Early Childhood Personnel Center

NATURALISTIC INTERVENTIONS FOR YOUNG CHILDREN WITH DISABILTIES

This is a product of the Early Childhood Personnel Center (ECPC) awarded to the University of Connecticut Center for Excellence in Developmental Disabilities and was made possible by Cooperative Agreement #H325B170008 which is funded by the U.S. Department of Education, Office of Special Education Programs.

However, those contents do not necessarily represent the policy of the Department of Education, and you should not assume endorsement by the Federal Government. 263 Farmington Avenue, Farmington, CT 06030-6222 • 860.679.1500 • infoucedd@uchc.edu

Introduction

Naturalistic interventions involve using intentional and systematic strategies to instruct young children with disabilities during ongoing activities. Many terms (names) have been used to describe or refer to naturalistic interventions including: (a) incidental teaching (e.g., Hart & Risley, 1975), (b) embedded instruction (e.g., Horn et al., 2000; Neef et al., 1984; Snyder et al., 2018), (c) naturalistic instruction (e.g., Grisham-Brown et al., 2005; Halle et al., 1981), (d) milieu teaching (e.g., Kaiser et al., 2000; Yoder et al., 1995), (e) transition-based teaching (e.g., Werts et al., 1992; Wolery et al., 1993), and (f) activity-based intervention (e.g., Botts et al., 2014; Losardo & Bricker, 1994). Although these interventions have different and unique names, often reflecting nuances in intervention characteristics or components, we will use the term "naturalistic interventions" to refer to the class of interventions in this paper for clarity unless otherwise noted.

As a class of intervention techniques, naturalistic interventions incorporate a repertoire of strategies designed to effectively engage young children with disabilities. These strategies include environmental arrangements, incidental teaching, time delay, prompting, modeling, responsive interactions, following the child's lead, and expansions of children's communication (Franzone, 2009; Snyder et al., 2018). Empirical evidence has shown positive effects of naturalistic interventions for children's language and communication skills (e.g., Coogle et al., 2018; Hester et al., 1996; Kaiser et al., 1995) and play skills (e.g., Ingersoll & Schreibman, 2000; Stahmer, 1995). Naturalistic interventions have been recognized as a recommended practice for young children with disabilities for many years (e.g., Odom & McLean, 1996; Sandall et al., 2000; Division for Early Childhood,

2014). Given the history of naturalistic interventions as a recommended practice for young children with disabilities, multiple reviews of the intervention technique have been conducted over the years.

One method of reviewing an accumulating body of evidence from multiple systematic reviews of interventions is an overview of reviews. An overview of reviews, also referred to as umbrella reviews and meta-reviews (e.g., Hennessy et al., 2019; Pollock et al., 2023), is an explicit and systematic method to summarize the current evidence on a subject in extant systematic reviews. Key elements (e.g., Gates et al., 2022) of an overview of reviews encompass the formulation of precise research inquiries; transparent and replicable procedures for identifying systematic reviews that meet predetermined inclusion criteria; and well-defined procedures for data extraction, analysis, and presentation of findings (Pollock et al., 2016). Pivotal data to be prioritized during the execution of an overview of reviews includes descriptive attributes of the systematic review methods, attributes and outcomes of the primary studies included in within and across reviews, assessments of the risks of bias of included reviews, and the degree of overlap among primary studies across systematic reviews (e.g., Gates et al., 2020; Hennessy et al., 2020). The application of overview of reviews methodology has been more frequently used in the fields of healthcare and psychology for assessing intervention efficacy (e.g., Becker & Oxman, 2011; Lecomte et al., 2020). Leveraging the methodological rigor inherent in an overview of reviews, our study employs this approach to investigate the impacts of naturalistic interventions on young children with disabilities by synthesizing and summarizing evidence from multiple existing reviews of the intervention technique.

We conducted an overview of reviews of naturalistic interventions for young children with or at risk of developmental disabilities to examine the following research questions:

(1) How have naturalistic interventions been operationalized in extant systematic reviews of the intervention method?;

(2) What are the characteristics of the methods, participants, interventions, and outcomes in systematic reviews of naturalistic interventions?;

(3) What intervention components or characteristics have been described in extant reviews of naturalistic interventions?; and

(4) What conclusions have extant systematic reviews (and meta-analyses) drawn on the effects of naturalistic interventions for young children with disabilities?

Method

Overview of Reviews Methodology

We conducted an overview of reviews of naturalistic interventions for young children with or at risk for delays or disabilities. This overview was conducted using contemporary guidelines for overview of reviews (e.g., Gates et al., 2020; Lunny et al., 2018; Pollock et al., 2023) and is reported consistent with contemporary standards set forth in the Preferred Reporting Items for Overview of Reviews (Gates et al., 2022).

Selection (Inclusion) Criteria

We included systematic reviews that reviewed naturalistic interventions that included at least one child with or at risk for delays or disabilities under the age of 5 years old. To meet the inclusion criteria for this overview, the primary focus of the review had to

have been explicitly stated as a review of naturalistic interventions (or a review of naturalistic interventions using the related terms outlined prior). To help locate reviews focused on naturalistic interventions as we defined the class of interventions in the introduction, we developed an inclusion criteria that specified one or both of the following must be met by each included review: (1) The title of the review included a term related to naturalistic instruction (i.e., naturalistic instruction, naturalistic teaching, naturalistic intervention, embedded instruction, embedded intervention, embedded teaching, incidental instruction, incidental teaching, incidental intervention, activity-based teaching, activity-based instruction, activity-based interventions, milieu teaching, milieu therapy, milieu interventions, or milieu instruction) or (2) the purpose of the review must have explicitly stated that the purpose of the review was to synthesize naturalistic interventions (or a related term). All included reviews also had to meet the following inclusion criteria: inclusion of one or more primary studies that utilized group comparative designs or single-case experimental designs; utilization of systematic review methods, including a replicable search strategy and selection criteria; published in a peer-reviewed journal; and written in English. We did not place a restriction on date of publication for a review to be included in this overview.

Selection Methods

We searched Medline, APA PsycINFO, Education Resource Information Center (ERIC), Cumulative Index of Nursing and Allied Health Libraries (CINAHL), and Academic Search Premier on December 19, 2023 using the search strategy shown in Supplemental Text 1. We also used “snowball methods” as recommended by Greenhalgh and Peacock

(2005) by searching titles from the reference lists of included reviews and searching for articles that had cited the included reviews. We exported the records from the electronic database searches into Covidence (Veritas Health Innovation, 2020) for screening and selection. Two reviewers independently screened records by title and abstract based on eligibility criteria, with disagreements resolved through consensus. The remaining records were then screened at the full-text stage, in which the same two screeners independently screened the full text of each record against the eligibility criteria. Disagreements between reviewers were resolved through discussion with a third party.

Data Collection

It is convention in systematic review methods (e.g., Cooper et al., 2019) to treat a "study" (i.e., an experimental comparison, not the primary publication or article) as the unit of analysis. In four of seven reviews (57%; Dubin & Lieberman-Betz, 2020; Lane et al., 2016; Lane et al., 2023; Snyder et al., 2015) reporting a study as the unit of analysis, there were more studies than primary publications (i.e., articles). For this overview, we extracted data for this overview, where necessary, with studies as the unit of analysis unless otherwise noted, and indicate a study in this manuscript with the symbol "*u*." Consistent with methodological standards for overview of reviews (e.g., Pollock et al., 2023), data were extracted primarily from the data reported in the published systematic reviews; when necessary, we examined the primary sources (i.e., primary publications and studies) to confirm or extract specific or missing data for some variables. For all data extraction, two reviewers extracted the data independently, with disagreements resolved through discussion and consensus. We extracted data on research characteristics (e.g., number of

primary studies, primary study research design), participant characteristics (e.g., number, age range, developmental characteristics), intervention characteristics (e.g., intervention name, setting, intervention agent, duration), and outcomes and results (e.g., dependent measure, number of participants showing positive treatment effect, effect size estimate).

To assess the risks of bias and rigor of the included reviews, we used the Johanna Briggs Institute's (JBI) Critical Appraisal Checklist for Systematic Reviews and Research Syntheses (Aromataris et al., 2015). The JBI Checklist contains 11 items that helps to assess the methodological rigor of a review and the extent to which the review has potentially addressed possible causes of bias. To complete the Checklist, two reviewers independently evaluated the 11 items for each review with disagreements resolved through discussion and reaching consensus. Data were analyzed descriptively by creating a summary figure across reviews for each of the 11 items on the JBI Checklist.

Data Analyses and Syntheses

We used the corrected covered area (CCA; Pieper et al., 2014) to quantify the degree of study overlap across included reviews. To calculate CCA, current recommendations are to use "primary publications" (i.e., articles; Pieper et al., p. 370) as the unit of analysis/calculation. The CCA was calculated as $CCA = \frac{N-u}{uc-u}$, where *N* was the number of included primary publications (including double counting), *u* was the number of primary publications (excluding duplicated reports), and *c* was the number of systematic reviews. We used Pieper and colleagues' guidelines for quantifying the level of CAA for slight (0 – 5%), moderate (5 – 10%), high (10 – 15%), or very high (> 15%) levels of overlap.

We also used graphical methods (e.g., Bougioukas et al., 2021; Bracchiglione et al., 2022) to explore overlap further.

We conducted descriptive and narrative syntheses of the outcomes reported in the seven included reviews. To synthesize the findings of the reviews, we first summarized the percentage of primary studies reporting positive findings within each review and aggregated this across the seven reviews. We then examined the study that used statistical synthesis to examine the average effects across studies reported in their analyses. Finally, we formulated conclusions regarding the evidence across reviews by exploring patterns in the aggregated data.

Results

Review Selection

The electronic database search yielded 2,945 records; 2,299 remained after removing 646 duplicates in Covidence. After screening out irrelevant records through title/abstract screening, 46 records remained. We screened the full text of these 46 records, of which seven met our inclusion criteria (Dubin & Lieberman-Betz, 2020; Gulboy et al., 2023; Lane et al., 2016; Lane et al., 2023; Rakap & Parlak-Rakap, 2011; Rakap & Rakap, 2014; Snyder et al., 2015). Our snowball selection process involved screening the titles of 609 articles that were included in the reference lists or had cited the seven included reviews. Examination of these titles yielded 32 additional records for full-text screening, with zero additional reviews meeting inclusion criteria.

Characteristics of Included Reviews

Inclusion criteria. The inclusion criteria of a systematic review relate directly to which primary studies are included in a review. Table 1 shows the inclusion criteria for each review, with the descriptions of the intervention quoted to provide maximum details of this critical component of each review. Given the specific language and terminology used to describe the intervention inclusion criteria, a summative synthesis of this was unable to be conducted; readers should see Table 1 for details regarding the specific way in which naturalistic interventions were defined in each review.

As shown in Table 1, all seven reviews each had inclusion criteria specifying a study must have been an empirical examination of the intervention practice published in a peer-review journal to be included in each review. Two of seven (29%) of reviews specified articles must have been published in English to be included in the review, however, all primary studies across reviews were published in English across all seven reviews, even the five in which this was not listed as an inclusion criteria. All seven reviews also had inclusion criteria related to participant characteristics, with all reviews specifying at least some of the children in each study had to be a child with or at risk of a disability or delay. Three reviews (Rakap & Parlak-Rakap, 2011; Rakap & Rakap, 2014; Snyder et al., 2015) specified children had to be under the age of 5 years, while three reviews (Dubin & Lieberman-Betz, 2020; Gulboy et al., 2023; Lane et al., 2016) indicating a maximum age between 6 and 9 years, and one review (Lane et al., 2023) specifying children needed to be between 3 and 12 years old to be included in the review. Finally, six of seven reviews (86%) contained an inclusion criterion related to the intervention delivery setting. Three reviews (Gulboy et al., 2023; Lane et al., 2016; Rakap & Parlak-Rakap, 2011) indicated interventions

must have occurred in inclusive classroom settings to be included in their reviews; two reviews (Lane et al., 2023; Snyder et al., 2015) and one review, the Rakap and Rakap (2014) review of parent implemented naturalistic interventions specified the intervention must have occurred in a natural setting or in a clinical setting if within the context of a developmentally appropriate activity.

Review characteristics. Characteristics of the seven included reviews are shown in Table 2. Most reviews (6 of 7; 86%) did not contain statistical syntheses across included studies; only one review (14%; Gulboy et al., 2023) included a statistical or meta-analytic synthesis across primary study findings. All reviews were published since 2010; four reviews (56%) were published in the decade of 2010-2019 and three reviews (43%) were published in the current decade (2020-current), with the two most current reviews (i.e., Gulboy et al., 2023; Lane et al., 2023) completed in the year in which this overview was conducted. With respect to the search methods, four reviews (57%) contained a date on which the searches were conducted and three reviews (43%) omitted this information.

Across the seven reviews, there were 130 cumulative primary publications[1]. This count (130) represents a gross count of primary publications inclusive of primary publications that were included in more than one review; the mean number of primary publications per review was 18.6 with a range of 10 to 37 primary publications per review. The total number of unique (unduplicated) primary publications was 102 (u = 102). Twenty-two primary publications (22%) were included in more than one review; 16 primary

[1] As stated in the method, the convention of examination of overlap in overview of reviews is conducted at the primary publication level, thus these figures represent publications, not studies, as with most of the other analyses reported in this overview.

publications were included in two reviews and six primary publications (Christensen-Sandfort & Whinnery, 2013; Daugherty et al., 2001; Fox & Hanline, 1993; Grisham-Brown et al., 2009; Harjusola-Webb & Robbins, 2012; Horn et al., 2000) were included in three of the seven reviews (43%). The amount of overlap estimated by the corrected covered area (CCA) was approximately 4.58%, indicating slight overlap (see Figure 2 for a GROOVE summary citation matrix and Supplemental Table 1 for a publication-level citation matrix). As seen in Figure 2, the majority overlap (12 of 21; 57%) of pairwise overlap comparisons between reviews had zero overlapping primary publications. There was moderate overlap between three sets of reviews (9.8%. - Lane et al., 2023 and Snyder et al., 2015; 9.3% - Gulboy et al., 2023 and Snyder et al., 2015; 8.3% - Lane et al., 2016 and Rakap & Rakap, 2014) and very high overlap between two sets of reviews (35.9% - Snyder et al., 2015 and Rakap & Parlak-Rakap, 2011; 23.8% - Gulboy et al., 2023 and Rakap & Parlak-Rakap, 2011).

Figure 3 shows a representation of the average ratings across reviews for the 11 JBI Appraisal Checklist items. The results of the appraisal suggest that the seven reviews had some risks of bias. As shown in Figure 3, two reviews (Lane et al., 2023; Snyder et al., 2015) were rated as not having any of the 11 JBI Checklist risks of bias or methodological concerns for the items for which we could assess (all items except those relating to meta-analytic syntheses). Across the other reviews, the JBI Checklist item for which bias had the highest risk was the provision of a clear description of the search strategy, with four reviews having concerns; three reviews were rated unclear (43%), and one review (14%) was rated as not meeting methodological criteria for this item. No other checklist criteria were rated as a risk of bias for more than one of the seven included reviews.

Participant characteristics. Table 3 presents the characteristics of child participants across the primary studies in the included reviews. There were 907 children with or at risk of a disability in 171 studies (Mean = 5.3 participants per study) from 102 unique publications. The age range of participants across reviews was reported to be 12 months and 156 months, with a gender distribution, when reported, of 552 males (77%) and 168 females (23%; gender was not reported for 187 children in 66 studies). Across reviews, the most common disability or diagnostic category was autism spectrum disorder. Studies also reported the involvement of children with other disabilities including attention deficit/hyperactivity disorder (ADHD), Down syndrome, intellectual and developmental disabilities, multiple disabilities, physical disability, pervasive developmental disorder, speech language impairment, visual impairment, cerebral palsy, spastic quadriplegia, chronological disease, pervasive developmental disorder-not otherwise specified, and other health impairment.

Table 4 provides an overview of the characteristics of adult participants derived from the 146 primary studies included in six reviews (Gulboy et al., 2023; Lane et al., 2016; Lane et al., 2023; Rakap & Parlak-Rakap, 2011; Rakap & Rakap, 2014; Snyder et al., 2015). These studies included 374 adults. Two reviews reported the mean age of adults as 33 (Rakap & Rakap, 2014) and 30 years old (Snyder et al., 2015), with the age range spanning from 20 to 55 years across reviews. Gender was reported in three reviews for 127 participants (127 of 374; 34%), with 115 females (91%) and 12 males (9%). The highest level of education achieved for the adult participants was reported in four reviews (Rakap & Parlak-Rakap, 2011; Rakap & Rakap, 2014; Snyder et al., 2015; Lane et al., 2023),

with the majority of adult participants (more than 50% in each review) having completed at least some college coursework; 68% of participants were reported as having achieved a bachelor's degree or higher. The professional roles of the adult participants varied, with the predominant roles being teachers (*n* = 122; 38%), caregivers (*n* = 74; 23%), and paraprofessionals (*n* = 58; 18%). Additional insights into the participants' experiences were provided by two reviews (Snyder et al., 2015; Lane et al., 2023), revealing the adult participants had an average of 7.7 years of experience (Snyder et al., 2015) with a range from 1 to 30 years or 0 to 22 years, respectively, in the Lane et al. and Snyder et al. reviews.

Intervention characteristics. Table 5 provides details on the intervention characteristics of the 171 studies from the 102 unique primary publications included across reviews. Across the primary studies, a range of terms were used to characterize naturalistic intervention, including embedded instruction/intervention (*u* = 41; 24%), milieu teaching/therapies (*u* = 26; 15%), naturalistic intervention (*u* = 22; 13%), incidental teaching (u = 5; 3%), and activity-based instruction (*u* = 3; 2%). Seventy-four studies (43%) used a term or name other than the five names listed in the previous sentence in which to refer to the naturalistic intervention. Most of the studies (118 of 171; 69%) were conducted in schools; 27 studies (16%) were conducted in clinics; 10 studies (6%) were conducted in home; 9 studies (5%) were conducted in multiple settings; 7 studies (4%) were conducted in other settings. There was great variability across studies with respect to the activity context where interventions occurred. Majority of the intervention occurred in center/free play (*u* = 82; 48%). Other activity contexts included mealtime (*u* = 15; 9%), small group (*u* = 21; 12%), routines (*u* = 6; 4%), and circle time (*u* = 4; 2%). The intervention density of

naturalistic intervention was reported differently across reviews, but the overall density was low.

Effects of naturalistic interventions. As shown in Table 6, all of the included reviews showed positive effects on skill acquisition for a majority of the participants represented in the primary studies included in each review. Examination of the primary end point (immediately after intervention) showed five reviews reported greater than 90% of participants in single case experimental design studies acquired the target skills, with an additional study showing 80% of participants acquired the target skills. The effects on generalization and maintenance, although reported in fewer studies, were also strong. For the four reviews reporting generalization for studies conducted using single case experimental designs, all five reviews showed at least 75% of the participants in the studies demonstrated generalization and at least 90% of participants demonstrated maintenance of skill acquisition in the four reviews reporting this endpoint. The effects from group experimental design studies were somewhat less robust, with the one review analyzing these studies separately (Dubin & Lieberman-Betz, 2020) showing 54% of the studies with sufficient rigor showed small to large positive effects, with good generalization and poor maintenance across the included studies.

Discussion

As shown in the included reviews, the majority of studies that have examined naturalistic instructions have been conducted using single case experimental designs and examined the effects of the intervention on communication and language skills, social-

emotional competence, and pre-academic or cognitive skills. An overwhelming majority of the studies included in the reviews had positive findings.

All of the reviews included in this overview were published since 2011. Earlier reviews of naturalistic interventions have been published (e.g., Halle et al., 1984, Hepting & Goldstein, 1996; Charlop-Christy et al., 1999), but did not meet the inclusion criteria of this overview in that a review must be systematic and replicable (with an operationalized search strategy). It is noteworthy that all of the reviews included in this overview were published in the past 15 years given that some of the first empirical reports of a naturalistic intervention, incidental teaching (e.g., Hart & Risley, 1975), were published nearly 50 years ago and textbooks describing the naturalistic intervention methods for practitioners have been available for over 30 years (e.g., Bricker & Cripe, 1992). Examination of the authors of the included reviews shows that one author (Rakap) was involved in over one-half (4 of 7, 57%) of the included reviews (Gulboy et al., 2023; Rakap & Parlak-Rakap, 2011; Rakap & Rakap, 2014; Snyder et al., 2015). Three researchers (Lane, Lieberman-Betz, Serife Rakap) were each involved in two reviews. However, even with the considerable overlap in review authors, the overall overlap of primary publications, as calculated by CCA, was low. As shown in Figure 2, the overlap between many of the reviews was zero percent. New reviews conducted by researchers outside of this circle will help increase the generality of the findings and elude possible biases (perceived or real) that may be present.

With continued attention to providing evidence-based instructional methods in inclusive settings, it is likely that additional reviews will be necessary to continue synthesizing the available evidence so that advances in research and practice can be

made. Differences in inclusion criteria may also have contributed to the small amounts of overlap in primary publications between reviews. Future systematic reviews should also consider, if deemed appropriate, the use of meta-analytic tools to provide summary effects and to explore if data are available that might provide details on the most effective methods or components of naturalistic interventions, on average.

Many terms or names have been used to describe naturalistic intervention techniques. While different interventions do have variations, these variations are often slight or nuanced and may not be an important distinction in practice. Furthermore, the multiple names given to a very similar set of intervention techniques is likely to cause confusion in both research and practice, hindering the advancement of the science and utilization of the intervention practices. To ensure that practitioners have the training and skills necessary to implement these evidence-based practices, a common terminology should be used moving forward.

Limitations

First, as shown by the CCA, there was minimal overlap of primary studies across the seven included reviews. Given the ubiquity of naturalistic interventions for young children with disabilities, it was surprising that such little overlap was found. With respect to using the CCA to examine study overlap, it is important to recall that this metric was calculated on primary publications (i.e., articles) but the level of analysis in most reviews, including this one, is at the study level. Given the treatment of studies within publications was not consistent across reviews, we were not able to examine the overlap of studies across reviews, and this disconnect should be considered a limitation. In addition, the influence

of extreme levels of overlap in an overview of reviews with relatively small sample of reviews is known to exist and may be present in the sample of reviews included in this overview. Issues regarding the overlap of studies within an overview of reviews has been identified as a methodological limitation of the method (e.g., Hennessy & Johnson, 2020) and should be considered when drawing conclusions from the findings of this review.

Second, we did not find consistent reporting or analysis of intervention components or active ingredients across reviews. Future work, both primary studies and systematic reviews, should aim to provide detailed information on the components of naturalistic instruction so that more fine-tuned findings about the effects and possible mechanisms underlying the effects can be made. Third, while naturalistic interventions have typically been used to help children acquire new appropriate skills, some studies have examined the effects of naturalistic instructions on challenging behavior (e.g., Mancil et al., 2009; Rakap & Balikci, 2016; Sigafoos et al., 2006). While some studies have explored the effects of the intervention on reducing challenging behavior, much less is known in this area and is worth future exploration. Finally, it was notable that social validity was examined in only one review (Rakap & Parlak-Rakap, 2011). Given the use of single case experimental designs in the extant literature, one might expect a greater emphasis reviewing the findings of the assessment of social validity in reviews of studies of naturalistic interventions. Future systematic reviews might look at this in greater detail to ensure that the intervention, which should have very high social validity, is seen as a socially valid practice.

Conclusion

Seven methodologically rigorous reviews were included in this overview of reviews of naturalistic interventions. Across these seven reviews, 102 primary publications with hundreds of child participants were included and synthesized, with the majority of the reviews showing greater than 90% of participants demonstrated skills acquisition associated with a naturalistic intervention. Collectively, the findings from this overview provides additional evidence of the positive and robust effects of the intervention for a majority of young children with or at risk of disabilities or delays.

References

* References with an asterisk indicate review included in overview

Aromataris, E., Fernandez, R., Godfrey, C. M., Holly, C., Khalil, H., & Tungpunkom, P. (2015). Summarizing systematic reviews: Methodological development, conduct and reporting of an umbrella review approach. *International Journal of Evidence-based Healthcare, 13*(3), 132-140. https://doi.org/10.1097/XEB.0000000000000055

Becker, L., & Oxman, A. (2011). Overviews of reviews. In J. P. T. Higgins & S. Green (Eds.), *Cochrane handbook for systematic reviews of interventions* (Version 5.1). Cochrane Collaboration.

Botts, D. C., Losardo, A. S., Tillery, C. Y., & Werts, M. G. (2014). A comparison of activity-based intervention and embedded direct instruction when teaching emergent literacy skills. *The Journal of Special Education*, *48*(2), 120–134. https://doi.org/10.1177/0022466912449652

Bougioukas, K. I., Vounzoulaki, E., Mantsiou, C. D., Savvides, E. D., Karakosta, C., Diakonidis, T., Tsapas, A., & Haidich, A. B. (2021). Methods for depicting overlap in overviews of systematic reviews: An introduction to static tabular and graphical displays. *Journal of Clinical Epidemiology, 132,* 34-45. https://doi.org/10.1016/j.jclinepi.2020.12.004

Bracchiglione, J., Meza, N., Bangdiwala, S. I., de Guzman, E. N., Urrutia, G., Bonfill, X., & Madrid, E. (2022). Graphical representation of overlap of overviews: GROOVE tool. *Research Synthesis Methods, 13,* 381-388. https://doi.org/10.1002/jsrm.1557

Bricker, D., & Cripe, J. (1992). *An activity-based approach to early intervention*. Brookes.

Charlop-Christy, M. H., LeBlanc, L. A., & Carpenter, M. H. (1999). Naturalistic teaching strategies (NATS) to teach speech to children with autism: Historical perspective, development, and current practice. *The California School Psychologist, 4,* 30-46.

Christensen-Sandfort, R. J., & Whinnery, S. B. (2013). Impact of milieu teaching on communication skills of young children with autism spectrum disorder. *Topics in Early Childhood Special Education, 32*(4), 211-222. https://doi.org/10.1177/0271121411404930

Coogle, C., Ottley, J. R., Rahn, N. L., & Storie, S. (2018). Bug-in-ear eCoaching: Impacts on novice early childhood special education teachers. *Journal of Early Intervention*, *40*(1), 87–103. https://doi.org/10.1177/1053815117748692

Cooper, H., Hedges, L. V., & Valentine, J. C. (Eds.). (2019). *The handbook of research synthesis and meta-analysis* (3rd ed.). Russell Sage.

Daugherty, S., Grisham-Brown, J., & Hemmeter, M. L. (2001). The effects of embedded skill instruction on the acquisition of target and nontarget skills in preschoolers with developmental delays. *Topics in Early Childhood Special Education, 21*(4), 213-221. https://doi.org/10.1177/112140102100402

Division for Early Childhood. (2014). *DEC recommended practices in early intervention/early childhood special education 2014*. http://www.dec-sped.org/dec-recommended-practices

*Dubin, A. H., & Lieberman-Betz, R. G. (2020). Naturalistic interventions to improve prelinguistic communication for children with autism spectrum disorder: A

systematic review. *Review Journal of Autism and Developmental Disorders, 7,* 151-167. https://doi.org/10.1007/s40489-019-00184-9

Fox, L., & Hanline, M. F. (1993). A preliminary evaluation of learning within developmentally appropriate early childhood settings. *Topics in Early Childhood Special Education, 13*(3), 308-327. https://doi.org/10.1177/027112149301300308

Franzone, E. (2009). *Overview of naturalistic intervention*. National Professional Development Center on Autism Spectrum Disorders. https://autismpdc.fpg.unc.edu/sites/autismpdc.fpg.unc.edu/files/imce/documents/Naturalistic-Intervention-Complete10-2010.pdf

Gates, M., Gates, A., Guitard, S., Pollock, M., & Hartling, L. (2020). Guidance for overviews of reviews continues to accumulate, but important challenges remain: A scoping review. *Systematic Reviews, 9*, 254. https://doi.org/10.1186/s13643-020-01509-0

Gates, M., Gates, A., Pieper, D., Fernandes, R. M., Tricco, A. C., Moher, D., Brennan, S. E., Li, T., Pollock, M., Lunny, C., Sepúlveda, D., McKenzie, J. E., Scott, S. D., Robinson, K. A., Matthias, K., Bougioukas, K. I., Fusar-Poli, P., Whiting, P., Moss, S. J., & Hartling, L. (2022). Reporting guideline for overviews of reviews of healthcare interventions: Development of the PRIOR statement. *British Medical Journal, 378*, e070849. https://doi.org/10.1136/bmj-2022-070849

Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Exceptional Children, 71*(2), 149-164. https://doi.org/10.1177/001440290507100202

Greenhalgh, T., & Peacock, R. (2005). Effectiveness and efficiency of search methods in systematic reviews of complex evidence: Audit of primary sources. *British Medical Journal, 331,* 1064–1065. https://doi.org/10.1136/bmj.38636.593461.68

Grisham-Brown, J., Hemmeter, M. L., & Pretti-Frontczak, K. (2005). *Blended practices for teaching young children in inclusive settings* (2nd ed.). Paul H. Brookes Publishing.

Grisham-Brown, J., Pretti-Frontczak, K., Hawkins, S. R., & Winchell, B. N. (2009). Addressing early learning standards for all children within blended preschool classrooms. *Topics in Early Childhood Special Education, 29*(3), 131-142. https://10.1177/0271121409333796

*Gulboy, E., Yucesoy-Ozkan, S., & Rakap, S. (2023). Embedded instruction for young children with disabilities: A systematic review and meta-analysis of single-case experimental research studies. *Early Childhood Research Quarterly, 63*, 181-193. https://doi.org/10.1016/j.ecresq.2022.12.014

Halle, J. W., Alpert, C. L., & Anderson, S. R. (1984). Natural environment language assessment and intervention with severely impaired preschoolers. *Topics in Early Childhood Special Education, 4*(2), 36-56. https://doi.org/10.1177/027112148400400204

Halle, J. W., Baer, D. M., & Spradlin, J. E. (1981). Teachers' generalized use of delay as a stimulus control procedure to increase language use in handicapped children. *Journal of Applied Behavior Analysis*, *14*(4), 389–409. https://doi.org/10.1901/jaba.1981.14-389

Harjusola-Webb, S., & Robbins, S. H. (2012). The effects of teacher-implemented naturalistic intervention on the communication of preschoolers with autism. *Topics in Early Childhood Special Education, 32*(2), 99-110. https://doi.org/10.1177/0271121410397060

Hart, B. M., & Risley, T. R. (1975). Incidental teaching of language in the preschool. *Journal of Applied Behavior Analysis, 8*(4), 411-420. https://doi.org/10.1901/jaba.1975.8-411

Hennessy, E. A., & Johnson, B. T. (2020). Examining overlap of included studies in meta-reviews: Guidance for using the corrected covered area index. *Research Synthesis Methods, 11*(1), 134–145. https://doi.org/10.1002/jrsm.1390

Hennessy, E. A., Johnson, B. T., & Keenan, C. (2019). Best practice guidelines and essential methodological steps to conduct rigorous and systematic meta-reviews. *Applied Psychology: Health and Well-Being*, *11*(3), 353–381. https://doi.org/10.1111/aphw.12169

Hepting, N. H., & Goldstein, H. (1996). What's natural about naturalistic language intervention? *Journal of Early Intervention, 20*(3), 249-265. https://doi.org/10.1177/105381519602000308

Hester, P. P., Kaiser, A. P., Alpert, C. L., & Whiteman, B. (1996). The generalized effects of training trainers to teach parents to implement milieu teaching. *Journal of Early Intervention*, *20*(1), 30–51. https://doi.org/10.1177/105381519602000105

Horn, E., Lieber, J., Li, S., Sandall, S., & Schwartz, I. (2000). Supporting young children's IEP goals in inclusive settings through embedded learning opportunities. *Topics in Early*

Childhood Special Education, *20*(4), 208–223. https://doi.org/10.1177/027112140002000402

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children, 71*(2), 165-180. https://doi.org/10.1177/001440290507100203

Ingersoll, B., & Schreibman, L. (2006). Teaching reciprocal imitation skills to young children with autism using a naturalistic behavioral approach: Effects on language, pretend play, and joint attention. *Journal of Autism and Developmental Disorders*, *36*(4), 487–505. https://doi.org/10.1007/s10803-006-0089-y

Kaiser, A. P., Hancock, T. B., & Nietfeld, J. P. (2000). The effects of parent-implemented enhanced milieu teaching on the social communication of children who have autism. *Early Education & Development*, *11*(4), 423–446. https://doi.org/10.1207/s15566935eed1104_4

Kaiser, A. P., Hester, P. P., Alpert, C. L., & Whiteman, B. C. (1995). Preparing parent trainers: An experimental analysis of effects on trainers, parents, and children. *Topics in Early Childhood Special Education*, *15*(4), 385–414. https://doi.org/10.1177/027112149501500401

Kaiser, A. P., & Trent, A. J. (2007). Communication intervention for young children with disabilities: Naturalistic approaches to promoting development. In S. L. Odom, R. H. Horner, M. E. Snell, & J. Blacher (Eds.), *Handbook of developmental disabilities* (pp. 224-246). Guilford Press.

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M, & Shadish, W. R. (2010). *Single-case designs technical documentation.* Retrieved from http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf

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

*Lane, J. D., Graley, D., Shepley, C., & Lynch, K. M. (2023). Systematic review of naturalistic language interventions in schools: Child- and adult-level outcomes for verbal communication. *Remedial and Special Education, 44*(4), 319-331. https://doi.org/10.11177/07419325221125887

*Lane, J. D., Lieberman-Betz, R., & Gast, D. L. (2016). An analysis of naturalistic interventions for increasing spontaneous expressive language in children with autism spectrum disorder. *The Journal of Special Education, 50*(1), 49-61. https://doi.org/10.1177/0022466915614837

Lecomte, T., Potvin, S., Corbière, M., Guay, S., Samson, C., Cloutier, B., Francoeur, A., Pennou, A., & Khazaal, Y. (2020). Mobile apps for mental health Issues: Meta-review of meta-analyses. *JMIR mHealth and uHealth*, *8*(5), e17458. https://doi.org/10.2196/17458

Ledford, J. R., Chazin, K. T., Lane, J. D., Zimmerman, K. N., Bennett, P. B., & Ayres, K. A. (2023, May). Single case analysis and review framework (SCARF). Retrieved from: http://ebip.vkcsites.org/scarfv2

Ledford, J. R., Lane, J. D., & Tate, R. (2018). Evaluating quality and rigor in single case research. In J. R. Ledford & D. L. Gast (Eds.), *Single case research in behavioral sciences* (3rd ed., pp. 365–392). Routledge.

Losardo, A., & Bricker, D. (1994). Activity-based intervention and direct instruction: A comparison study. *American Journal of Mental Retardation*, *98*(6), 744–765.

Lunny, C., Brennan, S. E., McDonald, S., & McKenzie, J. E. (2018). Toward a comprehensive evidence map of overview of systematic review methods: Paper 2—risk of bias assessment; synthesis, presentation and summary of the findings; and assessment of the certainty of the evidence. *Systematic Reviews, 7,* 159. https://doi.org/10.1186/s13643-018-0784-8

Mancil, G. R., Conroy, M. A., & Haydon, T. F. (2009). Effects of a modified milieu therapy intervention on the social communicative behaviors of young children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39*(1), 149-163. https://doi.org/10.1007/s10803-008-0613-3

Neef, N. A., Walters, J., & Egel, A. L. (1984). Establishing generative yes/no responses in developmentally disabled children. *Journal of Applied Behavior Analysis*, *17*(4), 453–460. https://doi.org/10.1901/jaba.1984.17-453

Odom, S. L., & McLean, M. E. (1996). *Early intervention/early childhood special education: Recommended practices.* Pro-Ed.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *British Medical Journal, 372,* n71. https://doi.org/10.1136/bmj.n71

Pieper, D., Antoine, S.-L., Mathes, T., Neugebauer, E. A. M., & Eikermann, M. (2014). Systematic review finds overlapping reviews were not mentioned in every other overview. *Journal of Clinical Epidemiology*, *67*(4), 368–375. https://doi.org/10.1016/j.jclinepi.2013.11.007

Pollock, M., Fernandes, R. M., Becker, L. A., Featherstone, R., & Hartling, L. (2016). What guidance is available for researchers conducting overviews of reviews of healthcare interventions? A scoping review and qualitative metasummary. *Systematic Reviews, 5,* 190. https://doi.org/10.1186/s13643-016-0367-5

Pollock, M., Fernandes, R., Becker, L., Pieper, D., & Hartling, L. (2023). Overviews of reviews. In J. P. T. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page, & V. A. Welch (Eds.), *Cochrane handbook for systematic reviews of interventions* (Version 6.4). www.training.cochrane.org/handbook

Rakap, S., & Balikci, S. (2016). Using embedded instruction to teach functional skills to a preschool child with autism. *International Journal of Developmental Disabilities, 63*(1), 17–26. https://doi.org/10.1080/20473869.2015.1109801

*Rakap, S. & Parlak-Rakap, A. (2011). Effectiveness of embedded instruction in early childhood special education: A literature review. *European Early Childhood Education Research Journal, 19*(1), 79-96. https://doi.org/10.1080/1350293X.2011.548946

*Rakap, S., & Rakap, S. (2014). Parent-implemented naturalistic language interventions for young children with disabilities: A systematic review of single-subject experimental

research studies. *Educational Research Review, 13,* 35-51. https://doi.org/10.10116/j.edurev.2014.09.001

Sandall, S., McLean, M. E., & Smith, B. J. (2000). *DEC recommended practices in early intervention/early childhood special education*. Sopris West.

Sigafoos, J., O'Reilly, M., Ma, C. H., Edrisinha, C., Cannella, H., & Lancioni, G. E. (2006). Effects of embedded instruction versus discrete-trial training on self-injury, correct responding, and mood in a child with autism. *Journal of Intellectual & Developmental Disability*, *31*(4), 196–203. https://doi.org/10.1080/13668250600999160

Snyder, P., Hemmeter, M. L., McLean, M., Sandall, S., McLaughlin, T., & Algina, J. (2018). Effects of professional development on preschool teachers' use of embedded instruction practices. *Exceptional Children, 84*(2), 213-222. https://doi.org/10.1177/0014402917735512

*Snyder, P. A., Rakap, S., Hemmeter, M. L., McLaughlin, T. W., Sandall, S., & McLean, M. E. (2015). Naturalistic instructional approaches in early learning: A systematic review. *Journal of Early Intervention, 37*(1), 69-97. https://doi.org/10.1177/1053815115595461

Stahmer, A. C. (1995). Teaching symbolic play skills to children with autism using pivotal response training. *Journal of Autism and Developmental Disorders*, *25*(2), 123–141. https://doi.org/10.1007/BF02178500

Tate, R. L., McDonald, S., Perdices, M., Togher, L., Schultz, R., & Savage, S. (2008). Rating the methodological quality of single-subject designs and n-of-1 trials: Introducing

the Single-Case Experimental Design (SCED) Scale. *Neuropsychological Rehabilitation, 18*, 385–401. https://doi.org/10.1080.09602010802009201

Veritas Health Innovation. (2020). *Covidence systematic review software* [Computer software]. www.covidence.org

Werts, M., Wolery, M., Holcombe-Ligon, A., Vassilaros, M., & Billings, S. (1992). Efficacy of transition-based teaching with instructive feedback. *Education and Treatment of Children*, *15*(4), 320–334. https://www.jstor.org/stable/42899280

What Works Clearinghouse. (2012). *Procedures and standards handbook* (Version 3.0). Retrieved from https://ies.ed.gov/ncee/wwc/docs/referenceresources/wwc_procedures_v3_0_standards_handbook.pdf

What Works Clearinghouse. (2017a). *Key criteria used in WWC reviews of single-case design research*. https://ies.ed.gov/ncee/wwc/Docs/ReferenceResources/wwc_scd_key_criteria_011017.pdf

What Works Clearinghouse. (2017b). *Standards handbook* (Version 4.0). U.S. Department of Education. https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_standards_handbook_v4.pdf

Wolery, M., Doyle, P. M., Gast, D. L., Ault, M. J., & Simpson, S. L. (1993). Comparison of progressive time delay and transition-based teaching with preschoolers who have

developmental delays. *Journal of Early Intervention*, *17*(2), 160–176. https://doi.org/10.1177/105381519301700207

Yoder, P. J., Kaiser, A. P., Goldstein, H., Alpert, C., Mousetis, L., Kaczmarek, L., & Fischer, R. (1995). An exploratory comparison of milieu teaching and responsive interaction in classroom applications. *Journal of Early Intervention*, *19*(3), 218–242. https://doi.org/10.1177/105381519501900306

TABLE CAPTIONS

Table 1. Inclusion and Exclusion Criteria (with operational definition of naturalistic intervention) for the Included Reviews

Table 2. Methodological Characteristics of Included Reviews

Table 3. Characteristics of Child Participants with Disabilities across Included Reviews

Table 4. Characteristics of Adult Participants (*u* indicates number of studies from which data were extracted)

Table 5. Intervention Characteristics of Included Reviews (*u* = number of studies with each characteristic)

Table 6. Child Participant Target Skill Acquisition Reported Across Reviews

Table 1. Inclusion and Exclusion Criteria (with operational definition of naturalistic intervention) for the Included Reviews

Review	Language	Publication status	Study design	Participants	Intervention Descriptor and *Intervention Inclusion Criteria* (copied from original text)	Setting	Outcome	Other
Rakap (2011)	Not specified	Peer-reviewed journal	Empirical	At least one preschool-aged child 36-60 months old with an identified disability	Embedded instruction: *"stud[ies] included in this review had to use the terms 'embedded', 'embedded instruction', 'embedding' or 'embedding instruction' throughout the study. Studies that used the terms 'naturalistic teaching', or 'activity based instruction' were also included … when they used 'embedding' or 'embedded instruction' in their introduction sections."* (Rakap & Parlak-Rakap, 2011, p. 82)	Inclusive preschool classrooms	Not specified	• Intervention occurred during activities, routines, or transitions • Sessions occurring in segregated sections of inclusive classrooms were excluded
Rakap (2014)	English	Peer-reviewed journal	Empirical study using single-subject experimental design	At least one child under 60 months old with an identified disability	Parent-implemented naturalistic language intervention: *"(1) the adult follows the child's lead; (2) the activities … are child led; (3) the targets chosen address skills needed by the child to participate in ongoing classroom activities; (4) the adult is highly responsive to the child's communicative attempts; and (5) instruction is provided by those adults who regularly interact with the child"* (Rakap & Rakap, 2014, p. 37)	Natural settings (clinical settings included if within context of developmentally appropriate activity	Parent implementation and child outcome data	• Age-appropriate activities • Studies investigating specific strategies were excluded • Studies of dialogic reading were excluded
Snyder (2015)	Not specified	Peer-reviewed journal	Empirically-based research	At least one child 36-60 months old with an identified disability	Naturalistic instruction: *"[naturalistic approaches] typically involve embedding (providing learning trials during naturally occurring or motivating activities) and embedded learning opportunities (providing intentional and systematic instructional episodes within and across activities based on children's individualized learning needs and outcomes)"* (Snyder et al., 2015, pp. 69-70)	Preschool classrooms (inclusive classroom or early childhood special education classroom)	Child learning	• Instruction during typical activities, routines, or transitions
Lane (2016)	Not specified	Peer-reviewed journal	Single-case design or quasi-experimental group design	At least one participant with ASD (or at-risk for ASD if under 8 years old)	Naturalistic language interventions: "*adults embed opportunities for expressive language in social exchanges and everyday routines and activities of the child … during typical activities, with adults providing learning opportunities across contexts.*" (Lane et al., 2016, p. 50)	Inclusive setting	Spontaneous verbalizations	• Studies with AAC or PECS were excluded • Stimuli selected based on child interests • Studies using external or social reinforcers were excluded

Dubin (2020)	Not specified	Peer-reviewed journal	Empirical study	Children under 8 years old diagnosed with or at-risk for ASD and significant communication delays	Naturalistic behavioral social communication interventions: *"intervention occurs in typical activities and routines, is based around child-interests and child initiations, is distributed throughout the day, and may involve behavioral … and social interaction strategies"* (Dubin & Lieberman-Betz, 2020, p. 153; quoted from Kaiser & Trent, 2007)	Not specified	Observational measure of a prelinguistic social communication skill	Not specified
Gulboy (2023)	English	Peer-reviewed journal	Empirical study using a single-case experimental design	At least one child under 72 months old with an identified disability	Embedded instruction: *"instructional procedures to teach a child's priority learning targets are implemented in the context of ongoing activities, routines, and transitions of inclusive preschool classrooms"* (Gulboy et al., 2023, p. 182, quoted from Snyder et al., 2013)	Inclusive settings	Child outcome data	Not specified
Lane (2023)	Not specified	Peer-reviewed journal	Experimental design (controlled group design or single-case design)	Children 3-12 years old with disability, provisional diagnosis of ASD, 3-5 years old and at-risk for a disability, or eligible for special education services	Naturalistic language interventions: *"[instructional] sessions occurred during typical activities, materials/toys were selected by the child or based on the child's interests, the adult introduced an intervention during the play-based activity"* (Lane et al., 2023, p. 321)	Classroom	Verbal social communication	• Intervention implemented by classroom teacher or other classroom staff • Intervention occurred during typical activities with materials/toys selected by child or based on child interests • Studies using AAC were excluded • Studies with early literacy or shared reading outcomes were excluded

Key: AAC = augmentative and alternative communication; PECS = picture exchange communication system

Table 2. Methodological Characteristics of Included Reviews

Review	Review type	Studies (*u*) in X articles	Search date (coverage)	Additional (supplemental) search methods	Rigor assessment method
Rakap (2011)	SR	*u* = 16 (16 articles)	n/r	• Hand search of reference lists of studies that passed initial screening	n/a
Rakap (2014)	SR	*u* = 15 (15 articles)	n/r	• Hand search of reference lists of studies that passed initial screening	**SCD:** Single-case Experimental Design Scale (Tate et al., 2008)
Snyder (2015)	SR	*u* = 43 (37 articles)	2013 (1980 to 2013)	• Ancestral search of reference lists of included studies • Author search through electronic databases • Additional search term generation through identified studies	**GDS:** Council for Exceptional Children Quality Indicators for Group Experiemental and Quasi-experimental Research (Gersten et al., 2005) **SCD:** WWC Single Case Design Technical Document (Kratochwill et al., 2010), Council for Exceptional Children Quality Indicators for Single Case Research (Horner et al., 2005)
Lane (2016)	SR	*u* = 24 (11 articles)	n/r	• Ancestral search of reference lists of included studies	**GDS:** WWC 3.0 (2012) **SCD:** WWC Single Case Design Technical Document (Kratochwill et al., 2010)
Dubin (2020)	SR	*u* = 25 (22 articles)	2018 (2001 to 2018)	• Hand search of reference lists of included studies • Hand search of prior review	**GDS:** WWC Single Case Design standards (Kratochwill et al., 2013) **SCD:** Single-case Analysis and Review Framework (SCARF; Ledford et al., 2023)
Gulboy (2023)	SR, MA	*u* = 10 (10 articles)	2020 (up to 2020)	• Footnote chasing method	**SCD:** WWC 4.0 (Kratochwill et al., 2017b)
Lane (2023)	SR	*u* = 38 (19 articles)	2020 (up to 2020)	• Hand search of relevant journals • Ancestral search of reference lists of included studies	**SCD:** WWC 4.0 (2017a, 2017b), Ledford et al. (2018) single case design chapter

Key: SR = systematic review; SCD = single case design; Obs. = Observational study; GD = group design; MA = meta-analysis; n/r = not reported; n/a = not applicable; GDS = group design study; SCD = single case design study; WWC = What Works Clearinghouse; SCARF = Single-case analysis and review framework

Table 3. Characteristics of Child Participants with Disabilities across Included Reviews

Review (number of studies)	Sample size	Age Mean (SD); range (in months)	Gender	Child Disability Categories
Rakap (2011)[1] *u* = 16	*n* = 67	Mean = n/r Range = 36 to 72	Male = 39 Female = 18 n/r = 10	ADHD, ASD, DD, DS, IDD, MD, PD, PDD, SLI, VI
Rakap (2014) *u* = 15	*n* = 70	Mean = 43.5 Range = 16 to 95	Male = 55 Female = 15	ASD, CP, DD, DS, PDD, SLI, SQ
Snyder (2015)[2] *u* = 43	*n* = 211[5]	Mean = 51.5 (7.8) Range = 24 to 84	Male = 135 Female = 66 n/r = 10	ASD, CP, DD, DS, MD, SLI
Lane (2016) *u* = 24	*n* = 45	Mean = n/r Range = 24 to 108	n/r	ASD, CD, DD
Dubin (2020)[3] *u* = 25	*n* = 387	Mean = n/r Range = 12 to 96	Male = 309 Female = 62 n/r = 16	ASD, PDD-NOS
Gulboy (2023) *u* = 10	*n* = 21	Mean = 54.7 Range – 36 to 66	Male = 14 Female = 7	ASD, DD, PD, SLI
Lane (2023)[4] *u* = 38	*n* = 106	Mean = n/r Range = 24 to 156	n/r	ASD, CP, DD, DS, IDD, OHI, SI, SLI

Note: [1] - *u* = 15 for age, gender, and disability category; [2] - *u* = 38 for mean age, *u* = 39 for age range and gender; [3] - *u* = 23 for gender; [4] - *u* = 36 for disability category; [5] - Review noted 26 children without disabilities reported in 3 studies

Key: n = total number of child participants; n/r = not reported; ADHD = attention deficit/hyperactivity disorder; ASD = autism spectrum disorder; DD = developmental delay; DS = Down syndrome; IDD = intellectual and developmental disabilities; MD = multiple disabilities; PD = physical disability; PDD = pervasive developmental disorder; SLI = speech language impairment; VI = visual impairment; CP = cerebral palsy; SQ = spastic quadriplegia; CD = chronological disease; PDD-NOS = pervasive developmental disorder-not otherwise specified; OHI = other health impairment

Table 4. Characteristics of Adult Participants (*u* indicates number of studies from which data were extracted)

Author and Year	Sample size	Age Mean (SD; Range) (in years)	Gender	Education attainment	Interventionist Role	Experience Mean (SD; Range) (in years)
Rakap (2011) *u* = 16	*n* = 49	n/r	Male = 1 Female = 15 n/r = 33	<HS = 6% HSD = 23% SC/AD = 14% UD = 37% GD = 20%	Teacher (*n* = 26) Para. (*n* = 15) Res. Staff (*n* = 8)	n/r
Rakap (2014) *u* = 15[1]	*n* = 74	Caregivers: M = 33 (R = 20 to 48)	Male = 9 Female = 65	HSD = 23% SC/AD = 32% UD = 30% GD = 15%	Caregivers (*n* = 74)	n/r
Snyder (2015) *u* = 43[2]	*n* = 122	M = 30 (4.8; R = 20 to 55)	Male = 2 Female = 35	HSD = 8% SC/AD = 7% UD = 85%	Teacher (*n* = 58) Para. (*n* = 39) Res. Staff (*n* = 18) Therapist (*n* = 7)	Mean = 7.7 (5.7; R = 0 to 22)
Lane (2016) *u* = 24	*n* = 57	n/r	n/r	n/r	n/r	n/r
Dubin (2020) *u* = 25	n/r	n/r	n/r	n/r	n/r	n/r
Gulboy (2023) *u* = 10	*n* = 15	n/r	n/r	n/r	Teacher (*n* = 7) Para. (*n* = 1) Res. Staff (*n* = 7)	n/r
Lane (2023) *u* = 38[3]	*n* = 57	R = 21 to 51	n/r	UD = 60% GD = 40%	Teacher (*n* = 31) Teacher/Para. (*n* = 23) Para. (*n* = 3)	R = 1 to 30

Note: [1] = *u* = 5 for age, *u* = 11 for education; [2] = *u* = 8 for age; *u* = 13 for gender, *u* = 19 for education, *u* = 18 for interventionist role and experience; [3] = *u* = 34 for sample size, *u* = 9 for age, *u* = 14 for education, *u* = 32 for interventionist role, *u* = 17 for experience

Key: *n* = number of participants; n/r = not reported; M = male; F = female; <HS = less than high school degree; HSD = high school degree; SC/AD = some college or associate degree; UD = undergraduate degree; GD = graduate degree; Para = paraprofessional; Res. Staff = research staff (includes undergraduate and graduate students); M = mean; R = range

Table 5. Intervention characteristics of included reviews (*u* = number of studies with each characteristic)

Review	Named Intervention	Intervention settings	Activity Context[1]	Intervention Density
Rakap (2011) *u* = 16	EI (*u* = 16)	School (*u* = 16)	Center/free play (*u* = 7) Small group (*u* = 3) Circle (*u* = 1) Mealtime (*u* = 2) Other (*u* = 9) n/r (*u* = 1)	n/r
Rakap (2014) *u* = 15	MT (*u* = 7) Other (*u* = 8)	School (*u* = 3) Home (*u* = 5) Clinic (*u* = 4) Multiple (*u* = 3)	Center/free play (*u* = 13) Other (*u* = 4)	Session duration: 0 to 30 minutes (*u* = 1) 30 to 60 minutes (*u* = 8) 60+ minutes (*u* = 3) n/r (*u* = 3)
Snyder (2015) *u* = 43[2]	EI (*u* = 15) NI (*u* = 14) MT (*u* = 5) ABI (*u* = 3) Other (*u* = 6)	School (*u* = 43)	Center/free play (*u* = 27) Small group (*u* = 9) Mealtime (*u* = 11)	Trials per session: M = 7.2 trials (R = 3 to 20; *u* = 21) Session duration: M = 23.3 min (R = 5 to 180; *u* = 18) Intervention duration: M = 13 weeks (R = 6 to 30; *u* = 5)
Lane (2016) *u* = 24	IT (*u* = 2) MT (*u* = 7) Other (*u* = 15)	School (*u* = 2) Home (*u* = 3) Clinic (*u* = 16) Multiple (*u* = 3)	n/r	Session frequency: 1-4 per week (*u* = 14) 5+ per week (*u* = 9) n/r (*u* = 1)
Dubin (2020) *u* = 25	MT (*u* = 2) Other (*u* = 23)	School (*u* = 7) Home (*u* = 2) Clinic (*u* = 7) Multiple (*u* = 3) Other (*u* = 6)	n/r	Intervention Duration: 0 to 2 months (*u* = 11) 2.1 to 6 months (*u* = 10) 6.1 to 12 months (*u* = 1) n/r (*u* = 2)
Gulboy (2023) *u* = 10	EI (*u* = 10)	School (*u* = 9) Other (*u* = 1)	Center/free play (*u* = 8) Small group (*u* = 3) Circle (*u* = 1) Mealtime (*u* = 2) Other (*u* = 4)	n/r
Lane (2023) *u* = 38	IT (*u* = 3) MT (*u* = 5) NI (*u* = 8) Other (*u* = 22)	School (*u* = 38)	Center/free play (*u* = 27) Small group (*u* = 6) Circle (*u* = 2) Routine (*u* = 6) Other (*u* = 9)	n/r

Note: [1] – may sum greater than number of studies due to reference of multiple activity types in some included studies; [2] - *u* = 21 for trials per session, *u* = 19 for session duration, *u* = 13 for intervention duration

Key: EI = embedded instruction/intervention; n/r = not reported; MT = milieu teaching/therapies; NI = naturalistic intervention; ABI = activity-based instruction; IT = incidental teaching

Table 6. Child Participant Target Skill Acquisition Reported Across Reviews

Review	Research design	Target Skills	Participants Reported to Have Acquired Target Skill(s) at **Primary Endpoint**	Participants Reported to Have Shown **Generalization**	Participants Reported to Have Shown **Maintenance**
Rakap (2011) *u* = 16	SCD (*u* = 15)	PA/C (*u* = 9) L/C (*u* = 5) M/A (*u* = 4) S-E (*u* = 4)	55 of 60 (92%) participants from 15 SCD studies	12 of 15 (80%) participants from 5 SCD studies	17 of 18 (94%) participants from 6 SCD studies
Rakap (2014) *u* = 15	SCD (*u* = 15)	L/C (*u* = 14) S-E (*u* = 1)	66 of 70 (94%) participants from 15 SCD studies	27 of 34 (79%) participants from 8 SCD studies	25 of 25 (100%) participants from 5 SCD studies
Snyder (2015) *u* = 43	SCD (*u* = 40) GD (*u* = 3)	PA/C (*u* = 18) L/C (*u* = 26) M/A (*u* = 12) S-E (*u* = 8)	207 of 211 (98%) participants from 40 SCD studies and 3 GD studies	47 of 50 (94%) participants from 18 SCD studies	56 of 61 (92%) participants from 20 SCD studies
Lane (2016) *u* = 24[1]	SCD (*u* = 12)	L/C (*u* = 24)	4 of 6 (67%) **SCD studies** using demonstration designs meeting WWC standards showed "strong effects" 3 of 6 (50%) **SCD studies** using comparison designs meeting WWC standards showed "differentiated effects"	n/r	n/r
Dubin (2020) *u* = 25[2]	SCD (*u* = 7) GD (*u* = 13)	L/C (*u* = 25)	20 of 25 (80%) participants from 7 SCD studies with "sufficient rigor" 7 of 13 (54%) **GD studies** with "sufficient rigor" showed small to large positive effect sizes	Some evidence in 5 of 7 (71%) SCD studies Some evidence in 11 of 13 (85%) GD studies	Some evidence in 3 of 7 (43%) SCD studies Some evidence in 2 of 13 (15%) GD studies
Gulboy (2023) *u* = 10	SCD (*u* = 10)	PA/C (*u* = 5) L/C (*u* = 5) M/A (*u* = 3)	21 of 21 (100%) participants from 10 SCD studies	20 of 20 (100%) participants from 9 SCD studies	8 of 8 (100%) participants from 4 SCD studies
Lane (2023) *u* = 38[3]	SCD (*u* = 23)	L/C (*u* = 34) S-E (*u* = 19)	4 of 4 (100%) participants from 1 SCD study	3 of 4 (75%) participants from 1 SCD study	n/r

Note: SCD = single case design; PA/C = pre-academic/cognitive; L/C = language/communication; M/A = motor/adaptive; S-E = social-emotional; GD = group design; WWC = What Works Clearinghouse; [1] – 12 SCD studies met WWC standards with or without reservations and were synthesized for review (11 SCD studies and 1 GD study did not meet standards); [2] – 7 SCD studies and 13 GD studies had "sufficient rigor" and were synthesized for review (4 SCD and 1 GD study did not meet standards); [3] – 1 SCD study met rigor standards and were synthesized for review (22 studies did not meet rigor standards)

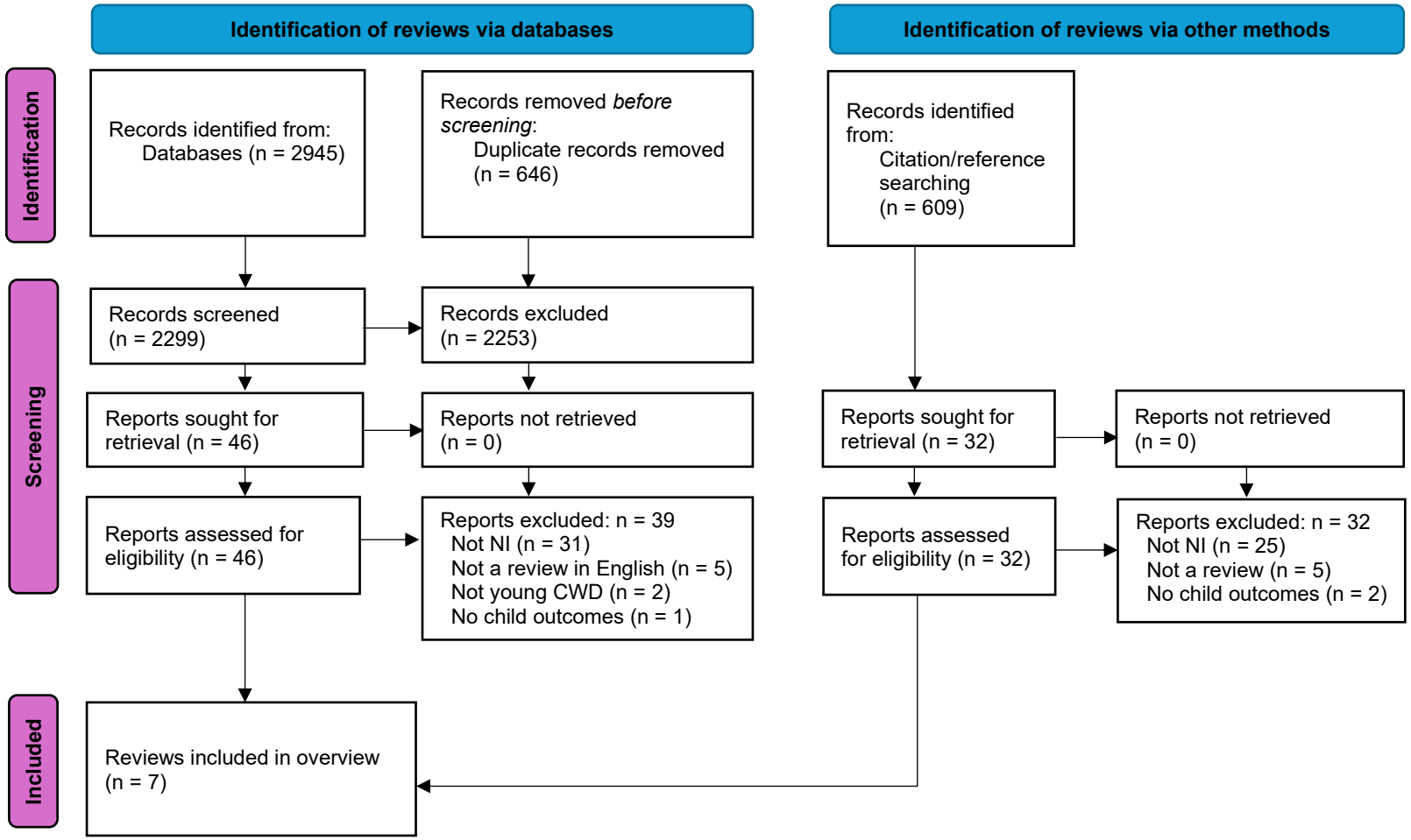
FIGURE CAPTIONS

Figure 1. Review Selection Flow Diagram

Figure 2. GROOVE Primary Study Overlap Summary Across Included Reviews

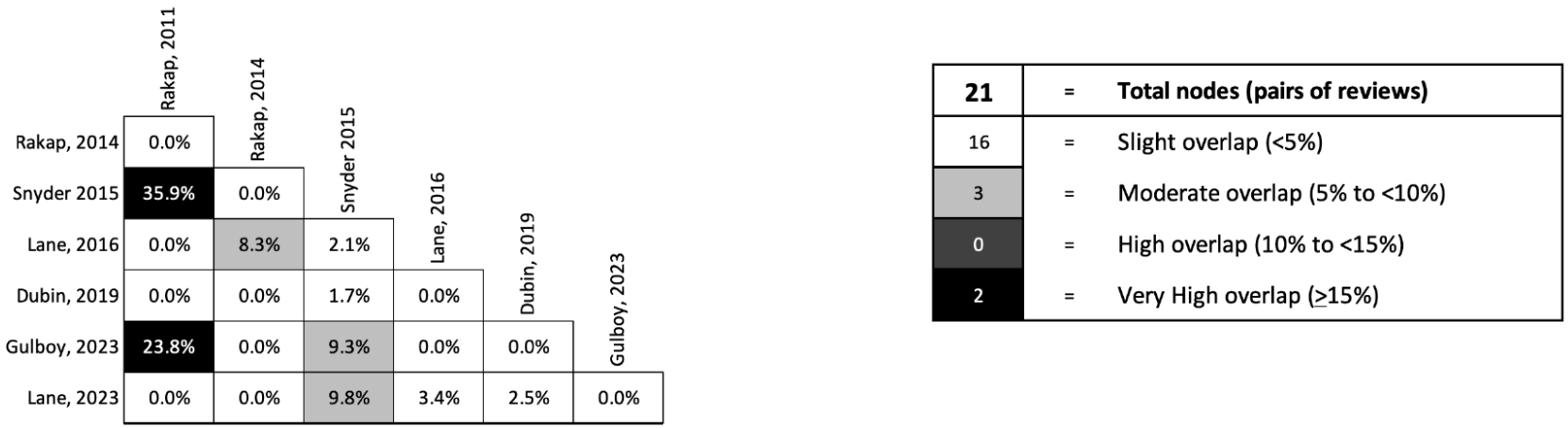
Figure 3. JBI Appraisal Checklist for Systematic Reviews and Research Syntheses Summary

Figure 1. Review Selection Flow Diagram



Flow Diagram adapted from: Page et al. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ, 372*, n71. https://doi.org/10.1136/bmj.n71

Figure 2. GROOVE Primary Study Overlap Summary Across Included Reviews

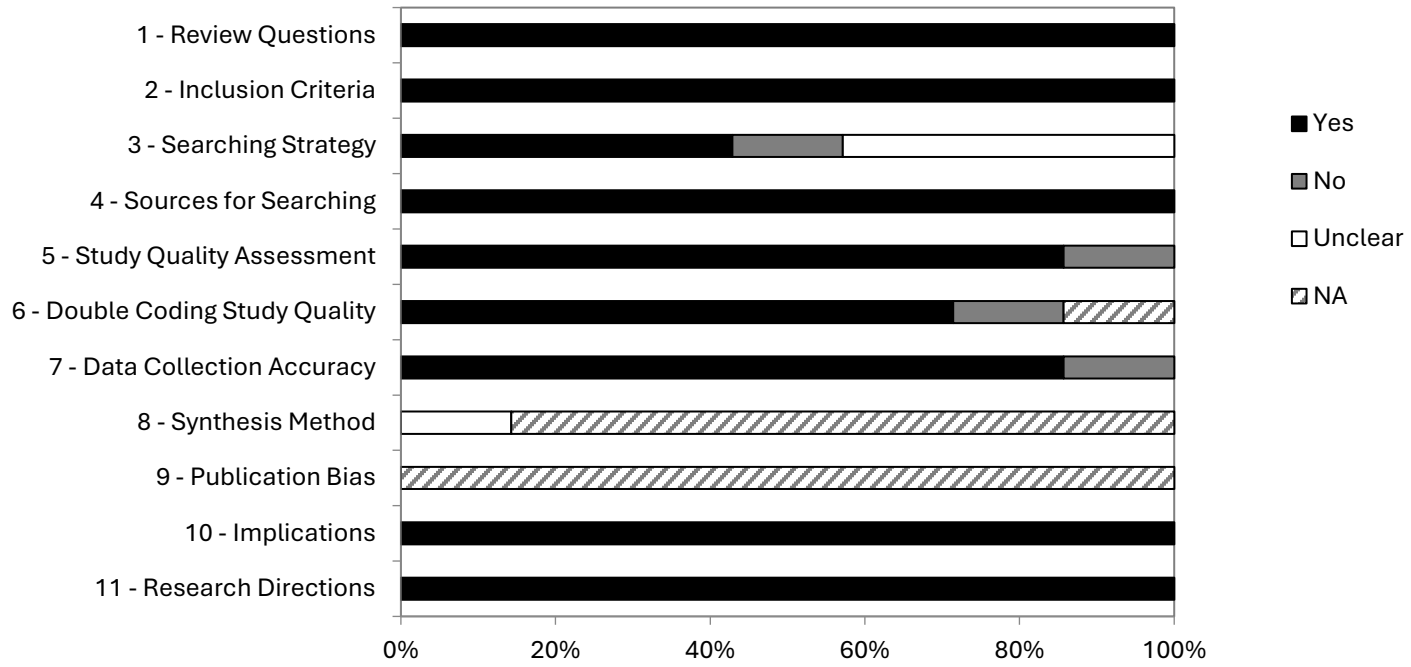


Rakap, 2014	0.0%	
Snyder 2015	35.9%	0.0%
Lane, 2016	0.0%	8.3%	2.1%
Dubin, 2019	0.0%	0.0%	1.7%	0.0%
Gulboy, 2023	23.8%	0.0%	9.3%	0.0%	0.0%
Lane, 2023	0.0%	0.0%	9.8%	3.4%	2.5%	0.0%

Rakap, 2011 | Rakap, 2014 | Snyder 2015 | Lane, 2016 | Dubin, 2019 | Gulboy, 2023

21	=	**Total nodes (pairs of reviews)**
16	=	Slight overlap (<5%)
3	=	Moderate overlap (5% to <10%)
0	=	High overlap (10% to <15%)
2	=	Very High overlap (≥15%)

Figure 3. JBI Appraisal Checklist for Systematic Reviews and Research Syntheses Summary



Supplemental Text 1: Search Strategy for Medline, APA PsycINFO, Education Resource Information Center (ERIC), Cumulative Index of Nursing and Allied Health Libraries (CINAHL), and Academic Search Premier

1. TX “systematic review” OR “systematic literature review” OR “literature review” OR “scoping review” OR “scoping literature” OR “scoping literature review” OR “rapid review” OR “comprehensive review” OR “comprehensive literature” OR “comprehensive literature review” OR “narrative review” OR “narrative literature review” OR “integrative review” OR “best evidence synthesis” OR “mapping review” OR “evidence map” OR “meta-analysis” OR “metaanalysis” OR “meta-analyses” OR “meta analyses” or “metaanalyses” OR “meta-syntheis” OR “metasynthesis” OR “quantitative review” OR “quantitative synthesis” OR “research synthesis” OR “research review” OR “review of research”
2. TX “embed* instruct*”
3. TX “embed* intervention*”
4. TX “embed* teach*”
5. TX “natural* instruct*”
6. TX “natural* intervention*”
7. TX “natural* teach*”
8. TX “milieu instruct*”
9. TX “milieu intervention*”
10. TX “milieu teach*”
11. TX “incidental instruct*”
12. TX “incidental intervention*”
13. TX “incidental teach*”
14. TX “activity-based instruct*”
15. TX “activity-based intervention*”
16. TX “activity-based teach*”
17. TX “intentional instruct*”
18. TX “intentional intervention*”
19. TX “intentional teach*”
20. 2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14 OR 15 OR 16 OR 17 OR 18 OR 19
21. 1 AND 20 (limiter – peer reviewed)