, SE = 0.134; low methodological quality group: $w\mathrm{ES}_1 = 0.897$, SE = 0.148), and they were medium to large for the improvement in adaptive behaviour (high methodological quality group: $w\mathrm{ES}_1 = 0.421$, SE = 0.154; low methodological quality group: $w\mathrm{ES}_1 = 0.474$, SE = 0.108). The differences between the groups were not statistically significant (see Table 3). Practically, these values of ESs imply that the behavioural EIPs were very effective in improving the intellectual and language abilities of children with ASD; and quite effective in improving the adaptive behaviour of those children. Although the effect sizes did differ (from −0.14 to 1.22) from study to study, most individual effect sizes were quite high, which affirms that the behavioural EIPs included in this analysis have a significant impact on the developmental trajectory of the children.

The above results for the ES_1s, in total, suggest reasons to be optimistic regarding the potential impact of behavioural EIPs, and are of some importance for both parents and professionals. An obvious potentially important parental concern is about the future of their children with ASD (Autism Society of America, 2007; Wilgosh, Waggoner, & Adams, 1988). The results from the ES_1s show that the behavioural EIPs can improve their children's language comprehension, communication skills, and socialization. At the same time, the behavioural EIPs improve the intellectual abilities of the children, which constituted one of the main measures for the determination of developmental progress of the children, and the effectiveness of the programs, in many of the early studies (Lovaas, 1987; Smith et al., 1997).

4.3.2. Effect size 2

The second approach to assessing the effectiveness of the behavioural EIPs, ES_2, gives the magnitude of the effect comparing the behavioural group with the control group (where one existed). Table 3 shows that the ES_2 for intellectual abilities is medium for both groups (high methodological quality: $w\mathrm{ES}_2 = 0.568$, SE = 0.192; low methodological quality studies: $w\mathrm{ES}_2 = 0.730$, SE = 0.177). The ES_2 for language abilities is medium in high quality studies ($w\mathrm{ES}_1 = 0.534$, SE = 0.244), are large in low quality studies ($w\mathrm{ES}_1 = 0.910$, SE = 0.177), and the opposite for adaptive behaviour (high methodological quality: $w\mathrm{ES}_2 = 0.971$, SE = 0.256; low methodological quality studies: $w\mathrm{ES}_2 = 0.656$, SE = 0.153). The differences between the groups are not statistical significant for intellectual abilities, and adaptive behaviour, $t < 1$, but they are significant for language abilities ($t = 3.87$, $p = 0.03$). Namely, the behavioural EIPs in low methodological quality studies are much more effective in comparison with the control groups than the behavioural EIP in the high methodological quality studies. In general, the above results show, in terms of practical considerations in delivering a program, that behavioural EIPs are much more effective than the eclectic (control) programs in improving the intellectual, language, and adaptive behaviour abilities of children with ASD.

At this point, and for the interpretation of the results for ES_2, it is expedient to note the variability of the eclectic (control) programs. In contrast with the behavioural programs, which share the same teaching principles and methods (applied behaviour analysis), the control programs are quite heterogeneous. They combine different teaching approaches and treatments. However, despite the heterogeneity of the eclectic programs, there is no serious reason for concern in interpreting the current results. The consistently positive ES_2s, as well as the high weighted mean ES_2s, show that the behavioural programs are more effective than whichever eclectic programs they were compared against. Of course, the variety of the eclectic programs may explain some of the range of ES_2, but these values were seldom negative, and this suggests that the eclectic program was rarely as effective as its behavioural contrast.

4.4. Effects of program characteristics

The analysis of the two forms of ESs showed that the behavioural EIPs are effective, and that they can improve the intellectual, language, and adaptive functioning of children with ASD. However, their effectiveness (ES_1) varied across the studies, from some slight negative effects (−0.14), to very positive effects (1.22; see Appendix A). Of course, this variance in effect sizes could simply be attributed to the differences between the samples studied in the included reports, or to a possible ceiling effect. However, there may also be an impact of some intervening factors, such as the characteristics of EIPs, and children's characteristics at intake. The identification of these factors could contribute to better matching of the different types of intervention programs to the children's needs, as well as to the development of better intervention programs. Thus, the next step of the present meta-analysis was to explore the relationships between the ESs and the factors that have been highlighted from the relevant bibliography, or from researchers, as possible predictors of effectiveness. Table 4 gives the summarised results from a series of partial correlations. Partial correlations were chosen in order to examine the relationship between different factors, while controlling the effect of methodological quality of the studies.

4.4.1. Intensity

This is one of the main characteristics of behavioural EIPs that has attracted the interest of investigators, from even the very early studies (e.g., Lovaas, 1987; Smith et al., 1997). The present analysis shows (see Table 4) that there are statistically significant correlations between the intensity of the EIP and ES_1 for children's intellectual, and adaptive behavioural, functioning. The correlation coefficients of 0.674 for intellectual abilities, and 0.855 for adaptive behavioural functioning, are characterised as moderate to high. Thus, it appears that more intensive programs, in general, have a higher impact on the gain in intellectual, and adaptive behavioural abilities of children with ASD. However, the intensity of the behavioural program does not seem to be correlated with progress on the children's language abilities. There was also a high, positive

Table 4
Partial correlations of the effect sizes with some of the characteristic of the programs and the children at intake, while controlling for the effects of the methodological quality of the studies.

	ES_1			ES_2		
	Intellectual	Language	Adaptive behaviour	Intellectual	Adaptive behaviour	
Intensity	Pearson	0.674	0.571	0.885	0.841	0.839
	Sign	0.012	0.237	0.002	0.004	0.037
	N	12	5	8	8	5
Duration	Pearson	−0.167	0.610	0.417	0.168	0.898
	Sign	0.551	0.146	0.230	0.666	0.015
	N	14	6	9	8	5
Age at intake	Pearson	−0.233	−0.736	−0.477	−0.656	−0.691
	Sign	0.404	0.059	0.164	0.055	0.129
	N	14	6	9	8	5
Parental training	Pearson	−0.214	0.150	−0.554	0.251	0.946
	Sign	0.443	0.748	0.121	0.515	0.004
	N	14	6	8	8	5
Intellectual abilities at intake	Pearson	0.229	0.370	0.583	0.355	0.261
	Sign	0.311	0.414	0.100	0.348	0.617
	N	14	6	8	8	5
Language at intake	Pearson	−0.020	0.081	0.428		
	Sign	0.965	0.863	0.397		
	N	6	6	5	2	2
Adaptive behaviour at intake	Pearson	0.242	0.924	0.218	0.635	0.859
	Sign	0.501	0.003	0.545	0.176	0.028
	N	9	6	9	5	5

correlation between program intensity and ES_2 for intellectual abilities ($r = 0.842$), and between program intensity and ES_2 for adaptive behaviour abilities ($r = 0.885$). This means that the more intensive the behavioural EIP, the more effective it is compared to the control program in improving the intellectual and adaptive behaviour abilities of children with ASD.

It should be noted that these results do not give an absolutely clear answer regarding an optimal range of hours per week that will deliver the best developmental achievement, without engaging in unnecessarily intensive training, and exhausting children and staff alike (Hastings & Johnson, 2001), and which is also feasible for parents in terms of time investment and financial expenses. All the behavioural EIPs but one (the experimental group in Reed et al., 2007a) which were implemented for 25 h per week or more, varied from quite, to very effective (ESs > 0.7), in improving all the studied developmental aspects, intellectual, language and adaptive behaviour abilities, although increases in the program intensity above 25 h did not produce increases in the outcomes. There was higher variability of the outcomes for the less intensive ($\leq$25 h/week) intervention programs, which could be attributed to various factors.

4.4.2. Duration

This aspect of the program constitutes another important characteristic of EIPs. In the present study, it describes the interval between the baseline and the follow-up assessment. Although it was expected that duration would be an important predictor of the outcome effectiveness, and, that the longer a child was in a behavioural program, the better progress would be made; it was found that there was only one statistically significant correlation ($r = 0.898$) for ES_2 for adaptive behaviour abilities. This correlation shows that the longer the children stay in a behavioural EIP, the more effective the behavioural EIP are compare to the control programs. No statistical significant relationship was found for any ES_1 of developmental aspects. Thus, the most conservative conclusion would be that prolonging an intervention program over several years does not necessarily entail the maintenance of the same progress rate in the developmental domains, and that the effectiveness of the program varies independently from the programs' duration.

It should be mentioned that a problem that was also noticed was that in some studies, (a) some of the children had just started the intervention program, and some were already receiving it, and (b) in some studies the implementation of the program was completed some weeks, or months, before the follow-up assessment, and in some other studies the program was still in progress when the follow-up assessment was carried out. This heterogeneity could be another reason of the absence of statistically significant correlations, as well as the ceiling effect. Thus, duration is a factor which has to be studied more, in order for its impact on the effectiveness to be assessed, if, for no other reason, than because this factor is correlated with the cost of the intervention.

4.4.3. Parent training

Although all the programs asked for parental involvement, only half of them (10/20) provided relevant training to parents. Parental training in understanding and use of the behavioural teaching techniques could potentially contribute to the creation of a learning environment for the child that has both consistency and continuity (Moroz, 1989). It is often suggested that this would facilitate the generalization, and the maintenance, of the teaching skills in children with ASD, and their development (Ingersoll & Dvortcsak, 2006). For this reason, it was expected that the variability of this factor might have impact on the effectiveness of the program. However, the analysis showed only one statistically significant association ($r = 0.946$) between parent training and ES_2 for adaptive behaviour. This correlation shows that the difference between behavioural and control groups was bigger when the behavioural programs include parental training. Of course, the latter comparison would be more accurate, if both the behavioural and the control programs had parental training.

4.5. Effects of child characteristics

4.5.1. Child's age at intake

Another factor which varies across the studies, and whose impact on the effectiveness of the program is of significant interest, is the child's age at intake. However, the current analyses did not find statistically significant correlations, only two trends between these variables. Firstly, the children's age at intake was negatively correlated ($r = 0.736$, $p = 0.059$) with the ES_1 for language abilities. This means that the younger the children are at intake, the greater the impact of the behavioural EIPs on their language abilities; while, when the children at the start the program were relatively older, the effect of the program was only moderate.

Apart from the above trends, there were no statistical significantly correlations found with the ES_1s. However, as this factor has been widely suggested as a potential predictor (e.g., Fenske et al., 1985; Green, 1996; Harris & Handleman, 2000), the scatter-plots were checked for possible extreme cases, or non-linear patterns in the relationship. These relationships between ES_1 for intellectual, language, and adaptive behavioural abilities, with age at intake, are presented in Fig. 1. Inspection of these data appears to suggest some potentially interesting relationships. Although there was no indication of any clear correlation between the pairs of variables, almost all of the programs that the children started to attend before their 35th month of age, seem to be moderately effective ($ES_1 > 0.5$; see Fig. 1, box C). From the studies included in this meta-analysis, the programs that work with very young children appear not to produce unsatisfactory results, while, for the children who started the program at an older age, there was greater variability in the outcomes, which could be attributed to various interfering factors.

The second trend ($r = -0.798$) that was found was with ES_2 for intellectual abilities. Specifically, the younger the children at the start of the program, the more effective the behavioural EIP, compared to the control programs. For the children who started their behavioural EIP later, the difference in the effectiveness between the behavioural and control interventions was relatively small. Concerning the other ES_2 relationships, there were not any statistically significant correlation, and it would be unwise to comments further, as the number of the programs providing the relevant ES_2s was quite small.

4.5.2. Intellectual abilities at intake

This factor was not correlated with any of the ES_1 and ES_2 for the studied developmental aspects. This implies that the children's intellectual abilities at intake have little impact on the effectiveness of the behavioural programs. The behavioural EIPs had the same effectiveness for children with very low, and medium, intellectual abilities, and there was no clear

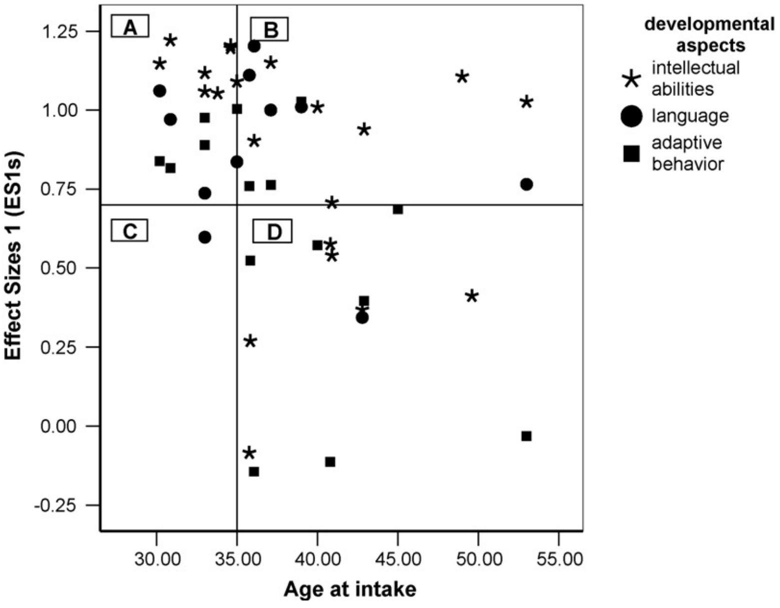


Fig. 1. The scatter-plot of age at intake with the ES_1s.

covariation between the relative effectiveness of the behavioural and the control groups, and the intellectual abilities of the children at the beginning of the program.

At this point, it should be clarified that the absence of significant correlation between the intellectual abilities of the children at the beginning of the program and the ES_1 for intellectual abilities does not contradict the findings from several studies regarding the correlation of the intellectual abilities of the children at the baseline and the follow-up assessment (e.g., Eldevik et al., 2006; Harris & Handleman, 2000; Sallows & Graupner, 2005). Actually, in the present study, the strength of the correlation is verified as very strong ($r = 0.895$), which shows that although the programs do not seem to be more effective for the low or high functioning children, their post-treatment performance is related to their pre-treatment one.

4.5.3. Language ability at intake

This factor is not correlated with any of the ESs. This means that the effectiveness of the behavioural EIPs does not depend on the language abilities of the children at intake. The behavioural EIPs are equally effective for both nonlinguistic children, and children who were able to speak at the beginning of the program, and their effectiveness compared to the control programs is not affected by language abilities of the children at intake.

4.5.4. Adaptive behaviour at intake

This factor is correlated significantly with ES_1 for language abilities ($r = 0.924$), and with ES_2 for adaptive behaviour ($r = 0.859$). The first relationship means that the higher the adaptive behaviour abilities of the children at intake, the more effective the behavioural EIPs are in improving the language abilities of the children. The second correlation shows that the better the adaptive behaviour abilities of the children at intake, the more effective the behavioural EIPs are compared to eclectic programs in improving these abilities of the children.

5. Conclusions

This meta-analysis represents a synthesis of 14 studies on the effectiveness of behavioural EIPs for children with ASD. The purpose of the analysis was to evaluate the magnitude of the effectiveness of the behavioural EIPs, comparing both pre–post-treatment performances, and behavioural-control outcomes, as well as to study the impact of various different factors on these outcomes. All the analysis carried out while controlling for the effect of the methodological quality of the studies.

The meta-analysis showed that the behavioural EIPs are very effective in improving the intellectual, language, communication and social abilities of children with ASD, while they had a moderate to high effect on the adaptive behavioural improvement of the children. From the other meta-analyses in the field, only that of Reichow and Wolery (2009) provided an estimation of effect sizes from the gains between baseline and follow-up assessments (ES_1). However, Reichow and Wolery (2009) used the same formula as for the effect sizes between treatment group differences (ES_2). Consequently, this means that the present study, and that of Reichow and Wolery (2009), used different formulas, which had some impact to the actual values of effect sizes. The values of ES_1 in the present meta-analysis were a bit higher that the values of the previous one. Nevertheless, despite the small differences, the general conclusion was the same. The behavioural early intervention programs are quite effective for children with ASD.

The present meta-analysis, apart from the effectiveness based on developmental gains, also measured the effectiveness of behavioural programs, comparing them with control (non-ABA) programs. Those findings indicated that the behavioural early intervention programs were much more effective than the control eclectic EIPs in improving the intellectual and language abilities, and adaptive functioning of children. That effect size (ES_2) was also calculated by the other meta-analyses (Eldevik et al., 2009; Reichow & Wolery, 2009), but each study included different studies in the analysis. This fact had some impact on the actual value of mean ES_2s, but the differences in general were small. Thus, all of the studies were in agreement that the behavioural intervention programs were effective compared to control treatments, and ES_2 ranged from medium to large. The exception was Spreckley and Boyd's (2009) study, which included *only four* studies, and found that there were not adequate evidence that behavioural intervention programs are more effective than standard care.

Returning to the current study, no statistical significant differences were found between high and low methodological quality studies. Only exception was the ES_2 for language abilities which showed that in the low quality studies the behavioural programs seem to be more effective compared to control program in improving the language abilities of the children, than the behavioural programs in the high quality studies.

Further analysis attempted to detect factors that impact on the developmental trajectory of the children. This analysis showed that the improvement of intellectual and adaptive behaviour abilities of the children in the behavioural EIPs is affected by the intensity of the program (at least up to a point). If parents want to have the best possible outcome on the intellectual functioning of their children with ASD, they appear to have to enroll their child in a reasonably intensive behavioural EIP. Although it should be noted that intensity is a relative concept, and for the present study, "intensive" could be applied to programs administered for about 25 h per week (which is just slightly higher than that suggested in the one single study of this factor, Reed et al., 2007a). The difference between the behavioural and control programs in improving the intellectual abilities is also affected by the intensity of the program. Namely, the more intensive the behavioural EIP is, the more effective, compared to control programs.

In terms of specific abilities, the development of children's language seems to be affected by the adaptive behaviour of the children at intake. The language abilities of the children improve more when the children start with high adaptive behavioural skills. The improvement of adaptive behaviour is impacted by the intensity of the program. The children's adaptive behaviour improves more when they attend an intensive behavioural EIP. Moreover, the difference between behavioural and control interventions in improving children's adaptive behaviour is affected by the intensity of the program, the duration, the parent training and the adaptive behaviour abilities of the children at intake. Consequently, the more intensive the behavioural EIPs are and the longer they last, the more effective they are compare to the control programs. Moreover, they are more effective when they have parent training and when the adaptive behaviour skills of the children at intake are high.

Thus, the results of this meta-analysis provide evidence that behavioural EIPs are quite effective in changing different developmental aspects of children with ASD, and they are much more effective than eclectic programs. Factors that have been distinguished as important for the efficacy of the EIPs are the intensity and the duration of the program, the age of the children at intake, the adaptive behavioural abilities of the children at intake, and parent training.

However, further studies and meta-analyses in the field are still necessary. In particular, these should address the limitation of the previous studies and meta-analysis. While the object of a meta-analysis is, in part, to overcome such problems, it should be noted that most of the included studies in this meta-analysis are characterised by some methodological limitations, whose impact on the result cannot be estimated. Some of the limitations of some of the included studies are: their small sample size; the lack of comparison group, matched groups, or random assignment of the children; the use of a variety of measures in the same study, which may have different emphasis on particular skills, as well as potential floor effects. Many tests have floors, around 40, which is the lowest score that the children could produce, and several children in the included studies had such a 'floor' score.

Another limitation of the present meta-analysis is based on the unit of analysis. Some studies with more than one experimental, or control, group gave more than one unit of analysis. In those studies, only one of these groups was the target group. Thus, the 'Rosenthal effect' may appear, according to which researcher's beliefs, biases, and expectations can have impact on the phenomenon under investigation. In order to overcome this problem in the present meta-analysis the behavioural group that was used as control groups were relegated to the group of low methodological quality studies. Finally many studies were excluded in an effort to create a homogeneous group of studies. Nevertheless, other meta-analysis can be conducted synthesizing single-subject designs or unpublished studies.

The inclusion of both studies with pre–post design, and studies with control groups, warrants some comment. It should be mentioned that all the 'high quality' studies had a control group, and so this analysis was conducted on a homogeneous set of studies. Nevertheless, there were some other studies which had a control group, but also some other limitations, which impacts on their quality. For this reason, it was decided that, apart from the existence of a control group, a number of methodological factors also to be checked (6th inclusion criterion). Moreover, if the studies with only pre–post design had been separated from the studies with control groups, there would not be enough studies to conduct such an analysis. Finally, the studies with control group differ a lot in the nature of the comparison group; some have a treatment as usual group, others have different ABA programs, and some others have another type of intervention which makes the rationale for separate analysis even more debatable.

Another issue which could raise discussion is the number of included studies, and the inclusion only of published studies. In fact, there is little agreement among researchers about the appropriate approach to these issues (see Lipsey & Wilson, 2001), but what is recommended is that meta-analysts should keep the methodological criteria strict, and accept the consequences in regards to the limitation that this will be imposed. The proponents of including only published studies argue this on the basis of the following ideas. The quality of the meta-analysis depends on the quality of the included studies. Usually, published studies have a higher quality since they have passed through a rigorous review. Additionally, meta-analytic techniques do not depend upon how many data points there are, and meta-analysis conducted on small research literature can be both practically and theoretically informative (Mullen, 1989, pp. 17, 30–31; Lipsey & Wilson, 2001, pp. 9–10). Nevertheless, at the beginning of this meta-analysis, there was an attempt to include unpublished studies, but, in the studied field, the unpublished studies, apart from their questionable quality, are difficult to locate. Moreover, their representativeness is uncertain, as the information provided about the study is often not extensive, and, usually, the people who conducted them either did not respond to our calls for clarifications, or they did not have good records of what they had done. Lastly, another limitation in the present study is the lack of a plan for the sampling procedure. Since it was not considered from the beginning such a plan to be reported, there are not records of the exact number of studies retrieved and excluded.

It is almost a truism to say that more study of the efficacy of behavioural EIPs needs to be performed in order to be able to draw secure conclusions about the factors that can affect their effectiveness. However, on the basis of the present meta-analytic review, some recommendations could be made concerning future studies. Given that it is very difficult to conduct studies with large samples, the efficacy of the different treatments should continue to be evaluated by employing further meta-analytic studies, as new data becomes available. For this reason, it would be useful if future investigators could choose common, and widely used instruments for the assessment of the children's developmental characteristics as they have started to do the last years, to give the data in a comparable form in all studies, and to provide SDs for all measures. Moreover, it would be useful if more studies carried out, which look for other factors (in the family and the school) that may have impact of on the outcomes.

Appendix A

Studies	Methodological quality	*N* of subjects in BGs	*N* of subjects in CGs	h/w	Duration	Mean age at intake	Parent training	Baseline assessments		
								IQ	Language (total)	Adaptive behaviour
Anderson et al. (1987)	Low	14		20	12	42.79		57.26	52.35	50.71
Ben-Itzchak and Zachor (2007)	Low	25			8	26.6	Training	70.67		
Cohen et al. (2006)	High	21	21	37.5	36	30.2	Training	61.60	52.30	69.80
Eldevik et al. (2006)	High	13	15	12.5	20.3	53	Training	41.00	35.55	52.50
Howard et al. (2005)	Low	29	16	32.5	14.21	30.86	No training	58.54	52.02	70.46
	Low	29*	16	32.5*	14.21*	30.86*	Training	58.54*	52.02*	70.46*
Lovaas (1987)	Low	19	21	40	30	34.6	Training*	53.00		
	Low	19	21**	10	30	40.9	Training	46.00		
Magiati et al. (2007)	Low	28	16	32.4	25.5	38	Training	83		59.6
Reed et al. (2007a)	Low	14		30.4	9.5	42.9	Training	57.20		59.30
	Low	13		12.6	9.5	40.8	No training	49.30		56.50
Reed et al. (2007b)	Low	12	20	20.4	9	40	No training	55.60	22.88	58.20
	Low	12*	16	20.4*	9*	40*	No training	55.60*	22.88*	58.20*
Remington et al. (2007)	High	23	21	25.6	24	35.7	No training*	61.43		
Sallows and Graupner (2005)	High	13		37.58	48	35	Training	50.85	43.39	59.54
	Low	10		31.28	48	37.1	No training	52.10	43.61	60.90
Smith et al. (1997)	Low	11		30	35	36	No training	27.81		
	Low	10		10	26	38	No training	27.00		
Smith, Groen et al. (2000)	High	15		24.52	58	36.07	No training	50.53	39.59	63.44
	Low	13			56.46	35.77	No training	50.69	41.94	65.17
Weiss (1999)	Low	20		40	24	41.5	Training			49.85

*: repetition in order this EG to be compared with both CGs; **: repetition in order this CG to be compared with both EGs; GAP: not applicable or not available.

Appendix B

Studies	Intellectual abilities		Language		Adaptive behaviour (total)		VABS communication		VABS daily living skills		VABS socialization		Intellectual abilities		Language		Adaptive behaviour	
	ES_1	SE	ES_1	SE	ES_1	SE	ES_1	SE	ES_1	SE	ES_1	SE	ES_2	SE	ES_2	SE	ES_2	SE
Anderson et al. (1987)	0.367	0.661	0.344	0.673														
Ben-Itzchak and Zachor (2007)	1.020	0.767																
Cohen et al. (2006)	1.160	0.972	1.038	1.011	0.812	0.711	0.864	0.840	0.547	0.579	1.095	0.826	0.88	0.32	0.66	0.32	1.37	0.34
Eldevik et al. (2006)	1.027	0.791	0.765	0.984	−0.031	0.147	0.582	0.689	−0.484	0.544	0.679	0.503	0.27	0.38	0.35	0.38	0.47	0.38
Howard et al. (2005)	1.221	0.891	0.970	0.912	0.816	0.703	1.169	0.729	0.667	0.508	0.883	0.603	1.33 1.08	0.34 0.33	0.97 1.11	0.33 0.33	1.01 1.20	0.33 0.34
Lovaas (1987)	1.204 0.576	1.064 0.654											1.29 −0.27	0.35 0.32				
Magiati et al. (2007)	−0.449	0.465					0.125	0.246	−0.280	0.279	0.246	0.275						
Reed et al. (2007a)	0.939 0.575	1.070 0.814			0.396 −0.113	0.435 0.207												
Reed et al. (2007b)	1.009	1.054			0.572	0.569	0.712	0.850	0.405	0.557	0.557	0.665	0.39 0.92	0.37 0.40	0.53 0.96	0.37 0.40	−0.02 −0.09	0.37 0.38
Remington et al. (2007)	0.856	0.773											0.48	0.31				
Sallows and Graupner (2005)	1.105 1.159	1.358 1.423	0.790 0.972	1.147 1.374	0.815 0.687	1.047 0.853	0.802 0.987	1.204 1.338	0.173 0.000	0.500 0.000	0.932 0.993	1.124 0.931						
Smith et al. (1997)	0.726 −0.481	0.944 0.620																
Smith, Groen et al. (2000)	0.903 −0.084	1.076 0.332	1.203 1.111	1.637 1.418	−0.144 0.729	0.459 0.795	0.574 −0.134	0.916 0.364	−0.519 −0.652	0.809 0.843	0.295 −0.024	0.604 0.158						
Weiss (1999)																		

GAP: not applicable or not available.

References

American Psychiatric Association (APA). (1994). *Diagnostic and Statistical Manual of Mental Disorders (DSM IV)* (4th ed.). Washington, DC: American Psychiatric Association.

Anderson, S. R., Avery, D. L., DiPietro, E. K., Edwards, G. L., & Christian, W. P. (1987). Intensive home-based early intervention with autistic children. *Education and Treatment of Children, 10*, 352–366.

Autism Society of America. (2007, July 26). *Stress on Families*. Retrieved July 26, 2006 from http://www.autism-society.org/site/PageServer?pagename=about_lwa_stress.

Bassett, K., Green, C. J., & Kazanjian, A. (2000). *Autism and Lovaas treatment: A systematic review of effectiveness evidence (BCOHTA 00:1T)*. Vancouver, BC: British Columbia Office of Health Technology Assessment. Retrieved from: http://www.chspr.ubc.ca/files/publications/2000/bco00-01T_autism3.pdf.

Beadle-Brown, J., Dorey, H., & Murphy, G. (2004). *Early intervention for autism study: Summary of findings* (University of Kent). Retrieved May 25, 2007, from http://www.earlyintervention.org.uk/Libraries/Local/125/Docs/Tizard%20research%20-%20findings.pdf.

Becker, B. J. (1988). Synthesizing standardized mean-change measures. *British Journal of Mathematical and Statistical Psychology, 41*, 257–278.

Ben-Itzchak, E., & Zachor, D. A. (2007). The effects of intellectual functioning and autism severity on outcome of early behavioral intervention for children with autism. *Research in Developmental Disabilities, 28*, 287–303.

Bibby, P., Eikeseth, S., Martin, N. T., Mudford, O. C., & Reeves, D. (2002). Progress and outcomes for children with autism receiving parent-managed intensive interventions. *Research in Developmental Disabilities, 23*, 81–104.

Birnbrauer, J. S., & Leach, D. J. (1993). The Murdoch early intervention program after 2 years. *Behaviour Change, 10*, 63–74.

Boyd, R. D., & Corley, M. J. (2001). Outcome survey of early intensive behavioral intervention for young children with autism in a community setting. *Autism, 5*, 430–441.

Cohen, H., Amerine-Dickens, M., & Smith, T. (2006). Early intensive behavioural treatment: Replication of the UCLA model in a community setting. *Developmental Behavioral Pediatrics, 27*, S145–S155.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Eikeseth, S. (2001). Recent critiques of the UCLA young autism project. *Behavioral Interventions, 16*, 249–264.

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

Eldevik, S., Hastings, R. P., Hughes, J. C., Jahr, E., Eikeseth, S., & Cross, S. (2009). Meta-analysis of early intensive behavioral intervention for children with autism. *Journal of Clinical Child and, Adolescent Psychology, 38*, 439–450.

Farrell, P., Trigonaki, N., & Webster, D. (2005). An exploratory evaluation of two early intervention programmes for young children with autism. *Educational and Child Psychology, 22*, 29–40.

Fenske, E. C., Zalenski, S., Krantz, P. J., & McClannahan, L. E. (1985). Age at intervention and treatment outcome for autistic children in a comprehensive intervention program. *Analysis and Intervention in Developmental Disabilities, 5*, 49–58.

Green, G. (1996). Early behavioral intervention for autism: What does research tell us? In C. Maurice, G. Green, & S. Luce (Eds.), *Behavioral intervention for young children with autism: A manual for parents and professionals* (pp. 29–44). Austin, TX: Pro-Ed.

Guilliksen, H. (1986). The increasing importance of mathematics in psychological research (part 3). *The Score, 9*, 1–5.

Harris, S. L., & Handleman, J. S. (2000). Age and IQ at intake as predictors of placement for young children with autism: a four- to six-year follow-up. *Journal of Autism and Developmental Disorders, 30*, 137–142.

Harris, S. L., Handleman, J. S., Gordon, R., Kristoff, B., & Fuentes, F. (1991). Changes in cognitive and language functioning of preschool children with autism. *Journal of Autism and Developmental Disabilities, 21*, 281–290.

Hastings, R. P., & Johnson, E. (2001). Stress in UK families conducting intensive home-based behavioural intervention for their young child with autism. *Journal of Autism and Developmental Disorders, 31*, 327–336.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.

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

Hoyson, M., Jamieson, B., & Strain, P. (1984). Individualized group instruction of normally developing and autistic-like children: A description and evaluation of the LEAP curriculum model. *Journal of the Division for Early Childhood, 8*, 157–181.

Ingersoll, B., & Dvortcsak, A. (2006). Including parent training in the early childhood special education curriculum for children with autism spectrum disorders. *Topics in Early Childhood Special Education, 26*, 179–187.

Jacobson, J. W., & Mulick, J. A. (2000). System and cost research issues in treatments for people with autistic disorders. *Journal of Autism and Developmental Disorders, 30*, 585–593.

Jacobson, J. W., Mulick, J. A., & Green, G. (1998). Cost benefit estimates for early intensive behavioural intervention for young children with autism-general model and single state case. *Behavioral Interventions, 13*, 201–226.

Kasari, C. (2002). Assessing change in early intervention programs for children with autism. *Journal of Autism and Developmental Disorders, 32*, 447–461.

Larsson, E. V. (2007, January 14). *Data-Based research in support of intensive early intervention*. Retrieved April 5, 2007 from http://rsaffran.tripod.com/IBIrefs.html.

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

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

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

Marcus, L. M., Rubin, J. S., & Rubin, M. A. (2000). Benefit-cost analysis and autism services: A response to Jacobson and Mulick. *Journal of Autism and Developmental Disorders, 30*, 595–598.

Matson, J. L., Benavidez, D. A., Compton, L. S., Paclawskyj, T., & Baglio, C. (1996). Behavioral treatment of autistic persons: A review of research from 1980 to the present. *Research in Developmental Disabilities, 17*, 433–465.

McEachin, J. J., Smith, T., & Lovaas, O. I. (1993). Long-term outcome for children with autism who received early intensive behavioral treatment. *American Journal on Mental Retardation, 97*, 359–372.

Moroz, K. J. (1989). Educating autistic children and. youths: A school-family–community partnership. *Social Work in Education,* Winter: 107–122.

Mulick, J. (2003, July 8). *Some autistic children make remarkable progress with intensive behavioral program, study suggests.* Retrieved May 25, 2006, from http://researchnews.osu.edu/archive/autism3.htm.

Mullen, B. (1989). *Advanced BASIC meta-analysis*. Hillsdale, NJ: Lawrence Erlbaum Associates.

O'Mara, A. J., Marsh, H. W., & Craven, R. G. (2005). *Meta-analytic methods in educational research: Issues and their solutions* (OMA05407). Retrieved September 20, 2009 from: http://www.aare.edu.au/05pap/oma05407.pdf.

Papps, I., & Dyson, A. (2004). *The costs and benefits of earlier identification and effective intervention — final report. DfES research report 505*. Nottingham: DfES Publications.

Reed, P., Osborne, L. A., & Corness, M. (2007). Brief Report: Relative effectiveness of different home-based behavioural approaches to early teaching intervention. *Journal of Autism and Developmental Disorders, 37*, 1815–1821.

Reed, P., Osborne, L. A, & Corness, M. (2007). The real-world effectiveness of early teaching interventions for children with autistic spectrum disorders. *Exceptional Children, 73*, 413–417.

Reichow, B., Volkmar, F. R., & Cicchetti, D. V. (2008). Development of the evaluative method for evaluating and determining evidence based practices in autism. *Journal of Autism Developmental Disorders, 38*, 1311–1319.

Reichow, B., & Wolery, M. (2009). Comprehensive synthesis of early intensive behavioral interventions for your children with autism based on the UCLA young autism project model. *Journal of Autism Developmental Disorders, 39*, 23–41.

Remington, B., Hastings, R., Kovshoff, H., degli Espinosa, F., Jahr, E., Brown, T., et al. (2007). Early intensive behavioral intervention: Outcomes for children with autism and their parents after two years. *American Journal on Mental Retardation, 112*, 418–438.

Sallows, G., & Graupner, T. (2005). Intensive behavioral treatment for children with autism: Four-year outcome and predictors. *American Journal On Mental Retardation, 110*, 417–438.

Shea, V. (2004). A perspective on the research literature related to early intensive behavioural intervention (Lovaas) for young children with autism. *Autism, 8*, 349–367.

Sheinkopf, S., & Siegel, B. (1998). Home-based behavioral treatment for young autistic children. *Journal of Autism and Developmental Disorders, 28*, 15–23.

Smith, T. (1999). Outcome of early intervention for children with autism. *Clinical Psychology: Science and Practice, 6*, 33–49.

Smith, T., Buch, G. A., & Gamby, T. E. (2000). Parent-directed, intensive early intervention for children with pervasive developmental disorder. *Research in Developmental Disabilities, 21*, 297–309.

Smith, T., Eikeseth, S., Klevstrand, M., & Lovaas, O. I. (1997). Intensive behavioral treatment for preschoolers with severe mental retardation and pervasive developmental disorder. *American Journal of Mental Retardation., 102*, 238–249.

Smith, T., Groen, A. D., & Wynn, J. W. (2000). Randomized trial of intensive early intervention for children with pervasive developmental disorder. *American Journal on Mental Retardation, 105*, 269–285.

Spreckley, M., & Boyd, R. (2009). Efficacy of applied behavioral intervention in preschool children with autism for improving cognitive, language and adaptive behavior: A systematic review and meta-analysis. *Journal of Pediatrics, 154*, 338–344.

Torgerson, C. J. (2006). Publication bias: The Achilles' heel of systematic reviews? *British Journal of Educational Issues, 54*, 89–102.

Weiss, M. J. (1999). Differential rates of skill acquisition and outcomes of early intensive behavioural intervention for autism. *Behavioral Interventions, 14*, 3–22.

Wilgosh, L., Waggoner, K., & Adams, B. (1988). Parent views on education and daily living concerns for children with mental handicaps. *Australian and New Zealand Journal of Developmental Disabilities, 14*, 225–259.

World Health Organisation (WHO). (1992). *The ICD-10 classification of mental and behavioural disorders – clinical descriptions and diagnostic guidelines*. Geneva: World Health Organisation.

Yamamoto, T. & Nakano, Y. (2002). *Early intensive behavioral intervention with children with autism: pilot study in Japan* (3rd Focus on Behavior Analysis in Education Conference. Ohio State University). Retrieved May 2006 from http://www.coe.ohio-state.edu/paes/bained02/documents/11-Yamamoto.pdf.

Zachor, D. A., BenItzchak, E., Rabinovich, A., & Lahat, E. (2007). Change in autism core symptoms with intervention. *Research in Autism Spectrum Disorders, 1*, 304–317.

Clinical Psychology Review 30 (2010) 387–399

Contents lists available at ScienceDirect

Clinical Psychology Review

Applied behavior analytic intervention for autism in early childhood: Meta-analysis, meta-regression and dose–response meta-analysis of multiple outcomes

Javier Virués-Ortega *

Instituto de Salud Carlos III, CIBERNED, Madrid, Spain
ABA España, Madrid, Spain

ARTICLE INFO

Article history:
Received 13 May 2009
Received in revised form 30 December 2009
Accepted 29 January 2010

Keywords:
Autism spectrum disorders
Applied behavior analysis
Language
Meta-analysis

ABSTRACT

A number of clinical trials and single-subject studies have been published measuring the effectiveness of long-term, comprehensive applied behavior analytic (ABA) intervention for young children with autism. However, the overall appreciation of this literature through standardized measures has been hampered by the varying methods, designs, treatment features and quality standards of published studies. In an attempt to fill this gap in the literature, state-of-the-art meta-analytical methods were implemented, including quality assessment, sensitivity analysis, meta-regression, dose–response meta-analysis and meta-analysis of studies of different metrics. Results suggested that long-term, comprehensive ABA intervention leads to (positive) medium to large effects in terms of intellectual functioning, language development, acquisition of daily living skills and social functioning in children with autism. Although favorable effects were apparent across all outcomes, language-related outcomes (IQ, receptive and expressive language, communication) were superior to non-verbal IQ, social functioning and daily living skills, with effect sizes approaching 1.5 for receptive and expressive language and communication skills. Dose-dependant effect sizes were apparent by levels of total treatment hours for language and adaptation composite scores. Methodological issues relating ABA clinical trials for autism are discussed.

Contents

1. Introduction

Applied behavior analysis is a behavioral science devoted to the experimental study of socially significant behavior as a function of environmental variables. Throughout the last four decades a number of procedures aimed at enhancing, reducing and maintaining significant human behaviors have been developed by applied behavior analysts (Cooper, Heron, & Heward, 2007a). This research has had a significant impact in the fields of severe problem behavior, developmental disabilities, organizational behavior, behavioral pharmacology, behavioral economics and others. The field of applied behavior analysis has shown a more significant growth in the area of behavioral intervention

---

* Centro Nacional de Epidemiología, Instituto de Salud Carlos III, Sinesio Delgado 6, 28029 Madrid, Spain. Tel.: +34 134 91 822 2651; fax: +34 134 91 387 7815.
E-mail address: jvortega@cienciaconducta.com.

0272-7358/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.cpr.2010.01.008

for children with autism and autism spectrum disorders as suggested by the increasing number of service providers and certified professionals in this field (Cooper, Heron, & Heward, 2007b; Shook, 2005). Since the mid-80s (Fenske, Zalenski, Krantz, & McClannahan, 1985; Lovaas, 1987) the evidence suggesting that applied behavior-analytic intervention (hereafter referred to as ABA intervention) is beneficial to the intellectual, verbal, and social functioning of children with autism and autism spectrum disorders has accumulated steadily (Foxx, 2008; Remington et al., 2007).

Although there are several models of ABA intervention in autism and developmental disabilities, all bonafide programs should share a common set of core features: (1) treatment may begin as early as 3 to 4 years of age, (2) intervention is intensive (20–40 weekly hours) and in addition, incidental teaching and practice goals may be operating during most waking hours, (3) intervention is individualized and comprehensive targeting a wide range of skills, (4) multiple behavior analytic procedures are used to develop adaptive repertoires, (5) treatment is delivered in one-to-one format with gradual transition to group activities and natural contexts, (6) treatment goals are guided by normal developmental sequences, and (8) parents are, to different extents, trained and become active co-therapists (Maurice, Green, & Foxx, 2001).

Positive results have been reported for daily living skills, academic performance and communication skills (Eikeseth, Smith, Jahr, & Eldevik, 2007; Remington et al., 2007). Studies suggest that with ABA intervention, children have a greater chance of integrating into school without additional specialist support whilst maintaining gains over long follow-up periods (McEachin, Smith, & Lovaas, 1993). These findings have had some effects on the social and health policies of different countries (New York State Department of Health, 1999; Ontario Ministry of Education, 2007; U.S. Public Health Service, 1999). However, dissemination of research findings may still be considered limited. For example, recent reviews on autism do not even acknowledge the very existence of ABA intervention (Hughes, 2008) or misrepresent its application and effects (Volkmar & Davies, 2003).

Although a number of studies have been conducted to explore the effectiveness of ABA intervention in children with developmental disabilities, the collective examination of this literature is hampered by a number of factors: (1) studies implement inconsistent methodological features with regard to research design, sampling methods and quality standards, (2) intervention features are highly variable including treatment intensity, duration, the intervention model itself and format of treatment delivery (e.g. clinic-based vs. parent-managed), (3) participants are highly variable with regard to their pre-intervention functioning and age and, finally (4) studies use of variety of different metrics when reporting outcomes making it difficult to implement standard meta-analytical procedures (Morris & DeShon, 2002). Furthermore, the fact that most literature in this area has been single-subject design research and that studies are often procedure-specific (in terms of approaches to treatment) has prevented wider dissemination of results through standard methods of clinical science. Although attempts have been made to summarize single-subject research, these methods are still controversial (Scruggs & Mastropieri, 1998; Severtson, Carr, & Lepper, 2009).

A precise quantification of ABA intervention effectiveness is not currently available. Previous reviews have focused on very specific aspects of ABA intervention (Delprato, 2001), or have failed to incorporate advanced meta-analytical procedures including quality assessment, meta-regression, dose–response meta-analysis, and meta-analysis of studies of different metrics (Eldevik, Hastings, Hughes, Jahr, & Eikeseth, 2009). The present study has the following goals: (1) ascertain the collective effectiveness of ABA intervention for autism, (2) estimate ABA intervention effectiveness in terms of as many outcome variables as possible in order to provide a comprehensive assessment of its effects, and (3), analyze the effect of study characteristics including intervention duration and intensity, study design, intervention model and intervention delivery format. This study pursues a comprehensive account of the effects of comprehensive, intensive and long-term ABA intervention over subjects' functioning in molar skills domains, therefore, studies targeting specific behaviors or procedures will be discarded.

2. Methods

2.1. Literature search and study selection

MEDLINE, PsycINFO, and the Cochrane Clinical Trials databases were searched for all studies reporting the effect of intensive, long-term ABA intervention with children with autism and pervasive developmental disabilities not otherwise specified. Although ABA intervention focuses on specific skills and behaviors at a time, as we examined the molar effects of long-term, comprehensive ABA intervention, no specific behavior or behavior procedure could be contemplated as an inclusion criterion in the assumption that they were many throughout the treatment process. Formal search strategies for randomized controlled trials were supplemented with less restrictive search strategies in order to enhance the detection of low impact journals and mid-to-low quality studies (Botella & Gambara, 2006; Robinson & Dickersin, 2002) (see search strategy in Appendix A). The search period was January 1985 through April 2009, with no language restrictions. The reference lists of selected review articles were also reviewed (British Columbia Office of Health Technology Assessment, 2001; Eldevik et al., 2009).

A number of pre-specified exclusion criteria were used to identify key studies. The 11 exclusion criteria were: (1) the study was non peer-reviewed, non-original, non-empirical, methodological or unpublished; (2) none of the intervention groups implemented ABA intervention for autism according to major features of comprehensive behavior-analytic intervention for autism (Maurice et al., 2001); (3) the focus of the intervention was for very specific areas (e.g., joint attention, problem behavior) or was restricted to a specific behavioral procedure (e.g., functional communication treatment, non-contingent reinforcement); (4) intervention did not meet the intensity and duration standards of ABA interventions (at least 10 weekly hours and no less than 45 weeks duration); (5) participants did not have a formal diagnosis of autism according to the Autism Diagnostic Interview-Revised (Lord, Rutter, & Le Couteur, 1994), the Autism Diagnostic Observation Schedule (Lord, Rutter, DiLavore, & Risi, 1999), the Diagnostic and Statistical Manual of Mental Disorders criteria for autism (American Psychiatric Association, 2000) or a combination of any of these methods; (6) the study utilized a single-subject study design or had an intervention group with less than five subjects; (7) the study was epidemiological; (8) the study reported anecdotal, qualitative or non-standardized outcome measures; (9) there was no pre-test measurement; (10) the study purposely biased subject selection (e.g., fast learners), and (11) mean and standard deviations were not available after attempts to contact authors and could not be calculated from descriptive data or statistical tests in the study manuscript. Exclusion criteria were implemented successively. Although a minimum of 10 weekly hours may be considered too low, this criterion may enable the more precise determination of the impact of intervention intensity on treatment effectiveness. Outcomes reported in less than three clinical trials were discarded from the meta-analysis. The selection process is summarized in the flow chart in Fig. 1.

2.2. Assessment of studies and data extraction

Two investigators (JV-O, MR-M) independently screened the titles and abstracts of the database searches and retrieved articles to determine eligibility (by virtue of the exclusion criteria) before extracting study data. Interrater agreement in the final number of

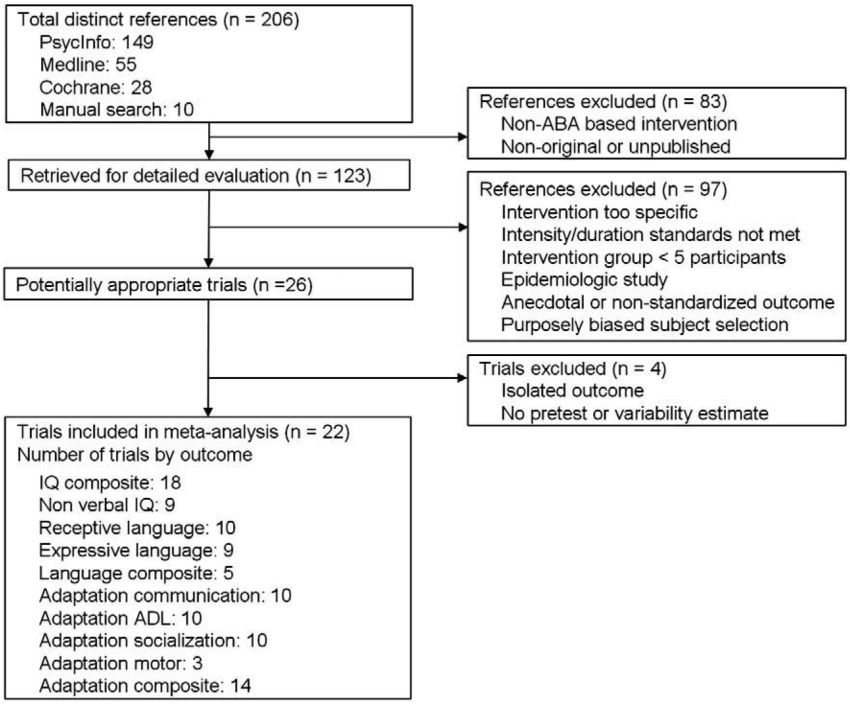


Fig. 1. Flow chart of trial selection process.

trials to be included in the meta-analysis reached 90.9%. Discrepancies were resolved by consensus. The authors of the original studies were contacted if relevant data were not available in the published reports. The service of an assistant translator, for studies published in languages other than English and Spanish, was used when necessary.

The following data were retrieved from all selected studies: (1) participant characteristic including mean pre-intervention age in months, percentage of male participants, pre-intervention IQ, (2) intervention characteristics including intervention intensity (weekly hours); intervention duration (weeks); total intervention duration (intensity multiplied by duration); intervention delivery format, whether clinic-based or parent-managed programs delivered at home and supervised by professionals (i.e., clinic-based vs. parent-managed programs), model of ABA intervention (UCLA model [Lovaas, 1981] vs. general applied behavior analytic model [e.g., Cooper et al., 2007a; Maurice, Green, & Luce, 1996]); study design (randomized controlled trial, non-randomized controlled trial, repeated measures study); sample size; outcome variables; assessment instruments; reported pre- and post-test outcome values (mean and standard deviation); and study quality. Two trained investigators (JV-O, MR-M) assessed the quality of the studies independently by means of the Downs and Black checklist for randomized and non-randomized studies of health care interventions (Downs & Black, 1998). Cohen's kappa for studies' total score reached 0.95. Discrepancies were resolved by consensus. Quality domains covered by the checklist are: Reporting, External validity, Internal validity-bias, Internal validity-confounding and Power. Domains were rated on a 0-1 scale in order to provide a 5-point total quality range and to avoid over-representation of scale domains holding more items (e.g., Reporting). As suggested by the original authors, the checklist was adapted specifically for the search topic by adding a list of confounders, adverse effects, and ranges for power assessment (see the scale and quality assessment in Appendix A). The quality checklist was selected because it was flexible enough to be applicable to both repeated measures and control group studies, whether randomized or not. Assessors' disagreements on these quality measures were resolved by consensus.

2.3. Statistical analysis

Because the instruments for evaluating a given outcome differed across studies (e.g., Wechsler Intelligence Scale for Children vs. Merrill-Palmer Scales of Mental Tests), we used effect sizes to obtain standardized measurements of the effect of the intervention on the outcome variable. For studies with a control group, effect sizes were calculated as the difference in outcome progression (that is, post- minus pre-test mean scores) between the intervention and control groups, divided by the pre-test standard deviation pooled across groups. For these studies, the intervention group comprised all participants receiving ABA intervention and the control group comprised all participants not receiving ABA intervention, irrespective of the concurrent use of other treatments and the alternative intervention assigned to the control group. For within-subjects designs, effect sizes were computed by dividing the mean difference between post- and pre-test outcomes by the pre-test standard deviation. Assuming that outcome changes at follow-up are the effect of treatment, effect size estimates from within-subjects studies are equivalent and comparable to those from controlled studies (Morris & DeShon, 2002), and they can be interpreted as the effect of the intervention on the outcome measured in pre-test within-group standard deviation units. Sensitivity analyses restricting meta-analysis to controlled studies were conducted in order to test this assumption. Once effect sizes were obtained from means and standard deviations results were combined across studies. The above effect size estimates were corrected for small-sample bias, and design-specific estimates of their sampling variance were computed (Becker, 1988; Morris, 2008). If not explicitly reported, outcome means and standard deviations were calculated from the available descriptive data or test statistics using standard methods (Morris & DeShon, 2002). Since the correlation between pre- and post-test outcomes is required to compute the effect size variance, a pooled correlation coefficient was estimated from studies in which sufficient data were available to calculate pre-post correlation coefficients for a given outcome (Morris & DeShon, 2002). The pooled estimate was then applied to all studies reporting the outcome. Interim measures were always discarded, selecting the pre-test and post-test measures closest to the beginning

Table 1
Studies reporting the effects of comprehensive and intensive applied behavior analytic intervention for autism.

First author, year	Country	Diagnosis	Male (%)	Mean age (months)	Pre-IQ	Control group	Sample size*	ABA Intervention			Outcome (instrument)	Quality score†
								Model	Intensity (h/week)	Duration (weeks)		
Clinic-based intervention programs												
Ben-Itzchak and Zachor (2007)	Israel	Autism	92.00	26.60	70.67	No	25	General	35.00	53.00	IQ composite (BSID-II, SBIS) Receptive (BO) Expressive (BO)	1.77
Ben-Itzchak et al. (2008)	Israel	Autism	97.67	27.29	74.84	Yes	39/37	General	45.00	53.00	IQ composite (BSID-II, SBIS)	3.04
Birnbrauer and Leach (1993)	Australia	Autism	55.56	39.00	51.28	Yes ‡	9/5	UCLA	18.72	105.12	IQ composite (BSID-II, SBIS, LEITER, PPVT) Language composite (RDLS, REEL)	2.30
Cohen et al. (2006)	United States	Autism, PDD NOS	83.33	31.70	60.50	Yes, R	21 /21	UCLA	37.50	141.00	IQ composite (BSID-II, WPPSI-R, WISC-III) Non-verbal IQ (MPSMT) Receptive (RDLS) Expressive (RDLS) Adaptation-C (VABS) Adaptation-DLS (VABS) Adaptation-S (VABS) Adaptation composite (VABS)	2.90
Eikeseth et al. (2002, 2007)	Norway	Autism	80.00	66.31	65.68	Yes, R	13/12	UCLA	23.50	148.10	IQ composite (BSID-II, WPPSI-R, WISC-III) Non-verbal IQ (MPSMT) Receptive (RDLS) Expressive (RDLS) Language composite (RDLS) Adaptation-C (VABS) Adaptation-DLS (VABS) Adaptation-S (VABS) Adaptation composite (VABS)	2.97
Eldevik et al. (2006)	Norway	Autism	85.71	50.86	44.32	Yes	13/15	General	12.00	88.91	IQ composite (BSID-II, SBIS, WPPSI-R, WISC-III) Non-verbal IQ (MPSMT) Receptive (RDLS) Expressive (RDLS) Adaptation-C (VABS) Adaptation-DLS(VABS) Adaptation-S (VABS) Adaptation composite (VABS)	2.26
Harris et al. (1991)	United States	Autism	88.24	47.40	65.56	Yes	16/12	UCLA	–	49.14	IQ composite (SBIS-IV) Language composite (PLS)	2.03
Harris and Handleman (2000)	United States	Autism	85.19	49.00	59.33	No	27	General	40.00	407.34	IQ composite (SBIS-IV)	3.25
Howard et al. (2005)	United States	Autism, PDD NOS	84.44	33.20	56.69	Yes	26/16	General	32.50	62.24	IQ composite (BSID-II, WPPSI, WISC, DP-II, SBIS) Non-verbal IQ (MPSMT) Receptive (RDLS) Expressive (RDLS) Adaptation-C (VABS) Adaptation-DLS (VABS) Adaptation-S (VABS) Adaptation-motor (VABS) Adaptation composite (VABS)	2.86
Lovaas (1987)	United States	Autism	84.21	34.60	54.34	No ‡	19	UCLA	40.00	106.00	IQ composite (WPPSI, WISC, SBIS, CIIS, BSID-II, MPSMT, LEITER)	2.38
Magiati et al. (2007)	United Kingdom	Autism	88.64	39.64	76.53	Yes	28/16	UCLA	32.40	109.50	Non-verbal IQ (MPSMT) Receptive (BPLS-II) Expressive (BPLS-II) Adaptation-C (VABS) Adaptation-DLS (VABS) Adaptation-S (VABS) Adaptation composite (VABS)	2.80

Matos and Mustaca, 2005	Argentina	Autism, PDD NOS	88.89	42.00	15.00	No ‡	9	General	40.00	48.18	IQ composite (BSID-II) Receptive (PPVT) Adaptation-C (VABS) Adaptation-DLS (VABS) Adaptation-S (VABS) Adaptation-motor (VABS) Adaptation composite (VABS)	1.17
Remington et al. (2007)	United Kingdom	Autism	–	37.10	61.90	Yes	23/21	General	25.60	105.12	IQ composite (BSID-II, SBIS) Adaptation-C (VABS) Adaptation-DLS (VABS) Adaptation-S (VABS) Adaptation-motor (VABS) Adaptation composite (VABS)	2.67
Sallows and Graupner (2005)	United States	Autism	84.62	33.23	50.85	No ‡	13/10	UCLA	37.58	211.25	IQ composite (BSID-II, WPPSI, WISC,) Non-verbal IQ (MPSMT, LEITER) Receptive(RDLS, CELF-III) Expressive (RDLS, CELF-III) Adaptation-C (VABS) Adaptation-DLS(VABS) Adaptation-S (VABS) Adaptation composite (VABS)	3.63
Smith et al. (1997)	United States	Autism	90.48	36.95	27.57	Yes, R	11/10	UCLA	30.00	53.00	IQ composite (BSID-II)	2.90
Smith et al. (2000)	United States	Autism, PDD NOS	82.14	35.93	50.87	Yes, R	15/13	UCLA	24.52	250.67	IQ composite (BSID-II, SBIS) Non-verbal IQ (MPSMT) Receptive(RDLS) Expressive (RDLS) Language composite (RDLS) Adaptation-C (VABS) Adaptation-DLS(VABS) Adaptation-S (VABS) Adaptation composite (VABS)	2.90
Weiss (1999)	United States	Autism, PDD NOS	95.00	41.50	–	No	20	General	40.00	106.00	Adaptation composite (VABS)	2.02
Parent-managed intervention programs												
Anan et al. (2008) §	United States	Autism, PDD NOS	84.70	44.00	–	No	72	General	20.00	12.00	Receptive (MSEL) Expressive (MSEL) Adaptation-C (VABS) Adaptation-DLS (VABS) Adaptation-S (VABS) Adaptation-M (VABS) Adaptation composite (VABS)	3.55
Anderson et al. (1987)	United States	Autism, PDD NOS	76.92	43.00	57.83	No	13	General	20.00	53.00	IQ composite (BSID-II, SBIS) Language (PLS, SPT, PPVT, SICD) Adaptation composite (VABS)	2.53
Baker-Ericzen et al. (2007) §	United States	Autism, PDD NOS	83.00	49.36	–	No	158	Pivotal training	–	12.00	Adaptation-C (VABS) Adaptation-DLS (VABS) Adaptation-S (VABS) Adaptation composite (VABS)	3.88
Bibby et al. (2001)	United Kingdom	Autism	83.33	43.40	50.80	No	22	UCLA	5.85	31.60	IQ composite (WPPSI-R, WISC-III) Adaptation composite (VABS)	3.11
Reed et al. (2007) §	United Kingdom	Autism	100.00	41.89	53.40	No ‡	14	General	12.20	41.61	IQ composite (PEP-R) Adaptation composite (VABS)	2.81
Reed et al. (2007) §	United Kingdom	Autism	100.00	41.89	53.40	No ‡	13	General	27.00	41.61	IQ composite (PEP-R) Adaptation composite (VABS)	2.81
Sallows and Graupner (2005)	United States	Autism	80.00	34.20	52.10	No ‡	10	UCLA	31.28	198.85	IQ composite (BSID-II, WPPSI, WISC,) Non-verbal IQ (MPSMT, LEITER) Receptive (RDLS, CELF-III) Expressive (RDLS, CELF-III) Adaptation-C (VABS)	3.63

(continued on next page)

Table 1 (*continued*)

First author, year	Country	Diagnosis	Male (%)	Mean age (months)	Pre-IQ	Control group	Sample size*	ABA Intervention			Outcome (instrument)	Quality score†
								Model	Intensity (h/week)	Duration (weeks)		
Sheinkopf and Siegel (1998)	United States	Autism, PDD NOS	–	34.66	62.00	Yes, R	9/10	UCLA	19.45	68.90	Adaptation-DLS (VABS) Adaptation-S (VABS) Adaptation composite (VABS) Non-verbal IQ (MPSMT, CIIS, WPPSI, WISC)	2.54

*Total number of subjects for repeated-measures designs, or number of subjects in intervention/control groups for between-group studies.

†Quality score based on Downs and Black (1998) quality scale; rank: 0 (lowest quality) to 5 (highest quality).

‡Control and comparison groups in the studies by Sallows and Graupner (2005) and Reed et al., (2007) were analyzed separately as the control group received more than 10 weekly hours of ABA based intervention. IQ effect size for Birnbrauer and Leach (1993) was computed as a within-subject study as no post-test values are provided for the control group. Lovaas (1987) was analyzed as a within-subject study due to insufficient data reporting for the control group. Matos and Mustaca (2005) did not provide a standardized estimate of pre-intervention IQ.

§Studies not meeting the duration and intensity inclusion criteria but selected for meta-regression analyses.

Adaptation-C, Communication; Adaptation-DLS, Daily living skills; Adaptation-M, motor functioning; Adaptation-S, Socialization; BO, Systematic behavioral observation; BPLS-II, British Picture Language Scale (2nd Ed.); BSID-II, Bayley Scales of Infant Development (2nd Ed.); CELF-III, Clinical Evaluation of Language Fundamentals (3rd Ed.); CIIS, Cattell Infant Intelligence Scale; DP-II, Developmental profile II; LEITER, Leiter International Performance Scale; MPSMT, Merrill-Palmer Scales of Mental Tests; MSEL, Mullen Scales of Early Learning; PEP-R, Psycho-educational Profile (revised); PLS, Preschool Language Scale; PPVT, Peabody Picture Vocabulary Test; R, Randomized assignment; RDLS, Reynell Developmental Language Scales; REEL, Receptive-Expressive Emergence Language Scale; SBIS, Stanford-Binet Intelligence Scales; SBIS-IV, Stanford-Binet Intelligence Scales (4th Ed.); SICD, Sequenced Inventory of Communication Development; SPT, Symbolic Play Test; VABS, Vineland Adaptive Behavior Scales; WISC-III, Wechsler Intelligence Scale for Children (3rd Ed.); WPPSI-R, Wechsler Preschool and Primary Scale of Intelligence-Revised.

and end of the entire treatment period, even when the last follow-up outcome measure was reported in a separate paper.

For each outcome of interest, pooled estimates and 95% confidence intervals (CI) of effect sizes were calculated by using an inverse–variance weighted random-effects meta-analysis (Cottrell, Drew, Gibson, Holroyd, & O'Donnell, 2007). Between-study outcome variation (i.e., heterogeneity) was quantified with the I^2 statistic, which describes the percentage of variation across studies due to heterogeneity rather than chance regardless of treatment effect metric (Higgins & Thompson, 2002). Values around 25%, 50% and 75% refer to low, medium and high heterogeneity, respectively. Although I^2 was developed to be independent of the number of studies, it should be interpreted cautiously when few studies are meta-analyzed (Huedo-Medina, Sanchez-Meca, Marin-Martinez, & Botella, 2006).

When two or more studies were available, sensitivity analyses were performed by restricting the analysis to control group designs. In addition, separate meta-analyses were conducted by intervention model (UCLA, general ABA) and delivery format (clinic-based, parent-managed) to check consistency of treatment effects. At least two studies needed to be available for a sensitivity analysis to be conducted. For brevity, only effect size differences of 0.50 or above across ABA intervention models and intervention delivery format will be reported. Random-effects meta-regression (Thompson & Sharp, 1999) was used to separately evaluate whether results were different by population and intervention features, such as pre-intervention age, pre-intervention IQ, and treatment duration and intensity. For the purpose of analyzing the effects of intervention duration and intensity more thoroughly, studies were rank-ordered by total intervention hours (duration multiplied by intensity). A dose–response meta-analysis was conducted by studies' levels of total intervention hours. In order to strengthen the power of the analysis, studies excluded solely on the basis of limited treatment duration were included in the meta-regression and dose–response meta-analyses. Finally, publication and small-study effects biases were assessed using the extended Egger's test (Egger, Smith, Schneider, & Minder, 1997; Thompson & Sharp, 1999). Statistical analyses were carried out with Stata v. 8.1 (Stata Corporation, College Station, Texas).

3. Results

3.1. Study characteristics

Twenty-six studies met the pre-specified inclusion criteria (Fig. 1). Two studies were excluded because the relevant outcome was present in less than three papers (Boyd & Corley, 2001; Zachor, Ben Itzchak, Rabinovich, & Lahat, 2007). Two studies were excluded because of limited data reporting, including failure to provide pre-test measures and estimates of random variability (Luiselli, Cannon, Ellis, & Sisson, 2000; McEachin et al., 1993). The remaining 22 studies were included in the meta-analysis (Anderson, Avery, DiPietro, Edwards, & Christian, 1987; Ben, Lahat, Burgin, & Zachor, 2008; Ben-Itzchak & Zachor, 2007; Bibby, Eikeseth, Martin, Mudford, & Reeves, 2001; Birnbrauer & Leach, 1993; Cohen, Merine-Dickens, & Smith, 2006; Eikeseth, Smith, Jahr, & Eldevik, 2002; Eikeseth et al., 2007; Eldevik, Eikeseth, Jahr, & Smith, 2006; Harris & Handleman, 2000; Harris, Handleman, Gordon, Kristoff, & Fuentes, 1991; Howard, Sparkman, Cohen, Green, & Stanislaw, 2005; Lovaas, 1987; Magiati, Charman, & Howlin, 2007; Matos & Mustaca, 2005; Remington et al., 2007; Sallows & Graupner, 2005; Sheinkopf & Siegel, 1998; Smith, Eikeseth, Klevstrand, & Lovaas, 1997; Smith, Groen, & Wynn, 2000; Weiss, 1999). Three additional studies excluded solely on the basis of insufficient intervention duration were included in meta-regression and dose–response analyses (Anan, Warner, McGillivary, Chong, & Hines, 2008; Baker-Ericzen, Stahmer, & Burns, 2007; Reed,

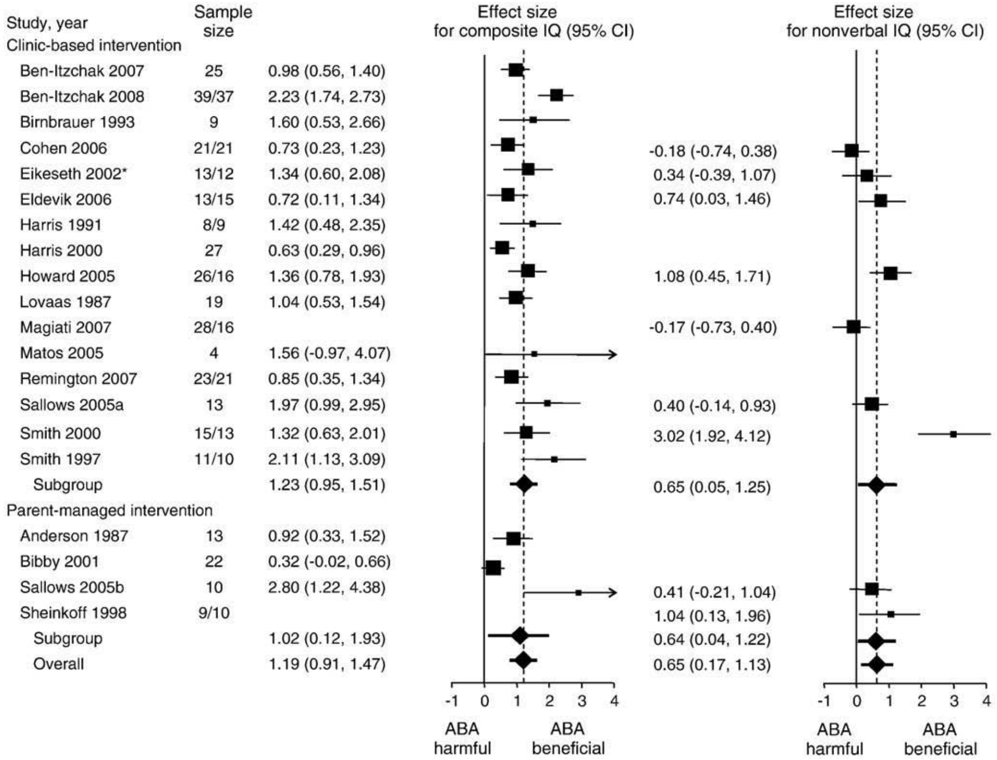


Fig. 2. Effect size for IQ and nonverbal IQ of applied behavior analysis intervention for participants with autism and pervasive developmental disabilities not otherwise specified. The area of each square is proportional to the study weight in the pooled analysis. Horizontal lines represent 95% confidence intervals (CI). Diamonds represent pooled estimates from inverse–variance weighted random-effects meta-analyses. Effect sizes and 95% CI are also presented numerically. Studies are classified by intervention delivery format (clinic-based, parent-managed). Sample sizes are total number of subjects for repeated-measures designs, or number of subjects in intervention/control groups for control group designs. Eikeseth et al., (2002) and Eikeseth et al., (2007) report data from the same cohort at different follow up periods; a single effect size was computed with the last follow up as post-intervention measure. Sallows and Graupner (2005) are reported as two independent repeated measures studies. Given that intervention and comparison groups at Sallows and Graupner (2005) received more than 10 weekly hours of intervention, each group was analyzed as an independent within subject study.

Table 2
Pooled effect sizes for IQ, language and adaptive behavior according to intervention features in studies of ABA intervention.[a]

	IQ composite			Language composite			Adaptive behavior composite		
Intervention feature	No. studies	Effect size (95% CI)	p value[b]	No. studies	Effect size (95% CI)	p value†	No. studies	Effect size (95% CI)	p value[b]
Pre-intervention age	20	−0.02 (−0.05 to 0.01)	0.157	5	−0.02 (−0.10 to 0.06)	0.621	18	0.01 (−0.04 to 0.06)	0.670
Pre-intervention IQ	19	0.00 (−0.03 to 0.03)	0.874	4	0.06 (−0.19 to 0.06)	0.317	14	0.03 (−0.02 to 0.07)	0.234
Duration, weeks	20	0.00 (−0.03 to 0.02)	0.500	5	0.01 (0.00 to 0.02)	0.001	18	0.00 (0.00 to 0.01)	0.346
Intensity, hours/week	19	0.01 (−0.01 to 0.04)	0.333	4	0.05 (−0.19 to 0.29)	0.705	17	0.05 (0.01 to 0.09)	0.015

CI, confidence interval.

[a] Pooled effect sizes were estimated from random-effects meta-regression models including indicator variables for each category of the intervention feature. Matos study was not included because only BSID-II raw scores were provided and in the absence of exact birth date, PDI scores under an IQ-equivalent scale could not be obtained.

[b] p value for heterogeneity of pooled effect sizes.

Osborne, & Corness, 2007). Two studies using control groups receiving more than 10 weekly hours of ABA intervention, were analyzed as independent repeated measures studies and will be referred to as separate studies (e.g., Reed et al., 2007; Sallows & Graupner, 2005). Two studies reporting data from the same cohort at different follow up periods (Eikeseth et al., 2002; 2007) were analyzed as a single study. In addition, Lovaas (1987) was analyzed as a within-subject study due to insufficient data reporting for the control group. The reader is referred to Table 1 for a systematic description of studies included in the meta-analysis. A summarization of study features is presented below.

The following outcomes were reported: full scale IQ (18 studies), nonverbal IQ (9 studies), receptive language (10 studies), expressive language (9 studies), language composite (5 studies), adaptive behavior–communication (10 studies), adaptive behavior–daily living skills (10 studies), adaptive behavior–socialization (10 studies), adaptive behavior–motor skills (3 studies), and overall composite adaptive behavior scores (14 studies). A complete listing of the instruments used to assess each of these outcomes is available in Table 1.

The mean quality score (of a possible maximum of 5) was 2.5 (range of 1.2 to 3.6). Studies tended to score higher in Reporting (0.8 out of 1.0) and Internal Validity-bias (0.7) as opposed to External Validity (0.4), Internal Validity-confounding (0.3) and Power (0.3). Quality scores by intervention model equaled 2.8 (range of 2.0 to 3.6) for UCLA model and 2.4 (range of 1.2 to 3.3) for general ABA intervention. Parent-managed programs obtained an average quality index of 3.0 (range of 2.5 to 3.6) while clinic-based programs scored 2.6 (range of 2.0 to 3.6). The reader is referred to Appendix A for the complete report of quality assessment.

A total of 323 subjects were included in intervention groups. The participants mean age ranged from 22.6 to 66.3 months. The percentage of male participants ranged from 55.6 to 97.7%. Fifteen studies reported results exclusively on children diagnosed with autism, while participants in 7 studies were both children diagnosed with autism and pervasive developmental disabilities not otherwise specified. With regard to intervention features, 13 studies followed the UCLA model, and 9 studies used the intervention model described as general ABA. Eighteen studies reported clinic- or school-based programs. Among them, two studies were delivered in the participants' home (Magiati et al., 2007; Weiss, 1999). Four trials reported data from parent-managed programs (Anderson et al., 1987; Bibby et al., 2001; Sallows & Graupner, 2005; Sheinkopf & Siegel, 1998). Intervention duration and intensity ranged from 48 to 407 weeks and from 12 to 45 weekly hours respectively.

There were 8 studies with within-subjects design included in the meta-analysis. Thirteen studies had control groups of which 6 used random or quasi-random assignment. Control groups of 9 studies comprised those having either an eclectic intervention or a combination of standard interventions including Treatment and Education of Autistic Children and related Communication Handicapped Children (TEACCH, see Piazza & Fadanni, 2002), special education classes and sensory integration therapy (Ben et al., 2008; Cohen et al., 2006; Eikeseth et al., 2002, 2007; Eldevik et al., 2006; Howard et al., 2005; Magiati et al., 2007; Sheinkopf & Siegel, 1998; Smith et al., 2000). One study used a public school special education group as control group (Remington et al., 2007) and in another study a group of typically developing children attending regular school participated as controls (Harris et al., 1991). The control group of the study by Smith et al. (1997) was comprised of children with autism receiving low intensity (i.e., <10 weekly hours) ABA intervention. Finally, Birnbrauer and Leach (1993) did not report any specific intervention in their control group.

3.2. Intelligence quotient

ABA intervention produced positive effects in all 18 studies reporting general IQ (Fig. 2). The pooled effect size across studies, covering a total of 311 participants, was 1.19 (95% CI 0.91 to 1.47, $p<0.001$). Effects tended to be stronger for clinic-based programs compared to parent-managed interventions with effect sizes of 1.23

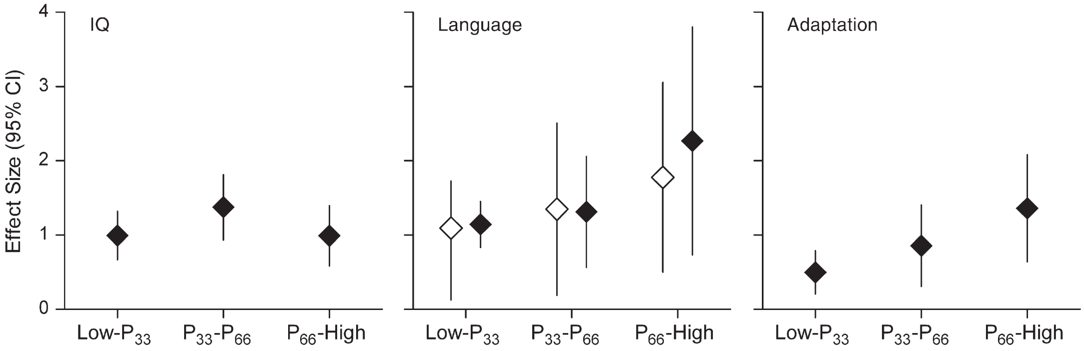


Fig. 3. Dose–response meta-analysis by levels of applied behavior analysis total intervention hours for IQ, language (receptive, expressive) and adaptive behavior (composite scores). Total intervention hours levels were established by percentile 33 ($P_{33}=1833.8$) and 66 ($P_{66}=4129.3$) of total intervention hours of all meta-analyzed studies (treatment intensity multiply by duration). Open diamonds in the central graph show expressive language and solid diamonds show receptive language. Horizontal lines represent 95% confidence intervals (CI). Diamonds represent pooled estimates from inverse–variance weighted random-effects meta-analyses by total intervention hour levels.

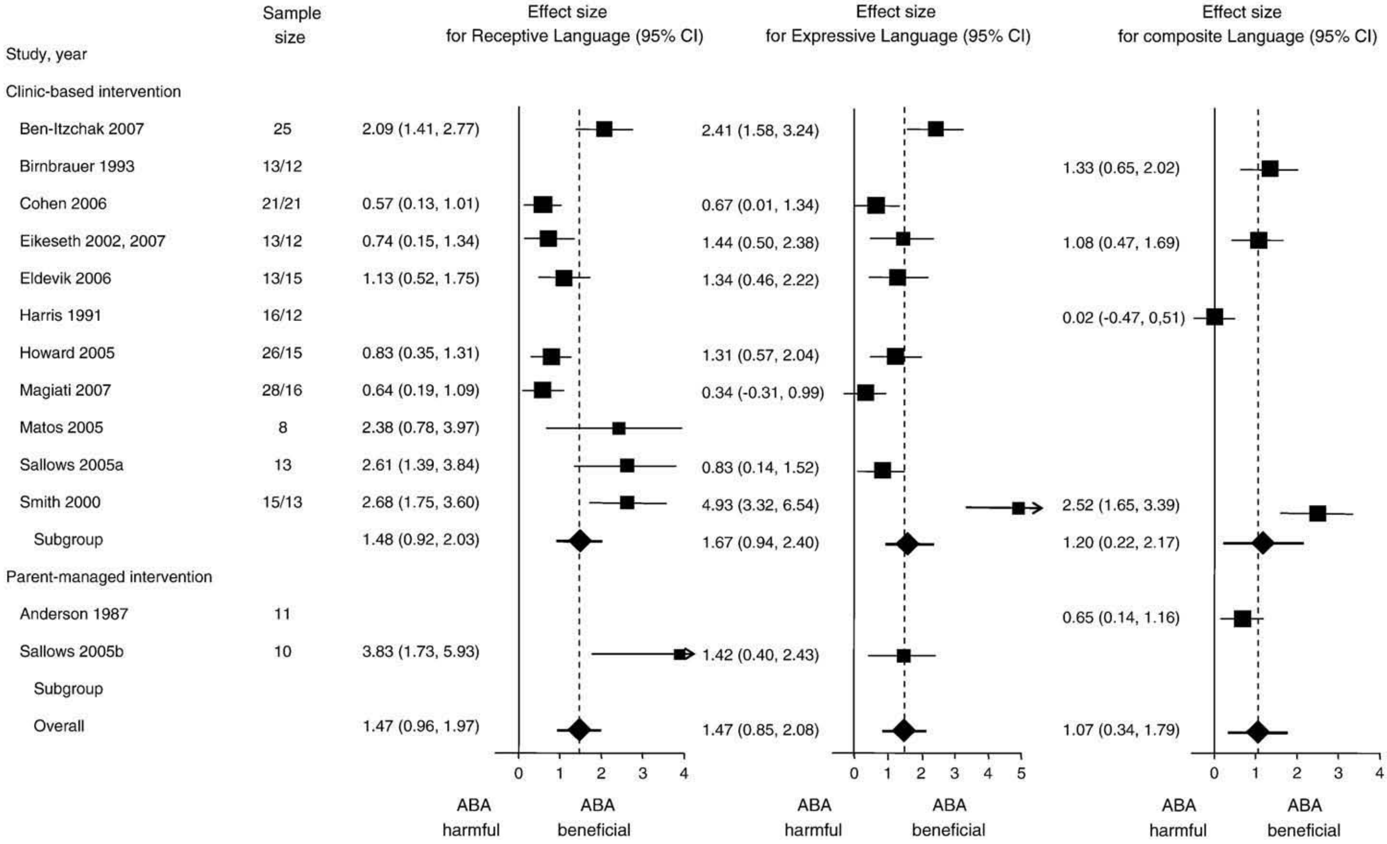


Fig. 4. Effect size for receptive language, expressive language and composite language of applied behavior analysis intervention for participants with autism and pervasive developmental disabilities not otherwise specified. See Fig. 2 notes.

(95% CI 0.95 to 1.51, $p<0.001$) versus 1.02 (95% CI 0.12 to 1.93, $p<0.027$) (see Fig. 2). Similar results were obtained when analysis was restricted to the 10 studies ($n=169$) that included a control group with an effect size of 1.31 (95% CI 0.92 to 1.70, $p<0.001$). Meta-regression did not show clear effects of intensity or duration (Table 2). Dose–response meta-analysis of studies' total treatment duration suggested that high total treatment duration did not improve treatment gains above average levels (Fig. 3). There was evidence of heterogeneity ($I^2=75\%$, 95% CI 60 to 84%) and publication bias ($p=0.012$).

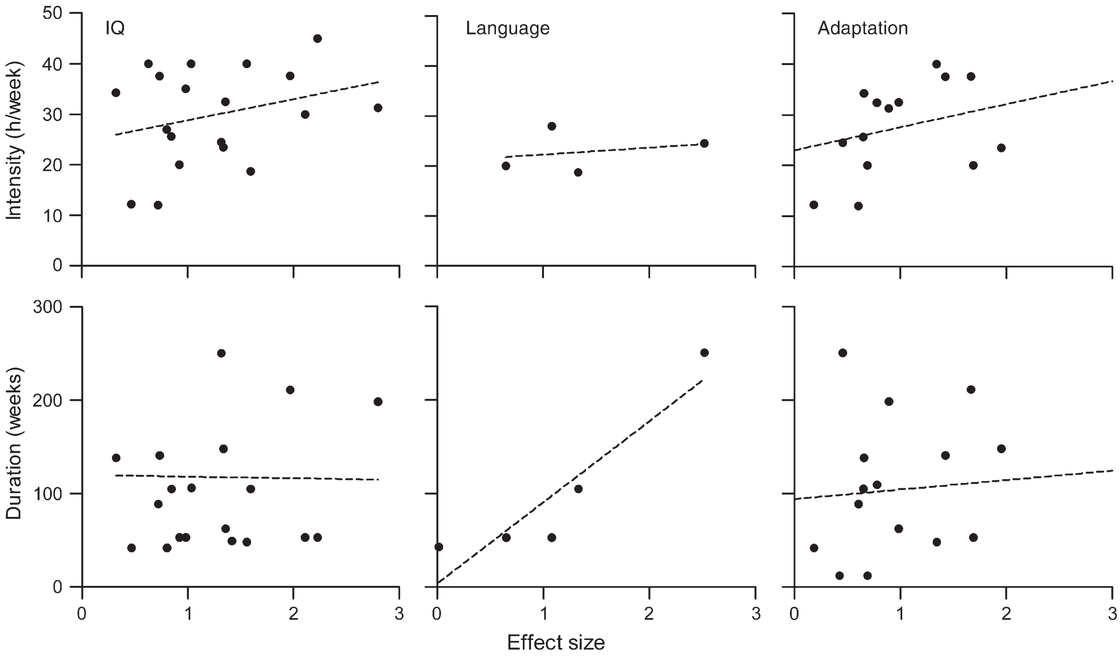


Fig. 5. Scatter plot and linear regression line of intervention intensity and duration by intervention effect size for IQ, language (composite score) and adaptive behavior (composite score) outcomes.

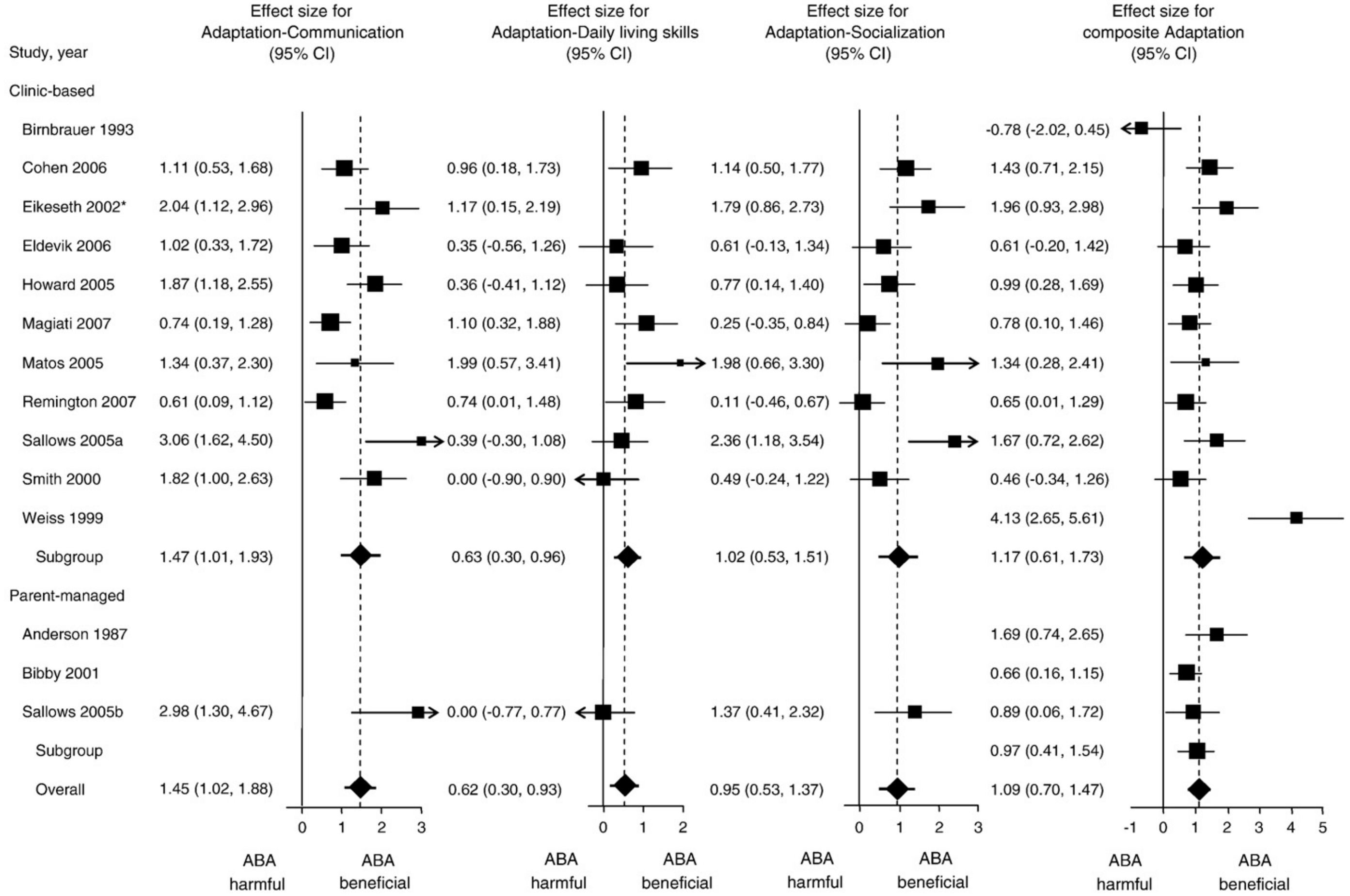


Fig. 6. Effect size for adaptive behavior domains including communication, daily living skills, socialization and adaptive behavior composite scores of applied behavior analysis intervention for participants with autism and pervasive developmental disabilities not otherwise specified. See Fig. 2 notes.

Non-verbal IQ was reported by 10 studies for a total of 146 participants. All but one reported positive effects of ABA intervention. The pooled effect size for this outcome was 0.65 (95% CI 0.17 to 1.13, $p = 0.008$). Effects were similar across intervention models: clinic-based programs showed an effect size of 0.65 (95% CI 0.05 to 1.25, $p = 0.033$) and the parent managed programs effect size was 0.65 (95% CI 0.05 to 1.25, $p = 0.034$). Eight of these studies included a control group covering 123 participants. When meta-analysis was restricted to controlled studies the effect size was 0.76 (95% CI 0.10 to 1.42, $p = 0.024$). There was evidence of heterogeneity ($I^2 = 78\%$, 95% CI 59 to 88%) and publication bias ($p = 0.013$).

3.3. Language skills

Receptive language was assessed in 11 different studies that provided ABA intervention to 172 participants. All of these studies reported favorable effects of the intervention for receptive language performance with a pooled effect size of 1.48 (95% CI 0.96 to 1.97, $p < 0.001$) (Fig. 4). When meta-analysis was restricted to the 7 studies that included a control group ($n = 116$) then an effect size of 0.99 (95% CI 0.56 to 1.42, $p < 0.001$) was obtained. There was evidence of heterogeneity ($I^2 = 81\%$ CI 65 to 89%). Egger's test suggested publication bias ($p = 0.048$).

Expressive language was reported in 10 studies covering 164 participants. All studies reported favorable effects of ABA intervention for expressive language skills as measures by standardized assessments. Pooled random meta-analysis effect size was 1.47 (95% CI 0.85 to 2.08, $p < 0.001$). There was evidence of heterogeneity ($I^2 = 80\%$, CI 62 to 89%) and publication bias was likely ($p = 0.003$). Both results for receptive and expressive language demonstrated clear dose–response trends for intervention total duration (Fig. 3).

General language skills were reported in 5 studies that provided ABA intervention to 64 participants with a pooled effect size of 1.07 (95% CI 0.34 to 1.79, $p = 0.004$). All but one of these studies showed distinctively favorable effects of ABA intervention (Fig. 4). When meta-analysis was restricted to the 4 studies that included a control group, effect size equaled 1.20 (95% CI 0.22 to 2.17, $p = 0.017$). Meta-regression indicated that effect size increased directly with intervention duration (Fig. 5, Table 2). There was evidence of heterogeneity ($I^2 = 86\%$ 95% CI 69 to 94%) and publication bias ($p = 0.009$).

3.4. Adaptive behavior domains

Adaptive behavior was assessed by means of standardized assessments of competence in the domains of communication, daily living skills, motor skills, socialization and an adaptive behavior composite measure. Communication, daily living skills and socialization were assessed in 11 studies ($n = 170$). All studies reported favorable effects of ABA intervention for these three domains (Fig. 6). Communication resulted in a pooled effect size of 1.45 (95% CI 1.02 to 1.88, $p < 0.001$) with all studies reporting favorable effect sizes. Sensitivity analysis restricted to the 8 controlled studies ($n = 138$) provided an effect size of 1.25 (95% CI 0.83 to 1.67, $p < 0.001$). The effect size for communication tended to be higher for the 6 studies implementing the UCLA model intervention (ES = 1.73, 95% CI 1.06 to 2.39, $p < 0.001$) as opposed to the 4 studies implementing general ABA intervention (ES = 1.17, 95%CI 0.59 to 1.76, $p < 0.001$). There was evidence of heterogeneity ($I^2 = 68\%$ CI 38 to 83%) and publication bias ($p = 0.002$). Effect size for daily living skills reached 0.62 (95% CI 0.30 to 0.93, $p<0.001$), whereas meta-analysis restricted to the 8 studies including a control group was 0.68 (95% CI 0.36 to 0.99, $p<0.001$). There was no evidence of heterogeneity ($I^2=27\%$ 95% CI 0 to 65%) or publication bias ($p=0.191$). Socialization produced a pooled effect size of 0.95 (95% CI 0.53 to 1.37, $p<0.001$). Sensitivity analysis restricted to the 8 controlled studies ($n=138$) resulted in an effect size of 0.68 (95% CI 0.29 to 1.06, $p=0.001$). There was strong evidence of heterogeneity ($I^2=66\%$ 95% CI 34 to 83%) and publication bias ($p=0.002$). Motor skills data were reported in three controlled studies for 51 participants. All three studies reported positive effects of ABA intervention when examined individually. The pooled effect size was 0.71 (95% CI 0.19 to 1.22, $p=0.008$). There was no evidence of heterogeneity ($I^2=0\%$ CI 0 to 90%) or publication bias ($p=0.109$). An adaptive behavior composite measure (combining all four domains described above) was reported in 15 distinct papers ($n=232$). Thirteen out of these 15 studies showed positive effects of ABA intervention (Fig. 6). The pooled effect size was 1.09 (95% CI 0.70 to 1.47, $p<0.001$). Results were consistent across program delivery formats; clinic-based programs had an effect size of 1.17 (95% CI 0.70 to 1.47, $p<0.001$) and parent-managed programs had a pooled effect size of 0.97 (95% CI 0.61 to 1.739, $p=0.001$). Meta-analysis limited to the 10 studies that included a control group ($n=165$) produced a somewhat smaller effect size (0.81; 95% CI 0.39 to 1.23, $p<0.001$). Effects increased with intervention intensity while duration did not affect effect size (Table 2). Dose–response meta-analyses demonstrated a clear increase in effect sizes by treatment total duration (Fig. 3). There was a strong heterogeneity effect ($I^2=68\%$ 95% CI 44 to 82%), while publication bias was not apparent ($p=0.091$).

4. Discussion

The overall appreciation of long-term, comprehensive ABA intervention effects for autism through standardized molar skills assessments has been hampered by the varying methods, designs and treatment features of published studies. In an attempt to fill this gap in the literature, state-of-the-art meta-analytical methods were implemented, including quality assessment, sensitivity analyses, meta-regression, dose–response meta-analysis and meta-analysis of studies of different metrics. Results suggest that long-term, comprehensive ABA intervention leads to (positive) medium to large effects in terms of intellectual functioning, language development, and adaptive behavior of individuals with autism. Although favorable effects were apparent across all outcomes, language-related outcomes (IQ, receptive and expressive language, communication) were distinctively superior to non-verbal IQ, social functioning and daily living skills, with effect sizes approaching 1.5 for receptive and expressive language and communication skills. This is particularly noteworthy as qualitative impairments in communication are one of the core features of autism. This finding is also consistent with the amount of time devoted by most ABA curricula to language and communication skills (Maurice et al., 1996). A trend strengthened in recent years through the development of novel ABA intervention procedures for language (Greer & Ross, 2007; Sundberg, 2007).

Pooled effects were consistent when analysis was restricted to controlled studies. In addition, potential confounding sources as pre-intervention age and pre-intervention IQ did not make a difference to treatment effectiveness (Table 2). Effects were also consistent for both clinic-based and parent-managed programs with slightly superior effect sizes found for clinic-based programs in terms of IQ, nonverbal IQ and adaptive behavior composite measures. Meta-analysis of all outcomes other than composite language skills and motor skills showed similar effects of both ABA intervention models (i.e., UCLA, general ABA) for all outcomes with the exception of communication, which showed stronger positive effects for the UCLA based programs. However, as these differential effects were not substantiated by expressive and receptive language outcomes, the significance of this finding remains unclear.

Although the comparison between ABA intervention models may be highly informative, more sophisticated distinctions including operational definitions of each intervention model and the addition of intervention fidelity measures should be employed to help discriminate models. Fidelity measures and standards cannot currently be assumed for studies in this field (McIntyre, Gresham, DiGennaro, & Reed, 2007; Wheeler, Baggett, Fox, & Blevins, 2006). The adoption of these quality standards would help to interpret any model-specific effects to be found in the future in terms of the curricula, programs and behavioral mechanisms that may be distinctive of any particular approach. On the other hand, this will also help professionals and clients to determine what these approaches may have in common.

Meta-regression analysis provided a clear account of the impact of intervention intensity and duration that is not obvious from the simple examination of individual studies. Overall language skills tended to benefit more from intervention duration while functional and psychosocial adaptive behaviors benefited more from intervention intensity. When meta-analysis was replicated for levels of total intervention duration, dose–response effects were evident for language performance and functional and psychosocial adaptive behaviors, while dose–response analysis for intellectual functioning showed, to some extent, an exhaustion of intervention effects. The highest magnitude of dose–response effects were demonstrated for receptive and expressive language. This finding, combined with the strong effects reported for language-related outcomes, suggest that verbal repertoires have a great potential for continuous treatment gains as opposed to other repertoires that may follow an asymptotic profile. Exhaustion of treatment effects by increasing levels of treatment intensity have been suggested before (Reed et al., 2007), however, our results indicate that this pattern may be different for intellectual functioning, verbal skills and functional and psychosocial repertoires.

Inclusion of repeated measures studies provides a preliminary external verification of controlled studies effects, particularly for those studies that did not have no-treatment controls, did not used randomization and reported small sample sizes. Namely, consistency of treatment effects across within-subjects and controlled studies strengthen the plausibility that control groups' composition was not severely biased and did not affect treatment outcomes to a high extent. Similarly, consistency of treatment effects across within-subjects and controlled studies provides also and external indication of within-subjects studies internal validity. Namely, consistency of effects across study designs suggests that within-subjects studies were not severely affected by design-specific threats including trend in baseline and over-reported effect sizes due to smaller variability.

As control groups were generally those receiving eclectic interventions (e.g., special education, sensory integration, TEACCH and others), meta-analysis also provides a preliminary comparison between ABA intervention and other forms of treatment for autism. This is an interesting extension of this study as there are few formal comparisons of ABA intervention effects to other treatment paradigms (Delprato, 2001; Reed et al., 2007). However, this comparison is only tentative; a formal comparison of intervention paradigms would require that two or more intervention groups have comparable treatment intensity and equal treatment fidelity requirements, which was not the case for the meta-analyzed studies herein reported. Nonetheless, the results of this meta-analysis are straightforward in their current form. Therefore, the superiority of ABA intervention suggested by these data shall not be discounted.

Randomization to group assignment was seldom implemented in the studies found, and the use of quasi-random assignment strategies (e.g., assignment to control or experimental group depending upon therapists' availability) raises various ethical and internal validity concerns. General quality standards of clinical studies including randomization, blindness, intention to treat analysis, and the use of prospective (as opposed to retrospective) designs, were inconsistently observed. In addition, quality standards specific to this field, e.g. comparable pre-intervention IQ across groups and treatment fidelity standards, were generally not followed (McIntyre et al., 2007; Wheeler et al., 2006). Although random effects meta-analysis and sensitivity analysis may partially compensate for this deficit, somewhat different results might well be found if studies employ such higher methodological standards. Moreover, publication bias was evident in all outcomes but daily living skills, motor functioning and composite adaptation. However, the limited sample size of most studies suggests that evidence of publication bias may simply be the byproduct of small sample size studies rather than genuine publication bias (Whitehead, 2002).

Recommendations for clinician and researchers planning to do controlled studies in this area include: (1) the observation of clinical trials quality standards including intention to treat analysis and randomization (see CONSORT guidelines for a complete listing of quality standards; Moher, Schulz, Altman, & the CONSORT Group, 2001), (2) use no-treatment controls or match treatment intensity and duration across groups, (3) monitor the degree to which therapist adhere to treatment protocols in the intervention group and also in the comparison group whenever controls follow an alternative treatment, (4) implement specific approaches to treatment in order to provide direct comparisons of different intervention paradigms both within ABA intervention (e.g., UCLA, CABAS, verbal behavior) and between ABA intervention and other forms of treatment (e.g., pharmacological, TEACCH).

A wide adoption of these standards may establish a clearer picture of the highly promising effects of ABA intervention and may constitute the basis for decision-making in public health and social policies relating to autism and developmental disabilities.

Acknowledgements

The author is in debt with Dr. Neil Martin and Dr. Stephen N. Haynes for their useful comments to earlier versions of this manuscript. Mónica Rodríguez-Mori replicated the quality assessment. The author is also grateful to Ming Tang for providing Chinese to English translations during the screening of references.

This study has not been funded nor subject to pre-submission approval by any institution including the ones to which the author is affiliated. The contents expressed in this article do not necessarily represent the position of the institutions to which the author is affiliated.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.pscychresns.2010.02.004.

References[1]

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders 4th Edition, Text Revised.* Arlington VA: American Psychiatric Publishing, Inc.

*Anan, R. M., Warner, L. J., McGillivary, J. E., Chong, I. M., & Hines, S. J. (2008). Group Intensive Family Training (GIFT) for preschoolers with autism spectrum disorders. *Behavioral Interventions*, *23*, 165−180.

*Anderson, S. R., Avery, D. L., DiPietro, E. K., Edwards, G. L., & Christian, W. P. (1987). Intensive home-based early intervention with autistic children. *Education and Treatment of Children*, *10*, 352−366.

*Baker-Ericzen, M. J., Stahmer, A. C., & Burns, A. (2007). Child demographics associated with outcomes in a community-based pivotal response training program. *Journal of Positive Behavior Interventions*, *9*, 52−60.

Becker, B. J. (1988). Synthesizing standardized mean-change measures. *British Journal of Mathematical and Statistical Psychology*, *41*, 257−278.

Ben-Itzchak, I. E., Lahat, E., Burgin, R., & Zachor, A. D. (2008). Cognitive, behavior and intervention outcome in young children with autism. *Research in Developmental Disabilities*, *29*, 447−458.

*Ben-Itzchak, E., & Zachor, D. A. (2007). The effects of intellectual functioning and autism severity on outcome of early behavioral intervention for children with autism. *Research in Developmental Disabilities*, *28*, 287−303.

*Bibby, P., Eikeseth, S., Martin, N. T., Mudford, O. C., & Reeves, D. (2001). Progress and outcomes for children with autism receiving parent-managed intensive interventions. *Research in Developmental Disabilities*, *22*, 425−447.

*Birnbrauer, J. S., & Leach, D. J. (1993). The Murdoch Early Intervention Program after 2 years. *Behaviour Change*, *10*, 63−74.

Botella, J., & Gambara, H. (2006). Doing and reporting meta-analysis. *International Journal of Clinical and Health Psychology*, *6*, 425−440.

*Boyd, R. D., & Corley, M. J. (2001). Outcome survey of early intensive behavioral intervention for young children with autism in a community setting. *Autism*, *5*, 430−441.

British Columbia Office of Health Technology Assessment. (2001). Autism and Lovaas treatment: A Systematic review of effectiveness evidence. *International Journal of Technology Assessment in Health Care*, *17*, 252. Complete reference list retrieved May 13, 2009, from http://www.chspr.ubc.ca/files/publications/2000/bco00-01T_autism3.pdf

*Cohen, H., Merine-Dickens, M., & Smith, T. (2006). Early intensive behavioral treatment: replication of the UCLA model in a community setting. *Development and Behavioral Pediatrics*, *27*, S145−S155.

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis*, 2nd ed. Upper Saddle River, NJ: Pearson.

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). Definition and characteristics of applied behavior analysis. *Applied Behavior Analysis* (pp. 2−23)., 2nd ed. Upper Saddle River, NJ: Pearson.

Cottrell, C., Drew, J., Gibson, J., Holroyd, K., & O'Donnell, F. (2007). Feasibility assessment of telephone-administered behavioral treatment for adolescent migraine. *Headache: The Journal of Head and Face Pain*, *47*, 1293−1302.

Delprato, D. J. (2001). Comparisons of discrete-trial and normalized behavioral intervention for young children with autism. *Journal of Autism and Developmental Disorders*, *31*, 315−325.

Downs, S. H., & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality of both randomised and non-randomised studies of health care interventions. *Journal of Epidemiology and Community Health*, *53*, 377−384.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, *315*, 629−634.

*Eikeseth, S., Smith, T., Jahr, E., & Eldevik, S. (2002). Intensive behavioral treatment at school for 4- to 7-year-old children with autism: A 1-year comparison controlled study. *Behavior Modification*, *26*, 49−68.

*Eikeseth, S., Smith, T., Jahr, E., & Eldevik, S. (2007). Outcome for children with autism who began intensive behavioral treatment between ages 4 and 7: A comparison controlled study. *Behavior Modification*, *31*, 264−278.

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

Eldevik, S., Hastings, R. P., Hughes, J. C., Jahr, E., & Eikeseth, S. (2009). Meta-analysis of early intensive behavioral intervention for children with autism. *Journal of Child and Adolescent Psychology*, *38*, 439−450.

Fenske, E. C., Zalenski, S., Krantz, P. J., & McClannahan, L. E. (1985). Age at intervention and treatment outcome for autistic children in a comprehensive intervention program. *Analysis and Intervention in Developmental Disabilities*, *5*, 49−58.

Foxx, R. M. (2008). Applied behavior analysis treatment of autism: The state of the art. *Child and adolescent psychiatric clinics of North America*, *17*, 821−834.

Greer, R. D., & Ross, D. E. (2007). *Verbal behavior analysis: Inducing and expanding new verbal capabilities in children with language delays.* Upper Saddle River, NJ: Allyn & Bacon.

*Harris, S. L., & Handleman, J. S. (2000). Age and IQ at intake as predictors of placement for young children with autism: a four- to six-year follow-up. *Journal of Autism and Developmental Disorders*, *30*, 137−142.

*Harris, S. L., Handleman, J. S., Gordon, R., Kristoff, B., & Fuentes, F. (1991). Changes in cognitive and language functioning of preschool children with autism. *Journal of Autism and Developmental Disorders*, *21*, 281−290.

Higgins, J. P., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*, 1539−1558.

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

Huedo-Medina, T. B., Sanchez-Meca, J., Marin-Martinez, F., & Botella, J. (2006). Assessing heterogeneity in Meta-Analysis: Q test or I^2 Index? *Psychological Methods*, *11*, 193−206.

Hughes, J. R. (2008). A review of recent reports on autism: 1000 studies published in 2007. *Epilepsy & Behavior*, *13*, 425−437.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, *24*, 659−685.

Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (1999). *Autism Diagnostic Observation Schedule.* Los Angeles: Western Psychological Services.

Lovaas, O. I. (1981). *Teaching developmentally disabled children: The me book.* Baltimore: University Park Press.

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

*Luiselli, J. K., Cannon, B. O., Ellis, J. T., & Sisson, R. W. (2000). Home-based behavioral interventions for young children with autism/pervasive developmental disorder: A

---

[1] References marked with an asterisk indicate studies included in the meta-analysis.

preliminary evaluation of outcome in relation to child age and intensity of service delivery. *Autism*, *4*, 426—438.

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

*Matos, M. A., & Mustaca, A. E. (2005). Análisis Comportamental Aplicado (ACA) y Trastornos Generalizados del Desarrollo (TGD): Su evaluación en Argentina [Applied behavior analysis and pervasive developmental disabilities: Assessment in Argentina]. *Interdisciplinaria*, *22*, 59—76.

Maurice, C., Green, G., & Luce, S. C. (Eds.). (1996). *Behavioral intervention for young children with autism: A manual for parents and professionals*. Austin, TX: Pro-Ed, Inc.

Maurice, C., Green, G., & Foxx, R. M. (2001). *Making a difference: Behavioral intervention for autism.* Austin, TX: Pro-Ed, Inc.

*McEachin, J. J., Smith, T., & Lovaas, O. I. (1993). Long-term outcome for children with autism who received early intensive behavioral treatment. *American Journal on Mental Retardation*, *97*, 359—372.

McIntyre, L. L., Gresham, F. M., DiGennaro, F. D., & Reed, D. D. (2007). Treatment integrity of school-based interventions with children in the Journal of Applied Behavior Analysis 1991–2005. *Journal of Applied Behavior Analysis*, *40*, 659—672.

Moher, D., Schulz, K. F., & Altman, D.the CONSORT Group. (2001). The CONSORT Statement: Revised recommendations for improving the quality of reports of parallel-group randomized trials. *Journal of the American Medical Association*, *285*, 1987—1991.

Morris, S. B. (2008). Estimating effect sizes from pretest–posttest-control group designs. *Organizational Research Methods*, *11*, 364—386.

Morris, S., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, *7*, 105—125.

New York State Department of Health. (1999). *Clinical practice guideline: Report of the recommendations. Autism/pervasive developmental disorders. Assessment and intervention for young children (age 0-3 years)* (Publication No. 4215). Albany, NY: Author. Retrieved January 1, 2009, from http://www.health.state.ny.us/community/infants_children/early_intervention/autism/9

Ontario Ministry of Education. (2007). *Questions and Answers: Students with autism spectrum disorders.* Toronto, Ontario: Author. Retrieved May 13, 2009, from http://www.edu.gov.on.ca/eng/general/elemsec/speced/autism.html

Piazza, S. D., & Fadanni, D. (2002). Autism, TEACCH and Lovaas methods: Towards a synthesis. *Revue Medicale de Liege*, *57*, 463—466.

*Reed, P., Osborne, L. A., & Corness, M. (2007). Brief report: Relative effectiveness of different home-based behavioral approaches to early teaching intervention. *Journal of Autism and Developmental Disorders*, *37*, 1815—1821.

*Remington, B., Hastings, R. P., Kovshoff, H., degli Espinosa, F., Jahr, E., Brown, T., et al. (2007). Early intensive behavioral intervention: Outcomes for children with autism and their parents after two years. *American Journal on Mental Retardation*, *112*, 418—438.

Robinson, K. A., & Dickersin, K. (2002). Development of a highly sensitive search strategy for the retrieval of reports of controlled trials using PubMed. *International Journal of Epidemiology*, *31*, 150—153.

*Sallows, G. O., & Graupner, T. D. (2005). Intensive behavioral treatment for children with autism: Four-year outcome and predictors. *American Journal of Mental Retardation*, *110*, 417—438.

Scruggs, T. E., & Mastropieri, M. A. (1998). Summarizing single-subject research. *Behavior Modification*, *22*, 221—242.

Severtson, J. M., Carr, J. E., & Lepper, T. L. (2009). Noncontingent reinforcement is an empirically supported treatment for problem behavior exhibited by individuals with developmental disabilities. *Research in Developmental Disabilities*, *30*, 44—57.

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

Shook, G. L. (2005). An examination of the integrity and future of the Behavior Analyst Certification Board credentials. *Behavior Modification*, *29*, 562—574.

*Smith, T., Eikeseth, S., Klevstrand, M., & Lovaas, O. I. (1997). Intensive behavioral treatment for preschoolers with severe mental retardation and pervasive developmental disorder. *American Journal of Mental Retardation*, *102*, 238—249.

*Smith, T., Groen, A. D., & Wynn, J. W. (2000). Randomized trial of intensive early intervention for children with pervasive developmental disorder. *American Journal of Mental Retardation*, *105*, 269—285.

Sundberg, M. (2007). Verbal behavior. In J. O. Cooper, T. E. Heron, & W. Heward (Eds.), *Applied behavior analysis* (pp. 525—547)., 2nd ed. Upper Saddle River, NJ: Pearson.

Thompson, S. G., & Sharp, S. J. (1999). Explaining heterogeneity in meta-analysis: A comparison of methods. *Statistics in Medicine*, *18*, 2693—2708.

U.S. Public Health Service. (1999). Mental health: A report of the surgeon general (chap. 3, section on autism). Retrieved May 13, 2009, from http://www.surgeongeneral.gov/library/mentalhealth/chapter3/sec6.html#autism

Volkmar, F. R., & Davies, M. (2003). Autism. *Lancet*, *362*, 133—141.

*Weiss, M. J. (1999). Differential rates of skill acquisition and outcomes of early intensive behavioral intervention for autism. *Behavioral Interventions*, *14*, 3—22.

Wheeler, J. J., Baggett, B. A., Fox, J., & Blevins, L. (2006). Treatment integrity: A review of intervention studies conducted with children with autism. *Focus on Autism and Other Developmental Disabilities*, *21*, 45—54.

Whitehead, A. (2002). *Meta-analysis of controlled clinical trials.* Hoboken, NJ: John Wiley and Sons.

*Zachor, D. A., Ben Itzchak, E., Rabinovich, A., & Lahat, E. (2007). Change in autism core symptoms with intervention. *Research in Autism Spectrum Disorders*, *1*, 304—317.

Til: Sundhedsstyrelsen, Evidens, Uddannelse og Beredskab

Ergoterapeutforeningen
Nørre Voldgade 90
DK-1358 København K
Tlf: +45 88 82 62 70
Fax: +45 33 41 47 10
cvr nr. 19 12 11 19
etf.dk

Den 15. september 2020
E-mail: mf@etf.dk
Direkte tlf: 53364923

Ergoterapeutforeningens høringssvar vedr. national klinisk retningslinje for behandling af autismespektrumforstyrrelser hos børn og unge

Ergoterapeutforeningen har med interesse læst udkastet til en national klinisk retningslinje (NKR) for behandling af autismespektrumforstyrrelser hos børn og unge, og er overordnet set positive over for retningslinjen. Det er dog beklageligt, at vi ikke har fået mulighed for at deltage i arbejdsgruppen, da mange ergoterapeuter arbejder med børn og unge med autismespektrumforstyrrelser, og har specialviden indenfor området. Særligt ift. picospørgsmålet omkring ”sanseintegrationsterapi” (nr. 6) ville vi med en faglig repræsentant i arbejdsgruppen kunne have kvalificeret arbejdet. Derfor har vi også særligt kommentarer ift. denne del.

Træk på arbejdet fra national klinisk retningslinje (NKR) for udredning og behandling af ADHD hos børn og unge

Ergoterapeutforeningen bidrog netop med praksis og teoretisk viden om betydningen af sanseintegration og sensorisk bearbejdning i arbejdsgruppen for national klinisk retningslinje (NKR) for udredning og behandling af ADHD hos børn og unge, som blev sendt i høring i foråret 2020. Det er vores anbefaling, at de to retningslinjer ensrettes ift. sprogbrug, beskrivelser og henvisninger ift. de respektive NKRs picospørgsmål om sansernes betydning i behandlingen af de to målgrupper.

Vi foreslår konkret at:

- Sanseintegrationsbehandling bruges konsekvent som betegnelse i begge NKR’er. Den sprogbrug bruges i forvejen i høringsudkastet for NKR for udredning og behandling af ADHD hos børn og unge. I NKR for behandling af børn og unge med autismespektrumforstyrrelser bruges derimod betegnelserne ”sanseintegrationsterapi” i overskriften og ”sanseintegrationsintervention” i underteksten (s. 5 og s. 32).

- Det indledende afsnit ift. picospørgsmålet om ”sanseintegrationsterapi” (nr. 6 s. 32) i NKR for behandling af autismespektrumforstyrrelser hos børn og unge, bør i højere grad afspejle afsnittet ”Baggrund for valg af spørgsmål” ift. picospørgsmål om ”sanseintegrationsbehandling” (nr. 5.7 s. 48) i NKR for udredning og behandling af ADHD hos børn og unge. Det skal selvfølgelig gøres med respekt for forskellene i opbygningen af de to NKR’er, som vi formoder skyldes, at de er blevet udarbejdet på forskellige tidspunkter.

Vi foreslår på den baggrund, at nedenstående tekst tilføjes til indledningen ift. picospørgsmålet om ”sanseintegrationsterapi” (nr. 6 s. 32) i NKR for behandling af autismespektrumforstyrrelser hos børn og unge:
”En overvejende del af børn og unge med ASF har komorbide sanseintegrationsforstyrrelser (Sensory Processing Disorder). Disse forstyrrelser medfører vanskeligheder med at registrere, modulere, integrere, fortolke og hensigtsmæssigt handle på sensorisk information fra egen krop og fra omgivelserne. En del forældre efterspørger behandling for at afhjælpe disse vanskeligheder hos barnet/den unge. Sanseintegrationsbehandling består som minimum af konsulterende rådgivning og vejledning med udgangspunkt i barnet/den unges sensoriske tolerancetærskler og behov på følgende områder: Det kropslige plan, aktivitetsdeltagelse, rammer- og struktur for læring og aktivitet, undervisnings- og hjemmemiljø og rådgivning omkring en regulerende sensorisk ramme der kan passe til barnet/ den unge”.

- Det er uhensigtsmæssigt, at det i NKR for behandling af børn og unge med autismespektrumforstyrrelser fremstår som meget ressourcetungt for forældre at støtte op om sanseintegrationsbehandlingen. Oplevelsen fra vores medlemmer som har specialiseret sig indenfor børne- og ungeområdet er tværtimod, at de møder børn, unge og forældre, der havde ønsket, at de langt tidligere blev tilbudt sanseintegrationsbehandling.
  Derfor foreslår vi, at teksten i afsnittet om ”patientpræferencer” (s. 32) i NKR for behandling af børn og unge med autismespektrumforstyrrelser erstattes med teksten i ”patientpræferencer” (s.48) fra NKR for udredning og behandling af ADHD hos børn og unge:
  ”*De fleste familier formodes at ville modtage rådgivning og vejledning, men der vil være forskel på hvorvidt familier vil have overskud til at igangsætte et behandlingsforløb*”.

- Der i den sidste linje i boksen ”svag anbefaling” ift. sanserne (nr. 6 s. 32), i NKR for behandling af autismespektrumforstyrrelser hos børn og unge, tilføjes ordene ”eksempelvis som” således, at sætningen bliver "*Tilstedeværelse af moderate til svære sanseintegrationsforstyrrelser kan defineres eksempelvis som en T-score > 70 på Sensory Processing Measure, svarende til Afgjort Dysfunktion*.”
  Den samme ændring er blevet foreslået af Ergoterapeutforeningen ift. høringen om NKR for behandling af ADHD for børn og unge.

- Under afsnittet ”Andre overvejelser” i NKR for behandling af børn og unge med autismespektrumforstyrrelser er der en litteraturhenvisning til Ayres, der ikke er tilstrækkelig og tidssvarende. Vi anbefaler, at man i stedet indsætter nedenstående tekst og henvisningerne fra afsnittet ”Andre overvejelser” i NKR for udredning og behandling af børn og unge med ADHD:
  ”Elementerne i sanseintegrationsbehandling kan tage udgangspunkt i behandlingsprincipperne beskrevet i "Sensory Integration Theory and Practice" 3rd. edition af Anita Bundy og Shelley Lane [56] eller "Sanseintegration hos børn" 2. udgave af Jean Ayres [69], og kan planlægges så det retter sig mod og trivsel i hverdagen”.

- I NKR for behandling af børn og unge med autismespektrumforstyrrelser ser man på spørgsmålet omkring brugen af melantonin skal bruges i behandlingen af søvnforstyrrelser hos børn og unge (nr. 8). Her kunne man med fordel have inddraget et spørgsmål omkring den gavnlige effekt af kæde/kugledyner, som der er blevet undersøgt i NKR for udredning og behandling af børn og unge med ADHD (nr. 5.8).

Øvrige kommentarer ift. picospørgsmål om sansernes betydning (nr. 6) i NKR for behandling af børn og unge med autismespektrumforstyrrelser:

- I udvælgelsen af det fokuserede spørgsmål (s. 33) fremgår to redskaber SPM og Sensory Profile 2. Fælles for de to redskaber er, at de beror på såkaldte rating scales, hvor scoringsværdier er udtryk for andre end undersøgerens opfattelse, og bør derfor ikke stå alene som udredningsmetoder. De bør suppleres med andre udredningsredskaber som eksempelvis kliniske observationer af sensorisk bearbejdning, standardiserede performancetests og øvrige indsamlede data der er relevante for vurdering af sensorisk bearbejdning.

- Ergoterapeutforeningen anbefaler, at man bruger samme formulering omkring mulige skadevirkninger forbundet med behandlingen i afsnit omkring ”Skadevirkninger” (s. 34) som under afsnittet om ”Gavnlige og skadelige bivirkninger” (s. 32) – dvs. følgende sætning:
  ”*Ingen af de inkluderede studier rapporterede skadevirkninger, men det er arbejdsgruppens vurdering, at der ikke er væsentlige skadevirkninger forbundet med behandlingen.”*

- Ergoterapeutforeningen vil gerne gøre opmærksom på, at der pågår et stort randomiseret studie i Storbritannien, der måler på effekt af Sanseintegrations Terapi (SIT), sammenlignet med Sensory Based Interventions (SBI), hos børn med autisme i alderen 4 - 11 år. Studiet inkluderer over 200 børn. SBI betyder, at man i barnets miljø sikrer at der er forskellige sanse-baserede tilbud, som barnet selv kan opsøge og benytte. SIT er derimod en en-til-en individualiseret terapeutisk behandling, tilrettelagt på baggrund af en grundig vurdering. Studiet er påbegyndt i 2019, og er derfor endnu ikke publiceret, men man kan følge med i studiet her: **https://trialsjournal.biomedcentral.com/articles/10.1186/s13063-019-3205-y**

Ergoterapeutforeningen står gerne til rådighed med uddybning af dette høringssvar.

Med venlig hilsen

Tina Nør Langager
Formand for Ergoterapeutforeningen

EEGInstitute.dk
Authorized BEE Medic / EEG Info Partner

Høringssvar vedrørende National klinisk retningslinje for behandling af autismespektrumforstyrrelser hos børn og unge

Fra:
EEGInstitute.dk
Stendyssevej 50, 9850 Hirtshals
Brenstrupgårdsvej 23, 8200 Aarhus
Tlf +45 20329901
Mail mail@eeginstitute.dk

Den 17. september 2020

EEGInstitute.dk er som partner i EEG Info Europe, ansvarlig for at uddanne Neurofeedbacktrænere efter Othmer-metoden i Danmark. Othmer-metoden er udviklet af Sue og Siegfried Othmer i USA, igennem ca 30 år. I dag har vi højtudviklet velfærdsteknologi der er godkendt i EU som medicinsk udstyr til bedre selvregulering såvel som klinisk gennemarbejdede og afprøvede protokoller. Udstyret sælges kun til uddannede Neurofeedbacktrænere. I EEGInstitute.dk har vi valgt at målrette uddannelsen mod fagpersoner der har ansvaret for mennesker der har brug for en ekstra indsats, for at håndtere livets udfordringer. Vi har siden uddannelsen kom til Danmark i januar 2018 uddannet 140 Neurofeedbacktrænere, der hver på deres felt oplever at deres klienter profiterer af Neurofeedback. Et Neurofeedbackforløb efter Othmer-metoden tilrettelægges individuelt, uden hensyntagen til eventuelle diagnoser, men med fokus på symptomer. Metoden er non-invasiv, non-verbal og ikke kravstillende.

Vi ved at Othmer-metoden kun er én af mange metoder til at træne Neurofeedback. Derfor finder vi det problematisk at SST i retningslinjen generaliserer Neurofeedback og bygger sine anbefalinger på studier der ikke er repræsentative for Neurofeedback som helhed. Vi ønsker derfor at SST vil kigge på brugen af ordet Neurofeedback, og præcisere retningslinjen efter de undersøgte metoder. Vi frygter at den generaliserende anvendelse af ordet Neurofeedback i retningslinjen underminerer brugen og dermed effekt og virkning af andre typer af Neurofeedback. Derfor finder vi det essentielt at anvendelsen af ordet Neurofeedback fremgår mere målrettet og præcist.

Høringssvar til Sundhedsstyrelsen:

Nationale kliniske retningslinjer for behandling af autismespektrumforstyrrelser hos børn og unge - høring

Generelle bekymringer:

Retningslinjerne er kun for aldersgruppen 0-17 år:

Generelt opleves mange udfordringer i voksengruppen, hvor støtte bortfalder, når de fylder 18 år, særligt blandt de normalt begavede.

Mange voksne med autismespektrumforstyrrelse (ASF) har lært en række kompenserende strategier, samt at maskere deres autisme. Dette gør udfordringerne sværere at se og forstå for andre, men ikke mindre opslidende, og til tider invaliderende, for personen med ASF.

Der er i høj også grad brug for kliniske retningslinjer for voksne med ASF. Dette skal bl.a. også ses i lyset af, at der ses et stadigt større antal sendiagnosticerede unge og voksne, som ikke har været i autismevenlige rammer som børn, og som derfor ikke har lært at leve med deres autisme på en måde, så de undgår en voldsom belastningsreaktion.

Der bør endnu større fokus på rammerne – og ikke kun på barnet

En del af behandlingen bør være at ændre på barnets overordnede rammer. Som nævnt side 65:

"Symptomerne skal være tilstede tidligt i barnets udvikling, men kan vise sig mere manifest når kravene til barnet/den unge overstiger resurserne og/eller når de tillærte kompenserende strategier ikke længere er tilstrækkelige."

Mange, børn som voksne, med ASF får en invaliderende belastningsreaktion pga. fraværet af autismevenlige rammer. Der er mange gode bud på adfærdsterapi m.m., men man bør undersøge og overveje effekten af en decideret *sygemelding* af børn / unge, som har fået en belastningsreaktion, indtil barnet / den unge er modtagelig for behandling. Som det er nu, får barnet ofte først tilbudt et specialskoletilbud, når belastningsreaktionen er så fremtrædende, at barnet har udviklet skolevægring. Dette betyder ofte, at barnet ikke er i stand til at benytte et ellers tilpasset og målrettet skoletilbud; det bliver dermed endnu et nederlag i rækken, og endnu et traume.

Brobygning mellem hjem og tilbud

Når et barn eller en ung er ramt af fx en belastningsreaktion, er det typisk en svær opgave at flytte sig fysisk mellem A og B. Ofte er det familien omkring barnet/den unge, der står med udfordringen med at skabe bro mellem hjem og skole/ungdomsuddannelse mm. Der mangler i den grad et støttetilbud til familien/barnet/den unge, som hjælper familien til at få motiveret personen med ASF til at træde ud af hjemmet og blive en del af det omkringliggende samfund.

Der bør være fokus på brobygning mellem hjem og tilbud - ikke kun med ord, men med støtte og kontakt til en person med viden inden for ASF.

Sanseintegrationstræning, forældre-psykoedukation og social-træning

Rapporten giver kun en "svag anbefaling" og vurderer en "lille netto-gevinst" ift. fx sanseintegrationstræning, forældre-psykoedukation og social-træning. Dette stemmer ikke overens med de familie- og forældrehistorier, som deles i foreningen og i autismenetværk, hvor disse indsatser vurderes til at have afgørende betydning, særligt forældre-psykoedukation.

Søvnforstyrrelser - Afsnit 8 samt Afsnit 12 Baggrund:

Afsnit 8, side 43:
"Overvej at tilbyde melatonin til søvnforstyrrelser hos børn og unge med autismespektrumforstyrrelser i alderen 2-17 år med søvnforstyrrelser, hvor søvnhygieniske tiltag ikke har haft tilstrækkelig effekt."

Det er bekymrende, at andre hjælpemidler, som fx tyngedyner, ikke er nævnt i fm. søvnforstyrrelser. Man går altså direkte fra søvnrutiner / søvnhygiejne til medicinering.

Dette er problematisk, da kommunerne kan bruge denne anbefaling til at kræve, at Melatonin forsøges, før der evt. bevilges tilskud til en tyngedyne.

Afsnit 12, side 65:
"[…] farmakologisk behandling aldrig er førstevalg ved søvnforstyrrelser hos børn og unge, og at man generelt skal man være meget tilbageholdende med medikamentel behandling."

Forslag til øvrige emner:

Medicinsk cannabis / cannabisolie

Der er blandt en del forældre fokus på virkningen af medicinsk cannabis i fm. angst på baggrund af autisme. Flere læger advarer mod dette, særligt til børn og unge, og en undersøgelse og anbefaling kunne være hensigtsmæssig.

Se bl.a. Lægemiddelstyrelsens hjemmeside:
https://laegemiddelstyrelsen.dk/da/special/medicinsk-cannabis/borgere/spoergsmaal-og-svar-om-medicinsk-cannabis/

Sammenholdt med bl.a. markedsføring af CBD på denne hjemmeside: https://cbd-priser.dk/cannabis-olie/
"CANNABIS OLIE OG AUTISME

Autisme er endnu et af de områder, som ser ud til at være særligt interessante i forhold til cannabis olie, men hvor forskningen på området desværre stadig er meget sparsom. Der er nogle forsøg med mus, som har vist positive resultater, men det er endnu ikke sådan at alverdens autisme-foreninger officielt anbefaler behandling med cannabis olie. Alligevel kaster desperate forældre til børn med autisme sig selv ud i forsøg med cannabis olie og resultaterne taler for sig selv. Hvis man skal tro de mange mirakelhistorier på internettet, så er der i hvert fald ingen tvivl om, at cannabis olie kan gøre en verden til forskel for både børn og voksne, som er ramt af autisme. Der er set eksempler på børn, *som aldrig har sagt et ord, der pludselig får et sprog og meget aggressive, udadreagerende og selvdestruktive autister, der bliver mere rolige og får det bedre ved hjælp af cannabis olie."*

Med venlig hilsen

Lilli Jørgensen
Kredsformand
Landsforeningen Autisme Kreds Fyn

DANSK
SELSKAB
FYSIOTERAPI

Dato:
18.09.2020

E-mail:
BDH@dsfys.dk

Høringssvar

Til: Sundhedsstyrelsen, Evidens, Uddannelse og Beredskab

Høringssvar vedrørende Sundhedsstyrelsens ”National klinisk retningslinje for behandling af autismespektrum forstyrrelser hos børn og unge”

Dansk Selskab for Fysioterapi har med interesse læst Sundhedsstyrelsens udkast til ”National klinisk retningslinje for behandling af autismespektrum forstyrrelser hos børn og unge”.

Vi takker for muligheden for at afgive høringssvar og kvitterer for et overvejende solidt arbejde, og har i nærværende brev samlet et høringssvar fra faglige selskaber i fysioterapi. Følgende høringssvar leveres på vegne af Dansk Selskab for Fysioterapi (DSF) i samarbejde med Dansk Selskab for Pædiatrisk Fysioterapi.

Vi håber, at arbejdsgruppen og Sundhedsstyrelsen finder kommentarerne i høringssvaret anvendelige at arbejde videre med i kvalificeringen af anbefalingerne.

Såfremt der er opklarende spørgsmål eller yderligere behov, er vi naturligvis behjælpelige.

Generelle bemærkninger

Generelt er det et meget relevant og aktuelt emne, og vi anerkender, at der i retningslinjen tilstræbes en højere grad af evidensbaseret tilgang på området, gennem f.eks. anbefalinger mod Neurofeedback og Særlige diæter.

Velvidende at der i retningslinjen ikke kan adresseres alle interessante aspekter, savner vi som faggruppe et fokus på effekten af ’Fysisk aktivitet’ (fx løb, cykling, svømning og ridning) som behandlingsstrategi for behandling af autismespektrum forstyrrelser hos børn og unge. Et sådanne fokus kunne med fordel målrettes outcomes som ’stereotyp adfærd’, ’academic performance’, ’ koncentration’, ’motorik’, ’social interaktion’, ’eksekutiv funktion’ etc.

Tekstnære bemærkninger

Vi har følgende enkelte konkrete kommentarer til *Afsnit 6 - Sanseintegrationsterapi til børn og unge i alderen 18 måneder -17 år med autisme:*

- Forældre- og klinikerbedømt adfærdsforstyrrelse og funktionsniveau er udvalgt som henholdsvis vigtige og kritiske outcomes. Det kunne imidlertid være interessant også at forholde sig til børnene selv og derigennem også give børnene en stemme.

- Den svage anbefaling for sanseintegrationsterapi er primært rettet mod børn med moderate eller svære sanseintegrationsforstyrrelser. Vi foreslår, at man i retningslinjerne forholder sig til om sanseintegrationsterapi kunne have effekt for børn med få sanseintegrationsforstyrrelser, som oftere har en normal intelligens, og dermed potentielt bedre forståelse af behandlingen, compliance og deraf afledt bedre effekt?

- Retningslinjen tager udgangspunkt i kvantitative effektstudier". Viden om børn og unges livskvalitet fra andre forskningsdesigns af mere kvalitativ karakter kunne med fordel have suppleret de kvantitative data med værdifulde perspektiver.

Danske Fysioterapeuter tilslutter sig bemærkningerne fra Dansk Selskab for Fysioterapi.

Danske Fysioterapeuter & Dansk Selskab for Fysioterapi står naturligvis til rådighed for uddybning af ovenstående og andet som måtte udspringe af høringssvaret.

Vi ser frem til at læse den endelige anbefaling for behandling af autismespektrum forstyrrelser hos børn og unge.

Med venlig hilsen

Sille Frydendal &
Afdelingschef, Profession og Karriere
Danske Fysioterapeuter

Lars Henrik Larsen
Næstformand,
Dansk Selskab for Fysioterapi

Høringssvar fra Center for Autisme ang rapporten: Behandling af autismespektrumforstyrrelser hos børn og unge: *Guideline national klinisk retningslinje*

Alle sidehenvisninger i parentes er til rapporten: Behandling af autismespektrumforstyrrelser hos børn og unge (Guideline national klinisk retningslinje)

Først og fremmest en stor tak til arbejdsgruppe og referencegruppen. Vores indtryk af de væsentligste dele af aktuelle NKR, de dele der dækker kap. 3-11, er overordnede positive. Det overordnede indtryk som efterlades Center for Autisme, efter læsning af særligt kapitel 2, er dog en udtalt bekymring idet det ikke forklares hvor begrænset et område NKR rettes mod og hvor begrænset et virke en NKR tiltænkes ifølge de gældende retningslinjer som alle NKR er underlagt. Vi er bekymrede fordi dette kan medføre misforståelser hos beslutningstagere (se målgruppe definition nedenfor) som kan lede til afvisning af behandlingsformer hørende under den generelle NKR-kategori "god praksis" som i øvrigt intetsteds nævnes i aktuelle NKR. Baggrunden for den udtalte bekymring uddybes nedenfor.

Målgruppedefinitionen, i aktuelle NKR, er ikke korrekt idet den mangler at medtage beslutningstagere fra den sekundære målgruppe som defineret i den generelle definition af en NKR som gælder for alle NKR:
sundhedsprofessionelle, f.eks. læger, sygeplejersker, fysioterapeuter, ergoterapeuter, tandlæger, tandplejere og kiropraktorer. En sekundær målgruppe er f.eks. patienter/borgere og beslutningstagere (https://www.sst.dk/-/media/Opgaver/Patientforl%C3%B8b-og-kvalitet/NKR/Hvad-er-en-national-klinisk-retningslinje-_NKR_.ashx?la=da&hash=3F0F71FF3D9822624BF1938BAECBBE7D529F9E66).
Målgruppen nævnes også indledningsvist under titlen "Ansvarsfraskrivelse" og også her udelades "beslutningstagere. Kategorien "Beslutningstagere" indgår ikke i aktuelle NKR og det er problematisk idet der derfor ikke kommunikeres til disse beslutningstagere direkte ang. hvorledes disse kan anvende aktuelle NKR. Hermed åbnes op for at beslutningstagere kan misbruge aktuelle NKR til at afvise behandlingsformer som ikke er en del af aktuelle NKR, for eksempel behandlingsformer der kan kategoriseres som "god praksis".

Når vi ser nærmere på hvad der egentlig karakteriserer en NKR så læser vi følgende:

En national klinisk retningslinje indeholder alene anbefalinger for de udvalgte 8-10 velafgrænsede problemstillinger i patientforløbet, hvorfor en national klinisk retningslinje ikke kan stå alene, men komplementeres og suppleres af andre retningslinjer – f.eks. tværfaglige- og tværsektorielle retningslinjer for andre dele af patientforløbet eller andre patientpopulationer, retningslinjer udarbejdet (monofagligt) af selskaber og faglige organisationer samt regionale og kommunale retningslinjer, vejledninger og instrukser.

Samt:

En national klinisk retningslinje indeholder 8-10 udvalgte og velafgrænsede kliniske problemstillinger ('punktnedslag i patientforløbet'). Disse kliniske problemstillinger er prioriteret af *den faglige arbejdsgruppe som de områder, hvor det - med en tidshorisont på ca. seks måneder til arbejdsgruppens læsning og vurdering af litteraturen - er vigtigst at få afklaret evidensgrundlaget. Indholdet i retningslinjen er således afgrænset til en systematisk og omfattende gennemgang af evidensen på disse udvalgte indsatsområder samt tilhørende anbefalinger* (se: https://www.sst.dk/-/media/Opgaver/Patientforl%C3%B8b-og-kvalitet/NKR/Hvad-er-en-national-klinisk-retningslinje-_NKR_.ashx?la=da&hash=3F0F71FF3D9822624BF1938BAECBBE7D529F9E66).

Det er efter læsning af ovenstående citater let at forstå at en NKR ikke er en anbefaling mod at anvende behandlingsformer som ikke er nævnt i NKR. For eksempel de mange behandlingsformer der kan kategoriseres som "god praksis". Denne beskrivelse af den meget begrænsede anvendelse af NKR, som gælder alle NKR, er meget nedtonet i aktuelle NKR. Ud over at der mangler informationer om hvor begrænset en anvendelighed NKR tiltænkes af sundhedsstyrelsen er der også en enkelt men afgørende fejl i aktuelle NKR's beskrivelse af formålet med NKR. Dette forhold uddybes nedenfor.

Formålet med de nationale kliniske retningslinjer er at sikre en evidensbaseret indsats af ensartet høj kvalitet på tværs af landet…..Målet med denne NKR er at sikre en ensartet kvalitet i sundhedsvæsnet i forbindelse med behandling af autismespektrumforstyrrelser hos børn og unge op til det fyldte 18. år gennem en række anbefalinger til sundhedsprofessionelle (s.11). Det kan ikke være formålet. Formålet er ikke at sikre men at understøtte (se: https://www.sst.dk/-/media/Opgaver/Patientforl%C3%B8b-og-kvalitet/NKR/Hvad-er-en-national-klinisk-retningslinje-_NKR_.ashx?la=da&hash=3F0F71FF3D9822624BF1938BAECBBE7D529F9E66).
Der er en meget stor forskel på at sikre og understøtte. Det er vigtigt at ordet sikres slettes fra NKR idet det ikke er i overensstemmelse med de generelle retningslinjer for en NKR. Fortolkningsforskellen af NKR som ordvalget sikre bevirker kan hos beslutningstagere i kommunalt regi medføre at man udelukker bevillinger til enhver interventionsform der ikke er anbefalet af NKR: beslutningstagere skal jo *sikre en evidensbaseret indsats*. Modsat: Når man bruger NKR til at understøtte behandling så orienterer man sig fortrinsvist i NKR i de tilfælde hvor netop den behandling som efterspørges er vurderet af NKR. Her afvises ikke alle behandlingsformer som ikke er udvalgt af NKR. Alternativt kan man fra kommunalt regi lade sig inspirere af de få behandlingsformer der behandles i NKR. Men man misforstår ikke hensigten med NKR og tror at NKR er et katalog over de eneste gode behandlingsformer.

Sundhedsstyrelsens nationale kliniske retningslinjer er systematisk udarbejdede udsagn (s.2). Ja, men det er ikke en systematisk udarbejdelse af alle mulige udsagn. Det er velbegrundede men få udpluk af forskellige interventionsformer hvor langt flere efterlades ubehandlet end behandlet. Det er efter vores mening derfor ikke korrekt at formålet med NKR er at sikre en evidensbaseret indsats. Det er det ikke! Formålet med NKR er at udvælge få af mange interventionsformer og gennemgå disse ud fra de på forhånd definerede kriterier som arbejdsgruppen der har udarbejdet NKR er underlagt. De udvalgte områder dækker en meget lille del af hvad der rent faktisk tilbydes og gennemføres og accepteres bredt som god praksis. Set i dette perspektiv er den upræcise beskrivelse af formålet med NKR bekymrende idet denne dermed ophøjes til mere end den blev tiltænkt. Dermed øges risikoen for at beslutningstagere/kommuner fremover vil afvise støtte til størstedelen af aktuelle interventioner, som nationalt set er bredt accepteret af alle eksperter indenfor autismeområdet og dermed kan kategoriseres utvetydigt som "god praksis", idet disse jo ikke sikrer en evidensbaseret indsats.

Samlet set vækker det således bekymring hos Center for Autisme at aktuelle NKR's begrænsede anvendelsesområde samt begrænsede virke ikke præciseres.
I aktuelle NKR bør tydeliggøres, at man ikke kan drage den konklusion, at begrænset forskning på området er lig med, at ingen af indsatserne hjælper. Det forholder sig således at der ikke er forsket i de fleste indsatser der kan kategoriseres som "god praksis". Der mangler evidens fordi der ikke er forsket i dem; ikke fordi forskning har vist at der ikke er evidens. En afgørende forskel der ikke tydeliggøres i aktuelle NKR. Hvis formålet med NKR reelt var at sikre en evidensbaseret tilgang ville de fleste interventionsformer, som alle vil kalde "god praksis", ikke længere kunne blive anvendt alene fordi der forskes for lidt i disse interventionsformer.

En måde hvorpå der kan rettes op på aktuelle NKR ville være at anvende kategorien "god praksis" i vid udstrækning som jo også er meningen. Her får panelet bag rapporten jo netop mulighed for at anbefale praksis som der ikke er forsket i men som der alligevel er enighed om er god praksis. Vi foreslår derfor at aktuelle NKR gør hyppig brug af kategorien "god praksis". Dette bør under alle omstændigheder ændres i aktuelle NKR for hvorfor defineres kategorien "god praksis" uden at anvende den? Kategorien "god praksis" anvendes nul gange i NKR (bortset fra de syv steder i rapporten hvor kategorien defineres.

Udover de mere generelle problemer nævnt ovenfor har vi opdaget nogle faktuelle fejl i beskrivelsen af diætinterventionsgennemgangen som vi retter opmærksomheden mod nedenfor:

I omtalen af studiet af Whiteley et al (2010) skrives at…*alvorlig risiko for bias på grund af manglende blinding af deltagere og bedømmere af effekten* (s.56) (effekten henviser til "klinikerbedømte kernesymptomer"= ADOS). Det er ikke korrekt at der er manglende blinding, da netop de klinikerne som bedømte kernesymptomerne var blindede. Whiteley et al studiet er netop et single-blindet studie, hvor klinikerne var blindede men ikke forældrene.

Der står desuden– samlet om de 6 studier der er analyseret – *Der var ingen studier, der rapporterede forældrebedømte autisme kernesymptomer, klinikerbedømt funktionsniveau, klinikerbedømt adfærdsvanskeligheder, livskvalitet eller forældretrivsel (s.14).* Whiteley et al (2010) rapporterede forældrebedømte autisme kernesymptomer (GARS skemaet) og klinikerne bedømte (blindede) funktionsniveau med Vineland.

Ovenstående aspekter ved Whiteley et al (2010) studiet var netop styrken ved studiet og når studiet så ikke fandt en effekt af diæten, så er konklusionen så meget desto stærkere, dvs. at man ikke kan anbefale diæten. Hvis fejlene der påpeges i rettes vil konklusionen i NRK være bedre begrundet.

Samlede anbefalinger til rettelser i NKR.

Anbefaling_1: Fremhæv, ikke bare i kap. 2 men gennemgående, at NKR er en samling af specifikke udvalgte behandlingsformer som på ingen måde kan opfattes som et fuldstændigt katalog over alle de gode interventionsformer der anvendes som del af "god praksis" i DK. Fremhæv at man kan bruge NKR når man skal tage stilling til lige præcis de metoder der nævnes i NKR men ikke til at tage stilling til metoder der ikke nævnes.

Anbefaling_2: Anvend kategorien "god praksis". Hvorfor defineres kategorien "god praksis" uden at anvende den? Kategorien "god praksis" anvendes nul gange i NKR. Brug kategorien "god praksis" til at understøtte **anbefaling_1** ovenfor.

Anbefaling_3: ret omtalen af Whiteley et al (2010) så den refereres faktuelt korrekt. Det vil kun styrke konklusionen om en stærk anbefaling mod diætintervention.

Anbefaling_4: Hvis der er planer om at lave en quick-guide som det er tilfældet med NKR for ADHD behandling så implementer anbefaling_1 i denne. Gør gerne dette mere tydeligt end tilfældet er for NKR for ADHD behandling hvor der blot lidt hengemt med småt står " *Den nationale kliniske retningslinje indeholder anbefalinger for udvalgte dele af området, og kan ikke stå alene, men skal ses i sammenhæng med øvrige retningslinjer, vejledninger, forløbsbeskrivelser mv. på området".* Det er ikke nok.

[redacted]

Ikast-Brande
Kommune

Sundhedsstyrelsen
Islands Brygge 67
2300 København S

18. september 2020

Børne- og Familieafdelingen
Bellisvej 2
8766 Nørre Snede
Tlf.: +4599604000

Sagsbehandler:
Anton Rasmussen
E-mail:
anras@ikast-brande.dk
Direkte telefon:
Tlf.: +4599605501

Sagsnr.:
27.00.00-A00-24-18

Hermed ønsker Ikast-Brande Kommune at afgive høringssvar på udkast til kliniske retningslinjer "Behandling af autismespektrumforstyrrelser hos børn og unge".

I den forbindelse stiller vi følgende spørgsmål:

- Da der er tale om nationale kliniske retningslinjer, er vi nysgerrige omkring hvorvidt kommunernes indsats over for børn og unge med autismespektrumforstyrrelse også er omfattet af disse retningslinjer? Eller er retningslinjerne udelukkende for personalet på sygehusene?

Høringssvaret skal ses i relation til, at det antages, at retningslinjerne er tænkt også at gælde i sammenhænge hvor kommunen har indsatser over for børn og unge med autisme.

Det er uklart hvilke institutioner og hvilket pædagogisk personale, psykologisk personale, socialfagligt personale eller sundhedspersonale de nationale kliniske retningslinjer retter sig imod.

Traditionelt vil "Nationale kliniske retningslinjer" være rettet mod sundhedspersonel. Vejledningen giver anbefalinger både inden for medicinsk behandling, familie og forældrerettede rådgivnings- og vejledningsforløb, individuel træning m.v. Der er således set fra et kommunalt perspektiv tale om en vejledning, der både kan være gældende inden for den regionale børne- og ungdomspsykiatri, almen medicin, undervisning, dagtilbud og børne- og familierettede indsatser.

Lovmæssigt kan der være tale om sundhedslov, folkeskolelov, dagtilbudslov og servicelov.

Det bør i vejledningen præciseres, hvilke institutioner og personale vejledningen omfatter.

Kommunerne anlægger i deres specialiserede indsats overfor udsatte og sårbare børn en helhedsbetragtning. Vejledningen ses at have en patologisk tilgang der strider mod den kommunale praksis, hvor indsatser og behandling ikke alene tilrettelægges ud fra diagnoser, men ud fra en helhedsvurdering af et barns, kompetencer og eventuelle udfordringer.

Et barn eller en ung kan have være udsat for flere risikofaktorer uden at der tales om komorbiditet. Inden for det kommunale sociale område vurderes de samlede risikofaktorer op mod de samlede beskyttelsesfaktorer i forhold til at tilrettelægge en eventuel specialiseret indsats.

De pædagogiske, psykologiske og sociale tilbud i kommunerne bygger i modsætning til sundhedsområdet i høj grad på erfaring, tradition, teori, kultur m.v. frem for evidens forståbet som resultat af kontrollerede forsøg. Det bør derfor overvejes om retningslinjerne skal begrænses til at gælde sundhedspersonale.

Der er derudover følgende kommentarer til enkelte dele af vejledningen:

I forhold til Pædagogisk Psykologisk Rådgivning sprogundervisning indgår børn med autisme som alle andre børn uden autisme i målgruppen for PPR` indsats, og er inkluderet i § 11 i dagtilbudsloven, hvorfor der i hvert enkelt tilfælde tages stilling til den nødvendige intervention. Her indgår en helhedsvurdering af barnets behov, der ikke kun tager udgangspunkt i autismediagnosen. Med henvisning til §11 i dagtilbudsloven ses ingen faglige argumenter eller anden evidens der tyder på, at der er behov for anden specifik indsats på autismeområdet.

I den forældrerettede intervention er det usikkert, om punktet skal forstås som vejledning til Børne og Ungdomspsykiatrien, eller den er en generel retningspinje for både regionale og kommunale indsatser.
I spørgsmålet om anbefalinger i forhold til neurofeedback findes det problematisk, at det i anbefalingen ikke fremgår tydeligt, hvilke neurofeedbeck metoder der ligger til grund for evalueringen. Der er flere forskellige metoder som har forskellige træningsprotokoller, hvorfor de ikke umiddelbart kan sammenlignes som en og samme metode.

Det ses at være problematisk at konkludere, at neurofeedback er én ting, da der er omkring 20 forskellige Neurofeedback metoder, som ikke kan sammenlignes en til en.

Neurofeedback fjerner ikke kernesymptomer for ASF. Hvis det diagnostiske arbejde er troværdigt, vil symptomerne altid være til stede, men der kan kompenseres med træning, hvor der er erfaring for at neurofeedback har sin berettigelse.
Neurofeedback kan have effekt på stress, angst og depression som ofte ses i sammenhæng med ASF.

Vi ser manglende evidens som en invitation til, at der foretages yderligere forskning af metoden, og vi vil gerne medvirke til at indsamle mere data til sådan en forskning.

Med venlig hilsen

Anton Rasmussen

Børne- og Familiechef

Med venlig hilsen

Anton Rasmussen

Børne-og familiechef

DANSK PSYKOLOG FORENING

Sundhedsstyrelsen
Sekretariatet for Nationale Kliniske Retningslinjer
Islands Brygge 67
2300 København S

København, den 18. september 2020

Høring vedr. national klinisk retningslinje for behandling af autismespektrumforstyrrelser hos børn og unge

Dansk Psykolog Forening takker for muligheden for at afgive høringssvar vedrørende national klinisk retningslinje for behandling af autismespektrumforstyrrelser hos børn og unge. Høringssvaret forholder sig til hhv. PICO 4, anbefalingen af sanseintegrationsintervention, og PICO 8, anbefaling mod gluten- og caseineliminationsdiæt samt en generel bemærkning til retningslinjen.

Udredning for og behandling af sanseintegrationsforstyrrelser
Dansk Psykolog Forening ser positivt på, at retningslinjen inkluderer PICO-spørgsmål 4 vedr. tilbud af sanseintegrationsintervention til børn og unge med autisme i alderen 18 måneder til 17 år med moderate til svære sanseintegrationsforstyrrelse. Da de nye diagnosekriterier for autismespektrumforstyrrelser hos børn og unge i DSM-5 og ICD-11 omfatter forstyrrelser af sanseintegration, opfordrer Dansk Psykolog Forening til, at vigtigheden af såvel udredning som behandling af sanseintegrationsforstyrrelser bliver understreget.

Der er dog en betænkelighed ved anbefalingen grundet en lav grad af evidens for sanseintegrationsintervention til børn og unge med autismespektrumforstyrrelse. Ud fra klinisk erfaring, kan der desuden være børn og unge, som reagerer voldsomt på sanseintegrationsintervention på grund af bivirkninger, som f.eks. overstimulering af sanser. Derfor er det vigtigt, at det er fagpersoner med særlig erfaring eller ekspertise, som foretager interventionen.

Nødvendig udredning, hvis børn eller unge udviser tegn på allergi eller intolerance
I PICO-spørgsmål 8 fremgår en stærk anbefaling MOD gluten- og caseineliminationsdiæt til børn og unge med autismespektrumforstyrrelse i aldersgruppen 3-17 år som ud fra diagnostisk anamnese, generel klinisk undersøgelse samt allergitestning ikke har indikationen cøliaki og/eller intolerance for mælkeprotein (casein). Dansk Psykolog Forening hæfter sig i den forbindelse ved, at der ikke er belæg i de rapporterede studier for anbefalingen mod gluten- og caseineliminationsdiæt. Det fremgår derimod af retningslinjerne, at denne anbefaling i stedet er baseret på arbejdsgruppens kliniske erfaring.

Vi er bekymrede for, at en stærk anbefaling MOD gluten- og caseineliminationsdiæt i praksis vil fjerne fokus fra undersøgelse og behandling af eventuelle vanskeligheder med kost og ernæring hos børn og unge med autismespektrumforstyrrelse, der har brug for det. Der er desuden stor usikkerhed på dette område, såvel hos fagfolk som forældre og interesseorganisationer, hvilket også fremgår af de refererede undersøgelser. Vi er derfor bekymrede for, at denne stærke anbefaling, der ikke er baseret på evidens, men på arbejdsgruppens kliniske erfaring, kan medvirke til en unødig polariseret og unuanceret debat. I værste fald kan det betyde, at børn og unge med problemer på dette område ikke får den relevante undersøgelse og behandling.

Sundhedsstyrelsen bør derfor i stedet anbefale, at der tilbydes en diagnostisk anamnese, generel klinisk undersøgelse samt allergitestning til børn og unge med autisme, hvis der er tegn på mulig cøliaki og/eller intolerance for mælkeprotein (casein) eller andre allergier og/eller intolerans for fødevarer. Denne udredning kan udelukke en eventuel medvirkende somatisk ætiologi for forværring af autismesymptomer og/eller følgesymptomer, sikre tilbud om korrekt somatisk behandling, evt. relevant kost, og forebygge fejlbehandling eller utilstrækkelig behandling, på grund af utilstrækkelig og for fagligt snæver (f.eks. udelukkende psykologisk) udredning og diagnostik.

Metodetilgang

I retningslinjen anbefales bestemte programmer, som f.eks. ESDM og PACT, som arbejdsgruppe-repræsentanterne blandt andet har skulle anbefale. Det er dog kun meget få psykologer i Danmark, som er certificerede i at kunne anvende f.eks. ESDM. Det er vigtigt med evidensbaserede tilgang, men med et fokus på indsnævrede programmer, kan de effektive tilgange i behandlingen i mindre grad blive anvendt i praksis. Det anbefales derfor, at der i højere grad er fokus på, hvilke faktorer, tilgange og processer, som fungerer godt i programmerne og derfor bør anbefales i behandlingen af autismespektrumforstyrrelse. Gennem en større grad af metodefrihed kan de anbefalede dele af programmerne bedre anvendes i psykologers behandling af børn og unge med autismespektrumforstyrrelse i Danmark.

Med venlig hilsen

Eva Secher Mathiasen
Formand, Dansk Psykolog Forening

Til NKR Sekretariatet
Sundhedsstyrelsen.

Dato 18. seprtember 2020
EMN-2017-02007
Annamaria Marrero Zwinge

Høringssvar vedr. national klinisk retningslinje for behandling af autismespektrumforstyrrelser hos børn og unge

Danske Regioner har den modtaget national klinisk retningslinje (NKR) for behandling af autismespektrumforstyrrelser hos børn og unge, i høring.

Danske Regioner takker for muligheden for at afgive bemærkninger til retningslinjen som har været i høring i regionerne. Nedenstående bemærkninger til udkastet er baseret på de regionale svar.

Som generel bemærkning til udkastet tager Danske Regioner forbehold for et eventuelt øget ressourceforbrug som følge af den nationale kliniske retningslinje.

Generelle faglige kommentarer

Generelt ses retningslinjen som vigtig med fokus på nogle relevante områder i behandling og håndtering af autisme.

Det kunne være ønskeligt at retningslinjen angav, hvorvidt den angivne behandling (punkt 3-5) specifikt skal tilbydes i hospitalsregi eller, hvorvidt denne kan tilbydes i andet regi (eksempelvis kommunalt) - særligt for tilbud der er målrettet hjem og/eller skole (f.eks Forældremedieret intervention og social færdighedstræning, der optimalt set bør foregå i barnets vante omgivelser).

Det kunne endvidere være hjælpsomt, om det af NKR fremgår, hvorvidt de angivne eksempler på manualer (igen i forbindelse med punkt 3, 4 og 5) forefindes i dansk udgave - således disse er lige til at implementere, eller om programmerne først skal oversættes og herefter valideres på en dansk/nordisk befolkning før disse redskaber kan implementeres.

DANSKE REGIONER
DAMPFÆRGEVEJ 22
2100 KØBENHAVN Ø
+45 35 29 81 00
REGIONER@REGIONER.DK
REGIONER.DK

Specifikke Kommentarer

Side 43

For melatonin angives ”svag anbefaling for” – dette bør være ”stærk anbefaling for”.

Side 49

for antipsykotika angives ”svag anbefaling mod” – dette bør ændres til ”svag anbefaling for”.

Med venlig hilsen

Rósa Víkingsdóttir
Konsulent
Center for sundhed og sociale indsatser (SUS)

Til Sekretariatet for Nationale Kliniske Retningslinjer

DASYS takker for muligheden for at afgive høringssvar til høring vedr. National klinisk retningslinje for behandling af autismespektrumforstyrrelser hos børn og unge.
Vi har spurgt vores medlemmer inden for området, men ikke fået tilbagemelding.

Vi har haft god repræsentation i arbejdsgruppen ved klinisk sygeplejespecialist Marianne Friberg Day.

Venlig hilsen

Helle Johnsen
DASYS' sekretariat

Tfl. 4695 4155
Sankt Annæ Plads 30,
DK-1250 København K
dasys@dasys.dk www.dasys.dk

"Behandling av autismespektrumforstyrrelser hos born og unge"
Utlåtande över NKR avgivet av Professor Christopher Gillberg 2020-09-15

Inledning

Jag har blivit ombedd av Sundhedsstyrelsen (Danish Health Authority) att uttala mig om den National Klinisk Retningslinje (NKR) gällande behandling av autism som sent i augusti 2020 överlämnades till mig för en referentbedömning.

Den aktuella NKR som har avgivits av en arbets- och referensgrupp till den danske Sundhedsstyrelsen har vissa styrkor, men också många svagheter som gör att dokumentet i sin nuvarande form, enligt min bedömning, *under inga omständigheter* kan slutpubliceras utan omfattande revisioner, revisioner som jag dessutom anser måste bli föremål för ny remissrunda innan dokumentet som helhet kan anses leva upp till gällande krav på NKR.

Styrkorna är delvis dolda i skrivningen på grund av de många bristerna ifråga om avgränsning av projektet, noggrannheten i beskrivningen av evidensprövningen, rekommendationer som ibland går stick i stäv med arbetsgruppens egna evidensprövning, och presentationen av de i arbets- och referensgrupperna ingående individerna.

Kritiska synpunkter

Jag kommer här att punktvis framföra min kritik:

1. Det är inte enkelt att inledningsvis få ett grepp om vad uppdraget från början har bestått i. Ingenstans framgår om avsikten har varit att ta fram ett helhetligt underlag för "behandling" av "all autism". Endast gradvis får läsaren en uppfattning om att **det rör sig om en samling begränsade frågeställningar, som tillsammans till ingen del motsvarar titeln "Behandling av autismespektrumforstyrrelser hos barn og unge"!**

2. **Det är inte någonstans klarlagt vad som avses med "behandling".** Utvärderat material hänför sig mestadels till interventioner/insatser och endast ibland till behandling. Det kan inte heller vara relevant att nästan genomgående redovisa effekter av interventioner/insatser/behandling för det breda åldersintervallet 18 månader till 17 år. – Se även nedan under punkten 8.

3. Arbetsgruppens medlemmar beskrivs med namn - ingenting mera; detta är oacceptabelt. **Man måste kunna vänta sig att få en beskrivning av akademisk grundutbildning, forsknings- och klinisk erfarenhet, när det gäller alla medlemmar i arbets- och referensgrupperna. Man måste kunna vara övertygad om att de rekommendationer som förs fram inte skulle kunna bottna i egenintresse hos vissa av arbetsgruppens medlemmar.**

4. **Komorbiditet/samsjuklighet och överlappning med t ex ADHD, DCD, tics, intellektuell funktionsnedsättning (IF), språkstörning, epilepsi och andra medicinska sjukdomar tas inte upp, annat än i bisatser.** Frågan om samsjuklighet är ett av de största områdena inom modern autismforskning och i ett upplägg som syftar till att lämna riktlinjer för behandling av *alla barn och unga 18 månader till 17 år* med autism måste detta tas i beaktande. På flera ställen i NKR sägs att slutsatser gäller barn med autism "utan komorbiditet". Mycket talar idag för att ytterst få - om ens några - barn med autism som kommer till utredning och "behandling" är helt utan komorbiditet. I "behandling av autism" måste ingå utredning (inklusive medicinsk-

neurologisk) både av autism och samsjuklighet och information till familjen om resultatet av dessa utredningar. Även om evidens för värdet av sådana utredningar skulle saknas är det inte acceptabelt att publicera behandlingsriktlinjer för autism utan att beröra nödvändigheten av adekvat diagnostisk utredning (utom i förbigående i en bisats på sid 65 av 87). Det tas i NKR inte heller upp något om vikten av att föräldrar och pedagoger i barnets förskola eller skola får information om barnets grundläggande kognitiva svårigheter relaterade till autism. Sådan kunskap behövs för att förstå särskilda beteendeproblem och för att kunna både förebygga och ge rätt hjälp i problemsituationer.

5. Den enda "behandling" som förbehållslöst anbefalles i NKR är den som gäller ångest vid autism. Det är mycket tveksamt om detta är att betrakta som *behandling* av ***autism!*** **Man får av NKR intrycket att det enda som överhuvudtaget har vetenskapligt stöd i behandlingen av autism är KBT mot ångest vid autism (dessutom endast för dem med normal begåvning). Detta är dock inte med sanningen överensstämmande.**

6. Avsnittet som beskriver hur man bestämt sig för att *endast* utreda de aktuella områdena (9 sk PICOs) måste fördjupas och också leda till att titeln "Behandling av...." ändras. **Den NKR som fastställts gäller bara *vissa* interventioner vid autism och bara för vissa åldrar, och den är inte på något sätt heltäckande.**

7. **Varför har litteraturen om ABA-IBT inte bedömts? Eller den om ADHD vid autism och behandlingen av ADHD-symptom vid autism? Eller, genomgående och inte bara i undantagsfall, i vad mån behandling som passar för personer med normalbegåvning och autism också passar för dem med IF? Och hur är det med TEACCH? Bumetanid? Vitamin D? Hur skall man behandla den stora gruppen av personer med autism som har epilepsi?**

8. Det råder inte fullständig transparens ifråga om sambandet mellan evidensvärdering och rekommendationer/anbefalningar. Som ett exempel kan nämnas avsnittet som gäller **Sanseintegrationsterapi till born och unge i alderen 18 måneder - 17 år med autism.** För det första konstaterar arbetsgruppen att studier på området uteslutande hänför sig till barn i åldern 3-12 år; det är därmed inte rimligt att dra några slutsatser om yngre eller äldre barn med autism. **Dessutom gör arbetsgruppen bedömningen att evidenskvaliteten är "meget lav". Hur kan man mot denna bakgrund ge någon form av (om också "svag") anbefalning?**

9. **I de flesta av de olika delavsnitten beskrivs i detalj hur artiklar utvalts för inklusion. Detta är föredömligt i sig självt, men när man kontrollräknar finner man att uppgifterna som lämnas i NKR inte stämmer**. För att ge två exempel. - På sid 13 av 87 uppges att "det blev anvendt 21 randomiserade forsog". I parentesen som följer räknas 22 referenser upp. När man sedan tittar tillbaka lite högre upp på sidan (under Rationale) anges att man bedömt 27 studier. 21 är inte detsamma som 22 och definitivt inte som 27. - På sid 19 av 87 uppges att man genom 2 systematiska reviews (av moderat/låg kvalitet) fått fram 10 randomiserade studier. Inom parentesen som följer efter detta påstående om 10 studier räknas dock endast 6 upp. Därtill hittade författarna ytterligare 4 studier och uppger att man därmed inkluderade 14 studier; detta trots att det verkar vara fråga om sammanlagt 10 studier. Därefter delar man upp de "14" studierna i sådana som levererades individuellt (5) och sådana som gavs i grupp (7). Det innebär att 12, inte 14 studier var inkluderade. 10, 12 och 14 är inte samma! **Slarvighet i rapporteringen av dessa siffror leder tyvärr till risk för att NKR inte kommer att tas på allvar.**

10. **Det är, enligt mig, inte bra att avsnittet "Bakgrund" hittas på en undanskymd plats på sid 65 av 87. Det är därtill mycket förvånande att på denna plats finna detaljerade rekommendationer om sömnfrämjande insatser och sömnhygien!**

Sammanfattning

Sammantaget blir min bedömning av aktuell NKR (Behandling av autismespektrumforstyrrelser hos born og unge) att ett stort arbete lagts ner, att man genomfört en omfattande evidensvärdering av vissa interventioner vid autism hos barn och unga, men att slutdokumentet inte bör publiceras utan omfattande revision.

Göteborg 2020-09-15

Christopher Gillberg, MD, PhD
Professor i barn och ungdomspsykiatri, Gillbergcentrum, Göteborgs universitet,
University of Glasgow, Pasteur Institute, and Kochi University Medical School

Karolinska Institutet
KIND
KAROLINSKA INSTITUTET
CENTER OF NEURODEVELOPMENTAL DISORDERS
Region Stockholm

2020-09-05

Institutionen för kvinnors och barns hälsa
Avdelningschef /neuropsykiatri
Professor, PhD
Sven Bölte

Review of the National Clinical Guidelines for the Treatment of Autism Spectrum Disorder in Children and Adolescents 2020, on behalf of Sundhetsstyrelsen

I had the pleasure and honor to receive the trust of the Danish Health Authority to evaluate the recently published Danish guidelines for the treatment of youth with autism spectrum disorder (ASD). The objective of these guidelines is to provide transparency and orientation to clinicians, caregivers and patients regarding ASD treatments, based on a systematic compilation of both scientific evidence, expert knowledge and interest organization views and information. Thus, these guidelines combine hard empirical data with expert opinion and user preferences and experience. The guidelines are not legally binding, as the universe of individual characteristics, demands and outcomes and clinical situations are deemed not to allow for stricter use of the guidelines.

The development of clinical guidelines in general, and for neurodevelopmental conditions in particular, is burdensome and complex, and always results in some form of compromise with limitations that limit validity and generalizability. It is also usually a challenge to have sufficient resources at hand for the given assignment, to find experts who are willing and qualified and at the same time free of conflicts of interests. There are many comprehensive searches and quality assessments of primary and secondary research needed to be performed, evidence to be summarized, and deadlines to be kept in order to generate guidelines sufficiently up-to-date. It is a great challenge to write guidelines that will satisfy all stakeholders. Often guidelines need to take a standardized form which might be unusual to encounter for clinicians or laymen. Therefore, before I mostly focus on concerns and points for improvement, I would like to congratulate on the work presented. I also understand that the guidelines are planned to be updated with intervals of 3 years, which is both necessary and laudable in a time where research is published in large amounts, and clinical views and routines may change fast. I also very much like that the reviews of the guidelines will be made publicly available – that is good scientific practice.

In order for the reader to be fully informed about the possible conflicts of interests of the author of this review, I would like to disclose that I work in research and clinical practice with several methods evaluated here (e.g. social skills training, neurofeedback), which might both mean that my knowledge is particularly high here, but also that there is a certain risk of bias in being objective regarding their usefulness. Based on my experience and knowledge, I might not agree with the conclusions of the work-group of these guidelines, especially where there is a good deal of space of subjectivity. I also want the readers to know that I have in the last 3 years acted as an author, consultant or lecturer for the companies Medice and Roche. I receive royalties for text books and diagnostic tools from Hogrefe, Kohlhammer and UTB. From Hogrefe, I receive royalties for a social skills training program for ASD (“KONTAKT”), which is also included in these guidelines.

1. Guidelines/report strengths

- Most of the guidelines are easy to read and follow. The text is well-structured and informative. The visual organization of the 10 PICOS using boxes and colors is excellent.

- I see as major strengths that the guidelines try to incorporate many different aspects related to the treatments, not exclusively study results, but also user views, expert knowledge and national prerequisites. I understand the complexity of this task and welcome the strategy.
- There is transparency in how the recommendations were achieved and there is a good sense of coproduction among members of the work group, the reference group and others involved. Therefore, it appears that a high degree of anchoring and consensus was achieved.
- The results are up-to-date, with studies until the year 2020.

2. Guideline/report weaknesses and limitations

General

- When I read the overarching description of the envisaged guideline and duties of the workgroup, it is not clear to me, if the current report is part of a larger guideline project on ASD or not. There is mention in the mission that there also should be guidelines on assessment and treatment of ASD and comorbidity. This current report includes no assessment guidelines, but some ASD with comorbidity treatments. Is this all, or is there more to come and when?
- It is not clear to me, why the guidelines are not more open, but focus on a limited amount of specific interventions. There is a lot more to evaluate relevant to ASD, which is probably equally important for Denmark's clinical services and service users.
- Related to the former, there is mention of the importance of certain comorbidities (ADHD, epilepsy, intellectual disability, but there is little to nothing on these in the report.
- There are no conflict of interest statements for the work group members, no indication of their academic degree or affiliations.
- There is an overly focus on the 10 PICOS questions and the evaluation of the treatments in the guidelines, but in order to grasp them, and why these and not others were chosen, there would need to be a proper introduction to the guideline background, ASD, interventions and all the methods applied to develop this report. The text is very poor here, and the information should also be placed in the start of the report, not the end, as in the present form. Overall, the 10 PICOS are too "stand alone" for a report that aims to give broad orientation to the field for a broad readership.
- If the report is really meant to be read by both clinicians and stakeholders and appear transparent to them, there must be a more accessible version or an explanation how it can be read and understood. The large majority is very technical, hard to access for laymen, and people not used to such documents. The presentation should be more educational for this audience.
- It is important to make clear that the way the guidelines are designed may also drive the results. What do the guidelines say, what can they and not say? People might experience them as absolute, which they are not (Bölte, 2015).
- The guidelines are not based on an independent own systematic review and meta-analysis, but on other's work and the evaluation of single additional primary studies. It is important to note that these are very different and not necessarily comparable approaches and data sources.
- I am not sure why the guidelines start with the reporting of the treatment of anxiety in ASD, not a treatment addressing ASD, and the logic of the ordering of the treatments is unclear to me. There could be a better sorting, e.g. after development/age, outcomes, comorbidity, pharma/non-pharma. It is a bit mixed now.
- These guidelines do not take into account any qualitative research, which should be made clear and might pose a limitation.
- The guidelines do not discuss environmental interventions in ASD, only child focused ones. In a time where inclusion is a crucial matter, this is disappointing. Although, such interventions may not be deemed "clinical", others included here, such as those directed to parents, are not necessary clinical, too.

- The very important sections on implementation, monitoring and description of methods are very sketchy and far from complete.

Specific

- I would be very careful with the concept of ”good practice”! It must be clearly delineated from empirical science. It happens often that good practice is confused with common practice, which is rather business as usual that good practice.
- The studies were not necessarily assessed for treatment quality and external validity.
- I deem that is also important to note that the National Danish Society of Autism may not be fully representative of the users. It is not unusual that there are multiple views on treatments, all from very positive to very negative. Caregivers may also have views different from children and youth or adults being diagnosed.
- Under the treatment categories, partly very different interventions are summarized, e.g. parent-mediated includes PACS and Triple-P, and neurofeedback is treated as one method. Social skills trainings might be individual or group-based.
- I have concerns if the workgroup judged each of the PICOS in a comparable and sober fashion and with the same degree of objectivity. For instance, it appears to me that there is an uncriticalness towards some, but not others. Concerning some of the treatments, where side effects are reported, these are weighted heavy (e.g. social skills training), but for others, where these were not assessed (e.g. CBT in ASD and anxiety), these are only weakly considered at all.
- Related, it is also discussed that social skills training might be burdensome for participants, but these same is true for CBT for anxiety. Actually, CBT is important in social skills training, and there are few to no interventions causing no burden. In fact, in the largest RCT in autism ever, an RCT on social skills group training, the side effects were rare and mostly minor.
- Neurofeedback is criticized for being resources demanding and complex, but no data on health economy or any objective information is provided to corroborate the claim.
- In the same way, the accessibility for youth with intellectual disability is commented for some treatments, but not others equally. This is unbalanced.
- Neurofeedback is not one technique, but a mix of very different ones, and there are differences regarding efficacy. There are standard protocols and non-standard ones. Neurofeedback may also be an intervention for ADHD in ASD, not an ASD specific one. This is not discussed.
- When it comes to the patient preferences, there is a reasoning that few would want to have the neurofeedback intervention. Our experience from a large trial is the complete opposite - so this conclusion seems to be speculative and not supported by any data.
- I acknowledge the concern under ”other considerations” that private service providers may exploit individuals with ASD and their families using neurofeedback. While I agree, there is a need for clarification. First, it is of paramount importance to mention that people in private practice mostly use non- standard protocols, not standard protocols, so there is a quite clear cut between scientific and more doubtful neurofeedback techniques. Second, I would argue that this risk might not at all specific to neurofeedback, but relevant to many other interventions, including those being part of the guidelines. This is a matter of quality of delivered treatments and serious clinical work, and stigmatization of a specific method is not helpful here. Scientifically-based neurofeedback is generally serious, although evidence in autism is limited.
- There is a split of the age ranges for social skills training used: 18 months to 5 years, and 6 to 17 years, but often there is an additional divide in age 6 to 12, and 13 to 17 years (child and adolescent trainings). The latter makes sense (pre-puberty & puberty), and has also led to different results.

- It was not clear to me how the working group arrived at the minimum lengths for interventions – this appears speculative, arbitrary and not helpful. Better to give an optimal lengths based on facts if possible
- I had a hard time to grasp why the anxiety treatment in ASD received highest ranking, as the outcomes were mostly parent- and clinician ratings, not self-report, which should be leading for an internalizing disorder.
- I would say that the guidelines would benefit from more taking into account Danish and Scandinavian perspectives of feasibility, accessibility, cultural adaptation and evaluation. Perhaps highlight more research for which Danish and Scandinavian data and research is available and report the origins of the research studies mentioned. Are Danish manuals available for the methods mentioned?
- I was not sure, why for melatonin the critical outcome was time to fall asleep, not other outcomes, like total sleeping time or night awakenings.
- The Clinical Global Impression scale is not a measure of functioning, but a pure measure of global clinical severity in terms of psychopathology. This is misleading and a wrong mix-up of basic concepts. It is recurring all over the report.
- The background section is brief and superficial; the bullet-point section on sleep is fragmented and not well integrated.

Minor

- Errors of spelling, Tipple-P must read Triple-P, Aspbergers must read Aspergers
- The male-female ratio is rather 3:1 than 4:1 (Loomes et al., 2017)
- There are formal errors, such as blanks or inadequate symbols (e.g. in "Referencer")

3. Suggestions for modifications/amendments

I understand that there is perhaps not space or time for a major change, but I think the guidelines would benefit from revisions in accord with the points raised above. The core of the guideline, the 10 PICOS are overall very solid, although I have a bit of concern whether all treatments where addressed with the same degree of neutrality. It appears that that there was a lot of focus and effort on the 10 treatment evaluations that everything else remained incomplete. The treatment evaluations are not well embedded, and the guidance of the reader regarding the methods and approach is not sound. These parts need to be elaborated and clearer. There are also many points that should be doable without a lot of challenges: the minor points and some to f the specific ones.

References

- Bölte S. The good, the bad and systematic reviews. Autism. 2015;19(1):3-5. doi:10.1177/1362361314561393
- Loomes R, Hull L, Mandy WPL. What Is the Male-to-Female Ratio in Autism Spectrum Disorder? A Systematic Review and Meta-Analysis. J Am Acad Child Adolesc Psychiatry. 2017;56(6):466-474. doi:10.1016/j.jaac.2017.03.013

Professor of Child & Adolescent Psychiatric Science, Director of KIND, Karolinska Institutet, Stockholm, Sweden

Høringssvar fra Børne- og Ungdomspsykiatrisk selskab, BUP-DK.

Generelt er vi meget positive overfor indholdet i retningslinjen, vi har dog enkelte bemærkninger:

Det problematiseres, at retningslinjen ikke tager stilling til, hvem der skal udføre de anbefalede behandlinger. Man kunne overveje følgende ændring under afsnit 2 Indledningen: At benævne, at indsatsen overfor målgruppen varetages af fagprofessionelle (og ikke kun sundhedsprofessionelle), såvel i kommuner som i regionerne, evt af praktiserende speciallæger, som også er en del af primær sektor. Baggrunden for dette forslag er, at retningslinjen anbefaler såvel sprogtræning og sanseintegrationsbehandling, og det bør ikke forventes, at ansvaret for disse opgaver ligger i børne- og ungdomspsykiatrien. Her kunne man med fordel henvise til SST's forløbsprogrammer i øvrigt.

Vedr punkt 9 om anbefaling af 2. generations antipsykotika:
Vi er generelt positive overfor, at retningslinjen anbefaler nøje overvejelser ift at anvende antipsykotika som eksempelvis Risperidon i længere perioder grundet bivirkningsprofilen. Det må dog også anføres, at det er anerkendt praksis af benytte antipsykotika til meget aggressive og udadreagerende børn og unge. Varigheden af behandlingen kan være nødvendig i længere perioder, så længe bivirkningerne ikke overstiger effekten og det kan være ganske nødvendigt at forsætte antipsykotika ved særligt udadreagerende børn og unge mhp at sikre både deres og pårørendes sikkerhed, ligesom det ofte kan støtte meget ængstelige børn og unge med autisme i svær grad at opretholde et nogenlunde funktionsniveau.
Vi kan være bekymrede for, at teksten i retningslinjen vil skabe usikkerhed hos de familier, hvor den pædagogiske indsats ikke har effekt på trods af relevant afprøvning, og hvor medicin bliver påkrævet. I gennemgangen af litteraturen er der holdepunkt for at medicinen kan have bivirkninger (det er velkendt), men der er også opregning af positive effektmål selvom metoderne i disse undersøgelser kan kritiseres. Anbefalingen fra BUP-DK vil derfor være, at teksten revideres således at behandlingen kan anvendes/anbefales og opretholdes over længere tid, når der er klar indikation for den.

Med venlig hilsen

Linda Hardisty Bramsen

Formand for BUP-DK.