Educational Research Review 26 (2019) 52–70

Contents lists available at ScienceDirect

Educational Research Review

journal homepage: www.elsevier.com/locate/edurev

Review

A meta-analysis on the effectiveness of phonics instruction for teaching decoding skills to students with intellectual disability

Rachel Sermier Dessemontet[a,*], Catherine Martinet[a], Anne-Françoise de Chambrier[a], Britt-Marie Martini-Willemin[b], Catherine Audrin[a]

[a] University of Teacher Education of the State of Vaud, Special Needs Education Unit, Av. de Cour 33, 1014, Lausanne, Switzerland
[b] University of Geneva, Department of Psychology and Educational Sciences, 40 Boulevard du Pont-d'Arve, 1211 Geneva 4, Geneva, Switzerland

ARTICLE INFO

Keywords:
Meta-analysis
Intellectual disability
Developmental disability
Reading instruction
Reading interventions
Phonics instruction

ABSTRACT

This meta-analysis examines the effects of phonics instruction on the decoding skills of students with intellectual disability using a random-effects model. Eight single-case experimental studies and six experimental or quasi-experimental group studies met the inclusion criteria, encompassing a total of 297 participants with intellectual disability. The overall effect of phonics instruction on the decoding skills of persons with intellectual disability was large: $g = 1.42$ (95% CI: 0.75, 2.10). Single-case studies yielded a larger average effect size ($g = 1.94$, 95% CI: 1.10, 2.78) than group studies ($g = 0.41$, 95% CI: 0.13, 0.69). Researcher-designed tests also yielded larger effect sizes than standardized tests. The type of interventionist was not a significant moderator. In the majority of the studies, phonics instruction was carried out using a systematic and direct instruction approach and a one-to-one format. Implications for practice and research are presented, and areas that require further investigation are identified.

1. Introduction

Literacy is a crucial competence in modern-day society. Literacy skills are also very important for persons with intellectual disability (ID) who display significant limitations both in intellectual functioning and adaptive behavior (American Psychiatric Association, 2013; Schalock et al., 2010). Literacy skills enhance their social participation, quality of life and self-esteem (Bochner, Outhred, & Pieterse, 2001; Forts & Luckasson, 2011). Many studies point out that a significant number of persons with ID experience difficulties in reading or are illiterate (Lemons et al., 2013; Ratz & Lenhard, 2013; Towles-Reeves, Kearns, Kleinert, & Kleinert, 2008).

Historically, researchers and professionals have tended to underestimate the competences of students with ID in literacy. For example, for many years learning to decode has been perceived as too complex for persons with moderate and severe ID, whose IQ was equal or inferior to 50–55. Therefore, reading instruction with these students tended to be underemphasized and to be restricted mainly to sight word instruction (Ahlgrim-Delzell & Rivera, 2015; Browder, Wakeman, Spooner, Ahlgrim-Delzell, & Algozzinexya, 2006; Roberts, Leko, & Wilkerson, 2013). Sight word instruction consists of teaching students to recognize words, as they would recognize images, without treating the letter/sound correspondences that compose these words. Even if sight word instruction is useful for enhancing the independence of persons with ID, it only allows them to read a limited number of words. It does not allow

* Corresponding author. University of Teacher Education of the State of Vaud, Av. de Cour 33, 1014, Lausanne, Switzerland.
E-mail addresses: rachel.sermier@hepl.ch (R. Sermier Dessemontet), catherine.martinet@hepl.ch (C. Martinet), anne-francoise.de-chambrier@hepl.ch (A.-F. de Chambrier), britt-marie.martini-willemin@unige.ch (B.-M. Martini-Willemin), catherine.audrin@hepl.ch (C. Audrin).

https://doi.org/10.1016/j.edurev.2019.01.001
Received 25 March 2018; Received in revised form 2 December 2018; Accepted 10 January 2019
Available online 21 January 2019
1747-938X/

them to read words that have not been taught.

Recently, this pessimistic view of the learning potential in reading of students with moderate and severe ID has been challenged. The number of studies showing that they can learn decoding skills as a result of receiving intensive and systematic phonics instruction has increased over the last twenty years (Ahlgrim-Delzell, Browder, & Wood, 2014; Allor, Mathes, Roberts, Cheatham, & Otaiba, 2014; Bradford, Shippen, Alberto, Houchins, & Flores, 2006; Finnegan, 2012; Fredrick, Davis, Alberto, & Waugh, 2013; Lemons et al., 2015; Lemons, Mrachko, Kostewicz, & Paterra, 2012; Tucker Cohen, Wolff Heller, Alberto, & Fredrick, 2008). In phonics instruction, children are taught to identify letter/sound correspondences and to use them to decode words (NICHHD, 2000). In real terms, they are taught to convert letters or letter combinations into sounds and to blend them into words. Systematic phonics instruction is considered to be an evidence-based practice for teaching decoding skills to typically developing children and children with reading disabilities (Castles, Rastle, & Nation, 2018; Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2001). In contrast, several questions remain unanswered as regards the efficiency of phonics instruction for students with ID. Can systematic phonics instruction be considered an evidence-based practice for teaching decoding skills to students with severe, moderate or mild ID? Which instructional components or implementation conditions enhance the effectiveness of phonics instruction for students with ID? Such information is greatly needed to help researchers and professionals plan and implement effective reading interventions for these students.

1.1. Effect of phonics instruction on typically developing children

The extensive meta-analysis of the National Reading Panel (NICHHD, 2000) showed that systematic phonics programs were more effective in teaching typically developing children, at-risk children, and children with reading disabilities to read words and non-words than programs, which provide unsystematic phonics instruction or no phonics instruction at all. In systematic phonics programs, letter-sound correspondences are taught explicitly and systematically in a planned and ordered sequence (NICHHD, 2000). The findings from a subsequent meta-analysis of experimental group studies with randomized controlled trials confirmed the positive effect of systematic phonics instruction on reading accuracy for typically developing children and at-risk children (Torgerson, Brooks, & Hall, 2006). Additionally, two recent meta-analyses concluded that phonics instruction was an effective intervention to improve the reading skills of children and adolescents with poor reading skills (Galuschka, Ise, Krick, & Schulte-Körne, 2014; McArthur et al., 2015).

Different phonics approaches exist, such as synthetic phonics, analytic phonics, analogy phonics, and phonics through spelling. In synthetic phonics programs, children are taught first to transform letters into sounds and then to blend these sounds to form words. In contrast, in analytic phonics programs children are taught to identify words, and in a second step to analyze letter-sound relations in these words. In analogy phonics programs, children are taught to use parts of known written words (e.g., onsets or rimes) to identify new words. In phonics through spelling programs, they are taught to convert sounds into letters in order to write words. The authors of the National Reading Panel (NICHHD, 2000) and of a subsequent meta-analysis (Torgerson et al., 2006) found that the effect sizes for synthetic phonics programs and larger-unit phonics programs (e.g., analytic programs) did not differ statistically from each other. Moreover, the impact of systematic phonics instruction on growth in reading was similar whether instruction was delivered individually, in a small group format or to an entire class of typically developing children and at-risk children (NICHHD, 2000), or individually versus small group format to poor readers (McArthur et al., 2015).

Furthermore, the meta-analysis of the National Reading Panel (NICHHD, 2000) highlighted the fact that training phonemic awareness also helped typically developing children, at-risk children, or children with reading disability to develop reading skills. Phonemic awareness refers to the ability to focus on, and manipulate, phonemes (i.e., the smallest units which constitute spoken language) in spoken words. The authors of the National Reading Panel found that phonemic awareness training was more effective when it was provided to a small group of pupils, rather than individually or to the entire class. It was also more efficient if a strong focus was placed on blending and segmenting skills. The National Reading Panel stressed that it was important to combine phonemic awareness instruction with instruction on letter-sound correspondences and to teach children explicitly how to apply phonemic awareness skills when performing reading tasks.

Lastly, to become successful readers, students must also learn to read texts with speed, accuracy, and with the appropriate expression and to understand what they read (Castles et al., 2018; NICHHD, 2000). Therefore, fluency, vocabulary and reading comprehension must also be taught (NICHHD, 2000).

1.2. Effects of phonics instruction on students with ID

Several systematic reviews of the literature on reading instruction for students with ID were carried out. Browder et al. (2006) found 128 studies conducted between 1975 and 2003 on reading interventions for students with moderate or severe ID. They observed that, in the majority of these studies, the reading interventions mainly targeted vocabulary and the acquisition of sight words. Approximately one quarter of the studies targeted reading comprehension and/or fluency. Only 10% of the studies investigated phonics instruction. Moreover, only one study on the effects of phonics instruction was of sufficient methodological quality. Studies whose intervention targeted the development of phonemic awareness were even scarcer (4%). A systematic review of research by Roberts et al. (2013), which focused on reading interventions for youths with moderate or severe ID, found 19 studies published between 1975 and 2011. Once again, in most of the studies, the reading intervention targeted vocabulary acquisition through sight word instruction. Two studies targeted fluency and one targeted reading comprehension. Phonics was taught in only one study. None of these studies included phonological awareness training in their intervention.

The authors of both reviews concluded that the interventions implemented for students with moderate or severe ID did not

adequately take into consideration the components of reading, and did not address the recommendations made in literacy instruction for typically developing students (Browder et al., 2006; Roberts et al., 2013). The neglect of phonological awareness training and phonics instruction in the reading interventions offered to students with moderate or severe ID has a detrimental effect, as both phonemic awareness and letter-sound knowledge have been found to be significant predictors of reading skills in persons with ID (Barker, Sevcik, Morris, & Romski, 2013; Klusek et al., 2015; Laing, Hulme, Grant, & Karmiloff-Smith, 2001; Lemons & Fuchs, 2010; Levy, Smith, & Tager-Flusberg, 2003; Menghini, Verucci, & Vicari, 2004; Saunders & DeFulio, 2007; Sermier Dessemontet & de Chambrier, 2015; Soltani & Roslan, 2013; Wise, Sevcik, Romski, & Morris, 2010).

Two systematic reviews focused more specifically on the effect of phonics-based interventions for persons with ID. Joseph and Seery (2004) identified only seven studies published between 1990 and 2002 on this topic. The majority of the reviewed studies had research designs that did not allow to draw solid conclusions about the effects of phonics-based interventions (e.g., studies without a control group, posttest-only studies). Moreover, the interventions used in most of these studies included strategies considered as phonics-based, but which did not correspond to systematic phonics instruction, as defined by the National Reading Panel (NICHHD, 2000).

In their recent systematic review, the purpose of which was to update the prior review of Joseph and Seery (2004), Hill (2016) found eleven studies published between 2001 and 2013. Six of these studies were experimental or quasi-experimental studies, and five were single-case experimental studies. In the majority of the reviewed studies, students received systematic phonics instruction aimed at teaching letter-sound correspondences and/or decoding skills. The author of the review concluded that there had been an increase in the number of published studies on the effects of phonics instruction for students with ID and that these students seemed to respond well to phonics instruction.

Although the systematic review by Hill (2016) provides useful information, no solid conclusions can be drawn as to whether phonics instruction can be considered an evidence-based intervention for students with ID. Neither can conclusions be drawn as to which instructional characteristics may render the interventions more efficient. This type of information can only be acquired through a meta-analysis, which typically includes an assessment of the methodological quality of the studies and the calculation of weighted effect sizes.

1.3. Present study

The purpose of this meta-analysis was to offer a more in-depth understanding of the effectiveness of phonics instruction for individuals with ID. More specifically, the following research questions were investigated:

1. Is phonics instruction effective in teaching decoding skills (word and non-word reading) to persons with mild, moderate or severe ID?
2. To what extent does the research design, the type of measure used, and the studies' methodological quality influence the effect sizes found across studies?
3. Which type of instructional approach (systematic vs. unsystematic, direct instruction vs. no direct instruction) and type of phonics instruction (synthetic, analytic, analogy, etc.) lead to the most effective outcomes for persons with ID?
4. What implementation conditions (instructional format and type of interventionist) lead to the most effective outcomes for persons with ID?

2. Method

2.1. Literature search procedure and criteria

The electronic database searches were conducted using PsycINFO and ERIC. Studies published between January 1990 and November 2016 in peer-reviewed journals were searched. The following descriptors for intellectual disability were used: “intellectual disabilit*”, “developmental disabilit*”, “mental retardation”, “significant cognitive disabilit*”, “Down syndrome”, “Williams syndrome”, “X Fragile syndrome”. All of these descriptors were cross-referenced with the following keywords: “reading intervention”, “reading instruction”, “phonics”, “phonological awareness”, “letter-sound knowledge”, and “letter knowledge”.

The following criteria were used to determine eligibility:

(a) *participants:* at least two-thirds of the study's participants were described as having an ID prior to the study or as presenting a syndrome known to be associated with an ID in the large majority of individuals, for example Down syndrome (Edgin, 2013), or a specific/separate analysis for the participants with ID was provided. When a single-case study had participants with a different diagnosis, only the findings related to the participants with ID were included in this meta-analysis.
(b) *intervention:* the intervention involved phonics instruction, i.e., teaching students to identify letter-sound correspondences and/or to use them systematically to read words (NICHHD, 2000).
(c) *design*: the study had either an experimental or quasi-experimental pretest posttest design with a control group of individuals with ID receiving another type of intervention, or a single-case experimental design. A group study was described as experimental if the participants were randomly assigned to the experimental and control group. If this randomization did not place, the study was described as quasi-experimental. Experimental single-case studies with AB designs were excluded because they do not provide sufficient experimental control to allow confirmation of a functional relationship between the manipulation of the independent

variable and change in the dependent variable (Cook et al., 2014; Horner et al., 2005; Smith, 2012). Moreover, single-case experimental studies had to have at least three measures during baseline and three measures during intervention (Cook et al., 2014; Smith, 2012). They also had to include at least three participants with ID in order to allow us to calculate an effect size (Zelinsky & Shadish, 2016).

(d) *dependent variables*: word reading and/or non-word reading were measured in the study.

In a first step, 409 records were identified using the PsycINFO and ERIC databases. In addition, the reference lists of systematic reviews were searched for further records (Browder et al., 2006; Hill, 2016; Joseph & Seery, 2004; Lemons & Fuchs, 2010; Machalicek et al., 2010; Roberts et al., 2013; Whalon, Al Otaiba, & Delano, 2009). A total of 23 new records were found.

In a second step, the title and abstracts of the 432 records were screened to examine if the article corresponded to the inclusion criteria. A total of 395 records were excluded. The remaining 37 articles appeared to correspond to the criteria. Inter-rater reliability between the authors of the present article in this screening process was 95%. When a disagreement between two authors occurred, the inclusion of the study was discussed until an agreement was reached.

In a third step, the reference lists of those 37 articles were searched for further records. Sixteen supplementary articles were found. Their abstracts were screened and four articles corresponding to the selection criteria were identified.

In a fourth step, these 41 articles were assessed for full-text eligibility independently by two authors of the present article. Interrater reliability in determining full-text eligibility was 97%. When a disagreement between two authors occurred, the inclusion of the study was discussed until an agreement was reached. Several articles were excluded because they did not meet the criterion related to participants' diagnosis (Coleman-Martin, Heller, Cihak, & Irvine, 2005; Hanser & Erickson, 2007; Swinehart-Jones & Heller, 2008; Travers et al., 2011; Truxler & O'Keefe, 2007). A few studies were excluded because the intervention did not involve phonics instruction (Basil & Reyes, 2003; Eikeseth & Jahr, 2001; Tripiana-Barbosa & de Souza, 2015). Several articles were excluded because they did not fulfill the criterion related to research design. Some single-case studies had a pretest–posttest design instead of an experimental design with at least three measures during baseline and intervention for word reading or non-word reading measures (Al Otaiba & Hosp, 2004; Beecher & Childre, 2012; Bradford et al., 2006; Grindle, Hughes, Saville, Huxley, & Hastings, 2013; Reed, 2013; Riepl, Marchand-Martella, & Martella, 2008; Tyler et al., 2015). One single-case study was excluded because it had fewer than three participants with ID (Boyle & Walker-Seibert, 1997). Another single-case study was excluded because it reported findings partially (individual data for only three participants among five) (Fredrick et al., 2013).

Studies by Browder, Ahlgrim-Delzell, Courtade, Gibbs, and Flowers (2008) and Browder, Ahlgrim-Delzell, Flowers, and Baker (2012) had to be excluded because the researchers used a composite score encompassing diverse early reading skills. No specific score for non-word reading or word reading was provided. Additionally, three single-case studies had to be excluded because word reading and non-word reading were not measured during the baseline, which did not allow us to calculate an effect size (Bailey, Angell, & Stoner, 2011; Flores, Shippen, Alberto, & Crowe, 2004; Waugh, Fredrick, & Alberto, 2009). Three group studies also had to be excluded because they did not report sufficient information for the calculation of an effect size (Cleave, Kay-Raining Bird, & Bourassa, 2011; Conners, Rosenquist, Sligh, Atwell, & Kiser, 2006; Goetz et al., 2007).

Lastly, several studies by Allor and colleagues were conducted with part of the same sample (Allor et al., 2014; Allor, Gifford, Al Otaiba, Miller, & Cheatham, 2013; Allor, Mathes, Roberts, Cheatham, & Champlin, 2010; Allor, Mathes, Roberts, Jones, & Champlin, 2010). Their study published in 2010 (Allor, Mathes, Roberts, Cheatham, et al., 2010) was chosen for inclusion in this meta-analysis instead of their more recently published study (Allor et al., 2014), because the latter added in the sample students in the borderline range of intellectual disability (IQ: 70–80).

A final total of 13 articles were selected for this meta-analysis. One of these articles reported a triple experiment (Lemons et al., 2012). Only the two first experiments reported by Lemons et al. (2012) were included in the meta-analysis because word reading was not measured in the third experiment. The total number of studies was therefore 14. The entire identification process is illustrated by a flowchart in Fig. 1.

2.2. Coding procedure

All 14 studies were coded descriptively. Two coders (authors of the article) coded them independently with a coding guide. Discrepancies in coding were discussed and resolved during group meetings. The following elements were coded: (1) country, (2) research design and type of control/baseline condition, (3) participants (diagnosis, age, IQ, associated impairments, setting), (4) intervention program (researcher-designed vs. existing program), (5) implementation of the intervention (instructional format, length of the intervention, intensity of the intervention, interventionist), and (6) dependent variables.

Furthermore, (7) the instructional characteristics of the phonics interventions were also coded with the following categories:

(a) *Reading instruction components involved in the program:* phonemic awareness training, phonics instruction, vocabulary instruction (instructional intervention designed specifically to increase students' word knowledge or comprehension), comprehension instruction (instructional intervention designed specifically to increase students' reading comprehension of connected texts; for example, questioning, strategy instruction, or text structure instruction), and fluency instruction (instructional intervention designed specifically to increase students' fluency in reading connected text; for example, repeated reading).

(b) *Systematic phonics instruction vs. unsystematic phonics instruction.* A key feature of systematic phonics instruction is that a planned and sequential set of letter-sound correspondences is taught explicitly, systematically and sequentially (NICHHD, 2000). In unsystematic phonics instruction approaches, the letter-sound correspondences to be taught are not prespecified. For example, in

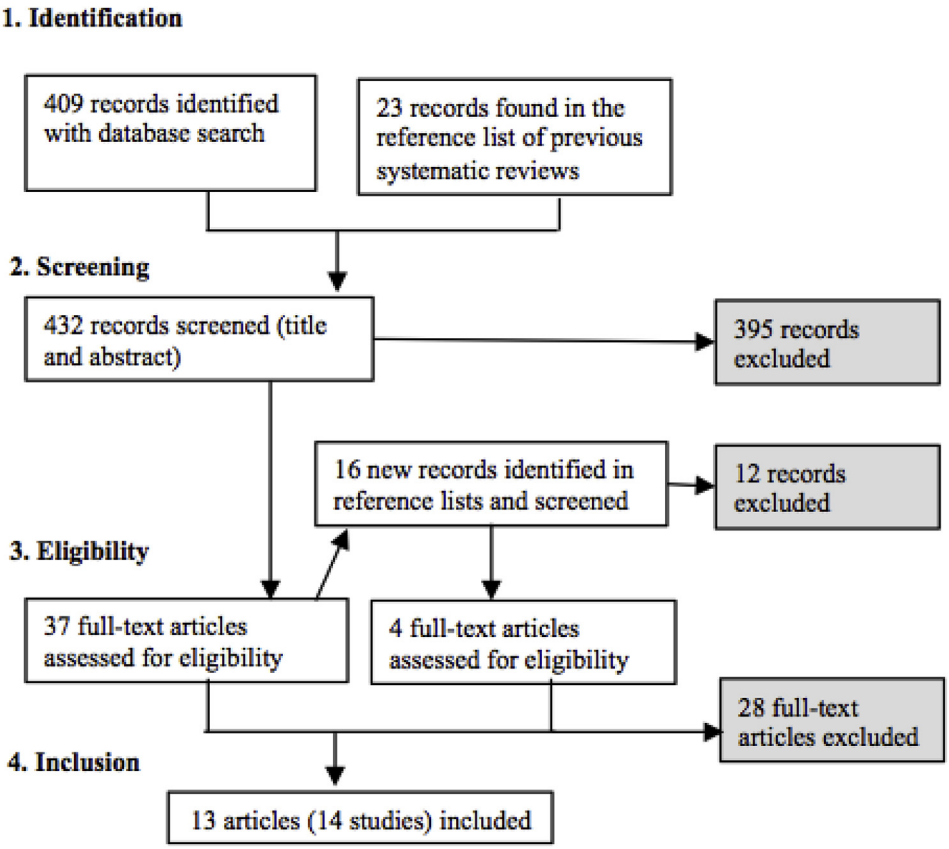

Fig. 1. Flowchart of the identification process.

the whole language approach, students are taught letter-sound correspondences unsystematically and incidentally in the context of oral or written language activities (NICHHD, 2000).

(c) *Direct instruction vs. no direct instruction.* In a direct instruction approach, teachers teach specifically and explicitly the letter-sound correspondences to their students and/or how they must be used to decode words. Direct instruction involves giving explicit instructions, modeling or demonstrating the required skills, and providing guided practice with feedback (Rupley, Blair, & Nichol, 2009).

(d) *Type of phonics instruction approach* (synthetic phonics, analytic phonics, phonics through spelling, analogy phonics, etc.)

(e) *Use of systems of prompts vs. no system of prompts.* These strategies make use of verbal, visual or physical prompts that are provided systematically and sequenced to increase or decrease the amount of information/help provided to the student. Their goal is to help students reach the highest level of independent performance (Downing, 2010). The following strategies were coded as systems of prompts: constant time delay (the instructor uses a fixed wait time prior to additional prompting), simultaneous prompting (the instructor guides the student as he makes a demand, leading to errorless learning), least-to-most prompts (the instructor begins with the less supportive/intrusive prompt and provides progressively more supportive/intrusive prompts following wait times and only if the student needs them), and most-to-least prompts (the instructor begins with the most supportive/intrusive prompts and fades them progressively as the student demonstrates greater efficiency (Downing, 2010). The use of systematic corrective procedures following students' mistakes using modeling was not coded as systems of prompts.

(f) Other relevant instructional characteristics

2.3. Assessment of the studies' methodological quality

Articles were reviewed to assess the studies' methodological quality using the quality indicators of the Council for Exceptional Children [CEC] for experimental or quasi-experimental studies and for single-case experimental studies (Cook et al., 2014). In developing these standards, the CEC workgroup (Cook et al., 2014) drew from a number of other sources, among them Gersten et al. (2005), Horner et al. (2005) and What Works Clearinghouse. The workgroup also took into account the feedback of anonymous special education researchers who participated in a Delphi study. The CEC identified 24 indicators that must be met in group studies and 22 indicators that must be met in single-case studies. One of the indicators is that the study "provides sufficient information to identify the population of participants to which results may be generalized and to determine or confirm whether the participants demonstrated the disability or difficulty of focus" (Cook et al., 2014, p. 3). In addition to meeting the inclusion criteria related to

diagnosis described previously, the studies were required to report IQ scores for the participants in order to be considered as meeting this indicator. Two authors of the present article independently coded each article with these quality indicators. Inter-rater reliability was high (97% agreement). Discrepancies in coding were discussed and resolved during group meetings.

2.4. Effect size estimates

Individual effect size estimates were computed for each study. Studies included in our meta-analyses were either single-case experimental studies or (quasi)experimental studies with control and treatment groups and pretest–posttest design. There is a consensus on the use of a standardized mean difference effect size (*d*-statistic) with a correction for small sample bias (often called Hedges' *g*) in meta-analysis of between-group studies reporting means and standard deviations collected using different measures (Borenstein, Hedges, Higgins, & Rothstein, 2009). In contrast, there is no consensus on the effect size that should be used in meta-analyses of single-case studies (Jamshidi et al., 2018; Shadish, Hedges, Horner, & Odom, 2015). Hedges, Pustejovsky, and Shadish (2012; 2013) developed a procedure for single-case designs that allows the researcher to calculate effect sizes in the same metric as the *d*-statistic from between-group designs, also correcting for small sample bias (Hedges' *g*). This procedure was used, as recommended by Zelinsky and Shadish (2016), because it allowed us to include in our meta-analysis studies with between-group designs, as well as studies with single-case designs, and to compare the magnitude of their effect sizes.

For the single-case studies, we used the Graph Digitizer program to digitize data from the graphs. Two authors also visually checked all the digitized data independently. A few inaccuracies in the digitized data were identified for four studies (Ahlgrim-Delzell et al., 2014; Lemons et al., 2012_1; Lemons et al., 2012_2; Lemons et al., 2015). The authors of these studies kindly agreed to send us their original data, which were used to calculate effect sizes. Next, we followed the procedure recommended by Hedges et al. (2012; 2013) to calculate a *d*-statistic effect size with a small sample bias correction. We used the SPSS DHPS effect size macro (Shadish et al., 2014) to calculate Hedges' *g* for the single-case studies. Because this macro is designed to estimate effect sizes for AB or multiple baseline designs, we excluded data collected during a maintenance phase or follow-up phase in a few studies (Ahlgrim-Delzell et al., 2014; Fallon, Light, McNaughton, Drager, & Hammer, 2004; Joseph, 2002). This also allowed us to keep as similar as possible measures of the dependent variable among the single-case studies, not all of them involving such supplementary phases.

For the group studies with control and treatment groups and pretest–posttest designs, we computed the *d*-statistic presented in Morris (2008). This effect size is defined as the mean pretest–posttest change in the experimental group minus the mean pretest–posttest change in the control group, divided by the pooled pretest standard deviation that pools the data from the experimental and control group. Morris' formula (2008) includes a small sample bias correction. Morris (2008) showed that this method for effect size calculation yielded more precise estimates of effect size and sampling variance than other methods. The sampling variance was computed following Morris and DeShon (2002)'s recommendation. We computed the variance for control and treatment groups and summed them to obtain the variance of our effect sizes. The study by Finnegan (2012) comprised two intervention groups receiving phonics instruction (synthetic vs. analogy phonics) and one control group. We calculated separate effect sizes for each of these interventions, comparing them to the control group following the recommendations made by Borenstein et al. (2009) for studies with several treatment groups.

In this meta-analysis, effect sizes corrected for small sample bias are referred to as *g*. Since *g* is a *d*-statistic corrected for small sample bias, it can be interpreted using the benchmarks defined by Cohen (1988). He defined a small effect size as $d = 0.2$, a medium effect size as $d = 0.5$, and a large effect size as $d = 0.8$.

2.5. Meta-analytical analysis

The approach described by Zelinsky and Shadish (2016), consisting in combining single-case studies and group studies in a meta-analysis, was adopted. This approach seemed especially interesting, as we expected to find only a few studies with small samples to include in our meta-analysis. Having only a small number of studies can potentially cause problems of power in a meta-analysis (Borenstein et al., 2009). A random-effects model was chosen because we assumed that study-level variance would be present as an additional source of random influence. Indeed, our meta-analysis included studies displaying several differences that could affect effect sizes (e.g., different research designs, type of measures or type of interventions).

Several group studies included in our meta-analysis had multiple outcomes: non-word reading and word reading measured with a standardized test, and sometimes also word reading measured with a researcher-designed test. This resulted in some studies having several effect sizes. To take into account this nested structure and to avoid losing information by averaging effect sizes, we used robust variance estimation when conducting the summary effect size analysis, heterogeneity test and the moderator analyses (Hedges, Tipton, & Johnson, 2010). This method allowed us to take into account the multiple effect sizes within studies while adjusting standard errors for the correlation between them. Robust variance estimation also allowed us to take into account the nested nature of the data in Finnegan's study (2002), which comprised two intervention groups and one control group. Moreover, robust variance estimation provides an adjustment of the variance and degrees of freedom for meta-analysis comprising only a small number of studies (Tipton, 2015; Tipton & Pustejovsky, 2015).

We used the robumeta package with the software R-statistics (Fisher & Tipton, 2015) to conduct the summary effect size analysis, heterogeneity test and the moderator analyses. The remaining analyses (influence analyses, publication bias analyses) cannot be performed with robust variance estimation (Zelinsky & Shadish, 2016). Therefore, we had to aggregate effect sizes in studies reporting multiple outcomes for the purpose of these analyses. These analyses were performed using the metafor package (Viecthbauer 2010).

Table 1
Experimental or quasi-experimental studies.

Study	Participants				Intervention			Implementation	Interventionist	Measures	Effect size
					Program	Components	Characteristics				
Ahlgrim Delzell et al. (2016)	N = 31	ID or DDel[1] AAC[2] users	IQ: 40-88	5-14y	Researcher designed	Phonemic awareness Phonics Comprehension	Systematic instruction Direct instruction Prompting strategies	1 : 1 format 8 months 15-20 minutes daily	Teacher	Word reading[3]	1.65
Allor et al. (2010)	N = 59	mild or moderate ID	IQ: 40-69	6-10y	*Early interventions in reading* (Mathes & Torgesen, 2005)	Phonemic awareness Phonics Vocabulary Comprehension Fluency	Systematic instruction Direct instruction Prompting strategies	1 : 1-4 format 2-3 years 40-50 minutes daily	Research-affiliated teacher	Non-word reading Non-word reading Word reading	0.69 1.03 0.43
Burgoyne et al. (2012)	N = 54	Down syndrome	IQ: NR	5-10y	Researcher designed	Phonemic awareness Phonics Vocabulary Comprehension	Systematic instruction	1 : 1 format 5 months 40 minutes daily	Teaching assistant	Non-word reading Word reading	0.25 0.22
Cohen et al. (2006)	N = 52	mild ID	IQ: 51-79	*M* = 33y	Researcher designed	Phonemic awareness Phonics Comprehension Fluency	Hierarchical adaptive intervention	1 : 1 format 15 months minutes: NR bi-weekly	Research-affiliated reading specialist	Word & non-word reading	0.50
Coyne et al. (2010)	N = 16	ID	IQ: NR	5-9y	Researcher designed + softwares	Phonemic awareness Phonics Vocabulary Comprehension Fluency	Scaffolded E-books and software programs	1 : 1 format 8 months 20-30 minutes 4-5 per week	Teacher - Computer	Non-word reading Word reading	0.37 0.07
Finnegan (2012)	N = 52	ID	IQ: *M* = 56 *SD* = 13 or NR	5-12y	Researcher designed	Phonics	Systematic instruction Direct instruction Synthetic phonics vs. Analogy phonics	1 : 1 format 1-2 months 15-20 minutes 2-3 per week	Researcher	Non-word reading Word reading[6] Word reading[7]: - trained words - transfer words	Sy[4]:0.82 An[5]: 0.97 Sy: 0.13 An: 0.15 Sy: 0.67 An: 0.76 Sy: 0.29 An: 0.16

Note 1. Ddel = developmental delay, Note 2. AAC = augmentative and alternative communication systems, Note 3. Reading a word and pointing to the right picture, Note 4. Synthetic phonics, Note 5. Analogy phonics, Note 6. Standardized test, Note 7. Researcher-designed test.

3. Results

3.1. General description of the studies

3.1.1. Research design

Of the 14 studies, six were group studies and eight were single-case experimental studies. Among the group studies, four had an experimental design with random allocation in the experimental and control groups (Ahlgrim-Delzell et al., 2016; Allor et al., 2010; Burgoyne et al., 2012; Finnegan, 2012) and two had a quasi-experimental design (Cohen et al., 2006; Coyne, Pisha, Dalton, Zeph, & Smith, 2010). All the group studies, with the exception of Ahlgrim-Delzell et al. (2016), used standardized tests to measure word reading. Non-word reading was also measured with standardized tests, except in the study by Burgoyne et al. (2012). Finnegan (2012) used a standardized test to measure word reading, as well as a researcher-designed test. The detailed content of the group studies is summarized in Table 1.

All the single-case experimental studies had a multiple baseline across participants design. In all the single-case experimental studies, the effect of phonics instruction on word reading was measured using a researcher-designed test. The detailed content of the single-case studies is summarized in Table 2.

The instruction provided to the control/baseline conditions varied from one study to the other and often from one class to the other within the same study: sight word instruction, no reading instruction, or education as usual that could include or not include some phonics instruction. The reading instruction provided to control/baseline conditions was not always reported in detail.

Table 2
Single-case studies.

Study	Participants				Intervention			Implementation	Interventionist	Measures	Effect size
					Program	Components	Characteristics				
Ahlgrim-Delzell et al. (2014)	N = 3	moderate or severe ID AAC[1] users	IQ: 31-54	7-10y	Researcher designed	Phonemic awareness Phonics Comprehension	Systematic instruction Direct instruction Prompting strategies	1: 1 format 19-23 sessions 15–20 min daily	Teacher	Word reading[2]	1.77
Fallon et al. (2004)	N = 3	ID AAC[1] users	IQ: NR	9-14y	Researcher designed	Phonemic awareness Phonics	Systematic instruction Direct instruction	1: 1 format 5-17 sessions 30 min 2-3 per week	NR	Word reading[2]	1.15
Joseph (2002)	N = 3	mild ID	IQ: NR	9-10y	Researcher designed	Phonemic awareness Phonics	Systematic instruction Direct instruction Phonics through spelling	1: 1 format 10-12 sessions 40 min daily	Researcher	Word reading	2.47
Lemons et al. (2012_1)	N = 6	Down syndrome	IQ: 42–43 or NR	7-13y	*Road to reading* (Blachman & Tangel, 2008)	Phonemic awareness Phonics	Systematic instruction Direct instruction	1: 1 format 13-35 sessions 30–40 min 4 per week	Teacher or reading specialist	Word reading	1.86
Lemons et al. (2012_2)	N = 5	Down syndrome	IQ: 42–70 or NR	7-9y	*Road to reading* (Blachman & Tangel, 2008) + Phonemic awareness activity	Phonemic awareness Phonics	Systematic instruction Direct instruction	1: 1 format 12-29 sessions 35–45 min 4 per week	Teacher or reading specialist	Word reading	2.57
Lemons et al. (2015)	N = 5	Down syndrome	IQ: 40-73	6-8y	*Road to the code* (Blachman et al., 2000) adapted	Phonemic awareness Phonics	Systematic instruction Direct instruction Prompting strategies	1: 1 format 24-45 sessions 33 min 4 per week	Research-affiliated teacher	Word reading	0.88
Tucker Cohen et al. (2008)	N = 5	mild and moderate ID	IQ: 40-61	9-14y	Researcher designed	Phonics	Systematic instruction Direct instruction Prompting strategies	1: 1 format 10-20 sessions minutes: NR daily	Researcher	Word reading	3.24
Wolff Heller et al. (2002)	N = 3	mild ID severe speech and physical impairment	IQ: NR	9-23y	Researcher designed	Phonics	Systematic instruction Direct instruction	1: 1 format 8-17 sessions minutes: NR intensity: NR	Teacher	Word reading[3]	3.97

Note 1. AAC = augmentative and alternative communication systems, Note 2. Reading a word and pointing to the right picture. Note 3. Reading a word and identifying it among words spoken orally by the teacher.

3.1.2. Participants

The meta-analysis comprised 297 participants. In most studies, the participants were described as having an ID. The study by Ahlgrim-Delzell et al. (2016) also included a minority of students diagnosed with a developmental delay. Three single-case studies (Lemons et al., 2012_1; Lemons et al., 2012_2; Lemons et al., 2015) and one group study (Burgoyne et al., 2012) were conducted exclusively with children with Down syndrome. One group study (Ahlgrim-Delzell et al., 2016) and three single-case studies (Ahlgrim-Delzell et al., 2014; Fallon et al., 2004; Wolff Heller, Fredrick, Tumlin, & Brineman, 2002) were conducted with students with ID who also had severe speech impairments and used augmentative and alternative communication systems. In several studies, some participants not only had an ID but also an autism spectrum disorder (Ahlgrim-Delzell et al., 2014, 2016; Allor et al., 2010; Cohen et al., 2006; Coyne et al., 2010).

Participants had IQs ranging from 31 to 88. The few participants having IQs exceeding 70–75 were found in the study by Ahlgrim-Delzell et al. (2016), which included a minority of children with developmental delays, and in the study by Cohen et al. (2006). IQs were not reported in four studies (Burgoyne et al., 2012; Coyne et al., 2010; Fallon et al., 2004; Joseph, 2002). In four other studies, IQs were reported for less than 60% of the participants (Lemons et al., 2012_1; Lemons et al., 2012_2; Finnegan, 2012; Wolff Heller et al., 2002). Scores in adaptive behavior were never reported. Most of the participants were between 5 and 14 years old. Two participants in the single-case study by Wolff Heller et al. (2002) were older. Only one study was conducted with adult participants (Cohen et al., 2006).

3.1.3. General description of the interventions

3.1.3.1. Reading instruction components involved in the intervention. In three studies, participants were exclusively taught phonics (Finnegan, 2012; Tucker Cohen et al., 2008; Wolff Heller et al., 2002). Tucker Cohen et al. (2008) and Wolff Heller et al. (2002) taught students to decode words with known letter-sound correspondences, naming the letter-sound correspondences and then blending them. Finnegan (2012) taught a group of students to decode words using a synthetic phonics approach and taught another group using an analogy phonics approach (teaching them common rimes and how to combine them with letter-sound correspondences to read words with similar patterns).

In five studies, interventions combined phonics with phonemic awareness training (Fallon et al., 2004; Joseph, 2002; Lemons et al., 2012_1; Lemons et al., 2012_2; Lemons et al., 2015). Lemons et al. (2012_1; 2012_2) used *Road to Reading* (Blachman & Tangel, 2008), a program designed for students in Grades K-5 struggling to learn to read. Lemons et al. (2015) used *Road to the Code* (Blachman, Ball, Black, & Tangel, 2000), a program created for kindergartners and first-graders demonstrating difficulties in phonological awareness and letter/sound knowledge. They made several adaptations to this program to align it with the characteristics of children with Down syndrome. Fallon et al. (2004) trained students with ID and severe speech impairments to decode words containing known letter-sound correspondences using internal speech. Joseph (2002) taught students to spell and decode words using word boxes (magnetic boards containing a drawn rectangle divided into connected boxes) and a word-sort activity.

In two studies, the interventions were described as comprising phonics, phonemic awareness, and comprehension instruction (Ahlgrim-Delzell et al., 2014, 2016). Ahlgrim-Delzell et al. (2014; 2016) used *Early Reading Skills Builder* (Browder, Ahlgrim-Delzell, & Wood, 2015), a program they created to teach phonemic awareness skills, letter-sound correspondences, decoding skills and comprehension of connected texts to students with ID and severe speech impairments with the help of an augmentative and alternative communication device.

In four studies, the interventions were described as comprising phonics, phonemic awareness, comprehension instruction, vocabulary instruction, and/or fluency instruction (Allor et al., 2010; Burgoyne et al., 2012; Cohen et al., 2006; Coyne et al., 2010). Allor et al. (2010) used a program created for children having difficulties in learning to read: *Early Interventions in Reading* (Mathes & Torgesen, 2005). They had to complete this program with a researcher-designed foundational level, many of their participants having not yet mastered the skills required to begin level 1 of the program. Burgoyne et al. (2012) used a program that they created for students with Down syndrome and which included a reading strand as well as a language strand aimed to teach new vocabulary and promote its use in expressive language. Coyne et al. (2010) used Universal Design for Learning scaffolded e-books and software programs (Island of adventures, Ocean adventures) that train phonological awareness and phonics. Cohen et al. (2006) used a hierarchical adaptive intervention starting with modules targeting phonemic awareness and phonics and ending with modules at the global-reading task level (narrative comprehension and information-seeking). Cohen et al. (2006) also used an oral-manual principle to help adults with difficulties in decoding.

3.1.3.2. Systems of prompts. In five studies, the intervention incorporated a system of prompting strategies (constant time delay, system of least-to-most prompts or/and system of most-to-least prompts) (Ahlgrim-Delzell et al., 2014, 2016; Allor et al., 2010; Lemons et al., 2015; Tucker Cohen et al., 2008). Other studies mentioned using a systematic corrective procedure using modeling (Fallon et al., 2004; Joseph, 2002).

3.1.3.3. Intensity and length of the intervention. In most studies, the intervention was implemented daily, or four times per week (Ahlgrim-Delzell et al., 2014; Ahlgrim-Delzell et al., 2016; Allor et al., 2010; Burgoyne et al., 2012; Coyne et al., 2010; Joseph, 2002; Lemons et al., 2012_1; Lemons et al., 2012_2; Lemons et al., 2015; Tucker Cohen et al., 2008). Sessions lasted between 30 and 50 min in some studies (Allor et al., 2010; Burgoyne et al., 2012; Fallon et al., 2004; Joseph, 2002; Lemons et al., 2012_1; Lemons et al., 2012_1; Lemons et al., 2015) and between 15 and 30 min in others (Ahlgrim-Delzell et al., 2014, 2016; Coyne et al., 2010; Finnegan, 2012). In the single-case studies, the intervention lasted approximately between 2 and 15 weeks (5–45 sessions). In group studies, the intervention lasted between five months and three school years, with the exception of the study by Finnegan (2012), where the

intervention lasted between one and two months.

3.2. Studies' methodological quality

3.2.1. Single-case studies

Two single-case studies (Lemons et al., 2015; Tucker Cohen et al., 2008) met all the quality indicators of the CEC (Cook et al., 2014). Three studies (Ahlgrim-Delzell et al., 2014; Lemons et al., 2012_1; Lemons et al., 2012_2) met all the quality indicators except one indicator related either to the intervention agent's description (the amount of training received by the professionals implementing the intervention was not reported; Ahlgrim-Delzell et al., 2014) or to the participants' description (IQ scores were reported partially; Lemons et al., 2012_1, 2012_2). The three other single-case studies did not meet the quality indicators related to internal validity (e.g., the baseline condition was not described) (Fallon et al., 2004; Joseph, 2002; Wolff Heller et al., 2002), to the participants' description (IQ scores were not reported or reported partially) (Fallon et al., 2004; Joseph, 2002; Wolff Heller et al., 2002) and to the intervention agent's description (Fallon et al., 2004; Wolff Heller et al., 2002).

3.2.2. Group studies

Two group studies (Ahlgrim-Delzell et al., 2016; Allor et al., 2010) met all the quality indicators. One group study met all the quality indicators except one indicator related to the participants' description (IQ scores were not reported) (Coyne et al., 2010). The other group studies did not meet indicators related to participants' description (IQ scores were not reported or reported partially) (Burgoyne et al., 2012; Finnegan, 2012), to the context's description (Finnegan, 2012), to the intervention agent's description (Cohen et al., 2006), to the outcome measures (evidence on the validity and/or the fidelity of the measures was not reported or reported partially) (Burgoyne et al., 2012; Cohen et al., 2006), or to implementation fidelity (no information was reported on implementation fidelity) (Cohen et al., 2006; Finnegan, 2012).

3.3. Preliminary analyses

Preliminary analyses were conducted prior to calculating the average effect size of phonics instruction on decoding skills. A Grubbs test revealed that the effect size found for the study by Wolff Heller et al. (2002) ($g = 3.97$) was an outlier ($G = 2.77$, $p = .03$). The extremely large effect size found in this study is very probably related to a problem of internal validity. Indeed, students had already had one year before the study began to learn and train the decoding strategy that was taught during the intervention phase. The intervention targeted words that were not yet read correctly by the students. Because of this previous training, it is not surprising that the effect of the intervention was extremely large. This study was therefore not retained in the meta-analytical analysis.

Influence analyses were conducted omitting the study by Wolff Heller et al. (2002). Since influence analyses cannot be performed with robust variance estimation, we conducted these analyses with the aggregated effect sizes for studies reporting several outcomes. Thus, influence analyses were performed on 13 effect sizes (one effect size per study). They showed that the effect sizes were normally distributed. The large effect size found by Tucker Cohen et al. (2008) ($g = 3.24$) was close to being an outlier and caused significant heterogeneity in the data. However, this study was not excluded from the meta-analytical analysis because it had a high methodological quality, and because the influence analyses revealed that it did not have significantly more influence on the average effect size than other studies.

3.4. Average effect of phonics instruction on word and non-word reading

The random-effects average effect size was computed on the 24 effect sizes nested in the 13 studies using robust variance estimation. A significant and large average effect size of phonics instruction on the decoding skills of students with ID was found: $g = 1.42$ (standard error of 0.31), CI 95% = [0.75, 2.10], and t (10.4) = 4.67, $p < .001$. This suggests that phonics instruction clearly improves the decoding skills of students with ID. These results are depicted in Fig. 2 (forest plot), which shows the estimated effect sizes (squares), the 95% CI for each effect size (horizontal bar for each square), the weight of each effect size (size of the square), and the overall average effect size (diamond at the bottom of the figure). The group studies are presented in the upper part of the forest plot and the single-case studies in the lower part.

A heterogeneity test was performed in order to determine if the set of effect sizes was sufficiently homogeneous to render the mean effect size representative of that set. The test indicated a significant heterogeneity in effect sizes across the studies: Q (12.02) = 33.08, $p < .001$. I^2 equaled 64%, indicating that almost two-thirds of the total variability among effect sizes was not caused by sampling error but by true heterogeneity between the studies. Consequently, there was reason to examine moderator variables that may explain this heterogeneity in the effects found.

A sensitivity analysis was performed to determine if the random-effects average effect size, standard error, and variance component estimates changed if a different value for within-study correlation had been used. The estimates for the standard error (0.305) did not change between the lowest possible correlation ($\rho = 0$) and highest possible correlation ($\rho = 1$). The average effect size and variance component varied only very slightly from $g = 1.425$ and $\tau^2 = 0.665$ for $\rho = 0$ to $g = 1.424$ and $\tau^2 = 0.667$ for $\rho = 1$, showing that our findings are robust.

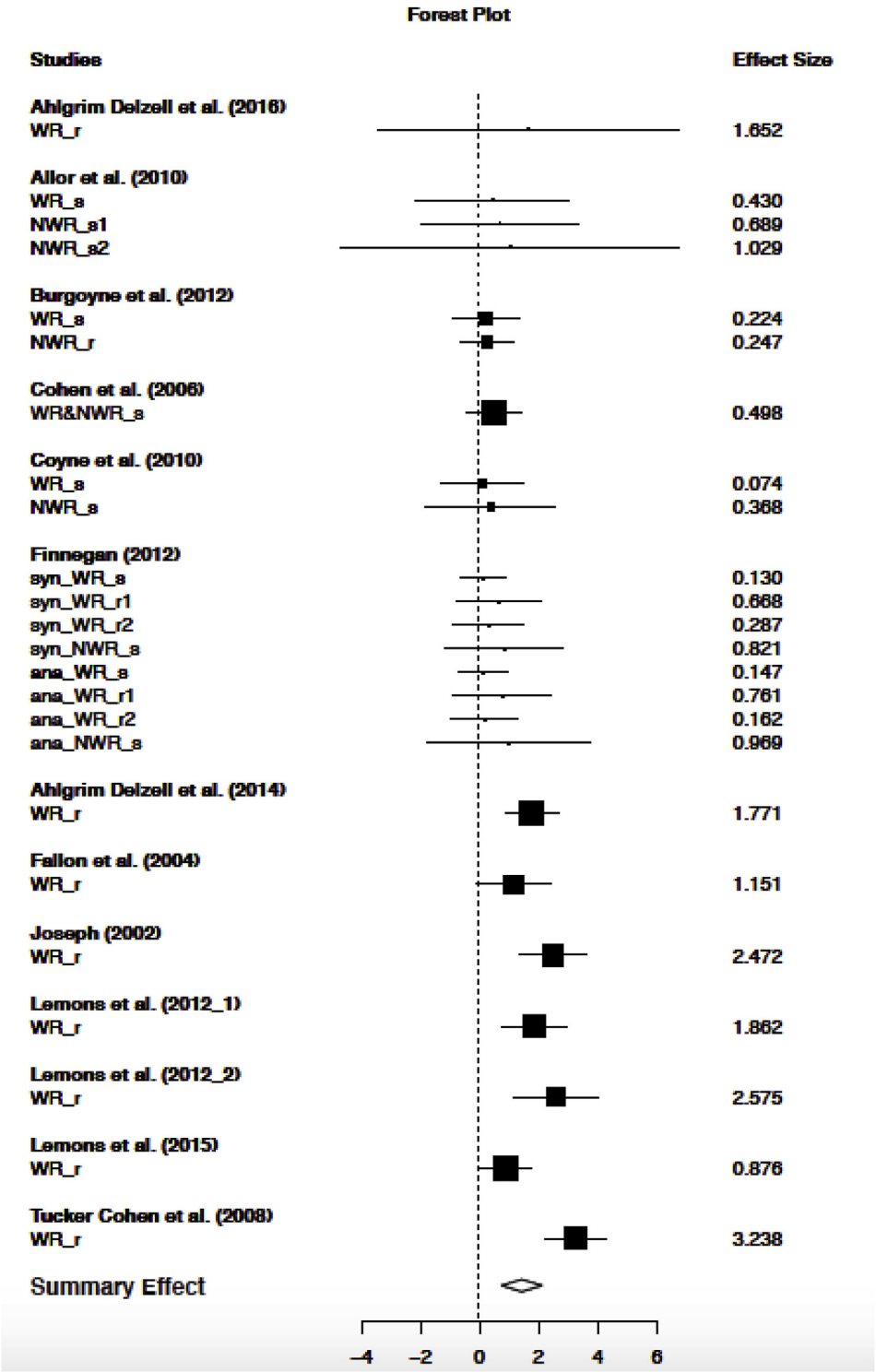

Fig. 2. Forest plot of effect sizes. WR_r = word reading measured with a researcher-designed test; WR_s = word reading measured with a standardized test; NWR_r = non-word reading measured with a researcher-designed test; NWR_s = non-word reading measured with a standardized test.

3.5. Influence of the studies' characteristics

3.5.1. Research design

The studies' research design (single-case study vs. group study) was a significant moderator: t (5.8) = −4.38, p = .005. Including this moderator in the model reduced by half the variability among effect sizes that is not caused by sampling error (from 64% to 31%). Single case studies yielded a larger overall average effect (g = 1.94, CI 95% = [1.10, 2.78], t (5.7) = 5.73, p = .001) than group studies (g = 0.41, CI 95% = [0.13, 0.69], t (2.9) = 4.81, p = .018). Both effect sizes were significant. In all the single-case studies, phonics instruction was found to have a large positive effect (g = 0.88–3.24). In contrast, the effect sizes found in group studies ranged from extremely small to large (g = 0.07–1.65).

3.5.2. Type of measure

The type of measure (researcher-designed test vs. standardized test) was a significant moderator: t (4.6) = −4.37, p = .009. Including this moderator in the model reduced the variability among effect sizes that was not caused by sampling error from 64% to 48%. Researcher-designed measures yielded a larger average effect size (g = 1.78, CI 95% = [1.04, 2.52], t (7.4) = 5.65, p < .001) than standardized measures (g = 0.40, CI 95% = [0.13, 0.67], t (2.8) = 4.97, p = .019). Both effect sizes were significant. Researcher-designed measures yielded small to extremely large effect sizes (g = 0.16–3.24). Standardized measures yielded extremely small to large effect sizes (g = 0.07–1.03).

It should be taken into account that this moderator may have a confounding influence with the type of research design, because all the single-case studies used a researcher-designed test. In contrast, only half of the group studies used a researcher-designed test for at least one of their outcome variables (Ahlgrim-Delzell et al., 2016; Burgoyne et al., 2012; Finnegan, 2012).

3.5.3. Studies' methodological quality

The studies' methodological quality (meeting all the quality indicators of the CEC vs. not meeting all the quality indicators) was not a significant moderator: t (2.8) = 0.59, p = .600. However, this finding must be considered with caution because the degree of freedom is inferior to 4, suggesting that it may be unreliable (Tipton, 2015).

3.6. Influence of the characteristics of the interventions

3.6.1. Instructional approach

3.6.1.1. Systematic vs. unsystematic phonics instruction. Systematic phonics instruction was used in all the studies reviewed in this meta-analysis, with perhaps the exception of the study by Coyne et al. (2010) for which there was insufficient information to determine if a systematic approach to teaching phonics was adopted. Therefore, it was not possible to conduct a moderator analysis yielding reliable findings on the influence of this variable.

3.6.1.2. Direct instruction. Direct instruction was used in the majority of the studies (Ahlgrim-Delzell et al., 2014; Ahlgrim-Delzell et al., 2016; Allor et al., 2010; Fallon et al., 2004; Finnegan, 2012; Joseph, 2002; Lemons et al., 2012_1; Lemons et al., 2012_2; Lemons et al., 2015; Tucker Cohen et al., 2008). Three group studies did not report using direct instruction (Burgoyne et al., 2012; Cohen et al., 2006; Coyne et al., 2010). The use of direct instruction was a significant moderator: t (3.1) = 4.41, p = .021. Studies that reported using a direct instruction approach yielded stronger effect sizes. However, this finding must be considered with caution because the degree of freedom is inferior to 4, suggesting that it may be unreliable (Tipton, 2015). It should also be taken into account that this moderator may have a confounding influence with the type of research design and the type of measure used, because the three studies that did not report using direct instruction were group studies using standardized tests.

3.6.1.3. Type of phonics approach. While some studies did not provide enough information to identify with certainty the type of phonics approach that was used, most of them seemed to use predominantly a synthetic approach. In this approach, students are taught to pronounce the sounds associated with letters and to blend the sounds to form words. One single-case study taught phonics through spelling to students with mild ID (Joseph, 2002). One group study used a synthetic phonics approach with one experimental group, and an analogy phonics approach with the other experimental group (Finnegan, 2012). There was an insufficient number of studies explicitly using other phonics instruction approaches to conduct a moderator analysis yielding reliable findings on the influence of type of phonics approach used.

3.7. Influence of the conditions of implementation of the interventions

3.7.1. Instructional format

The interventions were implemented in a one-to-one format in all the studies with the exception of the study by Allor et al. (2010), where the intervention was also implemented in small group format (2–4 students). There was an insufficient number of studies implementing the intervention in a small group format to conduct a moderator analysis yielding reliable findings on the influence of the instructional format.

3.7.2. Type of interventionist

Researchers or research-affiliated professionals implemented the intervention in several studies (Allor et al., 2010; Cohen et al.,

Table 3
Moderator analysis.

Moderator	β_0	SE	β_1	SE	*t*	*df*	*p*
Research design	3.48	0.68	−1.53	0.35	−4.38	5.77	.005
Type of measure	3.16	0.63	−1.38	0.32	−4.37	4.61	.009
Studies' methodological quality	1.29	0.31	0.59	1.00	0.59	2.77	.600[a]
Direct instruction	0.35	0.11	1.47	0.33	4.41	3.08	.021[a]
Type of interventionist	1.26	1.02	0.12	0.69	0.18	8.53	.865

Note. [a] signifies that the *p* value from robust variance estimation cannot be accurately computed because the associated moderator has less than 4 degrees of freedom.

2006; Finnegan, 2012; Joseph, 2002; Lemons et al., 2015; Tucker Cohen et al., 2008). In several other studies, the participants' teachers, reading specialists or teaching assistants implemented the intervention (Ahlgrim-Delzell et al., 2014; Ahlgrim-Delzell et al., 2016; Burgoyne et al., 2012; Coyne et al., 2010; Lemons et al., 2012_1; Lemons et al., 2012_2). A moderator analysis was conducted to assess if the type of interventionist had an impact on the effect sizes. We dichotomized the type of interventionist into two categories: researcher or researcher affiliated professional vs. participants' teacher, reading specialist or teaching assistant. The type of interventionist was not a significant moderator: $t\,(8.5) = 0.18, p = .865$. Findings from the moderator analysis are summarized in Table 3.

3.8. Publication bias analyses

Publication bias refers to the tendency of studies that report small or non-significant effects to be underrepresented in the published literature. Since publication bias analyses cannot be performed with robust variance estimation, these analyses were conducted with the aggregated effect sizes for studies reporting several outcomes. Thus, publication bias analyses were performed on 13 effect sizes (one effect size per study).

As recommended by Borenstein et al. (2009), we plotted a funnel plot and conducted a Trim and Fill analysis. In the funnel plot (Fig. 3), the effect sizes are plotted on the *X* axis and the variances on the *Y* axis. The studies reviewed in our meta-analysis appear as filled circles. In the absence of a publication bias, the studies will be distributed symmetrically about the average effect size. This seemed to be approximately the case with our data. We conducted a Trim and Fill analysis to estimate the number of studies missing in order to reach symmetry and thus obtain a less biased estimate of the overall average effect size. Only one study was added in the funnel plot (open circle in Fig. 3). The Trim and Fill analysis revealed that with this adjustment, the estimate of the unbiased average effect size would be $g = 1.50$, which is very close to the average effect size found in our meta-analysis ($g = 1.42$). These findings suggest that publication bias is probably not a large problem in this meta-analysis.

4. Discussion

The goal of this meta-analysis was to assess the effectiveness of phonics instruction for students with ID and to identify the instructional components and conditions of implementation that rendered it more effective. Eight single-case experimental studies and six group studies met the inclusion criteria, comprising a total of 297 participants with ID. Four of these studies met all the quality indicators of the CEC (Cook et al., 2014). Our findings suggest that there is an increase in the number of studies with high methodological quality conducted on the effects of phonics instruction on the decoding skills of students with ID since the review by

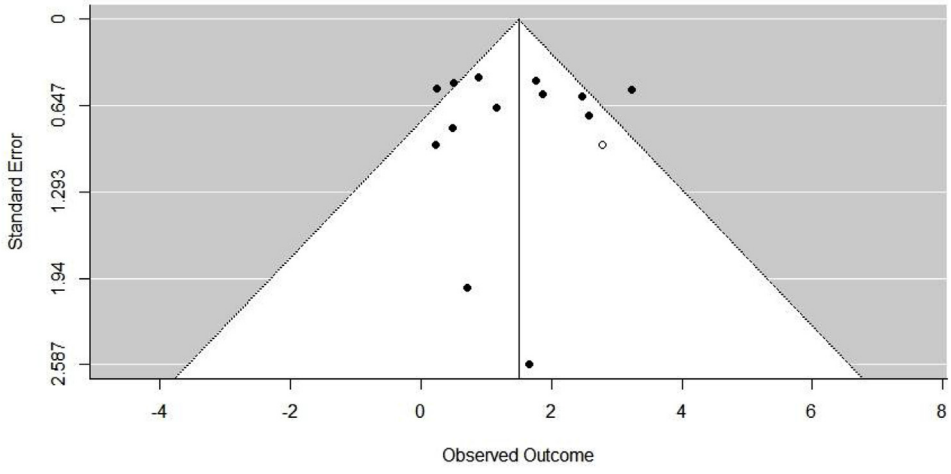

Fig. 3. Funnel plot with imputed studies (Trim and Fill analysis).

Joseph and Seery (2004). Furthermore, many of the studies conducted during these last two decades show an adequate consideration of the components of reading instruction described by the National Reading Panel and address their recommendations in the implemented interventions (NICHHD, 2000).

Phonics instruction was found to have a large overall effect on the decoding skills of students with ID (g = 1.42, 95% CI: 0.75, 2.10). Type of research design and type of measure were significant moderators. Type of interventionist was not a significant moderator. This suggests that teachers, reading specialists, or teaching assistants trained by researchers, can effectively implement systematic phonics instruction. In most of the studies, phonics instruction was carried out using a systematic and direct instruction approach and a one-to-one format. The number of studies with different instructional approaches and instructional formats (e.g., small-group format) was too small to conduct moderator analysis yielding reliable findings.

4.1. Is phonics instruction effective for persons with ID, taking into account the influence of the research design and the type of measure used?

Moderator analysis indicated that effect sizes varied depending on the study's research design. Single case studies yielded a larger average effect size (g = 1.94) than group studies (g = 0.41). This is not a surprising finding. Indeed, the effect sizes in single-case studies are known to be usually higher than in group studies with a control group (Durlak, 2009). The average effect size of phonics instruction found in this meta-analysis for group studies is similar to the average effect size of systematic phonics instruction reported in the meta-analysis of the NRP (d = 0.41) (NICHHD, 2000) and to the random-effects average effect size of phonics instruction reported in the meta-analysis by Torgersen et al. (ES = 0.38). It is also similar to the average effect size of phonics instruction on the reading skills of poor readers found in the meta-analysis conducted by Galuschka et al. (2014) (g = 0.32). It should be noted that even if the overall average effect size found for group studies in this meta-analysis (g = 0.41) would be considered as “small” according to the benchmarks provided by Cohen (1988), it does not mean that the effect is not important. Such “small” effect sizes have a practical value and indicate a meaningful improvement when they are found in group studies using standardized measures of academic achievement (Cooper, 2017; Durlak, 2009). Hattie (2009) considers that an average effect size in a meta-analysis equal or superior to 0.4 indicates that an intervention leads to an improvement in academic achievement that can be observed in real life.

Furthermore, effect sizes were on average lower when a standardized test was used to measure reading skills (g = 0.40) than when a researcher-designed test (g = 1.78) was used. In most of the studies reviewed, researcher-designed tests included only trained words. Only a few studies used a researcher-designed test that included transfer words (words with the same structure and letter-sound correspondences as the trained words) (Ahlgrim-Delzell et al., 2016; Finnegan, 2012; Tucker Cohen et al., 2008). In contrast, standardized tests of word and non-word reading are considered a measure of transfer performance because the students have to read non-words or words that are for the most part not trained during the intervention. The significant small average effect size found on reading skills measured with standardized tests (g = 0.40) suggests that after benefitting from phonics instruction, students with ID could transfer and generalize their decoding skills to untaught non-words and words.

4.2. What is the practical significance of the effect sizes found in the reviewed studies?

4.2.1. Single-case studies

In several single case studies, all the participants with mild to severe ID were able to read several trained words with a CV (consonant-vowel) and/or CVC (consonant-vowel-consonant) structure containing several letter-sound correspondences after 5 to 23 sessions of intervention (Ahlgrim-Delzell et al., 2014; Fallon et al., 2004; Joseph, 2002; Tucker Cohen et al., 2008). In the studies conducted by Lemons et al. (2012_1; 2012_2) the 11 participants, with the exception of one, were able to decode between 5 and 34 trained words after 12 to 35 sessions of intervention. In the study conducted by Lemons et al. (2015), the five participants, with the exception of one, were able to decode between 4 and 34 trained words after 24 to 45 sessions of intervention. In some of these studies, the participants already knew the letter-sound correspondences before the intervention (Fallon et al., 2004; Tucker Cohen et al., 2008). In other studies, the participants also learned several new letter-sound correspondences during the intervention (Ahlgrim-Delzell et al., 2014; Lemons et al., 2012_1; Lemons et al., 2012_2; Lemons et al., 2015).

Three single-case studies reviewed in this meta-analysis also measured if the students could read transfer/generalization words (words with the same structure and letter-sound correspondences) (Fallon et al., 2004; Tucker Cohen et al., 2008; Wolff Heller et al., 2002). Their findings suggest, that most of them needed, or would have needed, a little additional practice to generalize their decoding skills to at least 90% of novel words with the same structure and letter-sound correspondences (2–4 sessions in Tucker Cohen et al., 2008).

4.2.2. Group studies

Two group studies showed that phonics instruction allowed students with Down syndrome and adults with mild to moderate ID to read on average five to six new words of a standardized test after 5 months of daily interventions (Burgoyne et al., 2012) or 15 months of biweekly interventions (Cohen et al., 2006). One group study found that nonverbal students with ID were able to read on average nine new words of a researcher-designed test comprising trained and untrained words after 8 months of daily interventions (Ahlgrim-Delzell et al., 2016).

Finnegan (2012) found that students with ID made on average an improvement in word reading corresponding to a progress of one or two months in the age and grade equivalencies provided by the standardized test that they used. The students also made an improvement in non-word reading corresponding to a progress of approximately one school year after one to two months of intervention. With a standardized test used to monitor participants' progress in oral reading fluency when reading texts (a measure that

was not used as an outcome measure in this meta-analysis), Allor et al. (2010) found that half of the participants with mild or moderate ID benefiting from their intervention during two to three years attained or almost attained the benchmark score indicating that they were able to read connected text with the fluency expected at the end of first grade.

4.3. Can phonics instruction be considered an evidence-based practice for students with ID?

When using the findings reported in this meta-analysis to support conclusions about evidence-based practices, it seems important to take into account not only the strength of the effect sizes found but also the quality of evidence on which they are based. The CEC (Cook et al., 2014) recommends taking into account only studies meeting all the quality indicators when determining if a specific intervention can be considered evidence-based. Two of the experimental group studies (Ahlgrim-Delzell et al., 2016; Allor et al., 2010) (90 participants in total) and two of the single-case studies (Lemons et al., 2015; Tucker Cohen et al., 2008) (10 participants in total) reviewed in this meta-analysis met all the quality indicators. These four studies all documented positive effects (g = 0.43–3.24). Therefore, based on the criteria of the CEC (Cook et al., 2014), we can conclude from the current meta-analysis that there is enough evidence to consider phonics instruction as an evidence-based practice for teaching decoding skills to students with ID. However, the conditions under which the intervention was carried out should be taken into account. In these four studies, phonics was taught with a systematic and direct instruction approach. Interventions were implemented intensively (four times per week or daily). Moreover, they were conducted in a one-to-one format with the exception of one study, where instruction was also provided in small-group format (Allor et al., 2010).

4.4. Limitations and suggestions for future research

Although the results of the present meta-analysis have valuable implications for teachers and researchers, some limitations must be taken into account. The findings of this meta-analysis should be interpreted in the context of a potential publication bias (Sutton, Song, Gilbody, & Abrams, 2000; Thornton & Lee, 2000). Publication bias refers to the tendency of researchers and journal editors not to publish studies that fail to produce positive effects. Even if the findings from the publication bias analyses suggest that it is unlikely to be a large problem in this meta-analysis, a publication bias could still potentially have inflated the positive effects of phonics instruction found in this meta-analysis. Furthermore, the small number of studies included in this meta-analysis and their small sample sizes precluded more complex analysis (meta-regression) of the potential interaction effects of different moderators. If the number of high quality studies published on the effects of phonics instruction for students with ID continues to increase during the next decades, such thorough analysis may perhaps become possible. Lastly, combining single-case studies with group studies in a meta-analysis may be considered a limitation due to the differences in these two types of research designs. Our findings indicated that type of research design was a significant moderator, suggesting that readers should give more weight to the average effect sizes reported for these two types of designs separately, rather than to the summary effect size.

Several gaps in the existing research were identified, requiring caution in generalizing our conclusion about the effectiveness of systematic phonics instruction to all students with ID. First, only one group study meeting all the quality indicators of the CEC (Cook et al., 2014) found a positive effect of systematic phonics instruction on the decoding skills of students with ID with severe speech impairments (Ahlgrim-Delzell et al., 2016). According to the standards of the CEC, this is not enough to consider phonics instruction as an evidence-based practice for this group of children. Nevertheless, this type of intervention clearly seems promising in teaching decoding skills to students with ID and severe speech impairments. More high-quality studies should be conducted with this group of students.

Second, more studies are clearly needed on the effects of phonics instruction on the decoding skills of students with severe ID. Indeed, the majority of the participants in the studies reviewed were described as having mild or moderate ID. Participants with more severe limitations in intellectual functioning seemed to be scarce in the samples of the reviewed studies. However, it is difficult to know with certainty because most IQ-tests have a floor effect (Minimum IQ = 40 or 45). Future studies should also assess students' adaptive behavior in order to allow the reader to evaluate more clearly the severity of the participants' limitations and to know to which students their findings can be generalized.

Third, studies conducted with adults with ID are lacking. Only one study was conducted with adults (Cohen et al., 2006). High-quality studies are required to determine under what conditions systematic phonics instruction could lead adults with ID to become efficient decoders. Indeed, many adults with moderate ID probably did not benefit from such an instruction in the past (Ahlgrim-Delzell & Rivera, 2015). Because, lifelong learning is currently stressed as a right for persons with disabilities (United Nations, 2006), and because of the impact of literacy on everyday life, this seems to be an important area of investigation.

Fourth, most studies used a one-to-one instructional format. Because teaching students with ID in a one-to-one format sufficiently frequently and intensively may not always be possible in self-contained classrooms comprising several students with ID, more studies are clearly needed to verify if systematic and intensive phonics instruction is also efficient when provided in a small-group format.

Fifth, the majority of the studies were conducted with English-speaking students. Only one study was conducted with French-speaking adults with ID (Cohen et al., 2006). As the spoken language impacts the way reading is learned (Ziegler et al., 2010), having data collected in other languages may help to better identify the conditions that render phonics instruction effective for students with ID in a more situated way.

4.5. Implications for practice

Some caution is required in identifying implications for practice due to the relatively small number of studies reviewed in this meta-analysis. Nevertheless, our findings suggest that a systematic approach in teaching phonics to students with ID should be adopted. In other words, teachers should teach them letter-sound correspondences explicitly, systematically, and in a planned sequence taking into account their complexity and frequency, as is recommended for typically developing children (Castles et al., 2018; NICHHD, 2000).

Additionally, a direct instruction approach should also be favored. This approach stresses the importance of modeling or demonstrating the required skills and then providing guided practice with feedback (Rupley, Blair, & Nichols, 2009). Using systems of prompts (constant time delay, system of least-to-most prompts or most-to-least prompts) or using at the least systematic corrective procedures, also seemed to be a useful strategy for teaching decoding skills to students with ID efficiently.

Lastly, an examination of the practical significance of the effect sizes found after several months of intervention suggests that patience and perseverance are required when teaching decoding skills to students with ID. Nonetheless, it is worth the effort. Indeed, developing decoding skills allows students to read new words and to learn new vocabulary in an autonomous way. This skill is also necessary to store words in memory, and this storage is necessary to read with speed and accuracy. Fluency cannot be taught if children do not minimally automatize letter-sound correspondences. Similarly, satisfactory decoding skills are also crucial for reading comprehension. Thus, the benefits of systematic phonics programs for children, youths and adults with ID could be very important in terms of self-esteem, social participation and quality of life.

4.6. Conclusion

Systematic phonics instruction is effective to teach decoding skills to students with ID, as it is for typically developing children. Although learning to decode efficiently can take much more time for students with ID, especially for students with moderate or severe ID, they can learn decoding skills if they are provided with intensive systematic phonics instruction. Nevertheless, teaching reading skills to students with ID, especially students with moderate and severe ID, is very challenging. Despite the advances in knowledge observed in this meta-analysis, more studies are clearly needed to refine our understanding of the instructional characteristics and strategies that can render systematic phonics instruction the most effective for students with ID with different profiles, for example, depending on their age, language, communication skills, levels in reading skills or severity of limitations.

Acknowledgements

We sincerely thank the three anonymous reviewers whose suggestions helped improve this manuscript. This work was supported by the Swiss National Science Foundation [grant number 100019_173096].

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.edurev.2019.01.001.

References

Ahlgrim-Delzell, L., Browder, D., & Wood, L. (2014). Effects of systematic instruction and an augmentative communication device on phonics skills acquisition for students with moderate intellectual disability who are nonverbal. * *Education and Training in Autism and Developmental Disabilities, 49*(4), 517–532.

Ahlgrim-Delzell, L., Browder, D. M., Wood, L., Stanger, C., Preston, A. I., & Kemp-Inman, A. (2016). Systematic instruction of phonics skills using an iPad for students with developmental disabilities who are AAC Users. * *The Journal of Special Education, 50*(2), 86–97. https://doi.org/10.1177/0022466915622140.

Ahlgrim-Delzell, L., & Rivera, C. (2015). A content comparison of literacy lessons from 2004 and 2010 for students with moderate and severe intellectual disability. *Exceptionality, 23*(4), 258–269. https://doi.org/10.1080/09362835.2015.1064417.

Al Otaiba, S., & Hosp, M. K. (2004). Providing effective literacy instruction to students with Down syndrome. *Teaching Exceptional Children, 36*(4), 28–35.

Allor, J. H., Gifford, D. B., Al Otaiba, S., Miller, S. J., & Cheatham, J. P. (2013). Teaching students with intellectual disability to integrate reading skills. *Remedial and Special Education, 34*(6), 346–356. https://doi.org/10.1177/0741932513494020.

Allor, J. H., Mathes, P. G., Roberts, J. K., Cheatham, J. P., & Champlin, T. M. (2010). Comprehensive reading instruction for students with intellectual disabilities: Findings from the first three years of a longitudinal study. * *Psychology in the Schools, 47*, 445–466. https://doi.org/10.1002/pits.20482.

Allor, J. H., Mathes, P. G., Roberts, J. K., Cheatham, J. P., & Otaiba, S. A. (2014). Is scientifically based reading instruction effective for students with below-average IQs? *Exceptional Children, 80*(3), 287–306. https://doi.org/10.1177/0014402914522208.

Allor, J. H., Mathes, P. G., Roberts, J. K., Jones, F. G., & Champlin, T. M. (2010). Teaching students with moderate intellectual disabilities to read: An experimental examination of a comprehensive reading intervention. *Education and Training in Autism and Developmental Disabilities, 45*(1), 3–22. https://doi.org/10.1177/0741932513494020.

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders (DSM-5).* Arlington: American Psychiatric Association.

Bailey, R. L., Angell, M. E., & Stoner, J. B. (2011). Improving literacy skills in students with complex communication needs who use Augmentative/Alternative Communication Systems. *Education and Training in Autism and Developmental Disabilities, 46*(3), 352–368.

Barker, R. M., Sevcik, R. A., Morris, R. D., & Romski, M. (2013). A model of phonological processing, language, and reading for students with mild intellectual disability. *American Journal on Intellectual and Developmental Disabilities, 118*(5), 365–380. https://doi.org/10.1352/1944-7558-118.5.365.

Basil, C., & Reyes, S. (2003). Acquisition of literacy skills by children with severe disability. *Child Language Teaching and Therapy, 19*(1), 27–48. https://doi.org/10.1191/0265659003ct242oa.

Beecher, L., & Childre, A. (2012). Increasing literacy skills for students with intellectual and developmental disabilities: Effects of integrating comprehensive reading instruction with sign language. *Education and Training in Autism and Developmental Disabilities, 47*(4), 487–501.

Blachman, B. A., Ball, E. W., Black, R., & Tangel, D.,M. (2000). *Road to the code. A phonological awareness program for young children.* Baltimore: Paul H. Brookes Publishing Co.

Blachman, B. A., & Tangel, D.,M. (2008). *Road to reading. A program for preventing and remediating reading difficulties.* Baltimore: Paul H. Brookes Publishing Co.
Bochner, S., Outhred, L., & Pieterse, M. (2001). A study of functional literacy skills in young adults with Down syndrome. *International Journal of Disability, Development and Education, 48*(1), 67–90. https://doi.org/10.1080/10349120120036314.
Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2009). *Introduction to Meta-Analysis.* Chichester: John Wiley and Sons Ldt.
Boyle, J. R., & Walker-Seibert, T. (1997). The effects of a phonological awareness strategy on the reading skills of children with mild disabilities. *Learning Disabilities: A Multidisciplinary Journal, 8*(3), 145–153.
Bradford, S., Shippen, M. E., Alberto, P., Houchins, D. E., & Flores, M. (2006). Using systematic instruction to teach decoding skills to middle school students with moderate intellectual disabilities. *Education and Training in Developmental Disabilities, 41*(4), 333–343.
Browder, D. M., Ahlgrim-Delzell, L., Courtade, G., Gibbs, S. L., & Flowers, C. (2008). Evaluation of the effectiveness of an early literacy program for students with significant developmental disabilities. *Exceptional Children, 75*(1), 33–52. https://doi.org/10.1177/001440290807500102.
Browder, D. M., Ahlgrim-Delzell, L., Flowers, C., & Baker, J. N. (2012). An evaluation of a multicomponent early literacy program for students with severe developmental disabilities. *Remedial and Special Education, 33*, 237–246.
Browder, D., Ahlgrim-Delzell, L., & Wood, L. (2015). *ERSB. Early reading skills builder.* Verona: Attainment Company.
Browder, D. M., Wakeman, S. Y., Spooner, F., Ahlgrim-Delzell, L., & Algozzinexya, B. (2006). Research on reading instruction for individuals with significant cognitive disabilities. *Exceptional Children, 72*(4), 392–408. https://doi.org/10.1177/001440290607200401.
Burgoyne, K., Duff, F. J., Clarke, P. J., Buckley, S., Snowling, M. J., & Hulme, C. (2012). Efficacy of a reading and language intervention for children with Down syndrome: a randomized controlled trial. * *Journal of Child Psychology and Psychiatry, 53*(10), 1044–1053. https://doi.org/10.1111/j.1469-7610.2012.02557.x.
Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest, 19*(1), 5–51. https://doi.org/10.1177/1529100618772271.
Cleave, P. L., Kay-Raining Bird, E., & Bourassa, D. C. (2011). Developing phonological awareness skills in children with Down syndrome. *Canadian Journal of Speech-Language Pathology & Audiology, 35*(4), 332–343.
Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* New York: Academic Press.
Cohen, D., Plaza, M., Perez-Diaz, F., Lanthier, O., Chauvin, D., Hambourg, N., & Rivière, J. P. (2006). Individual cognitive training of reading disability improves word identification and sentence comprehension in adults with mild mental retardation. * *Research in Developmental Disabilities, 27*(5), 501–516. https://doi.org/10.1016/j.ridd.2004.07.008.
Coleman-Martin, M. B., Heller, K. W., Cihak, D. F., & Irvine, K. L. (2005). Using computer-assisted instruction and the Nonverbal Reading Approach to teach word identification. *Focus on Autism and other developmental disabilities, 20*(2), 80–90. https://doi.org/10.1177/10883576050200020401.
Conners, F. A., Rosenquist, C. J., Sligh, A. C., Atwell, J. A., & Kiser, T. (2006). Phonological reading skills acquisition by children with mental retardation. *Research in Developmental Disabilities, 27*(2), 121–137. https://doi.org/10.1016/j.ridd.2004.11.015.
Cook, B., Buysse, V., Klingner, J., Landrum, T., McWilliam, R., Tankersley, M., et al. (2014). Council for Exceptional Children: Standards for evidence-based practices in special education. *Teaching Exceptional Children, 46*(6), 206–212.
Cooper, H. (2017). *Research synthesis and meta-analysis. A step-by-step approach.* Los Angeles: SAGE.
Coyne, P., Pisha, B., Dalton, B., Zeph, L. A., & Smith, N. C. (2010). Literacy by Design: A Universal Design for Learning approach for students with significant intellectual disabilities. * *Remedial and Special Education, 33*(3), 162–172. https://doi.org/10.1177/0741932510381651.
Downing, J. E. (2010). *Academic instruction for students with moderate and severe intellectual disabilities in inclusive classrooms.* London: SAGE.
Durlak, J. A. (2009). How to select, calculate, and interpret effect sizes. *Journal of Pediatric Psychology, 34*(9), 917–928. https://doi.org/10.1093/jpepsy/jsp004.
Edgin, J. O. (2013). Cognition in Down syndrome: a developmental cognitive neuroscience perspective. *Wiley Interdisciplinary Reviews: Cognitive Science, 4*(3), 307–317. https://doi.org/10.1002/wcs.1221.
Eikeseth, S., & Jahr, E. (2001). The UCLA reading and writing program: an evaluation of the beginning stages. *Research in Developmental Disabilities, 22*(4), 289–307. https://doi.org/10.1016/S0891-4222(01)00073-7.
Fallon, K. A., Light, J., McNaughton, D., Drager, K., & Hammer, C. (2004). The effects of direct instruction on the single-word reading skills of children who require augmentative and alternative communication. * *Journal of Speech, Language, and Hearing Research, 47*(6), 1424–1439. https://doi.org/10.1044/1092-4388(2004/106.
Finnegan, E. G. (2012). Two approaches to phonics instruction: Comparison of effects with children with significant cognitive disability. * *Education and Training in Autism and Developmental Disabilities, 47*(3), 269–279.
Fisher, Z., & Tipton, E. (2015). *Robumeta: An R-package for robust variance estimation in meta-analysis.* Retrieved from https://arxiv.org/pdf/1503.02220.pdf.
Flores, M. M., Shippen, M. E., Alberto, P., & Crowe, L. (2004). Teaching letter-sound correspondence to students with moderate intellectual disabilities. *Journal of Direct Instruction, 4*(2), 173–188.
Forts, A.,M., & Luckasson, R. (2011). Reading, writing, and friendship: Adult implications of effective literacy instruction for students with intellectual disability. *Research and Practice for Persons with Severe Disabilities, 36*(3–4), 121–125. https://doi.org/10.2511/027494811800824417.
Fredrick, L. D., Davis, D. H., Alberto, P. A., & Waugh, R. E. (2013). From initial phonics to functional phonics: Teaching word-analysis skills to students with moderate intellectual disability. *Education and Training in Autism and Developmental Disabilities, 48*(1), 49–66.
Galuschka, K., Ise, E., Krick, K., & Schulte-Körne, G. (2014). Effectiveness of treatment approaches for children and adolescents with reading disabilities: A meta-analysis of randomized controlled trials. *PLoS One, 9*(2), e89900. https://doi.org/10.1371/journal.pone.0089900.
Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C. R., & Innocenti, M. S. (2005). Quality indicators for group experimental and quasi- experimental research in special education. *Exceptional Children, 71*, 149–164.
Goetz, K., Hulme, C., Brigstocke, S., Carroll, J. M., Nasir, L., & Snowling, M. (2007). Training reading and phoneme awareness skills in children with Down syndrome. *Reading and Writing, 21*(4), 395–412. https://doi.org/10.1007/s11145-007-9089-3.
Grindle, C. F., Hughes, C. J., Saville, M., Huxley, K., & Hastings, R. P. (2013). Teaching early reading skills to children with autism using MimioSprout Early Reading. *Behavioral Interventions, 28*(3), 203–224.
Hanser, G. A., & Erickson, K. A. (2007). Integrated word identification and communication instruction for students with complex communication needs: Preliminary results. *Focus on Autism and other developmental disabilities, 22*(4), 268–278. https://doi.org/10.1177/10883576070220040901.
Hattie, J. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement.* New York: Routledge.
Hedges, L. V., Pustejovsky, J. E., & Shadish, W. R. (2012). A standardized mean difference effect size for single case designs. *Research Synthesis Methods, 3*(3), 224–239. https://doi.org/10.1002/jrsm.1052.
Hedges, L. V., Pustejovsky, J. E., & Shadish, W. R. (2013). A standardized mean difference effect size for multiple baseline designs across individuals. *Research Synthesis Methods, 4*(4), 324–341. https://doi.org/10.1002/jrsm.1086.
Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods, 1*(1), 39–65. https://doi.org/10.1002/jrsm.5.
Hill, D. R. (2016). Phonics based reading interventions for students with intellectual disability: A systematic literature review. *Journal of Education And Training Studies, 4*(5), 205–214. https://doi.org/10.11114/jets.v4i5.1472.
Horner, R., Carr, E., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children, 71*(2), 165–179. https://doi.org/10.1177/001440290507100203.
Jamshidi, L., Heyvaert, M., Declercq, L., Fernández-Castilla, B., Ferron, J. M., Moeyaert, M., & Van den Noortgate, W. (2018). Methodological quality of meta-analyses of single-case experimental studies. *Research in Developmental Disabilities, 79*, 97–115. https://doi.org/10.1016/j.ridd.2017.12.016.
Joseph, L. M. (2002). Facilitating word recognition and spelling using word boxes and word sort phonic procedures. * *School Psychology Review, 31*(1), 122–129.
Joseph, L. M., & Seery, M. E. (2004). Where is the phonics?: A review of the literature on the use of phonetic analysis with students with mental retardation. *Remedial and Special Education, 25*(2), 88–94. https://doi.org/10.1177/07419325040250020301.
Klusek, J., Hunt, A. W., Mirrett, P. L., Hatton, D. D., Hooper, S. R., Roberts, J. E., et al. (2015). Reading and phonological skills in boys with Fragile X syndrome. *Journal*

of Autism and Developmental Disorders, 45(6), 1699–1711. https://doi.org/10.1007/s10803-014-2328-y.
Laing, E., Hulme, C., Grant, J., & Karmiloff-Smith, A. (2001). Learning to read in Williams syndrome: Looking beneath the surface of atypical reading development. *Journal of Child Psychology and Psychiatry, 42*(6), 729–739. https://doi.org/10.1111/1469-7610.00769.
Lemons, C. J., & Fuchs, D. (2010). Phonological awareness of children with Down syndrome: Its role in learning to read and the effectiveness of related interventions. *Research in Developmental Disabilities, 31*(2), 316–330. https://doi.org/10.1016/j.ridd.2009.11.002.
Lemons, C. J., King, S. A., Davidson, K. A., Puranik, C. S., Fulmer, D., Mrachko, A. A., & Fidler, D. J. (2015). Adapting phonological awareness interventions for children with Down Syndrome based on the behavioral phenotype: A promising approach? * *Intellectual and Developmental Disabilities, 53*(4), 271–288. https://doi.org/10.1352/1934-9556-53.4.271.
Lemons, C. J., Mrachko, A. A., Kostewicz, D. E., & Paterra, M. F. (2012). Effectiveness of decoding and phonological awareness interventions for children with Down syndrome. * *Exceptional Children, 79*(1), 67–90. https://doi.org/10.1177/001440291207900104.
Lemons, C. J., Zigmond, N., Kloo, A. M., Hill, D. R., Mrachko, A. A., Paterra, M. F., ... Davis, S. M. (2013). Performance of students with significant cognitive disabilities on early-grade curriculum-based measures of word and passage reading fluency. *Exceptional Children, 79*(4), 408–426. https://doi.org/10.1177/001440291307900402.
Levy, Y., Smith, J., & Tager-Flusberg, H. (2003). Word reading and reading-related skills in adolescents with Williams syndrome. *Journal of Child Psychology and Psychiatry, 44*(4), 576–587. https://doi.org/10.1111/1469-7610.00146.
Machalicek, W., Sanford, A., Lang, R., Rispoli, M., Molfenter, N., & Mbeseha, M. K. (2010). Literacy interventions for students with physical and developmental disabilities who use aided AAC devices: A systematic review. *Journal of Developmental and Physical Disabilities, 22*(3), 219–240. https://doi.org/10.1007/s10882-009-9175-3.
Mathes, P. G., & Torgesen, J. K. (2005). *Early interventions in reading.* Columbus: SRA/McGraw-Hill.
McArthur, G. M., Castles, A., Kohnen, S., Larsen, L., Jones, K., Anandakumar, T., et al. (2015). Sight word and phonics training in children with dyslexia. *Journal of Learning Disabilities, 24*, 391–407. https://doi.org/10.1177/0022219413504996.
Menghini, D., Verucci, L., & Vicari, S. (2004). Reading and phonological awareness in Williams syndrome. *Neuropsychology, 18*(1), 29–37. https://doi.org/10.1037/0894-4105.18.1.29.
Morris, S. B. (2008). Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods, 11*(2), 364–386. https://doi.org/10.1177/1094428106291059.
Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods, 7*(1), 105.
National Institute of Child Health and Human Development [NICHHD] (2000). *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction (NIH Publication No. 00-4769)*Washington, DC: Government Printing Office.
Ratz, C., & Lenhard, W. (2013). Reading skills among students with intellectual disabilities. *Research in Developmental Disabilities, 34*(5), 1740–1748. https://doi.org/10.1016/j.ridd.2013.01.021.
Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D., & Seidenberg, M. S. (2001). How psychological science informs the teaching of reading. *Psychological Science in the Public Interest, 2*(2), 31–74. https://doi.org/10.1111/1529-1006.00004.
Reed, D. K. (2013). The effects of explicit instruction on the reading performance of adolescent english language learners with intellectual disabilities. *Tesol Quarterly, 47*(4), 743–761. https://doi.org/10.1002/tesq.94.
Riepl, J. H., Marchand-Martella, N. E., & Martella, R. C. (2008). The effects of "Reading Mastery Plus" on the beginning reading skills of students with intellectual and developmental disabilities. *Journal of Direct Instruction, 8*(1), 29–39.
Roberts, C. A., Leko, M. M., & Wilkerson, K. L. (2013). New directions in reading instruction for adolescents with significant cognitive disabilities. *Remedial and Special Education, 34*(5), 305–317. https://doi.org/10.1177/0741932513485447.
Rupley, W. H., Blair, T. R., & Nichols, W. D. (2009). Effective reading instruction for struggling readers: The role of direct/explicit teaching. *Reading & Writing Quarterly, 25*(2/3), 125–138.
Saunders, K. J., & DeFulio, A. (2007). Phonological awareness and rapid naming predict word attack and word identification in adults with mild mental retardation. *American Journal on Mental Retardation, 112*(3), 155–166. https://doi.org/10.1352/0895-8017(2007)112[155:paarnp]2.0.co;2.
Schalock, R. L., Borthwick-Duffy, S. A., Bradley, V. J., Buntinx, W. H., Coulter, D. L., Craig, E. M. et al. Reeve, A. (2010). *Intellectual disability: Definition, classification, and systems of supports.* Washington: American Association on Intellectual and Developmental Disabilities.
Sermier Dessemontet, R., & de Chambrier, A.-F. (2015). The role of phonological awareness and letter-sound knowledge in the reading development of children with intellectual disabilities. *Research in Developmental Disabilities, 41–42*, 1–12. https://doi.org/10.1016/j.ridd.2015.04.001.
Shadish, W. R., Hedges, L. V., Horner, R. H., & Odom, S. L. (2015). *The role of between-case effect size in conducting, interpreting, and summarizing single-case research (NCER 2015-002).* National Center for Education Research.
Shadish, W. R., Hedges, L. V., Pustejovsky, J. E., Boyajian, J. G., Sullivan, K. J., Andrade, A., et al. (2014). A d-statistic for single-case designs that is equivalent to the usual between-groups d-statistic. *Neuropsychological Rehabilitation, 24*(3–4), 528–553.
Smith, J. D. (2012). Single-case experimental designs: A systematic review of published research and current standards. *Psychological Methods, 17*(4), 510–550 10.1037/a0029312. doi:10.1037/a0029312.
Soltani, A., & Roslan, S. (2013). Contributions of phonological awareness, phonological short-term memory, and rapid automated naming, toward decoding ability in students with mild intellectual disability. *Research in Developmental Disabilities, 34*(3), 1090–1099. https://doi.org/10.1016/j.ridd.2012.12.005.
Sutton, A. J., Song, F., Gilbody, S. M., & Abrams, K. R. (2000). Modelling publication bias in meta-analysis: a review. *Statistical Methods in Medical Research, 9*(5), 421–445. https://doi.org/10.1177/096228020000900503.
Swinehart-Jones, D., & Heller, K. W. (2008). Teaching students with severe speech and physical impairments a decoding strategy using internal speech and motoric indicators. *The Journal of Special Education, 43*(3), 131–144. https://doi.org/10.1177/0022466908314945.
Thornton, A., & Lee, P. (2000). Publication bias in meta-analysis. *Journal of Clinical Epidemiology, 53*(2), 207–216. https://doi.org/10.1016/S0895-4356(99)00161-4.
Tipton, E. (2015). Small sample adjustments for robust variance estimation with meta-regression. *Psychological Methods, 20*(3), 375–393. https://doi.org/10.1037/met0000011.
Tipton, E., & Pustejovsky, J. E. (2015). Small-sample adjustments for tests of moderators and model fit using robust variance estimation in meta-regression. *Journal of Educational and Behavioral Statistics, 40*(6), 604–634. https://doi.org/10.3102/1076998615606099.
Torgerson, C., Brooks, G., & Hall, J. (2006). *A systematic review of the research literature on the use of phonics in the teaching of reading and spelling (Research report RR711)* UK Department for Education and Skills. Retrieved from https://core.ac.uk/download/pdf/5911476.pdf.
Towles-Reeves, E., Kearns, J., Kleinert, H., & Kleinert, J. (2008). An analysis of the learning characteristics of students taking alternate assessments based on alternate achievement standards. *The Journal of Special Education, 42*(4), 241–254. https://doi.org/10.1177/0022466907313451.
Travers, J. C., Higgins, K., Pierce, T., Boone, R., Miller, S., & Tandy, R. (2011). Emergent literacy skills of preschool students with autism: A comparison of teacher-led and computer-assisted instruction. *Education and Training in Autism and Developmental Disabilities, 46*(3), 326–338.
Tripiana-Barbosa, A., & de Souza, S. R. (2015). A board game for the teaching, reading, and writing to intellectually disabled people. *Behaviour Analysis: Research and Practice, 15*(1), 90–106. https://doi.org/10.1037/h0101073.
Truxler, J. E., & O'Keefe, B. M. (2007). The effects of phonological awareness instruction on beginning word recognition and spelling. *Augmentative and Alternative Communication, 23*(2), 164–176. https://doi.org/10.1080/07434610601151803.
Tucker Cohen, E., Wolff Heller, K., Alberto, P., & Fredrick, L. D. (2008). Using a three-step decoding strategy with constant time delay to teach word reading to students with mild and moderate mental retardation. * *Focus on Autism and other developmental disabilities, 23*(2), 67–78. https://doi.org/10.1177/1088357608314899.
Tyler, E. J., Hughes, J. C., Wilson, M. M., Beverley, M., Hastings, R. P., & Williams, B. M. (2015). Teaching early reading skills to children with intellectual and developmental disabilities using computer-delivered instruction: A pilot study. *Journal of International Special Needs Education, 18*(1), 1–11. https://doi.org/10.

9782/2159-4341-18.1.1.
United Nations (2006). *Convention on the rights of persons with disabilities. New York: United Nations.* Retrieved from http://www.un.org/disabilities/documents/convention/convoptprot-e.pdf.
Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1–48.
Waugh, R. E., Fredrick, L. D., & Alberto, P. A. (2009). Using simultaneous prompting to teach sounds and blending skills to students with moderate intellectual disabilities. *Research in Developmental Disabilities, 30*(6), 1435–1447. https://doi.org/10.1016/j.ridd.2009.07.004.
Whalon, K. J., Al Otaiba, S., & Delano, M. E. (2009). Evidence-based reading instruction for individuals with autism spectrum disorders. *Focus on Autism and other developmental disabilities, 24*(1), 3–16. https://doi.org/10.1177/1088357608328515.
Wise, J. C., Sevcik, R. A., Romski, M., & Morris, R. D. (2010). The relationship between phonological processing skills and word and nonword identification performance in children with mild intellectual disabilities. *Research in Developmental Disabilities: A Multidisciplinary Journal, 31*(6), 1170–1175. https://doi.org/10.1016/j.ridd.2010.08.004.
Wolff Heller, K., Fredrick, L. D., Tumlin, J., & Brineman, D. G. (2002). Teaching decoding for generalization using the nonverbal reading approach. * *Journal of Developmental and Physical Disabilities, 14*(1), 19–35. https://doi.org/10.1023/a:1013559612238.
Zelinsky, N. A. M., & Shadish, W. (2016). A demonstration of how to do a meta-analysis that combines single-case designs with between-groups experiments: The effects of choice making on challenging behaviors performed by people with disabilities. *Developmental Neurorehabilitation,* 1–13. https://doi.org/10.3109/17518423.2015.1100690.
Ziegler, J. C., Bertrand, D., Toth, D., Csepe, V., Reis, A., Faisca, L., & Blomert, L. (2010). Orthographic depth and its impact on universal predictors of reading: A cross-language investigation. *Psychological Science, 21*, 551–559. https://doi.org/10.1177/0956797610363.