A Systematic Review of Frequency Building and Precision Teaching with School-Aged Children

Corinne Gist & Andrew J. Bulla

ONLINE FIRST

Journal of Behavioral Education

ISSN 1053-0819

J Behav Educ
DOI 10.1007/s10864-020-09404-3

Journal of Behavioral Education
https://doi.org/10.1007/s10864-020-09404-3

ORIGINAL PAPER

A Systematic Review of Frequency Building and Precision Teaching with School-Aged Children

Corinne Gist[1] · Andrew J. Bulla[2]

Abstract
This paper presents a systematic review of the literature that assessed the effectiveness of frequency building and precision teaching with school-aged children. The authors evaluated studies in accordance with the What Works Clearinghouse standards and the council for exceptional children standards for evidence-based practices in special education. A total of 11 studies examining the effectiveness of frequency building and precision teaching for 170 participants were included in this review. Additionally, effect sizes were calculated for eligible studies. Small to large effects were found for all included variables. Overall, results indicated that the combination of frequency building and precision teaching is an effective method for increasing a variety of academic skills with school-aged children. Limitations, recommendations for future research, and implications for practitioners are discussed.

Keywords Precision teaching · Frequency building · School-aged children · Academic skills

Introduction

Ogden Lindsley established precision teaching (PT) in 1964 as a response to his frustrations with the US education system (Lindsley 1964, 1990) and aspired to create classrooms that used scientific methods and allowed teachers and learners to make data-based discoveries (Lindsley 1990). With this goal in mind, Lindsley developed a precise system that he believed would help teachers become more precise in their teaching decisions—thus the name, precision teaching (Johnson and

✉ Corinne Gist
gist.22@osu.edu

Andrew J. Bulla
abulla@georgiasouthern.edu

1 The Ohio State University, Columbus, USA

2 Georgia Southern University – Armstrong, Savannah, USA

Published online: 12 July 2020

Street 2013). This program has been successfully applied in several educational settings over the past 50 years (Binder 1996).

Contrary to the name, PT is not one specific teaching strategy, but a measurement system that involves four basic steps: Pinpoint, Record, Change, and Try Again (Kubina and Yurich 2012). Step one, *Pinpoint*, refers to precisely defining the target behavior in observable and measurable terms. Pinpoints are created to help teachers accurately count specific behaviors in an objective way. Pinpoints typically specify the mode in which the learner contacts the material, and the behavior made in response to the specific stimulus (e.g., Hear word/Say definition); precision teachers refer to these as learning channels. In addition to specifying the learning channel, pinpoints often state the performance standard for the skill being targeted (e.g., 60–50 per minute). For example, instead of setting a goal of “completing math facts,” a pinpoint would state, “See Math Fact/Write answer, 60–50 answers per minute.” Specific pinpoints facilitate seamless data collection.

The second PT step, *Record*, refers to daily measurement and graphing of the learner’s behavior on the standard celeration chart (SCC). The SCC is a standardized six-cycle ratio (i.e., multiple/divide) graph for charting behavior frequency. A ratio graph differs from a linear (aka equal interval) graph in that it can show relative change in behavior. Relative change refers to how the behavior frequency multiples or divides over time. For instance, last Monday a learner could say two words, and the following Monday could say four words: a doubling or $\times 2$ (said “times two”) growth in performance.

Measuring relative change allows for the learner’s learning history to be considered when assessing progress (Johnson and Street 2013). It is more significant, for example, when children advance from speaking one to three words (i.e., 200 percent change) than when they progress from speaking 101 to 103 words (i.e., 1.98 percent change); however, the comparative growth would look the same on an equal-interval graph. It is easier to make progress once basic skills are mastered and using a graph that demonstrates relative change (e.g., SCC) makes decision-making easier and more precise for teachers (Lefebre et al. 2008; Johnson and Street 2013). Additionally, the SCC allows teachers to evaluate and revise instruction after every learning opportunity. In most cases, the daily measurement is the learner’s rate of responding.

The third step, *Change*, refers to the inductive process of analyzing the data and making decisions. There are many decisions a precision teacher can make. For example, a *continue decision* means the learner’s performance has met expectation for sufficient growth and the intervention should continue as is (Kubina 2019). *Complete decisions* are rendered when the data have reached the terminal criteria, oftentimes a “fluency aim” in PT parlance (discussed later). A third decision once can make is to *make a change* to some aspect of the instruction. A precision teacher would make this decision if the learner has not made adequate progress and requires an intervention. Teachers are able to make continue/change/complete decisions after every data point charted on the SCC. The frequent data-based decisions enable teachers to evaluate progress toward individualized goals.

Lastly, *Try Again* refers to analyzing the environmental changes and systematically intervening until the desired outcome is achieved. At the core of PT is standardized measurement, and by consistently measuring behavior, precision teachers are able to isolate variables that may place a ceiling on a learner's performance. By addressing each variable and lifting the ceiling on performance, precision teachers accelerate growth and positive learning outcomes. For a description of common performance ceilings, see Binder (2010).

Many discoveries and educational strategies have been derived from the PT literature. Frequency building is one commonly used technique in PT. Frequency building is timed repetition of pinpointed behavior followed by performance feedback (Kubina and Yurich 2012). Frequency building consists of having a learner practice a particular behavior as many times as they are able in a specific interval of time, and reinforcement is contingent on growing frequencies. For example, a learner may be asked to write digits 0–9 as many times as they can during a 30-s timed interval. Reinforcement is delivered for increases in overall frequencies across time. Frequency building continues daily until the learner reaches a specified goal. The pinpoint goal or aim appears in the format of a frequency range (e.g., 80–70 per minute). The ranges are considered mastery levels of responding and suggest fluency of performance; thus, the term is fluency aim.

Fluency, or more precisely *behavioral fluency*, is defined as accurate and well-paced behavior (Binder 1996; Datchuk and Kubina 2017; Johnson and Street 2004). "Fluency is the PT concept for performance that is flowing, flexible, effortless, errorless, automatic, confident, second nature, and masterful" (Johnston and Street 2004 p. 23). Developing fluent behavior is the goal of precision teachers at every level of the curriculum.

Practitioners have been successful applying frequency building and PT to produce significant results for over 50 years (Johnson and Street 2013). One of the earliest and largest applications of PT was The Precision Teaching Project of Great Falls, MT. The project was conducted in general education classrooms and consisted of 30 min of fluency practice (i.e., frequency building) and daily charting, which resulted in increases in elementary school learner standard assessment scores of 20–40 percentage points over a 3-year period (Beck and Clement 1991). Today, a number of private schools and learning centers employing PT continue to report impressive gains in learner learning. One example is Morningside Academy in Seattle, WA. Morningside offers a money-back guarantee of two years growth in one school year in the skill of greatest deficit (e.g., math, reading). Over 23 years, Morningside returned less than one percent of learner tuition (Johnson and Street 2004). Another example includes Fit Learning™ centers. Children receiving instruction in reading fluency and reading comprehension at Fit Learning™ centers regularly demonstrate growth that is eight times better than the national average in just 40 h of instruction (Fit Learning 2018).

Despite the positive outcomes produced when using frequency building and PT, the number of teachers and public schools using the method is minimal. This may be related to the fact that the body of published research on PT is relatively small, especially when compared to other educational practices, limiting the number of educators exposed to the practice. Binder (1996) reports several reasons for the lack of publications in the field of PT. First, practitioners, not researchers, developed PT and often practitioners have less interest, incentives, and resources to produce empirical research. Additionally, Lindsley discouraged early precision teachers from publishing their findings because he did not believe that publications changed professional behavior (Binder 1996).

Although Binder's (1996) article was written 20 years ago, his recounting of PT's historical influences against publication may still hold true. Instead of publishing in peer-reviewed journals, precision teachers often shared their data through chart shares, conferences presentations, and newsletters. *The Journal of Precision Teaching* was created in 1980 and published chart shares, application articles, and empirical studies. After 30 years, the journal was discontinued because precision teachers began to shift their focus to other forms of communication, including online special interest group forums and webinars (Standard Celeration Society 2018).

The dearth of research studies published by precision teachers in peer-reviewed journals has brought criticism and skepticism from practitioners and behavior analysts (Cihon 2007; Doughty et al. 2004). Cihon (2007) suggested, "the lack of data shared in peer reviewed journals may have been one of the most critical mistakes made by precision teachers" (p. 128). When precision teachers do publish, their research has been criticized for demonstrating weak methodology and experimental control (Cooper 2005; Doughty et al. 2004).

In a review on the effects of rate building on fluent behavior, Doughty et al. (2004) stated the majority of published PT literature consists of anecdotal evidence from practitioner programs or conceptual papers rather than a critique of empirical studies. Doughty et al. (2004) reviewed 48 peer-reviewed research studies that reported on the use of PT or similar "rate building" techniques. The authors stated that the empirical research supports claims that rate building (i.e., frequency building) and PT enhances learning and demonstrates positive educational outcomes. Doughty et al. (2004) suggested, however, that there was little empirical support for the claim that fluency is the result of PT or any of its component procedures. Two experts in the field of PT published responses to Doughty et al.'s (2004) review (i.e., Kubina 2005; Binder 2004), presenting the case that several of the terms and definitions used by the authors were incorrect, which may have led to inaccurate conclusions. Although Kubina (2005) and Binder (2004) disagreed with pieces of Doughty et al.'s (2004) analysis, all three authors agreed that additional methodologically sound research is needed. The current empirical state of the degree to which PT is more effective than other methods of measurement and instruction is unknown. Thus, systematic reviews of the literature are warranted to allow educators and behavior analysts objective information when looking to adopt evidence-based practices.

In 2016, Ramey et al. conducted a systematic literature review of 55 studies that assessed the effects of PT for individuals with developmental disabilities. The studies were evaluated for methodological quality and treatment effects. Results of the review indicated PT was an *emerging* treatment for persons with developmental disabilities based on the National Autism Council's (NAC) Strength of Evidence Classification System (NAC 2009); however, studies of all methodological levels were included in the review, including chart shares and case studies. Eighty percent of the studies included in the review received ratings of 1 or 0 (out of 5) on the Scientific Merit Rating Scale (SMRS), indicating low methodological rigor. Additionally, 72% of the studies did not provide enough evidence to determine whether the treatment effect was beneficial, ineffective, or adverse according to the NAC standards (NAC 2009). The authors of the review stated that the methodological rigor of the studies likely impacted the strength of evidence conclusions across all skill categories. It appears that the *emerging* rating may have been impacted more by the low methodological rigor of the studies, then by a lack of treatment effects.

Brosnan et al. (2018) published a multilevel analysis demonstrating the effects of ten methodologically sound studies she completed as part of her dissertation research. All ten studies focused on the effects of frequency building and PT on reading fluency in school-aged children. The authors used the quality standards developed by What Works Clearing House (WWC; Kratochwill et al. 2010), to determine the methodological quality of each of the study; only those studies that met standards were included in the analysis. Results of the analysis indicated that PT could be considered a promising intervention for increasing reading fluency with individuals at risk of reading failure (Brosnan et al. (2018)). Despite the rigorous methodologies employed by Brosnan et al. (2018), individual graphs for each study were not published, making it difficult for an independent analysis of the results. Additionally, all ten studies were conducted by the primary author, thus limiting the generality of the findings because these findings were not replicated independently.

Reviewing the current literature on PT offers unique challenges, because PT, as described above, does not specify *how* or *what* to teach. It is a measurement system that allows for more precise decisions. Two approaches can be taken to evaluate the effectiveness of PT as a system: (1) review all studies that used PT as their measurement and decision making system or (2) evaluate individual interventions commonly used with PT. Ramey et al. (2016) used the first approach to evaluate the effects of PT on outcomes associated with learners with disabilities. Within their review, multiple interventions were mentioned being used (e.g., frequency building, repeated reading, differential reinforcement, backward chaining, etc.). Thus, reviews of this nature would need to not only quantify the effects of each intervention, but also evaluate the degree to which decisions being made were done so in a more efficient way than other measurement systems. The second approach allows for a more thorough analysis of specific interventions commonly used by precision teachers. For example, Quigley et al. (2017) reviewed the available research on the say all fast a minute every day shuffled (SAFMEDS) intervention, a form of flashcards that uses 1-min timings and measurement with PT to improve memorization of facts. Since frequency building is a common intervention used within PT, the current authors opted to take the second approach to evaluate the effects of frequency building with PT. Thus, any study using frequency building without measuring and making decisions with PT was excluded from the review.

The primary purpose of this review was to objectively synthesize the available research on frequency building with the use of precision teaching. Several steps were taken to ensure that the findings accurately reported the current state of empirical evidence for the interventions. First, only studies deemed to be methodologically sound were included in the current review. Second, studies that employed both single-case design and group design were examined. Third, studies including children with and without disabilities were included; thus, any school-aged application of frequency building and PT was included. Lastly, the authors also used WWC and standards set forth by the Council for Exceptional Children (CEC) to determine whether frequency building and PT can be considered an evidence-based practice for school-aged children.

Method

Study Eligibility Criteria

To be included in this review, studies had to meet the following criteria. First, studies had to be available in English and published in a peer-reviewed journal. Second, studies had to use Precision Teaching (i.e., measurement decisions based on the SCC) to increase or decrease performance of an observable skill or behavior. If it was unclear whether the authors used a SCC, the authors were e-mailed for clarification (i.e., Roberts and Norwich 2010). Third, the studies had to include school-aged participants (ages 3–22 years). The review was limited to school-aged children, as the purpose was to evaluate frequency building and precision teaching in an educational context. Fourth, the experimental design had to be a single-case design (SCD) or a group design (i.e., random control trial, quasi-experiment). If the study used a SCD, the paper must have included charts or graphs for visual analysis. If the authors did not provide charts or graphs (e.g., Brosnan et al. 2018) or results were demonstrated using bar graphs or tables only (e.g., Cihon et al. 2017), the study was excluded from the review. Lastly, only studies that *met* or *met with reservation* the WWC quality standards (Kratochwill et al. 2010) were included in this review.

WWC has three categories for identifying the design rigor of a study: (1) *Meets Standards,* (2) *Meets Standards with Reservations,* and (3) *Does Not Meet Standards.* The criteria assessed for determining the assigned category for SCD studies were scored using a dichotomous scale (i.e., present or not present) and included the following: (a) the independent variable was systematically manipulated with the researcher determining when and how the independent variable conditions changed, (b) each outcome variable was measured repeatedly over time by more than once assessor, (c) a measure of interobserver agreement was reported for each eligible outcome variable for no less than 20% of sessions, (d) the reported level of interobserver agreement was at least 80% for percentage agreement indices, and (e) the study included at least three attempts to demonstrate an intervention effect at three different points in time or with three different phase repetitions.

If a study did not meet any one of the previously listed criteria, it was rated as *Does Not Meet Standards* and was not included in the review. Remaining studies were furthered assessed to determine if they had a sufficient amount of data points to meet WWC standards. The following classifications were used to categorize the studies: (a) *Meets Standards* if the study included five or more data points per phases, (b) *Meets Standards with Reservations* if the study included three or four data points per phase, and (c) *Does Not Meet Standards* if there were fewer than three data points per phase. As recommended by the WWC, studies that *Meet Standards* or *Meet Standards with Reservations* were included in this review.

The group design studies were also assessed against the WWC evidence standards, as described in the WWC Procedures and Standards Handbook Section III: Screening and Reviewing Studies (2017, p. 8–21). The following criteria were assessed: (a) The attrition rate was below 20% for all participants, and the different attrition was below 5 percentage points; (b) for studies using a quasi-experimental design, equivalence was demonstrated for the intervention and comparison groups on a baseline measure of the outcome or on another measure that is correlated to the outcome measure; if baseline differences exceeded 0.25 standard deviations for any of the measures with a domain, the study did not meet group design standards within that domain; (c) the study examined at least one eligible outcome measure that meets review requirements (i.e., demonstrates face validity, demonstrates reliability, is not over aligned with the intervention, or is collected in the same manner for both intervention and comparison groups); if the researchers did not examine these reliability standards, the reviewer determined the outcome measure was reliable if the responses could be scored by a single coder with low error (e.g., a multiple-choice test, counts of words spelled correctly); and (d) the study was free from confounding factors. If a study did not meet any one of the criteria listed above, it was categorized as *Not Meeting Design Standards* and was excluded from this review. The remaining studies were classified as *Meets Standards* if a random control trial design was used and *Meets Standards with Reservations* if a quasi-experimental design was used, as stated by WWC group standards.

Search Strategy

In May of 2018, the first author completed a comprehensive search to ensure all studies that met the inclusion criteria were included in this review. First, an electronic search was conducted and completed using four databases: ERIC, PsychInfo, Education Complete, and Academic Complete. The key phrases "precision teach*" and "frequency build* OR "rate build*" were searched separately. Limiters used included published in English and published in a peer-reviewed journal. After duplicates were removed from the individual searches, the search using "Precision Teach*" resulted in 282 studies and the search using "frequency build*," or "rate build*" resulted in 27 studies. When the searches were combined and duplicates were removed, the searches produced a total of 292 studies.

Next, abstracts and titles were screened and those that did not meet inclusion criteria were discarded, resulting in 87 remaining articles. A full-text review was conducted on the remaining articles, and seven qualifying studies were identified. Then, a backward search (i.e., the reference list of each identified study was reviewed) and a forward search (i.e., all articles that cited the identified studies in Google Scholar were reviewed) were conducted. One study was added from the backward search, and three additional studies were added from the forward search. Upon additional examination, the author discovered that no articles from the *European Journal of Behavior Analysis* appeared in the database search; however, two were identified through the forward and backward search. In order to confirm that all eligible studies were included in the review, an additional search within the *European Journal* *of Behavior Analysis* was conducted using the same terms used for the database search. No additional studies meeting the eligibility criteria were identified. Altogether, 11 studies were identified for this review. A total of nine new studies that were not included in previous reviews were identified from this search (i.e., Seevers et al. 2003 and Twarek et al. 2010 were included in Ramey et al. 2016). See Fig. 1 for a detailed description of the search, inclusion, and exclusion processes.

Study Coding

The studies were reviewed to assess methodological rigor, determine evidential strength, and report essential characteristics. Coding procedures varied depending on the design type (i.e., SCD or group design). The SCD studies were categorized

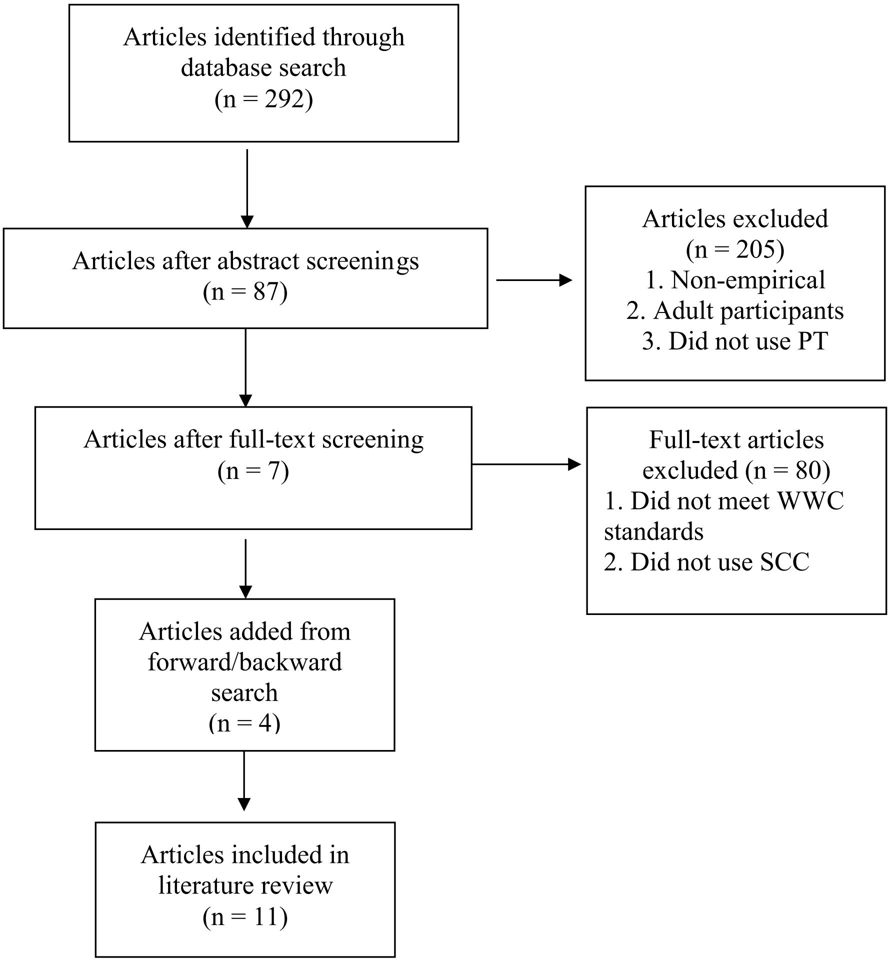

Fig. 1 Flow diagram of the literature search process

based on the designs outlined by Ledford and Gast (2018), and the group designs were categorized based on the WWC standards manual.

CEC Quality Indicators

In addition to the WWC standards, the studies were also measured against the Council for Exceptional Children (CEC) quality indicators for evidence-based practices. This was done for two reasons: (1) to further evaluate the studies' methodology and (2) to determine whether frequency building and PT qualify as an evidence-based practice. WWC does not provide criteria for establishing if a practice is evidence based when both SCD and group designs are included. Therefore, the CEC quality indicators and criteria were used. The quality indicators include criteria in eight domains: (a) context and setting, (b) participants, (c) intervention agent, (d) description of practice, (e) implementation fidelity, (f) internal validity, (g) outcome measures/dependent variables, and (h) data analysis. Within the eight domains, they are 22 indicators for SCD studies and 24 indicators for experimental group design studies. Each indicator was coded on a dichotomous scale (i.e., "yes" or "no").

Study Characteristics

In addition to the WWC and CEC standards, each study was reviewed and coded along several dimensions including data regarding (a) participants (i.e., number in study, gender, age or grade, disability type), (b) independent variable (i.e., group size, frequency of intervention, implementer, fidelity), (c) dependent variable, (d) maintenance data, and (e) publication data (i.e., publishing journal, date of publication).

Study Effects

SCD Effects

As recommended by WWC, studies that *Meet Standards* or *Meet Standards with Reservations* were evaluated using visual analysis to determine whether a functional relation was demonstrated. Visual analysis is the most frequently used data analysis method for SCD research and is the preferred method for many researchers (Ledford and Gast 2018). For this review, visual analysis was conducted based on the visual analysis recommendations by WWC and the summary of visual analysis described in Ledford and Gast (2018). Six features were evaluated to examine within-phase and between-phase data patterns: level, trend, variability, immediacy of the effect, overlap, and consistency of data patterns across similar phases (Kratochwill et al. 2010). Studies were rated as having no evidence of a causal relation, moderate evidence of a causal relation, or strong evidence of a causal relation (Kratochwill et al. 2010).

Although it is commonly agreed upon that visual analysis should be the primary analysis for determining an effect in SCD (Ledford and Gast 2018), several researchers have objected to using visual analysis as the sole method for determining experimental effects (DeProspero and Cohen 1979; Kazdin 2011). The WWC guidelines for SCD strongly recommend the use of one or more nonparametric effect size estimate for studies which demonstrate a moderate or strong evidence of a functional relationship (Kratochwill et al. 2010).

One effect size estimate used for SCD research is Tau-*U*. Tau-*U* was developed by Parker, Vannest et al. (2011) and can be explained as the percentage of nonoverlap between phases or by percent of data showing improvement between phases. Tau-*U* was chosen for this review due to its specific strengths as an effect size estimate for SCD.

Due to the advantages discussed, Tau-*U* scores were calculated in addition to visual analysis for all eligible SCD studies. Tau-*U* scores were calculated for each dependent variable measure across five studies involving 22 participants, resulting in eight calculations. Tau-*U* was not calculated for one SCD study (Seevers et al. 2003) because necessary information regarding y- and x-axis data was not provided. This study was included in the review because the graphs allowed for visual analysis. Tau-*U* was calculated separately for each participant's baseline–intervention contrast. The effect size across all participants was then calculated by averaging the individual Tau-*U* scores (Rakap 2015). Data were extracted from the SCC charts using the WebPlotDigitizer software (Rohatri 2018; Drevon et al. 2017). Effect size estimates were calculated using the online Tau-*U* calculator (Vannest et al. 2011). Effects are measured as *small* (0–0.65), *medium* (0.66–0.92), and *large* (0.93–1.00). The larger the effect size, the greater the difference between baseline and intervention.

Tau-*U* scores were only calculated for intervention phases. Maintenance and generalization data were not included in the effect size calculation. For studies that had multiple dependent variables, Tau-*U* was calculated for each dependent variable separately.

Group Effect Size

In order for a group effect size to be calculated, means, standard deviations, and number of participants in each group in the study must have been reported (or additional information that could lead to the calculation of a standard mean difference). Effects for group designs were calculated by entering the above information into an online calculator (Ellis 2009). The online calculator computed the standard mean difference, *d*, and converted those results into Hedges' *g* statistic. Hedges' *g* was determined to be the most appropriate effect measure because it offers small-sample bias correction (Borenstein et al. 2009). All of the included studies had sample sizes smaller than 80. Hedges' *g* was calculated separately for each dependent variable; however, only dependent variables that were collected by direct observation were included. For example, two studies (Roberts and Norwich 2010; Greene et al. 2018) included measures of self-esteem or self-concept as dependent variables, and these variables were not included in the analysis. Hedges' *g* effects are measured as *small* (0.2 or smaller), *medium* (0.5-0.7), and *large* (0.8 or larger). Effect sizes above

0.25 are considered large enough to be educationally significant (Slavin and Fashola 1998).

Inter-coder Agreement

For 55% ($n=6$) of the studies, a second rater independently coded all variables. The second rater was a doctoral-level graduate learner with training and experience in conducting systematic literature reviews. The rater was trained on the coding procedures through oral instructions and feedback on practice studies until 90% agreement was reached. Point-by-point agreement (i.e., exact agreements divided by opportunities for agreement and multiplied by 100) was calculated for each variable (Cooper et al. 2007). Overall agreement came to 94%. Within the coding categories, agreement was 97% for study characteristics, 96% for CEC standards, 80% for visual analysis, and 100% for effect size calculation.

Results

Eleven studies published in seven different peer-reviewed journals were included in this review. Sixty-six percent ($n=7$) of the studies included in this review were published within the last five years, suggesting that methodologically strong studies in the area of PT have increased in the last few years. A complete list of authors, journals, publication dates, and the studies settings is provided in Table 1.

Experimental Design

Six of the 11 studies included in this review used a SCD and five incorporated a group design. All six of the SCD studies used a variation of multiple baseline across participants design. The group studies included random control treatment ($n=2$), quasi-experimental ($n=2$), and mixed factorial experimental designs ($n=1$). See Table 1 for a complete overview of these findings.

Participants and Settings

A total of 235 learners participated with 170 learners receiving the intervention. The following data are only a reflection of learners that received the intervention. Two group design studies did not report gender, specifically for the intervention group (Mc Tiernan et al. 2018; Mannion and Griffin 2018). For the remaining studies, 59% ($n=86$) of the participants were male and 41% ($n=60$) were female. The age of participants ranged from 6 to 19 years. One study (Stromgren et al. 2014) did not reference the participants' age but did report grade level (fifth and sixth grades). Eighty-five percent ($n=144$) of the participants were learners without disabilities. Eighty-three percent ($n=5$) of the SCD studies included participants with disabilities, and one (20%) group study (Roberts and Norwich 2010) included participants with disabilities; however, the authors did not indicate the number of participants with disabilities and therefore it is not included

Table 1 Key characteristics of studies using precision teaching with school-aged children

Study	Journal	Research design	Number of participants (intervention)	Participant age range	Participant disabilities	Setting description	Intervention implementer
Chapman et al. (2005)	Behavioral Interventions	MBD	5	11–19	TBI (5)	A school within residential neuro-rehabilitation program	N/R
Datchuk and Kubina (2017)	Education and Treatment of Children	MBD	4	13–18	ID (1), SLD (2), none (1)	Urban charter School	Researcher
Datchuk et al. (2015)	Exceptionality	MBD	4	9–11	SLD and EBD (2), SLD (1), ASD (1)	Two elementary schools	Special education teachers
Greene et al. (2018)	Journal of Behavioral Education	RCT	12	8–12	N/R	All girls urban elementary school	Researcher and learner tutors
Lambe et al. (2015)	Behavioral Interventions	MBD	7	7–8	None	Primary school	Researcher
Mannion and Griffin (2018)	Irish Educational Studies	Mixed factorial experimental design	18	10–12	None	Primary school	Classroom teachers with support from researcher
Mc Tiernan et al. (2018)	European Journal of Behavior Analysis	RCT	14	9–11	None	Public elementary school	Master's level learners
Roberts and Norwich (2010)	Educational Psychology in Practice	Quasi-experimental design	77	11–16	At risk (N/R); EBD (N/R); LD (N/R)	Three elementary schools and two specialty schools	Trained teaching assistants
Seever et al. 2003	Journal of Precision Teaching and Celeration	MBD	7	11	SLD (7)	Urban elementary school	Researcher

Table 1 (continued)

Study	Journal	Research design	Number of participants (intervention)	Participant age range	Participant disabilities	Setting description	Intervention implementer
Stromgren et al. 2014	European Journal of Behavior Analysis	Quasi-experimental design	19	N/R (5^{th}–6th grade)	At risk (48)	Primary school	Researcher
Twarek et al. (2010)	Behavior Interventions	Noncurrent MBD	3	3–5	ASD (3)	Participant's home	Researcher

ID intellectual disability, *ASD* autism spectrum disorder, *TBI* traumatic brain injury, *EBD* emotional and behavioral disorder, *RCT* random control trial in the following data. Within the group of participants with disabilities, 12 (46%) were diagnosed with a specific learning disability (SLD), seven (27%) were diagnosed with a traumatic brain injury, four had diagnoses of autism (15%), two (8%) were diagnosed with an emotional and behavior disability, and one (4%) was diagnosed with an intellectual disability.

Ten of the 11 studies (91%) were conducted in a school setting, and one study (Twarek et al. 2010) was conducted in the participants' home. Interestingly, all of the group studies ($n=5$) were conducted outside of the USA: three (60%) in Ireland, one (20%) in Norway, and one (20%) in the South of England. One (17%) of the SCD study was conducted in Ireland, with the remainder (83%) taking place in the USA.

Characteristics of Interventions

This review included studies that targeted several different academic skill areas. Six (55%) studies focused on reading skills, four (36%) math skills, four (36%) writing skills, one (9%) fine motor skills, and one (9%) functional/communication skills. One study (Chapman et al. 2005) included four different academic areas, and therefore is included multiple times in the data above.

The interventions for all 11 studies included frequency building and decision-making based on data graphed on the SCC. Frequency building is defined as timed repetition of pinpointed behavior followed by performance feedback. Two studies (18%) included the use of the Morningside Math Facts Curriculum (Green et al. 2018; Mc Tiernan et al. 2018), which is a frequency-building program. In addition, the following interventions were also utilized: simple sentence instruction (Datchuk and Kubina 2017; Datchuk et al. 2015), instruction in paragraph writing (Datchuk and Kubina 2017), self-management (Seevers et al. 2003), and peer tutoring (Greene et al. 2018).

All six of the SCD studies were performed in a one-on-one setting, whereas the group studies interventions ranged from groups of 1 to 10 learners. One study (Roberts and Norwich 2010) did not report intervention group size. The majority of the SCD interventions were implemented by the researcher ($n=4$; 67%). One study (17%) had special education teachers implement the intervention (Datchuk and Kubina 2017), and one (17%) study did not report who implemented the intervention (Chapman et al. 2005). In contrast, the researcher implemented the intervention in only one (20%) group design study (Stromgren et al. 2014). For the remaining group design studies, the intervention was implemented by master's level learners (Mc Tiernan et al. 2018), trained teaching assistants (Robert and Norwich 2010), classroom teachers (Mannion and Griffin 2018), and peer tutors (Greene et al. 2018). The average intervention length was five weeks (*range*, 3–14 weeks), with sessions occurring three to five times a week.

Maintenance Data

Seven (66%) of the studies programmed for and reported data on maintenance measures with four (36%) of the studies collecting the last data point 3–4 weeks after the intervention, and two (29%) studies collecting the last data point after 1–2 weeks. One study (14%) did not report when the maintenance data were collected (Seever et al. 2003). SCD studies were more likely to report maintenance data ($n=5$, 83%) than group design studies ($n=2$; 40%).

WWC Standards and Visual Analysis

Four of the six (67%) SCD studies met the WWC design standards without reservations (Table 2). The remaining two (33%) studies had phases with only three data points and therefore met the standards with reservations. All six of the SCD studies were evaluated using the visual analysis standards outline by Kratchowill et al. (2010). Nine outcome variables were analyzed, and all demonstrated either strong or moderate evidence of a causal relationship. Two (40%) of the group design studies met the WWC design standards without reservations and three (60%) met with reservations (Table 3). Only studies that employed a random control treatment design were eligible to meet standards without reservations.

Effect Size

Tau-*U* was calculated for eight dependent variables across five of the six (83%) SCD studies. One participant from Chapman et al. (2005) was not included in the effect size measure because the dimension of behavior measured in baseline was different than the dimension measured in the intervention phase and the participant was not included as a participant within the multiple baseline design. Overall, two (25%) large effects, four (50%) medium effects, and two (25%) small effects were demonstrated. Table 2 presents the confidence intervals for all Tau-U calculations. Hedges' *g* was calculated for eight dependent variables across four of the five (80%) group designs. Three (38%) large effects, four (50%) medium effects, and one (13%) small effect was demonstrated. Large or moderate effect sizes were demonstrated for all academic skill areas and for learners with and without disabilities. The statistical software used to calculate Hedges' *g* did not yield the confidence interval values; thus, they are not included in Table 3.

CEC Standards

See Table 4 for a detailed list of CEC standards met for each study. Five (46%) of the 11 studies met all of the CEC quality indicators—four (67%) SCD studies and one (20%) group design study. The remaining studies met the majority of the indicators (range 77–96%; $M=85\%$). In all cases, the studies that did not meet all the indicators failed to report on implementation fidelity. In addition, two studies failed to provide sufficient information regarding who implemented the intervention (Chapman et al. 2005; Mannion and Griffin 2018). See Table 4 for detailed information on which indictors were met for each study.

Overall, in accordance with CEC standards (2014), findings from this review demonstrate that frequency building and PT can be considered an evidence-based practice for school-aged children. The conclusion is based on meeting the criteria of having one methodologically sound group comparison study with random

Table 2 Intervention outcomes for studies using SCD

Study	Dependent variable (s)	WWC design standards outcomes	WWC visual analysis outcomes	CEC standards outcomes (total 22)	Tau-*U* 90% CI	Maintenance/retention data
Chapman et al. (2005)	Number of basic auto-biographical questions answered per minute; number of math SAFMEDS completed correctly per minute; number of functional ASL signs completed per 30 s	Meets With Reservations	Strong	17	.79 (.47–1)	Yes
Datchuk and Kubina (2017)	Word sequences written per minute; simple sentences written per minute	Meets	Moderate; moderate	22	.63 (.42–.85); .64 (.42-.85)	Yes
Datchuk et al. (2015)	Word sequences written per minute; simple sentences written per minute	Meets	Moderate; strong	22	.81 (.57–1); .90 (.66–1)	Yes
Lambe et al. (2015)	Number of words read correctly per minute on SAFMEDS; number of words read correctly per minute on Dolch stories	Meets	Strong; strong	19	.81 (.59–1); 1.06 (.78–1)	No
Seevers et al. 2003	Detection and correction of mechanical errors in writing samples per minute	Meets	Strong	22	N/A	Yes

Table 2 (continued)

Study	Dependent variable (s)	WWC design standards outcomes	WWC visual analysis outcomes	CEC standards outcomes (total 22)	Tau-*U* 90% CI	Maintenance/ retention data
Twarek et al. (2010)	Percentage of big 6 + 6 composite skills completed	Meets with reservations	Strong	22	1 (.64–1)	Yes

All raw data are available upon request from the first author

Table 3 Intervention outcomes for studies using group designs

Study	Dependent variable(s)	WWC design standards outcomes	CEC standards outcomes (total 24)	Hedges (g)	WWC effect size classification	Maintenance data
Greene et al. (2018)	Number of math flashcards completed correctly per minute; standardized math fluency test; standardized math calculations test	Meets	24	1.49; 0.78; 0.18	Positive; positive; neutral	No
Mannion and Griffin (2018)	Number of sight words read correctly per minute	Meets with reservations	20	1.33	Positive	No
Mc Tiernan et al. (2018)	Number of math problems correctly completed per minute; standardized math fluency test	Meets	23	0.8234; .4422	Positive; positive	Yes
Roberts and Norwich (2010)	Words read correctly	Meets with reservations	23	N/A	N/A	No
Stromgren et al. 2014	Number of correct math problems completed per minute	Meets with reservations	20	.5097; .4162	Positive; positive	Yes

All raw data are available upon request from the first author

Table 4 CEC quality standards met for each study

CEC Quality standard/ study	Chapman et al. (2005)	Datchuk et al. (2015)	Datchuk and Kubina 2017	Greene et al. (2018)	Lambe et al. (2015)	Mannion and Griffin 2018	Mc Tiernan et al. (2018)	Roberts and Norwich 2010	Seevers et al. 2003	Strømgren et al. (2014)	Twarek et al. (2010)
1.1 B*	Met	Met	Met	Met	Met	Met	Met	Met	Met	Met	Met
2.1 B	Met	Met	Met	Met	Met	Met	Met	Met	Met	Met	Met
2.2 B	Met	Met	Met	Met	Met	Met	Met	Met	Met	Met	Met
3.1 B	Not met	Met	Met	Met	Met	Met	Met	Met	Met	Met	Met
3.2 B	Not met	Met	Met	Met	Met	Not met	Met	Met	Met	Met	Met
4.1 B	Met	Met	Met	Met	Met	Met	Met	Met	Met	Met	Met
4.2 B	Met	Met	Met	Met	Met	Met	Met	Met	Met	Met	Met
5.1 B	Not met	Met	Met	Met	Not met	Not met	Not met	Met	Met	Not met	Met
5.2 B	Not met	Met	Met	Met	Not met	Not met	Not met	Not met	Met	Not met	Met
5.3 B	Not met	Met	Met	Met	Not met	Not met	Not met	Met	Met	Not met	Met
6.1 B	Met	Met	Met	Met	Met	Met	Met	Met	Met	Met	Met
6.2 B	Met	Met	Met	Met	Met	Met	Met	Met	Met	Met	Met
6.3 B	Met	Met	Met	Met	Met	Met	Met	Met	Met	Met	Met
6.4 G**	N/A	N/A	N/A	Met	N/A	Met	Met	Met	N/A	Met	N/A
6.5 S***	Met	Met	Met	N/A	Met	N/A	N/A	N/A	Met	N/A	Met
6.6 S	Met	Met	Met	N/A	Met	N/A	N/A	N/A	Met	N/A	Met
6.7 S	Met	Met	Met	N/A	Met	N/A	N/A	N/A	Met	N/A	Met
6.8 G	N/A	N/A	N/A	Met	N/A	Met	Met	Met	N/A	Met	N/A
6.9 G	N/A	N/A	N/A	Met	N/A	Met	Met	Met	N/A	Met	N/A
7.1 B	Met	Met	Met	Met	Met	Met	Met	Met	Met	Met	Met
7.2 B	Met	Met	Met	Met	Met	Met	Met	Met	Met	Met	Met
7.3 B	Met	Met	Met	Met	Met	Met	Met	Met	Met	Met	Met
7.4 B	Met	Met	Met	Met	Met	Met	Met	Met	Met	Met	Met

Table 4 (continued)

CEC Quality standard/study	Chapman et al. (2005)	Datchuk et al. (2015)	Datchuk and Kubina 2017	Greene et al. (2018)	Lambe et al. (2015)	Mannion and Griffin 2018	Mc Tiernan et al. (2018)	Roberts and Norwich 2010	Seevers et al. 2003	Strømgren et al. (2014)	Twarek et al. (2010)
7.5 B	Met	Met	Met	Met	Met	Met	Met	Met	Met	Met	Met
7.6 G	N/A	N/A	N/A	Met	N/A	Met	Met	Met	N/A	Met	N/A
8.1 G	N/A	N/A	N/A	Met	N/A	Met	Met	Met	N/A	Met	N/A
8.2 S	Met	Met	Met	N/A	Met	N/A	N/A	N/A	Met	N/A	Met
8.3 G	N/A	N/A	N/A	Met	N/A	Met	Met	Met	N/A	Met	N/A
Total	17	22	22	24	19	20	21	23	22	20	22

*Applicable to both group design and single-case design

**Applicable only to group design

***Applicable only to single-case design assignment, positive effects, and at least 30 total participants, as well as three methodologically sound single-subject research studies with positive effects and at least ten total participants.

Discussion

The primary purpose of this review was to extend the work of Ramey et al. (2016) and assess the empirical support for the use of frequency building and precision teaching with school-aged children. This review differed from Ramey et al. (2016) in that the authors of the current paper focused on only one intervention frequently used with PT (i.e., frequency building). Additionally, the current review included children of all abilities levels, whereas Ramey et al. (2016) focused on students with disabilities. Lastly, nine new studies were included in the current review that were not included in the review conducted by Ramey et al. (2016).

Results of WWC and CEC standards, visual analysis, and effect size estimates indicate that frequency building and PT are effective methods for increasing academic skills with school-aged children. All of the studies included in this review *Met*, or *Met with Reservations*, the WWC standards, indicating strong methodological control. In addition, five of the studies met all of the CEC quality indicators, with the rest meeting the majority of the indicators. WWC examines criteria in five areas to determine a study's rating, whereas CEC examines 22 indicators for SCD and 24 indicators for group designs. The CEC standards examine factors such as training of implementer, measures of treatment fidelity, and information provided on the study's setting and participants, which are not included in the WWC standards. Even with the stringent standards laid out by the CEC, the practice of frequency building and PT meets the criteria as an evidence-based practice for school-aged children.

When examining the specific academic areas and learner populations, the group studies demonstrated strong effect sizes in the areas of math and reading for learners with and without disabilities (i.e., SLD, EBD, at risk). The SCD studies demonstrated strong effects in the areas of reading, writing, fine motor skills, and functional living/communication skills for learners with disabilities. Only one study evaluated fine motor, functional living, and communication skills. These results suggest that frequency building and PT can and should be considered in the selection of evidence-based procedures for learners with and without disabilities.

The current analysis of the existing literature reveals that researchers have used frequency building to teach academic skills across different domains to learners with and without disabilities. Further, the findings suggest that researchers, teaching assistants, and classroom teachers can implement frequency-building procedures with success. Similarly, the results show the generality of the procedure across settings including in the student's home, primary school, private school, and residential schooling programs. Taken together, these findings from the current analysis suggest that a variety of clinicians can use frequency building coupled with precision teaching to produce academic gains for learners of all kinds.

This review also offers an in-depth analysis on the current state of PT and frequency-building procedures. The findings provide a basis for addressing several concerns and criticisms in regard to use of these procedures. Doughty et al. (2004) remarked that the current state of literature does not support the claims made by precision teachers. Although the sample size of studies included in this review was relatively low, it seems that in the 14 years from the publication of the Doughty et al. (2004) paper that precision teachers have begun to listen to the call for peer-reviewed empirical evidence. All but one of the studies included in this review were published after 2004.

Limitations and Future Research Directions

When considering the results of this review, readers should take into consideration that only studies that were published in a peer-reviewed journal were included in this review. A potential for publication bias exists given that unsuccessful studies are less likely to be published than studies that show a positive effect (Lipsey and Wilson 2001). Additionally, this review excluded multiple studies that met the requirements for *Does Not Meet Standards* according to WWC. While this decision was made to examine the effectiveness of the intervention with a high degree of rigor, future researchers may want to examine the breadth of research on frequency building and PT by conducting a review that does not exclude studies based on WWC standards. Alternatively, all studies that met the remaining inclusionary criteria could be included. This would allow for multiple levels of analysis and would allow readers to see the full range of literature currently out there along with the rigor of each study.

Although this review demonstrates that the package of frequency building and PT is an effective technology for school-aged children, additional research is needed. To start, researchers in the field of PT should continue to produce studies of high methodological control. The number of high-quality studies has increased over the last few years; however, the number is small in comparison with other educational strategies. Additional controlled studies may increase the acceptance of PT and benefit the field of education by yielding more efficient and effective techniques for learning (Doughty et al. 2004).

Furthermore, several of the interventions included in this review had multiple components, making it difficult to determine which pieces contributed, or are needed, to generate effective outcomes. Research focused on component analyses may provide valuable information on the most efficient way for practitioners to produce strong academic gains. More specifically, Doughty et al. (2004) remarked that there are multiple components not accounted for in the available literature including (1) controlling for the overall number of practice opportunities and (2) controlling for the frequency of reinforcement. The current review did not measure the degree to which studies measured and controlled for these variables. Thus, future research and reviews should investigate not only the pragmatic nature of PT, but also the mechanistic nature of why these interventions produce behavior change. Studies directly comparing the outcomes of restricted operant teaching approaches (e.g., discrete trial training) with those of free operant approaches (e.g., frequency building) that controls for the same amount of instructional presentations are one area of research that should garner more attention (c.f. Holding et al. 2011; Nopprapun and Holloway 2014). Similar research should also control for the relative frequency of reinforcement delivered during frequency-building interventions compared to other instructional strategies.

Implications for Practitioners

Practitioners have successfully applied the science of behavior through PT for over 50 years; however, as stated earlier, PT is rarely used in mainstream education (Binder and Watkins 2013). It has been hypothesized that this may be due to a lack of high-quality research in the field (e.g., Cihon 2007; Ramey et al. 2016). Teachers are now required by law, and called on by professional organizations, to use evidence-based strategies in their classrooms (CEC 2014). Both the Individuals with Disabilities Education Act (IDEA 2004) and The Every Learner Succeeds Act (ESSA 2015) require schools to use programs, curricula, and practices based on "scientifically based research." School teams often turn to government and professional organization databases (e.g., WWC, CEC) to assess the quality of a teaching practice. The current review suggests that frequency building and PT can be included in the discussions for selection and implementation of evidence-based practices.

The results of this review indicate the use of frequency building and PT is effective for increasing academic skills across a variety of populations and subject areas. In addition, the research shows that large gains can be made in a short amount of time. The average intervention length was five weeks, with some studies demonstrated results in less than three weeks. Results of the research indicate that progress can be made in as little as ten minutes per day. Systems that can produce expeditious results are beneficial to teachers who are already overwhelmed with the amount of content they are expected to teach in a school day. Despite the empirical question of whether it is the number of exposures to material versus the speed at which they are presented, from a pragmatic standpoint, it may be wise to include frequency-building interventions if it speeds up the presentations of instructional episodes. Such a practice would enable teachers to spend less time per each instructional objective and help accelerate learning for children who are already behind.

Lastly, frequency building and PT are not specific to any particular curriculum or subject matter; therefore, any teacher can benefit from the precise measurement and graphical analysis that allows for individualized decision making (Kubina 2019). Teachers can begin to augment specific curricula their schools require by including an additional 15 min of frequency building a day. Similar methods were employed in the Great Falls Precision Teaching Project (Beck and Clement 1991) with much success. Incorporation of these techniques is being made easier for practitioners with the availability of several textbooks, blogs, webinars, and Web sites for novice precision teachers; PT is more accessible than ever before.

Conclusion

The current review includes strong evidence for the effectiveness of using frequency building and PT to increase academic skills for school-aged children. Precision teachers have produced impressive improvements in learning for several years; however, frequency building and PT have yet to become a staple in the public school classroom. Precision teachers should continue to publish methodologically strong studies in peer-reviewed journal in order to increase the likelihood that mainstream educators will learn about and implement frequency building and PT in their classrooms.

Acknowledgements This paper was completed as part of the dissertation requirements for the first author's doctoral degree. We would like to extend a thank you to the committee members Dr. Terri Hessler, Dr. Moria Konrad, Dr. Sheila Morgan, Dr. Christin McDonald-Fix, and in particular Dr. Donny Newsome, for their feedback on earlier drafts of this manuscript. In addition, we would like to thank Dr. Rick Kubina for his thoughtful review of this manuscript.

Funding This study was supported by OSEP Grant # H325D140084.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

Human and Animal Rights This research did not include human or animal subjects.

References

References marked with an asterisk denote articles included in the systematic review.

Beck, R., & Clement, R. (1991). The great falls precision teaching project: An historical examination. *Journal of Precision Teaching, 8*, 8–12.

Binder, C. (1996). Behavioral fluency: Evolution of a new paradigm. *The Behavior Analyst, 19*, 163–197.

Binder, C. (2004). A refocus on response-rate measurement: Comment on Doughty, Chase, and O'Shields (2004). *The Behavior Analyst, 27*, 281–286.

Binder, C. (2010). Building fluent performance: Measuring response rate and multiplying response opportunities. *The Behavior Analyst Today, 11*(4), 214–225.

Binder, C., & Watkins, C. L. (2013). Precision Teaching and direct instruction: Measurably superior instructional technology in schools. *Performance Improvement Quarterly, 26*, 73–115.

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

Brosnan, J., Moeyaert, M., Brooks Newsome, K., Healy, O., Heyvaert, M., Onghena, P., et al. (2018). Multilevel analysis of multiple-baseline data evaluating Precision Teaching as an intervention for improving fluency in foundational reading skills for at risk readers. *Exceptionality, 26*, 137–161.

*Chapman, S. S., Ewing, C. B., & Mozzoni, M. P. (2005). Precision Teaching and fluency training across cognitive, physical, and academic tasks in children with traumatic brain injury: A multiple baseline study. *Behavioral Interventions, 20*, 37–49.

Cihon, T. M. (2007). A review of training intraverbal repertoires: Can precision teaching help? *The Analysis of Verbal Behavior, 23*, 123–133.

Cihon, T. M., White, R., Zimmerman, V. L., Gesick, J., Stordahl, S., & Eshleman, J. (2017). The effects of precision teaching with textual or tact relations on intraverbal relations. *Behavioral Development Bulletin, 22*, 129–146.

Cooper, J. O. (2005). Applied research: The separation of applied behavior analysis and Precision Teaching. In W. Heward, T. Heron, N. Neef, S. Peterson, D. Sainato, G. Cartledge, R. Gardner, L. Peterson, S. Hersh, & J. Dardig (Eds.), *Focus on behavior analysis in education: Achievements, challenges, and opportunities* (pp. 295–303). Upper Saddle River: Prentice hall.

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

Council for Exceptional Children (2014). *Standards for evidence-based practices in special education.* Retrieved from https://www.cec.sped.org.

Doughty, S. S., Chase, P. N., & O'Shields, E. M. (2004). Effects of rate building on fluent performance: A review and commentary. *The Behavior Analyst, 27*, 7–23.

*Datchuk, S. M., & Kubina, R. M. (2017). A writing intervention to teach simple sentences and descriptive paragraphs to adolescents with writing difficulties. *Education and Treatment of Children, 40*, 303–326.

*Datchuk, S. M., Kubina, R. M., & Mason, L. H. (2015). Effects of sentence instruction and frequency building to a performance criterion on elementary-aged learners with behavioral concerns and EBD. *Exceptionality, 23*, 34–53.

DeProspero, A., & Cohen, S. (1979). Inconsistent visual analysis of intrasubject data. *Journal of Applied Behavior Analysis, 12*, 573–579.

Drevon, D., Fursa, S. R., & Malcolm, A. L. (2017). Intercoder reliability and validity of WebPlotDigitizer in extracting graphed data. *Behavior Modification, 41*(2), 323–339.

Ellis, P.D. (2009). *Effect size calculators*, website *[*https://www.polyu.edu.hk/mm/effectsize/calculator *]* Accessed on June 1, 2018.

Every Child Succeeds Act (ESSA) of 2015, Public Law No. 114-95, S.1177, 114th Cong. (2015). Retrieved from https://www.congress.gov/114/plaws/publ95/PLAW-114publ95.pdf.

Fit Learning™ (n.d.). Retrieved July 1, 2018, from https://fitlearners.com.

*Greene, I., Mc Tiernan, A., & Holloway, J. (2018). Cross-age peer tutoring and fluency based instruction to achieve fluency with mathematics computation skills: A randomized controlled trial. *Journal of Behavioral Education, 27*, 145–171.

Holding, E., Bray, M. A., & Kehle, T. J. (2011). Does speed matter? A comparison of the effectiveness of fluency and discrete trial training for teaching noun labels to children with autism. *Psychology in the Schools, 48*, 166–183. https://doi.org/10.1002/pits.20535.

Individuals with Disabilities Education Improvement Act [IDEA 2002]. (2004). Public Law 108-446 (20 U.S.C. 1400 et seq.).

Johnson, K., & Street, E. M. (2004). *The Morningside model of generative instruction: What it means to leave no child behind*. Concord, M.A.: Cambridge Center for Behavioral Studies.

Johnson, K., & Street, E. M. (2013). *Response to intervention and precision teaching: Creating synergy in the classroom*. New York, NY: Guilford Press.

Kazdin, A. E. (2011). *Single-Case Research Design: methods for clinical and applied settings*. New York, NY: Oxford University 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 https://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf.

Kubina, R. M. (2005). The relations among fluency, rate building, and practice: A response to Doughty, Chase, and O'Shields (2004). *The Behavior Analyst, 28*, 73–76.

Kubina, R. M. (2019). *The precision teaching implementation manual*. Pittsburgh, PA: Greatness Achieved Publishing Company.

Kubina, R. M., & Yurich, K. K. (2012). *The precision teaching book*. Lemont, PA: Greatness Achieved Publishing Company.

Lefebre, L., Fabrizio, M., & Merbitz, C. (2008). Accuracy and efficiency of data interpretation: A comparison of data display methods. *Journal of Precision Teaching and Celeration, 24*, 2–20.

Ledford, J. R., & Gast, D. L. (2018). *Single case research methodology: Applications in special education and behavioral sciences*. New York: New York, Routledge.

*Lambe, D., Murphy, C., & Kelly, M. E. (2015). The impact of a Precision Teaching intervention on the reading fluency of typically developing children. *Behavioral Interventions, 30*, 364–377.

Lindsley, O. R. (1964). Direct measurement and prosthesis of retarded behavior. *Journal of Education, 147*, 62–81.

Lindsley, O. R. (1990). Precision teaching: By teachers for children. *Teaching Exceptional Children, 22*, 10–15.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks: Sage Publications, Inc.
*Mannion, L., & Griffin, C. (2018). Precision Teaching through Irish: effects on isolated sight word reading fluency and contextualised reading fluency. *Irish Educational Studies, 37*, 391.
*Mc Tiernan, A., Holloway, J., Leonard, C., & Healy, O. (2018). Employing Precision Teaching, frequency building, and the Morningside Math facts curriculum to increase fluency with addition and subtraction computation: A randomised-controlled trial. *European Journal of Behavior Analysis, 19*, 90.
National Autism Center. (2009). *National standards report*. Randolph MA: National Autism Center.
Nopprapun, M., & Holloway, J. (2014). A comparison of fluency training and discrete trial instruction to teach letter sounds to children with ASD: Acquisition and learning outcomes. *Research in Autism Spectrum Disorders, 8*, 788–802.
Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B. (2011). Combining non-overlap and trend for single case research: Tau-U. *Behavior Therapy, 42*, 284–299.
Quigley, S. P., Peterson, S. M., Frieder, J. E., & Peck, K. M. (2017). A review of SAFMEDS: Evidenced for procedures, outcomes, and directions for future research. *Perspectives on Behavior Science, 41*(1), 283–301. https://doi.org/10.1007/s40614-017-0087-8.
Rakap, S. (2015). Effect sizes as result interpretation aids in single-subject experimental research: Description and application of four nonoverlap methods. *British Journal of Special Education, 42*, 11–33.
Ramey, D., Lydon, S., Healy, O., McCoy, A., Holloway, J., & Mulhern, T. (2016). A systematic review of the effectiveness of Precision Teaching for individuals with developmental disabilities. *Review Journal of Autism and Developmental Disorders, 3*, 179–195.
*Roberts, W., & Norwich, B. (2010). Using Precision Teaching to enhance the word reading skills and academic self-concept of secondary school learners: A role for professional educational psychologists. *Educational Psychology in Practice, 26*, 279–298.
Rohatgi, A. (2018). WebPlotDigitizer. Web-based software for data extraction (Version 4.0)[Web-based application]. San Francisco, California, USA. Retrieved May 20, 2018.
*Seevers, R., Malanga, P., & Cooper, J. (2003). The use of self-managed proofreading for detecting and correcting mechanical errors by learners with a learning disability. *Journal of Precision Teaching and Celeration, 19*, 2–17.
Standard Celeration Society (2018). Retrieved July 1, 2018, from https://celeration.org/precision-teaching-journal.
Slavin, R. E., & Fashola, O. S. (1998). *Show Me the Evidence! Proven and Promising Programs for America's Schools*. Thousand Oaks, CA: Corwin Press Inc.
*Strømgren, B., Berg-Mortensen, C., & Tangen, L. (2014). The use of precision teaching to teach basic math facts. *European Journal of Behavior Analysis, 15*, 225–240.
*Twarek, M., Cihon, T., & Eshleman, J. (2010). The effects of fluent levels of big 6+ 6 skill elements on functional motor skills with children with autism. *Behavioral Interventions, 25*, 275–293.
Vannest, K. J., Parker, R. I., & Gonen, O. (2011). *Single case research: Web-based calculators for SCR analysis. (Version 1.0)*[Web-based application]. College Station, TX: Texas A&M University. Retrieved June 1, 2018.
What Works Clearinghouse, Institute of Education Sciences, U.S. Department of Education. (2017). *What Works Clearinghouse: Procedures Handbook (Version 4.0)*. Retrieved from https://ies.ed.gov/ncee/wwc.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.