Review Journal of Autism and Developmental Disorders
https://doi.org/10.1007/s40489-022-00314-w

REVIEW PAPER

Considering Instructional Contexts in AAC Interventions for People with ASD and/or IDD Experiencing Complex Communicative Needs: a Single-Case Design Meta-analysis

Jay B. Ganz[1] · James E. Pustejovsky[2] · Joe Reichle[3] · Kimberly J. Vannest[4] · Margaret Foster[5] · April N. Haas[6] · Lauren M. Pierson[7] · Sanikan Wattanawongwan[1] · Armando Bernal[1] · Man Chen[2] · Rachel Skov[1] · S. D. Smith[8]

Received: 11 August 2021 / Accepted: 11 March 2022

Abstract

For children with autism or intellectual and developmental disabilities who also have complex communication needs, communication is a necessary skill set to increase independence and quality of life. Understanding the how, where, and communication style being taught is important for identifying deficits in the field as well as which interventions are most effective. This meta-analysis sought to identify effectiveness among different settings, behavioral strategies, and moderator variables. A systematic search and screening process identified 114 eligible studies with 330 participants; overall outcomes indicate that augmentative and alternative communication interventions were effective with Tau effects ranging from 0.53 to 1.03 and log response ratio effects ranging from 0.21 to 2.90. However, no instructional context variables systematically predicted differences in intervention effectiveness.

Keywords Autism · ASD · Intellectual disability · IDD · Complex communication needs · Minimally or non-verbal · Intervention · Setting · Instructional features · Behavioral strategies

Introduction

The Centers for Disease Control and Prevention (CDC) reports 1 in 54 children in the USA has autism spectrum disorder (ASD; CDC, 2018) with concomitant intellectual disabilities in 30% of this population placing students at high risk for complex communication needs (CCNs, Luyster et al., 2008). Social communication deficits represent a core feature of ASD (Reichle & Wacker, 2017; Tager-Flusberg & Kasari, 2013). Communication challenges for students with ASD are particularly prevalent in aspects of social communication (pragmatics) but can also occur in any or a combination of phonological, semantic, and/or syntactic domains (Reichle & Wacker, 2017; Wetherby & Prizant, 1993). Social communicative challenges can negatively impact a learner across their lifespan (Branson & Demchak, 2009). For learners with more severe social communicative challenges, problem behaviors (Reichle & Wacker, 2017) can further contribute to communication challenges.

Communication interventions often result in improved outcomes for individuals with ASD and/or intellectual and

✉ Sanikan Wattanawongwan
swattanawongwan@tamu.edu

1 Department of Educational Psychology, Texas A&M University, 4225 TAMU, College Station, TX 77843, USA

2 Department of Educational Psychology, University of Wisconsin-Madison, Madison, WI, USA

3 Department of Speech-Language-Hearing Sciences, University of Minnesota, Minneapolis, MN, USA

4 Department of Education, University of Vermont, Burlington, VT, USA

5 University Libraries, Texas A&M University, College Station, TX, USA

6 Life Skills Autism Academy, Plano, TX, USA

7 Department of Social Work and Communication Disorders, Tarleton State University, Stephenville, TX, USA

8 Department of Elementary, Early and Special Education, Southeast Missouri State University, Cape Girardeau, MO, USA

Published online: 20 April 2022

developmental disabilities (IDD; Paul, 1997). Instructional environments that emulate situations in which the learner is likely to use the new communicative behavior being taught have become increasingly prevalent in the intervention literature (Gevarter & Zamora, 2018; Logan et al., 2017; Shire & Jones, 2015). The implementation of augmentative and alternative communication (AAC) has proven effective for persons with ASD and/or IDD that results in complex communication needs (CCNs; Reichle et al., 2021) Positive outcomes of recent communication intervention research demonstrate the efficacy of a range of interventions including behavioral and naturalistic interventions, including AAC interventions in authentic settings (Holyfield et al., 2017; Logan et al., 2017).

Only one meta-analysis has compared authentic communication settings with more contrived contexts (Ganz et al., 2017) finding too few experiments to comprehensively examine this variable. Further, most reviews of context as a potential moderator in communication interventions have not examined *effect magnitude* as a function of authentic compared to non-authentic approaches to intervention (e.g., Gevarter & Zamora, 2018; Holyfield et al., 2017). Much of the extant research has been implemented in clinical settings, such as universities or schools (Biggs et al., 2018; Gevarter & Zamora, 2018; Holyfield et al., 2017; Logan et al., 2017; Therrien et al., 2016) rather than homes or other community settings (Shire & Jones, 2015). Consequently, there is a need to update and more systematically quantify the effects of instructional practices as moderators of intervention outcomes.

Prior meta-analyses and systematic reviews that provide insight regarding instructional contexts have not disaggregated individuals with ASD from other participants with IDD (Crowe et al., 2021; Ganz et al., 2017; Gevarter & Zamora, 2018). Given some potential qualitative differences in social communication skills among individuals with ASD compared to others with IDD, it is possible that these populations have different responsiveness to particular naturalistic versus more structured settings or instructional strategies and features. For example, children with Down syndrome have a greater propensity to engage in joint attention when compared to many with ASD (Hahn et al., 2018). Consequently, they may be more receptive to learning in the milieu of a natural environment. Compared to typically developing populations and others with developmental disabilities, persons with ASD have a greater tendency to be poor imitators, but at the same time may be better in learning visual discriminations (Johnston et al., 2012). This has significant implications for settings and response prompts chosen for initial intervention. Little past research has examined the extant literature to explore these variables with respect to population.

One limitation of the extant research has been the relatively modest role that authentic educators and parents have played as communicative partners during the implementation of research procedures in authentic settings (Crowe et al., 2021; Sutton et al., 2019). As a result, fully translational research is limited. Issues such as contextual fit, treatment acceptability, and treatment feasibility in authentic settings have received limited attention (Monzalve, 2016).

There seems to be widespread agreement that intervention opportunities blended within naturalistic contexts are desirable. Within the past 20 years, the term “blended approach” has appeared in the literature with increasing frequency (Reichle et al., 2019). Blended approaches incorporate some discrete trial opportunities within natural contexts (Hong et al., 2014). Naturalistic Developmental Behavioral Intervention models include interventions that blend substantial methodology supported by applied behavior analysis with elements of developmental relevance for young learners (Schreibman et al., 2015). Thus, Naturalistic Developmental Behavioral Interventions rely on ABA behavioral principles and methods (e.g., systematic prompting and fading; Schreibman et al., 2015); however, interventions are also viewed through a more social-pragmatic lens. Strategies may include (a) implementation within settings in which AAC skills would naturally be used, enabling generalization of skills into a range of settings (Light, 1997; Ogletree, 2012); (b) use of instructional prompts matched to learner needs; (c) expansion of current communication skills (e.g., verbal, AAC, gestures); (d) implementation of behavioral techniques including time delay, positive reinforcement, and prompting (Reichle et al., 2002) in natural contexts; and (e) inclusion of natural communication partners as key interventionists. Naturalistic Developmental Behavioral Interventions adhere to recommended early intervention practices of naturalized, routine-based, and family-centered interventions, as discussed by Schreibman et al. (2015). In spite of interventionists increasing embrace of blended approaches, there has been no large-scale effort to examine the relative contributions of the operationalized components of this approach. Doing so could provide information to enhance the customization of intervention strategies for a particular learner.

Purpose and Research Questions

There is a need to more systematically quantify the instructional practices, such as implementation setting, instructional features, and behavioral and naturalistic intervention strategies implemented. Our knowledge of the extent to which a range of instructional strategies are effective and the degree to which the strategies are actually used in translational settings would better inform practitioners. The analyses addressed in this article are designed to capture practices that address core intervention strategies, and resulting AAC outcomes that are effective in serving individuals with ASD and/or IDD.

In this article, research questions include:

1. What are the magnitude of intervention effects for AAC outcomes and other communication outcomes for participants with ASD and/or participants with IDD?
2. For AAC outcomes, to what extent does instructional setting (e.g., classroom, clinic, home) moderate treatment effects for AAC implementation, for participants with ASD and/or participants with IDD?
3. For AAC outcomes, are specific instructional features (i.e., systematic arrangement, preference assessment, reinforcement, modeling, prompts, and prompt fading) associated with the magnitude of intervention effects?
4. For AAC outcomes, what potential moderator variables are related to behavioral intervention strategies (i.e., child or interventionist initiated, dispersed versus massed teaching opportunities, contrived versus embedded activity contexts, group versus one-on-one instructional formats, limited versus varied teaching stimuli, controlled versus natural instructional environments) for children with ASD and/or IDD and what is their effect on intervention outcomes?

Finally, in further exploratory analysis, we also investigated how use of instructional features and use of behavioral intervention strategies varied based on participant characteristics. Doing so provided insight regarding differentiation in implementation of particular instructional features and strategies based on participant age and whether the participants were minimally or non-verbal.

Methodology

A comprehensive meta-analysis was undertaken between 2018 and 2020 to answer research questions related to potential moderators of instructional variations and setting variations. We report findings in a manner consistent with PRISMA 2020 guidelines (Page et al., 2021). All screening processes for the current study were part of a larger systematic review project. (See Figure 1 for a PRISMA flowchart describing the literature search and screening process.) This project received financial support from the Institute of Education Sciences (IES) and was pre-registered in the PROSPERO system (CRD42018112428).

Literature Search

A professional reference librarian conducted the literature search in the following databases: Academic Search Complete, Conference Proceedings Citation Index – Social Science & Humanities (Web of Science), ERIC, Proquest Dissertations & Theses Global, and PsycINFO. Keywords included words within the categories AAC, individuals with ASD/IDD with complex communication needs, and social communication and behavioral outcomes. The thesaurus within each database identified additional keywords associated with each category. The search string comprised the following: [((augmentative or alternative) within one word (w1) communicat*) or "sign language" or manual sign* or speech-generating device* or SGD or "voice output communication aid" or VOCA* or PECS or "picture exchange communication system" or AAC or "visual scene display" or "functional communication training"] AND [(down* w1 syndrome) or ((develop* or intellectual) w1 (delay* or disabil* or impair*)) or autis* or retard*].

Inclusion/Exclusion Criteria

The final search included 7,384 documents to be screened for inclusion. Documents were included if the following criteria were met: (a) the study included an AAC intervention (excluding AAC interventions that have been discredited such as facilitated communication or supported typing), (b) one or more participants were diagnosed with ASD and/or IDD and complex communication needs, (c) social communicative or challenging behaviors served as the dependent variable, (d) the study used a single-case experimental design, and (e) documents were written in English. After screening title, abstract, and full texts of documents meeting inclusion criteria, documents were screened for basic design quality standards. Documents were included if the following standards were met: (a) interobserver agreement was recorded and was at least 0.60 kappa or 80%, (b) phase changes occurred at least three times to demonstrate effects, and (c) each phase contained a minimum of three data points (four for each treatment in an alternating treatment).

For this systematic review and meta-analysis, we further restricted the sample according to the following inclusion criteria: (a) participants were of school age, less than 22 years old; (b) participants were diagnosed with ASD and/or IDD; (c) interventions primarily targeted communication; and (d) dependent variables measured outcomes related to AAC (e.g., percentage of intervals in which the participant used AAC; number or rate of communicative acts). Studies of functional communication training interventions were excluded because the purpose of those studies was primarily to decrease challenging behavior rather than increase specific communicative behaviors, although those are taught as replacement behaviors. Individual data series assessing outcomes for implementers were excluded. Finally, the sample used for some analysis was restricted based on requirements of the effect size metrics that we applied, as we explain subsequently. After completing title and abstract, full-text reviews, and quality screening, 114 study reports remained.

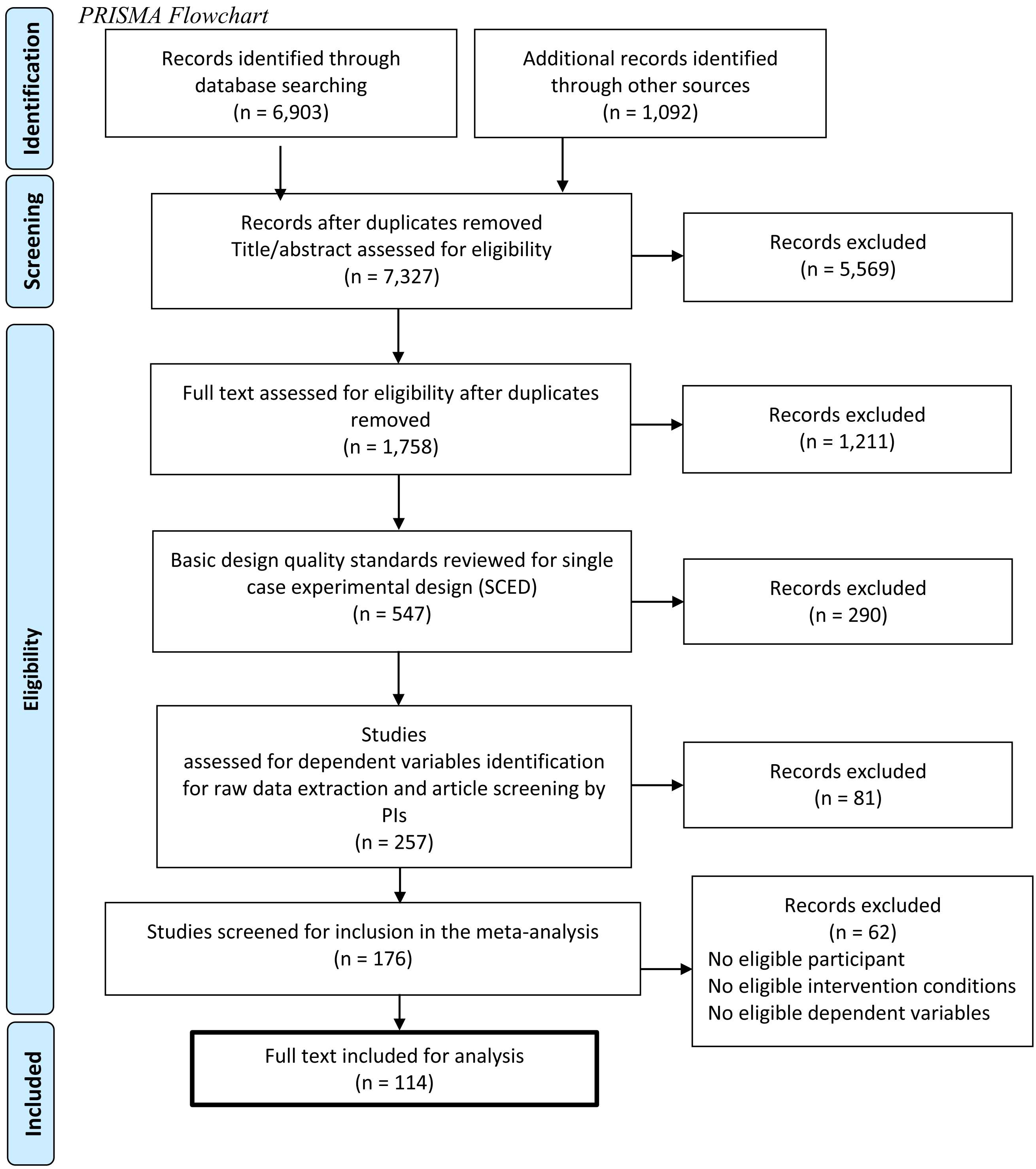

Fig. 1 PRISMA flowchart. *Note*: flowchart based on Moher et al. (2009)

Variable Coding

Studies were coded by one of four independent reviewers who were graduate students in special education. Inter-rater reliability was assessed for a minimum of 20% of the studies using a Google Form specifically created for these variables. The setting for each study was coded as home, classroom, employment, clinic, recreation, or other. With respect to teaching stimuli, we were interested in whether or not each of several strategies or approaches was implemented.

Reviewers coded each study for use of behavioral instructional features that included (a) environmental arrangement, (b) models, (c) verbal prompts, (d) physical/partial physical prompts, (e) graphic prompts, (f) preference assessments, or (g) prompt fading strategies. Finally, studies were coded to ascertain whether the intervention was implemented in the natural environment with others present or if the learning environment was controlled with distractions minimized. They first identified whether the teaching opportunities were child-initiated or interventionist-initiated. Then, they examined whether teaching opportunities were dispersed throughout the day or were implemented in a massed trial format. Additionally, each study was coded to identify if the intervention was embedded during functional activities or activities were contrived. Reviewers identified whether the intervention was implemented in one-on-one or group format. Next, they categorized the settings in which intervention was implemented along with who served as interventionist(s). Variables coded were not mutually exclusive. Rather, for behaviors targeted, setting, and strategies implemented, studies could be conducted in multiple settings, have multiple communication skills targeted, or use multiple behavioral strategies.

Data Extraction

In order to calculate effect size estimates for meta-analysis, we used the program Engauge Digitizer (Mitchell et al., 2017) to extract outcome data and design information from graphs reported in primary sources. Data extraction was conducted by four reviewers with at least 20% of the studies being reviewed by a second reviewer for inter-rater reliability.

Effect Size Calculations

Many different effect size metrics have been proposed to quantify intervention effects in single-case experimental studies, each with strengths and limitations. Because of this, scholars have recommended using multiple metrics when conducting meta-analysis of single-case studies (Pustejovsky & Ferron, 2017; Vannest et al., 2018). In this meta-analysis, we therefore applied two effect sizes with complementary strengths: Tau(AB) and the log response ratio.

Proposed by Parker et al. (2011), Tau(AB) is a non-overlap measure that describes intervention effects in terms of the proportion of all possible comparisons between baseline and intervention data where the intervention data point represents an improvement over the baseline data point. It ranges from -1 to 1, with a value of 0 indicating no difference between the distributions of data points in baseline versus intervention and a value of 1 indicating improvement in intervention that results in complete non-overlap. Because it is based on ordinal (non-overlap) comparisons, Tau(AB) is appropriate for a wide range of outcome types. However, it has limited sensitivity and suffers from range restriction when intervention effects are strong enough to create complete non-overlap (Pustejovsky, 2019).

To complement Tau(AB), we also used the log response ratio (LRR; Pustejovsky, 2018), a parametric effect size metric that describes intervention effects in terms of proportional change in the level of the dependent variable from baseline to intervention. LRR values of 0 correspond to no change in level; positive values correspond to increases in the outcome relative to baseline level. Although readily interpretable because of its connection to percentage change, LRR has the drawbacks that it is undefined for data series where the behavior is entirely absent during baseline and its magnitude is not meaningful when baseline levels are very near zero (Pustejovsky, 2018). We used the form of LRR appropriate for outcomes where increase is desirable (i.e., LRR-i).

We used the SingleCaseES package (Pustejovsky & Swan, 2019) for the R programming environment to calculate Tau(AB) and LRR estimates. For Tau(AB), we used the "null" standard error estimator. We estimated effect sizes for pairs of adjacent phases in multiple baseline, multiple probe, and treatment reversal designs. For alternating treatment designs, we calculated effect sizes comparing a specific intervention condition within an alternating treatment phase to the preceding baseline phase. For series that included multiple contrasts, we aggregated effect sizes to the level of the data series (Pustejovsky & Ferron, 2017). To account for the main limitations of LRR, we excluded data series where the baseline data were all at or near zero.

Meta-analysis and Moderator Analysis

For purposes of analysis, we merged data on effect size estimates and study-level moderating characteristics. Moderator analyses investigated the AAC-related outcomes only, because they were considered the most likely to be primary targets of AAC intervention studies. Data from the other communication outcomes category were excluded from moderator analyses because it represented a catch-all category and included any non-AAC outcomes and those outcomes that the primary study authors defined poorly or aggregated across broad communication behaviors.

We then conducted summary meta-analyses and moderator analyses for all included participants and for the sub-groups of participants with ASD and participants with IDD who did not also have ASD. We used the same statistical approach for both effect size measures, based on a multi-level meta-analysis (MLMA) model. The MLMA provides summary estimates of average effect sizes, along with estimates of variation in effects across included cases, participants, and studies (Pustejovsky & Ferron, 2017; Van den Noortgate & Onghena, 2008). Specifically, we used the MLMA model with random effects at the study level, at the participant level, and at the case level (for designs where a single participant was measured across multiple settings or behaviors), estimated using restricted maximum likelihood methods via the metafor package (Viechtbauer, 2010). Using the clubSandwich package (Pustejovsky, 2020), we calculated robust standard errors and confidence intervals for overall average effects, which allow for the possibilities that the standard errors of individual effect size estimates could be mis-estimated or that the structure of the MLMA model could be mis-specified.

We conducted moderator analyses to investigate possible differences in intervention effectiveness based on instructional setting, use of specific instructional features (systematic arrangement, preference assessment, reinforcement, modeling, verbal prompts, physical prompts, prompt fading, and graphic prompts), and use of specific behavioral intervention strategies. Additionally, we investigated whether variation in effect size was related to the *number* of instructional features used (without differentiating between specific features) or to the number of behavioral intervention strategies used (without differentiating between strategies). Moderator analyses were conducted for the subset of effect size estimates corresponding to AAC-related outcome variables, excluding other communication-related outcomes.

We analyzed these potential moderating characteristics using MLMA meta-regression models. Just as in the summary meta-analysis, we conducted separate analyses for each effect size metric (Tau(AB) and LRR) and for the overall set of participants, participants with ASD, and participants with IDD. For categorical variables, we used models with separate intercept terms (i.e., indicator variables) for each category. For numerical predictor variables, we used models with an overall intercept and a linear term for the predictor variable; the predictor was grand mean centered so that the intercept term was interpretable as the overall average effect size. Just as in the summary meta-analysis, all models included random effects at the study level, participant level, and case level; we again used robust variance estimation to insure against mis-specified assumptions.

For our main analyses, we conducted separate analyses for each potential moderator. As sensitivity analysis, we also estimated meta-regression models that simultaneously control for participant characteristics and instructional features. Specifically, the sensitivity analysis included predictors for participant age group, communication modes prior to intervention, word use prior to intervention, and imitation use prior to intervention, as well as instructional setting, instructional features, and use of behavioral intervention strategies. All control variables were grand mean centered to facilitate calculation of marginal average effect sizes.

Inter-rater Reliability

Inter-rater reliability (IRR) was collected for 100% of title review and 39% of abstract and full text review with 88.43% and 88.39% agreement, respectively. During variable coding, IRR was obtained on 20% of the studies with 93% agreement. IRR ranged from 83 to 100% for specific instructional features, 89 to 97% for specific intervention characteristics, and 0 to 93% for setting. Low range was indicative of a limited number of studies rated for a specific characteristic (i.e., $n = 1–2$). IRR was collected for 30% of the studies during data extraction with 98% agreement.

Open Data and Replicability

All data and computer code used to conduct the analyses are available on the Open Science Framework. Raw data and data cleaning code are available on the main project page at https://osf.io/b5ydr/. Codes for replicating all reported analyses are available at https://osf.io/2j5ay/.

Results

This study examined the effects of instructional contexts on the communication behavior of individuals with ASD and/or IDD who experienced CCN. In total, the analysis included $k = 114$ studies with 330 participants, including 247 participants with ASD (from 92 studies) and 83 participants with IDD (from 41 studies). Participants ranged in age from 1 to 21 years old, with a median age of 5 (interquartile range: 4–9.8 years old). Prior to intervention, most participants used multiple modes of communication. Levels of word use and imitation use prior to intervention were reported infrequently. Supplementary Table S1 provides further details about the characteristics of included participants.

Table 1 reports the distribution of instructional settings, instructional features, and behavioral intervention strategies examined in the included studies (see Supplementary Table D1 for information about each included study). Most studies occurred in classroom or clinical settings. The most common instructional features were reinforcement, prompt fading, and systematic arrangement, each of which was used in about 90% of included studies. Fewer studies used verbal prompts, modeling, physical prompts, or preference assessment, and only 6 studies made use of graphic prompts. Studies implemented between 1 and 7 of these instructional features, with a median of 5 features (see Supplementary Figure S1). Regarding behavioral intervention strategies, the vast majority of studies used interventionist initiation, massed teaching opportunities, contrived (rather than naturalistic)

Table 1 Distribution of instructional characteristics

Level	Studies (*k*)	Participants (*P*)	Effect size estimates (*N*)
Instructional setting			
Classroom	40	115	216
Clinic	31	92	243
Home	12	31	65
Other	5	15	24
Multiple	25	74	213
Not reported	1	3	6
Instructional features			
Graphic prompt	6	21	28
Modeling	75	216	483
Physical prompts	63	183	466
Preference assessment	57	168	431
Prompt fading	103	295	679
Reinforcement	106	304	715
Systematic arrangement	102	294	709
Verbal prompts	76	216	540
Initiator			
Child	17	49	71
Interventionist	97	281	696
Teaching opportunities			
Dispersed	12	36	65
Massed	102	294	702
Activity context			
Contrived	93	270	646
Embedded	21	60	121
Instructional format			
Group	10	30	46
One-on-one	104	300	721
Teaching stimuli			
Limited	41	124	279
Varied	73	206	488
Instructional environment			
Controlled	54	160	439
Natural	60	170	328

activity context, and a one-on-one instructional format (see Supplementary Figure S1).

Summary Meta-analysis

Across all included participants and outcomes, the average Tau(AB) effect size was 0.72, 95% confidence interval (CI) [0.67, 0.77]. Effects were highly heterogeneous, with an estimated total SD of 0.22. Across all participants, average effects were slightly larger for AAC-related outcomes (average Tau(AB) = 0.75, 95% CI [0.69, 0.81], total SD = 0.22). Effects were smaller and more heterogeneous for other communication-related outcomes (average Tau(AB) = 0.65, 95% CI [0.56, 0.73], total SD = 0.30). As shown in Table 2, these trends were very similar for the subsets of participants with ASD and participants with IDD.

When using the LRR effect size metric, the overall average effect size across all participants and outcomes was 1.86, 95% CI [1.58, 2.13]. This average LRR corresponds to an improvement of 541% from baseline to intervention levels, 95% CI [386, 744]. LRR effects were also highly heterogeneous at the study, participant, and contrast levels, with an estimated total SD of 1.62. Across all participants, effects were larger for AAC-related outcomes, with an average LRR of 2.01, 95% CI [1.64, 2.39] (corresponding to a 649% improvement, 95% CI [413, 993]), and smaller for other communication outcomes, with an average LRR of 1.60, 95% CI [1.20, 2.00] (corresponding to a 394% improvement, 95% CI [232, 635]). As shown in Table 3, these trends were similar when limited to the subset of participants with ASD. However, when limited to participants with IDD, average effect sizes for AAC-related outcomes were smaller and those for other communication-related outcomes were larger (Table 3).

Effects of Setting on AAC Instruction

To understand variation in effects across instructional settings, we conducted moderator analyses of Tau(AB) and LRR effect sizes for the AAC-related outcomes. For the Tau(AB) metric, differences in average effects across instructional settings were not statistically distinguishable, nor were differences apparent for participants with ASD or participants with IDD. Across all participants, estimated Tau(AB) effect sizes were largest (in terms of numerical magnitude) for studies conducted in home settings or clinical settings and were relatively smaller for studies conducted in classroom settings, although these differences were not statistically distinct (see Table 2). Similarly, for the LRR metric, differences across instructional settings were not statistically distinguishable, although average effect size estimates were numerically largest for studies conducted in home or clinical settings and smallest for studies conducted in classroom settings or other settings.

Associations of Instructional Features and AAC Outcomes

We examined associations between effect sizes for AAC-related outcomes and instructional feature use in terms of the number of features used and the use of specific features. The number of instructional features was not significantly associated with Tau(AB) effect sizes in the overall sample, nor in the subsamples of participants with ASD or participants with IDD (Table 2). Of the specific instructional features

Table 2 Average effect sizes based on Tau(AB) metric

Category	All participants		Participants with ASD		Participants with IDD	
	Est. (SE)	95% CI	Est. (SE)	95% CI	Est. (SE)	95% CI
Overall average	0.719 (0.025)	[0.670, 0.768]	0.707 (0.028)	[0.652, 0.761]	0.728 (0.050)	[0.628, 0.829]
AAC outcomes	0.746 (0.031)	[0.685, 0.807]	0.731 (0.035)	[0.661, 0.801]	0.774 (0.052)	[0.666, 0.882]
Other communication outcomes	0.648 (0.043)	[0.562, 0.734]	0.643 (0.047)	[0.547, 0.739]	0.650 (0.086)	[0.469, 0.830]
Instructional setting[ab]	$F(3,26.9) = 1.9$	$p = .150$	$F(3,22.6) = 2.3$	$p = .108$	$F(3,3.3) = 2.2$	$p = .248$
Classroom	0.654 (0.068)	[0.513, 0.795]	0.633 (0.087)	[0.450, 0.816]	0.609 (0.109)	[0.352, 0.866]
Clinic	0.824 (0.035)	[0.752, 0.896]	0.835 (0.034)	[0.765, 0.906]	0.812 (0.070)	[0.643, 0.981]
Home	0.821 (0.037)	[0.731, 0.912]	0.797 (0.048)	[0.676, 0.917]	0.905 (0.011)	[0.769, 1.042]
Other	0.688 (0.403)	[−4.433, 5.810]	0.684 (0.404)	[−4.451, 5.820]		
Multiple	0.733 (0.073)	[0.576, 0.890]	0.659 (0.087)	[0.469, 0.848]	0.927 (0.017)	[0.879, 0.976]
Not reported	0.622 (0.048)	[0.007, 1.237]	0.622 (0.051)	[−0.021, 1.265]		
Number of instructional features[a]	$F(1,17.1) = 1.4$	$p = .251$	$F(1,10.9) = 1.7$	$p = .221$	$F(1,7.2) = 0.1$	$p = .796$
Overall average effect	0.749 (0.030)	[0.690, 0.808]	0.736 (0.034)	[0.668, 0.803]	0.772 (0.051)	[0.666, 0.878]
Instructional features	−0.027 (0.023)	[−0.075, 0.021]	−0.035 (0.027)	[−0.094, 0.024]	0.011 (0.040)	[−0.084, 0.105]
Systematic arrangement[a]	$F(1,8.4) = 1.0$	$p = .347$	$F(1,4.4) = 5.6$	$p = .071$	$F(1,3.6) = 0.4$	$p = .560$
No	0.815 (0.070)	[0.650, 0.981]	0.891 (0.063)	[0.713, 1.069]	0.707 (0.108)	[0.329, 1.084]
Yes	0.739 (0.033)	[0.673, 0.805]	0.719 (0.037)	[0.645, 0.793]	0.786 (0.058)	[0.663, 0.908]
Preference assessment[a]	$F(1,66.6) = 0.9$	$p = .346$	$F(1,52.7) = 0.5$	$p = .484$	$F(1,20.5) = 0.4$	$p = .521$
No	0.779 (0.043)	[0.692, 0.867]	0.760 (0.053)	[0.652, 0.869]	0.809 (0.051)	[0.695, 0.922]
Yes	0.722 (0.042)	[0.636, 0.808]	0.711 (0.047)	[0.616, 0.806]	0.742 (0.089)	[0.544, 0.939]
Reinforcement[a]	$F(1,4.1) = 0.9$	$p = .390$	$F(1,2.0) = 0.9$	$p = .448$	$F(1,2.4) = 0.1$	$p = .762$
No	0.652 (0.099)	[0.363, 0.941]	0.561 (0.185)	[−0.324, 1.447]	0.796 (0.048)	[0.574, 1.019]
Yes	0.752 (0.032)	[0.688, 0.816]	0.738 (0.036)	[0.667, 0.810]	0.770 (0.059)	[0.647, 0.893]
Modeling[a]	$F(1,62.4) = 5.6$	$p = .021$	$F(1,54.2) = 6.6$	$p = .013$	$F(1,15.3) = 0.4$	$p = .547$
No	0.824 (0.033)	[0.756, 0.893]	0.825 (0.037)	[0.749, 0.901]	0.812 (0.068)	[0.651, 0.973]
Yes	0.692 (0.045)	[0.601, 0.782]	0.661 (0.052)	[0.555, 0.767]	0.751 (0.073)	[0.594, 0.907]
Verbal prompts[a]	$F(1,48.7) = 1.1$	$p = .304$	$F(1,41.2) = 2.9$	$p = .094$	$F(1,11.0) = 1.5$	$p = .251$
No	0.785 (0.039)	[0.704, 0.867]	0.804 (0.045)	[0.711, 0.896]	0.693 (0.068)	[0.524, 0.862]
Yes	0.726 (0.042)	[0.642, 0.809]	0.692 (0.047)	[0.596, 0.788]	0.809 (0.067)	[0.666, 0.952]
Prompt fading[a]	$F(1,5.1) = 0.6$	$p = .487$	$F(1,2.0) = 0.0$	$p = .949$	$F(1,2.3) = 1.4$	$p = .352$
No	0.820 (0.102)	[0.548, 1.092]	0.748 (0.242)	[−0.402, 1.897]	0.875 (0.082)	[0.450, 1.301]
Yes	0.740 (0.032)	[0.677, 0.804]	0.730 (0.035)	[0.659, 0.801]	0.758 (0.059)	[0.635, 0.880]
Graphic prompt[a]	$F(1,1.1) = 0.8$	$p = .520$	$F(1,1.1) = 14.3$	$p = .142$	$F(1,1.9) = 0.1$	$p = .812$
No	0.744 (0.031)	[0.681, 0.807]	0.728 (0.035)	[0.657, 0.799]	0.774 (0.055)	[0.660, 0.888]
Yes	0.820 (0.077)	[−0.150, 1.790]	0.899 (0.028)	[0.541, 1.257]	0.752 (0.062)	[−0.035, 1.539]
Physical prompts[a]	$F(1,50.3) = 1.3$	$p = .252$	$F(1,28.1) = 2.0$	$p = .172$		
No	0.694 (0.058)	[0.574, 0.814]	0.637 (0.082)	[0.464, 0.810]		
Yes	0.773 (0.035)	[0.702, 0.844]	0.763 (0.038)	[0.686, 0.840]		
Behavioral intervention strategies (scale)[a]	$F(1,18.9) = 3.5$	$p = .077$	$F(1,13.9) = 6.3$	$p = .025$	$F(1,7.5) = 0.0$	$p = .865$
Overall average effect	0.749 (0.029)	[0.691, 0.808]	0.737 (0.032)	[0.672, 0.802]	0.773 (0.053)	[0.663, 0.883]
Strategies	0.055 (0.029)	[−0.007, 0.117]	0.084 (0.033)	[0.012, 0.155]	0.005 (0.029)	[−0.063, 0.073]
Initiator[a]	$F(1,6.9) = 0.6$	$p = .460$	$F(1,1.1) = 0.6$	$p = .573$	$F(1,6.1) = 0.2$	$p = .705$
Child	0.660 (0.117)	[0.370, 0.949]	0.437 (0.389)	[−4.506, 5.380]	0.743 (0.076)	[0.529, 0.957]
Interventionist	0.755 (0.032)	[0.692, 0.818]	0.740 (0.034)	[0.673, 0.808]	0.782 (0.063)	[0.648, 0.916]
Teaching opportunities[a]	$F(1,5.5) = 0.2$	$p = .675$	$F(1,4.4) = 0.6$	$p = .468$	$F(1,1.2) = 7.5$	$p = .184$
Dispersed	0.683 (0.153)	[0.280, 1.086]	0.598 (0.180)	[0.083, 1.113]	0.942 (0.038)	[0.464, 1.421]
Massed	0.752 (0.031)	[0.691, 0.813]	0.743 (0.035)	[0.673, 0.813]	0.756 (0.056)	[0.639, 0.873]
Activity context[a]	$F(1,18.0) = 0.1$	$p = .779$	$F(1,13.2) = 0.6$	$p = .444$	$F(1,5.7) = 2.9$	$p = .142$

Table 2 (continued)

Category	All participants		Participants with ASD		Participants with IDD	
	Est. (SE)	95% CI	Est. (SE)	95% CI	Est. (SE)	95% CI
Contrived	0.751 (0.032)	[0.688, 0.815]	0.746 (0.035)	[0.676, 0.817]	0.743 (0.064)	[0.608, 0.877]
Embedded	0.724 (0.092)	[0.525, 0.922]	0.651 (0.116)	[0.390, 0.912]	0.882 (0.051)	[0.731, 1.034]
Instructional format[a]	F(1,4.3) = 0.6	p = .470	F(1,3.1) = 0.2	p = .700	F(1,2.1) = 0.2	p = .732
Group	0.678 (0.086)	[0.432, 0.924]	0.683 (0.114)	[0.298, 1.068]	0.731 (0.105)	[0.191, 1.272]
One-on-one	0.751 (0.032)	[0.687, 0.815]	0.734 (0.037)	[0.661, 0.807]	0.778 (0.057)	[0.660, 0.896]
Teaching stimuli[a]	F(1,59.1) = 3.9	p = .052	F(1,52.9) = 3.5	p = .067	F(1,11.9) = 2.3	p = .157
Limited	0.814 (0.036)	[0.740, 0.888]	0.802 (0.042)	[0.716, 0.889]	0.855 (0.023)	[0.799, 0.911]
Varied	0.701 (0.044)	[0.613, 0.790]	0.678 (0.052)	[0.573, 0.783]	0.741 (0.072)	[0.587, 0.895]
Instructional environment[a]	F(1,68.3) = 3.7	p = .059	F(1,51.2) = 4.5	p = .038	F(1,20.5) = 0.3	p = .568
Controlled	0.800 (0.034)	[0.731, 0.868]	0.795 (0.036)	[0.721, 0.868]	0.804 (0.061)	[0.669, 0.940]
Natural	0.682 (0.051)	[0.577, 0.786]	0.639 (0.064)	[0.508, 0.770]	0.744 (0.084)	[0.557, 0.931]

Est., average effect size estimate; *SE*, standard error; *CI*, confidence interval; *F*, *F*-statistic for test that average effect sizes are equal across categories; numerator and denominator degrees of freedom are reported in parentheses after the test statistic

[a]Average effect sizes for AAC outcomes only

[b]*F*-tests are for differences between the categories classroom, clinic, home, or multiple settings (excluding other and not reported)

examined, systematic differences in Tau(AB) effects were apparent only for the use of modeling (Table 2). Studies implementing modeling had average effects of 0.69, 95% CI [0.60, 0.78], whereas studies not implementing modeling had average effects of 0.82, 95% CI [0.76, 0.89], a statistically significant difference. This systematic difference was also apparent in the subsample of participants with ASD, but not in the subsample of participants with IDD. Differences by use of modeling were of similar magnitude but were not statistically significant in models controlling for participant characteristics and other aspects of the instructional context (see Supplementary Table S2). No other specific instructional features were significantly associated with Tau(AB) effect sizes, and the numerical estimates of differences in average effect size were often counterintuitive. For example, studies implementing systematic arrangement had lower estimated effect sizes than studies not implementing this feature, and a similar pattern held for use of preference assessment, modeling, verbal prompts, and prompt fading.

The number of instructional features was not significantly associated with LRR effect sizes in the overall sample, nor in the subsamples (Table 3). Of the specific features examined, systematic differences only emerged for use of physical prompts, where studies implementing physical prompts had average LRR effects that were significantly larger than studies not using physical prompts. This systematic difference was also evident in the subsample of participants with ASD, and the differences remained statistically significant in models controlling for participant characteristics and other aspects of the instructional context (see Supplementary Table S3). No other specific instructional features were significantly associated with LRR effect sizes. However, unlike with Tau(AB), studies that implemented a specific feature generally tended to have estimated LRR effect sizes numerically larger than those in studies not implementing the feature.

Behavioral Versus Naturalistic AAC Instructional Features

Finally, we examined the association between effect size for AAC-related outcomes and use of behavioral versus naturalistic intervention strategies, both in terms of the overall number of behavioral strategies and in terms of specific strategies. Tau(AB) effect sizes had a small, positive association with the use of more behavioral intervention strategies across the full sample of participants, but this association was not statistically distinguishable from null ($\beta = 0.055$, 95% CI [−0.007, 0.117], $p = .077$). In the subsample of participants with ASD, the association was stronger and statistically distinct from null: use of one additional behavioral strategy was associated with a difference of $\beta = 0.084$ in average Tau(AB), 95% CI [0.012, 0.155], $p = .025$. In the subsample of participants with IDD, the association was near zero ($\beta = 0.005$, 95% CI [−0.063, 0.073], $p = .865$). These estimated associations were of similar magnitude but less precise in models controlling for participant characteristics and other features of the instructional context (Supplementary Table S2). Of the specific intervention strategies, none had statistically distinguishable associations with Tau(AB) effect sizes in the full participant sample or in the subsample of participants with IDD. Among participants with ASD, only one strategy had a statistically distinct association: interventions conducted in controlled environments

Table 3 Average effect sizes based on log response ratio metric

Category	All participants		Participants with ASD		Participants with IDD	
	Est. (SE)	95% CI	Est. (SE)	95% CI	Est. (SE)	95% CI
Overall average	1.857 (0.139)	[1.581, 2.133]	1.824 (0.164)	[1.498, 2.150]	1.872 (0.225)	[1.413, 2.331]
AAC outcomes	2.013 (0.189)	[1.635, 2.392]	2.125 (0.219)	[1.685, 2.565]	1.683 (0.298)	[1.053, 2.312]
Other communication outcomes	1.597 (0.198)	[1.199, 1.995]	1.374 (0.230)	[0.909, 1.840]	1.962 (0.292)	[1.344, 2.579]
Instructional setting[ab]	$F(3,20.4) = 0.2$	$p = .872$	$F(3,15.3) = 0.3$	$p = .850$	$F(3,3.1) = 0.3$	$p = .805$
Classroom	1.779 (0.357)	[1.011, 2.546]	1.833 (0.423)	[0.887, 2.780]	1.768 (0.635)	[0.211, 3.324]
Clinic	2.119 (0.271)	[1.554, 2.684]	2.167 (0.292)	[1.555, 2.780]	1.985 (0.651)	[0.137, 3.834]
Home	2.299 (0.980)	[−0.223, 4.822]	2.594 (1.152)	[−0.610, 5.799]	1.055 (0.470)	[−4.915, 7.025]
Other	0.405 (0.369)	[−4.289, 5.098]	0.417 (0.291)	[−3.283, 4.117]		
Multiple	2.140 (0.356)	[1.360, 2.919]	2.347 (0.452)	[1.316, 3.378]	1.534 (0.246)	[0.743, 2.325]
Not reported	1.146 (-)		1.146 (-)			
Number of instructional features[a]	$F(1,15.3) = 4.5$	$p = .051$	$F(1,9.3) = 3.5$	$p = .093$	$F(1,5.9) = 0.2$	$p = .648$
Overall average effect	1.992 (0.182)	[1.628, 2.357]	2.097 (0.212)	[1.668, 2.526]	1.669 (0.287)	[1.058, 2.281]
Instructional features	0.201 (0.095)	[−0.001, 0.403]	0.193 (0.103)	[−0.039, 0.426]	0.145 (0.302)	[−0.598, 0.889]
Systematic arrangement[a]	$F(1,7.7) = 0.2$	$p = .666$	$F(1,3.6) = 0.0$	$p = .993$	$F(1,4.6) = 0.0$	$p = .950$
No	1.804 (0.490)	[0.602, 3.006]	2.118 (0.742)	[−0.258, 4.495]	1.648 (0.595)	[−0.264, 3.559]
Yes	2.043 (0.205)	[1.630, 2.455]	2.125 (0.232)	[1.658, 2.593]	1.694 (0.352)	[0.932, 2.457]
Preference assessment[a]	$F(1,50.6) = 0.6$	$p = .461$	$F(1,37.5) = 0.7$	$p = .392$	$F(1,15.6) = 0.1$	$p = .747$
No	2.178 (0.296)	[1.566, 2.790]	2.370 (0.385)	[1.560, 3.181]	1.584 (0.240)	[1.028, 2.139]
Yes	1.892 (0.247)	[1.389, 2.394]	1.967 (0.261)	[1.431, 2.503]	1.784 (0.560)	[0.485, 3.083]
Reinforcement[a]	$F(1,4.7) = 4.5$	$p = .091$	$F(1,2.2) = 27.5$	$p = .028$	$F(1,2.8) = 0.2$	$p = .696$
No	1.395 (0.246)	[0.707, 2.083]	0.964 (0.059)	[0.699, 1.230]	1.825 (0.157)	[1.143, 2.507]
Yes	2.072 (0.204)	[1.662, 2.482]	2.204 (0.229)	[1.743, 2.666]	1.656 (0.358)	[0.887, 2.425]
Modeling[a]	$F(1,43.4) = 0.7$	$p = .419$	$F(1,35.1) = 1.3$	$p = .270$	$F(1,9.4) = 1.7$	$p = .223$
No	1.834 (0.224)	[1.368, 2.301]	1.841 (0.265)	[1.281, 2.402]	2.194 (0.441)	[1.040, 3.347]
Yes	2.122 (0.272)	[1.570, 2.674]	2.299 (0.311)	[1.662, 2.936]	1.437 (0.377)	[0.606, 2.268]
Verbal prompts[a]	$F(1,32.6) = 0.2$	$p = .639$	$F(1,27.5) = 0.1$	$p = .743$	$F(1,4.8) = 1.3$	$p = .303$
No	1.871 (0.387)	[1.053, 2.688]	2.011 (0.452)	[1.039, 2.984]	1.229 (0.349)	[0.114, 2.344]
Yes	2.080 (0.213)	[1.649, 2.510]	2.181 (0.243)	[1.686, 2.676]	1.814 (0.369)	[1.015, 2.613]
Prompt fading[a]	$F(1,4.6) = 3.3$	$p = .133$	$F(1,2.2) = 3.4$	$p = .197$	$F(1,1.3) = 0.0$	$p = .920$
No	1.456 (0.266)	[0.705, 2.206]	1.319 (0.406)	[−0.500, 3.138]	1.629 (0.417)	[−3.666, 6.923]
Yes	2.067 (0.204)	[1.657, 2.477]	2.176 (0.229)	[1.714, 2.639]	1.693 (0.340)	[0.968, 2.419]
Physical prompts[a]	$F(1,46.3) = 9.4$	$p = .004$	$F(1,25.4) = 6.3$	$p = .019$		
No	1.387 (0.222)	[0.926, 1.847]	1.378 (0.335)	[0.656, 2.099]		
Yes	2.424 (0.255)	[1.905, 2.943]	2.442 (0.262)	[1.908, 2.977]		
Graphic prompt[a]	$F(1,1.1) = 0.0$	$p = .930$			$F(1,1.3) = 0.6$	$p = .557$
No	2.013 (0.192)	[1.627, 2.399]			1.659 (0.317)	[0.986, 2.332]
Yes	2.050 (0.289)	[−1.617, 5.718]			2.062 (0.418)	[−3.250, 7.374]
Behavioral intervention strategies (scale)[a]	$F(1,16.5) = 1.5$	$p = .233$	$F(1,12.2) = 1.1$	$p = .306$	$F(1,6.2) = 0.0$	$p = .935$
Overall average effect	2.012 (0.188)	[1.635, 2.389]	2.124 (0.218)	[1.684, 2.563]	1.686 (0.309)	[1.030, 2.342]
Strategies	0.144 (0.117)	[−0.102, 0.391]	0.162 (0.151)	[−0.168, 0.491]	0.016 (0.185)	[−0.434, 0.465]
Initiator[a]	$F(1,7.7) = 0.3$	$p = .577$	$F(1,1.1) = 0.4$	$p = .634$	$F(1,7.4) = 0.8$	$p = .387$
Child	1.748 (0.476)	[0.582, 2.914]	2.872 (1.210)	[−12.496, 18.239]	1.318 (0.402)	[0.198, 2.438]
Interventionist	2.050 (0.206)	[1.637, 2.463]	2.093 (0.224)	[1.642, 2.544]	1.831 (0.387)	[0.985, 2.678]
Teaching opportunities[a]	$F(1,6.1) = 0.1$	$p = .745$	$F(1,4.9) = 0.0$	$p = .911$	$F(1,1.3) = 0.2$	$p = .699$
Dispersed	2.308 (0.947)	[−0.142, 4.758]	2.250 (1.175)	[−1.046, 5.546]	1.910 (0.403)	[−3.207, 7.027]
Massed	1.980 (0.185)	[1.608, 2.352]	2.110 (0.209)	[1.687, 2.533]	1.657 (0.333)	[0.946, 2.368]
Activity context[a]	$F(1,15.0) = 0.3$	$p = .576$	$F(1,10.0) = 0.7$	$p = .425$	$F(1,7.2) = 0.5$	$p = .489$

Table 3 (continued)

Category	All participants		Participants with ASD		Participants with IDD	
	Est. (SE)	95% CI	Est. (SE)	95% CI	Est. (SE)	95% CI
Contrived	2.054 (0.225)	[1.601, 2.507]	2.192 (0.251)	[1.684, 2.699]	1.583 (0.403)	[0.704, 2.463]
Embedded	1.847 (0.283)	[1.216, 2.478]	1.801 (0.397)	[0.859, 2.742]	1.944 (0.288)	[1.135, 2.753]
Instructional format[a]	$F(1,3.4) = 0.3$	$p = .630$	$F(1,2.2) = 0.3$	$p = .631$	$F(1,2.7) = 0.2$	$p = .667$
Group	1.794 (0.398)	[0.517, 3.072]	1.824 (0.532)	[−0.511, 4.159]	1.918 (0.456)	[−0.062, 3.898]
One-on-one	2.030 (0.201)	[1.626, 2.433]	2.145 (0.231)	[1.679, 2.611]	1.644 (0.344)	[0.904, 2.383]
Teaching stimuli[a]	$F(1,34.1) = 1.2$	$p = .274$	$F(1,28.5) = 1.7$	$p = .204$	$F(1,6.5) = 0.0$	$p = .939$
Limited	2.346 (0.384)	[1.538, 3.154]	2.570 (0.444)	[1.620, 3.519]	1.719 (0.446)	[0.446, 2.992]
Varied	1.859 (0.211)	[1.430, 2.287]	1.912 (0.242)	[1.418, 2.406]	1.673 (0.378)	[0.849, 2.497]
Instructional environment[a]	$F(1,50.2) = 0.3$	$p = .584$	$F(1,33.6) = 0.1$	$p = .786$	$F(1,12.3) = 0.1$	$p = .780$
Controlled	2.106 (0.228)	[1.640, 2.571]	2.174 (0.247)	[1.668, 2.679]	1.792 (0.460)	[0.655, 2.930]
Natural	1.888 (0.323)	[1.221, 2.555]	2.038 (0.429)	[1.130, 2.946]	1.618 (0.404)	[0.716, 2.519]

k, number of studies; *P*, number of participants; *N*, number of data series; *Est.*, average effect size estimate; *SE*, standard error; *CI*, confidence interval; *F*, *F*-statistic for test that average effect sizes are equal across categories; numerator and denominator degrees of freedom are reported in parentheses after the test statistic

[a]Average effect sizes for AAC outcomes only

[b]*F*-tests are for differences between the categories classroom, clinic, home, or multiple settings (excluding other and not reported)

were associated with higher Tau(AB) effects than interventions conducted in naturalistic environments (Table 2). However, this difference was smaller and not statistically distinguishable from null when controlling for participant characteristics and instructional context variables (Supplementary Table S2).

When measuring effect size using the LRR metric, the number of behavioral intervention strategies was not significantly associated with effect sizes for AAC-related outcomes in the full sample ($\beta = 0.144$, 95% CI [−0.102, 0.391], $p = .233$), nor in either subsample. When considering use of behavioral intervention strategies taken singly, none had statistically distinguishable associations with LRR effect sizes in the full participant sample or in either subsample. This pattern was similar in models with additional control variables (Supplementary Table S3).

Exploratory Analysis

In further, exploratory analysis, we considered how instructional features and intervention strategy use varied depending on participant background characteristics. When considering use of instructional features as a function of participant age (Supplementary Figure S2), it was apparent that graphic prompt use and systematic arrangement were used more frequently with younger participants, while decreasing in frequency with older participants. When considering behavioral intervention strategies by participant age (Supplementary Figure S3), we noted that use of limited stimuli decreased at older ages; other strategies were used at fairly consistent levels across age bands. Analysis of instructional features by participant word use (Supplementary Figure S4) showed that use of preference assessment decreased as participants use more words, possibly because they can request items verbally or have moved to secondary reinforcers such as tokens that they may use later to select earned reinforcement, or are intrinsically reinforced by learning new material. Behavioral intervention strategy use did not appear to vary depending on participant word use (Supplementary Figure S5). Considering instructional features by participant verbalization/vocalization (Supplementary Figure S6), we noted that there was less use of physical prompts and preference assessments among vocal/verbal participants than those who were minimally or non-verbal and a greater use of verbal prompts for verbal participants, which aligns with expectations. Finally, considering behavioral intervention strategy use by participant verbalization/vocalization (Supplementary Figure S7), we noted that use of limited stimuli and use of controlled environments were less frequent with vocal/verbal participants than with those who were minimally or non-verbal; all other strategies were used at high levels with both groups of participants.

Discussion

In this comprehensive systematic review and meta-analysis, we synthesized a large body of single-case experiments evaluating AAC interventions for school-aged individuals with ASD and/or IDD, and considered the associations between instructional features and AAC-related outcomes. Average effects expressed as Tau and LRR indicated that AAC interventions were effective. However, few differences were found between potential moderators related to instructional contexts, including instructional setting; number of instructional features (i.e., graphical/physical/verbal prompts, prompt fading, modeling, preference assessments, reinforcement, systematic arrangement); or functional-behavioral versus social-behavioral intervention strategies (i.e., embedded versus contrived contexts, interventionist- versus child-initiated, natural versus controlled instruction, one-on-one versus group instruction, massed versus dispersed instructional opportunities, varied versus limited stimuli).

Overall effects were divided into those outcomes that reported AAC-related social communication behaviors (i.e., use of aided AAC, unaided AAC, or gestures) and those that reported other social communication behaviors (i.e., verbalizations, vocalizations, and dependent variables that measured aggregated communication behaviors for which the reader could not determine rates of AAC-related behaviors independently). Tau(AB) analyses found significantly higher effects for AAC-related outcomes than for other social communication outcomes, and also found that heterogeneity was somewhat smaller and few of the studies in the AAC-related DVs group had ES scores of 0 or lower.

Given that participants with ASD may respond differently to instructional features than those with IDD, outcomes were further divided into groups of participants with ASD, including those with multiple diagnoses, and those with IDD who did not have an ASD diagnosis. Results for participants with ASD closely tracked the overall results. One exception is that the participants with ASD had statistically higher effects for higher numbers of behavioral strategies implemented for Tau(AB). Further, for participants with ASD, interventions conducted in controlled environments had higher Tau(AB) effects than those in naturalistic environments, although this difference was not statistically significant when controlling for participant characteristics and instructional context variables. The results for participants with IDD demonstrated some differences from the overall results, which are discussed below. However, these differences may simply reflect fewer participants in those categories, resulting in a magnification of performance variability.

Some of the outcomes resulting from the Tau(AB) analyses were counterintuitive, although most were not significantly different with regard to whether or not a particular feature was implemented (e.g., systematic arrangement, preference assessment, reinforcement). For instance, although not significantly different for the overall participant pool, among learners with ASD, studies that did not implement systematic arrangement had larger effects than those that did involve systematic arrangement. Further, studies that did not implement modeling had significantly larger effects, for the overall pool and participants with ASD, than those that did report implementing a modeling strategy. This was not the case for individuals with IDD without ASD. It is possible that participants with ASD are less likely to have imitation skills than those with IDD who do not have ASD. Similarly, for participants with ASD, studies that involved verbal prompts had lower effects than those that did not. It is less likely that participants with ASD had advanced verbal repertoires. Consequently, interventions that involved verbal prompts may have been less effective compared to those that utilized physical prompts for teaching AAC skills. This also aligns with the topography of AAC, which involves a physical act versus a verbal act; thus, reliance on physical prompts may make more sense than verbal prompts. Physical prompts did have larger effects across diagnoses, although not statistically different from studies that did not use physical prompts.

Results for moderator analyses using LRR differed somewhat from results obtained using Tau(AB). Reinforcement procedures had larger effects for both the overall participant pool and for those with ASD. Although not statistically significant, modeling also had a larger effect than not modeling with respect to the overall pool and for those with ASD. Physical prompts were calculated for the overall pool and those with ASD, with systematically larger effects for studies using physical prompts than for studies that did not use them with either group.

There were no systematic differences in the number of behavioral intervention strategies or use or non-use of specific strategies for Tau(AB) or LRR, with one exception. For Tau(AB), controlled contexts produced a larger effect than natural instructional environments. This is understandable given that the data reported were for initial acquisition, which may be responsive in the short term, while generalization may be addressed better given natural instructional strategies. Although beyond the scope of this meta-analysis, evaluating generalization and maintenance data rather than acquisition data might lead to different findings.

Results in Relation to Prior Reviews and Meta-analyses

As noted above, the extant literature has lacked meta-analyses related to implementation of AAC in naturalistic contexts (Crowe et al., 2021; Ganz et al., 2017; Sutton et al., 2019). We conducted analyses on factors related to natural contexts and instructional strategies, but we also found that the small number of primary studies that allowed for coding of these variables likely impacted our ability to detect any differences that may exist. Although some reviews have narratively described naturalistic instructional strategies (Gevarter & Zamora, 2018; Holyfield et al., 2017), this study expanded on that through quantitative analyses.

Research and Practice Implications

This meta-analysis included calculation of two metrics, Tau(AB) and LRR, with distinct interpretations. For instance, Tau(AB) measures the extent of overlap of data between phases, but LRR measures, in proportional terms, the size of a shift in the level of the outcome from baseline to intervention. Additionally, Tau(AB) has a ceiling effect, meaning that, beyond a point at which the data do not overlap, Tau(AB) does not further distinguish the size of effects. However, LRR is limited to use with data series that are above floor levels, which resulted in the exclusion of a sizeable number of studies from the calculations. This suggests that researchers should use multiple metrics, rather than relying on one, given the features and limitations of each one. With respect to setting, we found no statistically significant differences between settings. Thus, there appears to be little justification for researchers to conduct the bulk of AAC intervention research in lab or research settings rather than authentic settings.

Limitations and Future Research

Limitations in the primary literature base impacted our ability to detect differences in potential moderators. For instance, use of some of the behavioral versus naturalistic instructional features was heavily lopsided, which limited our ability to investigate moderating effects of these variables. Few studies embedded instruction within natural contexts, included primarily child-lead instruction, or involved dispersing learning opportunities across time, activities, and other contexts. Such naturalistic strategies would have been apt to have greater contextual fit in typical classrooms and homes, allowing for more natural and efficient implementation of instruction/maintenance of AAC use within natural daily routines.

The population sampled in the current investigation was large and heterogeneous. This heterogeneity could not be well quantified as a result of the limited methodological detail provided by many authors. There were a number of omissions in the details reported by the original study authors that contributed to an inability to detect differences between moderators, if they existed. For example, cognitive status data were often missing from participant descriptions. Among those with intellectual delay, it was impossible to apply standard quantifiable descriptions related to severity. Communication status at the outset of intervention was often vaguely or inadequately described. Often key information was omitted related to skills that could have a significant influence on communicative performance such as vocabulary comprehension skills, presence of receptive and/or expressive joint attention, and imitation. Notably, in studies teaching productive use of AAC, it was rare to see a description of comprehension skills—even though comprehension is a likely mediator of productive communication skills in AAC. Further, participant descriptions typically lacked information about the size of the learner's graphic, gestural, and vocal/verbal repertoire at the outset of intervention. Often, prior use of AAC applications was not addressed in describing participants. Only 77 of the 114 identified studies included AAC-related dependent variables, and 53 did not; in many cases, this was because the primary study authors' measures were broad communication behaviors and combined vocalizations/verbalizations with AAC outcomes. All of these factors present limitations for any detailed analysis of moderators of communication skill acquisition.

Another limitation of the current meta-analysis is that we did not investigate treatment intensity (e.g., by Warren et al., 2007). This is particularly important in evaluating intervention procedures. However, in the studies examined in this investigation, rarely were the majority of treatment intensity parameters delineated. This, in turn, may have introduced an uncontrolled variable relevant for control in future systematic reviews and meta-analyses.

Reinforcement was almost universally implemented, which could be a confounding factor related to determining the efficacy of various AAC packages and procedures. That is, reinforcement is universally known to be effective in increasing and maintaining behaviors, which may have a heavy impact on the outcomes of given AAC intervention procedures. This calls to question whether procedures could be streamlined, such as by relying heavily on the use of reinforcement to increase learning in natural communication partners and individuals with CCN.

A final limitation is that the evidence identified in this systematic review may have been influenced by publication biases, such as the preference of journals to publish single-case studies with visually apparent and consistent functional relations (Shadish et al., 2016; Sham & Smith, 2014; Tincani & Travers, 2018). In another component of this project, we used statistical tools for investigating publication bias, finding few indications of bias. However, available tools were developed for meta-analyses of group designs and are not ideally suited for detecting publication bias in single-case research designs. Thus, the true extent of publication bias remains difficult to gauge. Effectively mitigating the inferential threat posed by publication bias may require changes in journal publication practices (Kittelman et al., 2018) and wider adoption of pre-registration and registered reports of single-case research (Cook et al., 2021; Johnson & Cook, 2019).

Conclusion

The current investigation represents one of the largest meta-analyses examining instructional strategies implemented in serving school-aged learners experiencing

ASD and/or IDD who also have a CCN. Average effects expressed as Tau and LRR indicated that AAC interventions are effective. However, we were not able to identify specific moderators that had a distinct effect on treatment outcomes. In part, this may be influenced by the heterogeneity of the population. Unfortunately, heterogeneity of the participants could not be easily addressed.

Overall, analyses for AAC-related social communication behaviors (i.e., use of aided AAC, unaided AAC, or gestures) resulted in significantly greater effects for AAC-related outcomes than other social communication behaviors (i.e., verbalizations, vocalizations, and dependent variables that measured aggregated communication behaviors for which the reader could not determine rates of AAC-related behaviors independently). Related to diagnosis, at the outset, it was reasonable to hypothesize that persons with ASD may respond differently to certain instructional features than those with IDD. Although the results for participants with IDD resulted in some differences, we cannot rule out that differences may reflect fewer participants in that category.

Going forward, there is a need for more detailed and better controlled analyses of potential moderator variables. This may be best addressed by identifying homogeneous populations and directly comparing a variable while controlling remaining variables. Unfortunately, this work requires that primary studies are thoroughly and completely reported. Particularly important is carefully describing participant characteristics. It is important to know more about specific learner characteristics at the outset of an intervention study so that their contribution to the learning process can be evaluated. We are hopeful that future research will focus on the impact of components of intervention strategies and their effectiveness. Doing so is our best means of developing individualized intervention protocols to meet the needs of a very heterogeneous population of learners with CCN.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s40489-022-00314-w.

Funding The research described here is supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R324A180110 to Texas A&M University. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

Data Availability Supplementary materials for this study, including raw data and code for replicating all results, are available on the Open Science Framework at https://osf.io/2j5ay/.

Declarations

Conflict of Interest The authors declare no competing interests.

References

Biggs, E. E., Carter, E. W., & Gilson, C. B. (2018). Systematic review of interventions involving aided AAC modeling for children with complex communication needs. *American Journal on Intellectual and Developmental Disabilities, 123*(5), 443–473. https://doi.org/10.1352/1944-7558-123.5.443

Branson, D., & Demchak, M. (2009). The use of augmentative and alternative communication methods with infants and toddlers with disabilities: A research review. *Augmentative and Alternative Communication, 25*(4), 274–286. https://doi.org/10.3109/07434610903384529

Centers for Disease Control and Prevention. (2018). *Autism prevalence slightly higher in CDC's ADDM network*. Retrieved from Center for Disease Control and Prevention: https://www.cdc.gov/media/releases/2018/p0426-autism-prevalence.html

Cook, B. G., Johnson, A. H., Maggin, D. M., Therrien, W. J., Barton, E. E., Lloyd, J. W., Reichow, B., Talbott, E., & Travers, J. C. (2021). Open science and single-case design research. *Remedial and Special Education*. Advance online publication. https://doi.org/10.1177/0741932521996452

Crowe, B., Machalicek, W., Wei, Q., Drew, C., & Ganz, J. (2021). Augmentative and alternative communication for children with intellectual and developmental disability: A mega-review of the literature. *Journal of Developmental and Physical Disabilities*. Advance online publication. https://doi.org/10.1007/s10882-021-09790-0

Ganz, J. B., Morin, K. L., Foster, M. J., Vannest, K. J., Tosun, D. G., Gregori, E. V., & Gerow, S. L. (2017). High-technology augmentative and alternative communication for individuals with intellectual and developmental disabilities and complex communication needs: A meta-analysis. *Augmentative and Alternative Communication, 33*(4), 224–238. https://doi.org/10.1080/07434618.2017.1373855

Gevarter, C., & Zamora, C. (2018). Naturalistic speech-generating device interventions for children with complex communication needs: A systematic review of single-subject studies. *American Journal of Speech-Language Pathology, 27*(3), 1073–1090. https://doi.org/10.1044/2018_AJSLP-17-0128

Hahn, L. J., Loveall, S. J., Savoy, M. T., Neumann, A. M., & Ikuta, T. (2018). Joint attention in Down syndrome: A meta-analysis. *Research in Developmental Disabilities, 78*, 89–102. https://doi.org/10.1016/j.ridd.2018.03.013

Holyfield, C., Drager, K. D. R., Kremkow, J. M. D., & Light, J. (2017). Systematic review of AAC intervention research for adolescents and adults with autism spectrum disorder. *Augmentative and Alternative Communication, 33*(4), 201–212. https://doi.org/10.1080/07434618.2017.1370495

Hong, E. R., Ganz, J. B., Gilliland, W., & Ninci, J. (2014). Teaching caregivers to implement an augmentative and alternative communication intervention to an adult with ASD. *Research in Autism Spectrum Disorders, 8*(5), 570–580. https://doi.org/10.1016/j.rasd.2014.01.012

Johnson, A. H., & Cook, B. G. (2019). Preregistration in single-case design research. *Exceptional Children, 86*(1), 95–112. https://doi.org/10.1177/0014402919868529

Johnston, S. S., Reichle, J., Feeley, K. M., & Jones, E. A. (2012). *AAC strategies for individuals with moderate to severe disabilities*. Brookes Publishing Company.

Kittelman, A., Gion, C., Horner, R. H., Levin, J. R., & Kratochwill, T. R. (2018). Establishing journalistic standards for the publication of negative results. *Remedial and Special Education, 39*(3), 171–176. https://doi.org/10.1177/0741932517745491

Light, J. (1997). "Let's go star fishing": Reflections on the contexts of language learning for children who use aided AAC. *Augmentative*

and Alternative Communication, 13(3), 158–171. https://doi.org/10.1080/07434619712331277978
Logan, K., Iacono, T., & Trembath, D. (2017). A systematic review of research into aided AAC to increase social-communication functions in children with autism spectrum disorder. *Augmentative and Alternative Communication, 33*(1), 51–64. https://doi.org/10.1080/07434618.2016.1267795
Luyster, R. J., Kadlec, M. B., Carter, A., & Tager-Flusberg, H. (2008). Language assessment and development in toddlers with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 38*(8), 1426–1438. https://doi.org/10.1007/s10803-007-0510-1
Mitchell, M., Muftakhidinov, B., & Winchen, T. (2017). *Engauge Digitizer* (Version 12.1) [computer software]. http://markummitchell.github.io/engauge-digitizer
Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine, 6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097
Monzalve, M. A. (2016). *Examining the relation between contextual fit and implementation fidelity on behavior support plans and student outcomes* (Doctoral dissertation, University of Oregon).
Ogletree, B. T. (2012). Stakeholders as partners: Making AAC work better. *Perspectives on Augmentative and Alternative Communication, 21*(4), 151–158. https://doi.org/10.1044/aac21.4.151
Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., … & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *International Journal of Surgery, 88*, 105906.
Parker, R. I., Vannest, K. J., & Davis, J. L. (2011). Effect size in single-case research: A review of nine nonoverlap techniques. *Behavior Modification, 35*(4), 303–322. https://doi.org/10.1177/0145445511399147
Paul, R. (1997). Facilitating transitions in language development for children using AAC. *Augmentative and Alternative Communication, 13*(3), 141–148. https://doi.org/10.1080/07434619712331277958
Pustejovsky, J. E. (2018). Using response ratios for meta-analyzing single-case designs with behavioral outcomes. *Journal of School Psychology, 68*, 99–112. https://doi.org/10.1016/j.jsp.2018.02.003
Pustejovsky, J. E. (2019). Procedural sensitivities of effect sizes for single-case designs with directly observed behavioral outcome measures. *Psychological Methods, 24*(2), 217–235. https://doi.org/10.1037/met0000179
Pustejovsky, J. E. (2020). *clubSandwich: Cluster-robust (sandwich) variance estimators with small-sample corrections* (Version 0.4.2) [R package]. https://cran.r-project.org/package=clubSandwich
Pustejovsky, J. E., & Ferron, J. M. (2017). Research synthesis and meta-analysis of single-case designs. *Handbook of Special Education*, 168-186. Routledge.
Pustejovsky, J. E. & Swan, D. M. (2019). *SingleCaseES: A calculator for single-case effect sizes* (Version 0.4.3) [R package]. https://cran.r-project.org/package=SingleCaseES
Reichle, J., & Wacker, D. P. (2017). *Functional communication training for problem behavior*. Guilford Publications.
Reichle, J., Drager, K., & Davis, C. (2002). Using requests for assistance to obtain desired items and to gain release from nonpreferred activities: Implications for assessment and intervention. *Education and Treatment of Children, 25*(1), 47–66.
Reichle, J., Simacek, J., Wattanawongwan, S., & Ganz, J. (2019). Implementing aided augmentative communication systems with persons having complex communicative needs. *Behavior Modification, 43*(6), 841–878. https://doi.org/10.1177/0145445519858272
Reichle, J., O'Neill, R. E., & Johnston, S. S. (2021). Advances in AAC intervention: Some contributions related to applied behavior analysis. *Augmentative and Alternative Communication, 37*(3), 206–216. https://doi.org/10.1080/07434618.2021.1962405
Schreibman, L., Dawson, G., Stahmer, A. C., Landa, R., Rogers, S. J., McGee, G. G., Kasari, C., Ingersoll, B., Kaiser, A. P., Bruinsma, Y., McNerney, E., Wetherby, A., & Halladay, A. (2015). Naturalistic developmental behavioral interventions: Empirically validated treatments for autism spectrum disorder. *Journal of Autism and Developmental Disorders, 45*(8), 2411–2428. https://doi.org/10.1007/s10803-015-2407-8
Shadish, W. R., Zelinsky, N. A., Vevea, J. L., & Kratochwill, T. R. (2016). A survey of publication practices of single-case design researchers when treatments have small or large effects. *Journal of Applied Behavior Analysis, 49*(3), 656–673. https://doi.org/10.1002/jaba.308
Sham, E., & Smith, T. (2014). Publication bias in studies of an applied behavior-analytic intervention: An initial analysis. *Journal of Applied Behavior Analysis, 47*(3), 663–678. https://doi.org/10.1002/jaba.146
Shire, S. Y., & Jones, N. (2015). Communication partners supporting children with complex communication needs who use AAC: A systematic review. *Communication Disorders Quarterly, 37*(1), 3–15. https://doi.org/10.1177/1525740114558254
Sutton, B. M., Webster, A. A., & Westerveld, M. F. (2019). A systematic review of school-based interventions targeting social communication behaviors for students with autism. *Autism, 23*(2), 274–286. https://doi.org/10.1177/1362361317753564
Tager-Flusberg, H., & Kasari, C. (2013). Minimally verbal school-aged children with autism spectrum disorder: The neglected end of the spectrum. *Autism Research, 6*(6), 468–478. https://doi.org/10.1002/aur.1329
Therrien, M. C. S., Light, J., & Pope, L. (2016). Systematic review of the effects of interventions to promote peer interactions for children who use aided AAC. *Augmentative and Alternative Communication, 32*(2), 81–93. https://doi.org/10.3109/07434618.2016.1146331
Tincani, M., & Travers, J. (2018). Publishing single-case research design studies that do not demonstrate experimental control. *Remedial and Special Education, 39*(2), 118–128. https://doi.org/10.1177/0741932517697447
Van den Noortgate, W., & Onghena, P. (2008). A multilevel meta-analysis of single-subject experimental design studies. *Evidence-Based Communication Assessment and Intervention, 2*(3), 142–151. https://doi.org/10.1080/17489530802505362
Vannest, K. J., Peltier, C., & Haas, A. (2018). Results reporting in single case experiments and single case meta-analysis. *Research in Developmental Disabilities, 79*, 10–18. https://doi.org/10.1016/j.ridd.2018.04.029
Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(1), 1–48. https://doi.org/10.18637/jss.v036.i03
Warren, S. F., Fey, M. E., & Yoder, P. J. (2007). Differential treatment intensity research: A missing link to creating optimally effective communication interventions. *Mental Retardation and Developmental Disabilities Research Reviews, 13*(1), 70–77. https://doi.org/10.1002/mrdd.20139
Wetherby, A. M., & Prizant, B. M. (1993). Profiling communication and symbolic abilities in young children. *Journal of Childhood Communication Disorders, 15*(1), 23–32. https://doi.org/10.1177/152574019301500105