Article:

Jiang, Y. orcid.org/0000-0002-3805-0201, Gai, X. orcid.org/0000-0001-6836-5442, Üstün-Yavuz, M.S. orcid.org/0000-0002-0117-9044 et al. (2 more authors) (2023) A meta-analysis of training effects on English phonological awareness and reading in native Chinese speakers. PsyCh Journal, 12 (5). pp. 599-617. ISSN 2046-0252

https://doi.org/10.1002/pchj.675

Received: 6 July 2022 | Accepted: 30 June 2023

DOI: 10.1002/pchj.675

REVIEW ARTICLE

PsyCh Journal

A meta-analysis of training effects on English phonological awareness and reading in native Chinese speakers

Ying Jiang[1] | Xiaosong Gai[1] | Meryem S. Üstün-Yavuz[2] | Mingzhe Zhang[3] | Jenny M. Thomson[2]

[1]School of Psychology, Northeast Normal University, Changchun, China

[2]Division of Human Communication Sciences, Health Sciences School, University of Sheffield, Sheffield, UK

[3]Faculty of Psychology, Tianjin Normal University, Tianjin, China

Correspondence
Dr. Jenny M. Thomson, Division of Human Communication Sciences, Health Sciences School, University of Sheffield, 362 Mushroom Lane, Sheffield, S10 2TS, UK.
Email: j.m.thomson@sheffield.ac.uk

Abstract

Enhancing English phonological awareness is critical in promoting native English speakers' reading development. However, less attention has been paid to the role of phonological awareness development for English language learners in a logographic context. This meta-analysis aims to evaluate the effectiveness of training native Chinese speakers' English phonological awareness and reading across age groups. Thirty-three articles, including 37 independent samples, were identified as training studies that reported English phonological awareness as an outcome measure, and 16 articles, including 17 independent samples, featured training studies that reported reading as an outcome measure. Results based on a random-effect model revealed the effect sizes for overall English phonological awareness (including English syllable awareness, phoneme awareness, and rhyme awareness) and overall reading (including word reading and pseudoword reading) were $g = 0.651$ ($n = 3137$) and $g = 0.498$ ($n = 1506$), respectively. Specifically, instructional training exerted a small impact on word reading ($g = 0.297$), moderate effects on syllable awareness ($g = 0.468$) and pseudoword reading ($g = 0.586$), a medium to large effect on phoneme awareness ($g = 0.736$), and a large impact on rhyme awareness ($g = 0.948$). The moderator analyses yielded several significant findings. Regarding the English phonological awareness outcome, programs integrating lexical semantic knowledge exhibited the largest trend in enhancing native Chinese speakers' skills. Among all age groups, upper elementary students benefited most from instructional training. Furthermore, more intensive training had a greater impact than less intensive training. In terms of the reading outcome, similar to English phonological awareness findings, upper elementary students realized the greatest improvements. Additionally, unpublished articles indicated a larger training effect on reading than published ones. These findings provide practitioners with guidelines for delivering effective instruction to promote phonological awareness and reading ability for English language learners in a logographic language context.

KEYWORDS

native Chinese speakers, phonics, phonological awareness, training, word reading

"Phonological awareness" refers to an individual's ability to perceive and manipulate the sound structure of a spoken word. Research with native English speakers (NESs) indicates that phonological awareness is the main factor associated with successful reading performance, accounting for much of the unique variance in reading accuracy (Jongejan et al., 2007; Parrila et al., 2004). In addition, English phonological awareness (EPA) can help English learners (either with an alphabetic or logographic background) develop their letter knowledge, word reading, spelling, and oral vocabulary ability (Guan et al., 2006; Lin et al., 2017; Vadasy & Sanders, 2011). For native Chinese speakers (NCSs), EPA has been found to uniquely predict their real word reading ability after controlling for age, intelligence, cognitive ability, vocabulary, or Chinese character identification (Chung & Lam, 2020; Guan et al., 2006; B. L. Li et al., 2011).

However, NCSs' ability in EPA significantly lags behind that of NESs (Bialystok et al., 2003; Pan & Gai, 2013; Yeong et al., 2014). Pan and Gai (2013) revealed that NESs outperformed NCSs in all EPA subcomponents, including syllable awareness, phoneme awareness, and onset-rime awareness, with the performance of Chinese university students below that of native English children in primary school. Since phonological awareness plays a vital role in NCSs' development of word reading skills and NCSs' EPA levels are relatively low, it appears valuable to explore the impact of EPA training for NCSs.

TRAINING IN DEVELOPING ENGLISH LANGUAGE LEARNERS' PHONOLOGICAL AWARENESS AND READING

There is still a relative paucity of research concerning phonological awareness and reading training for English language learners (ELLs). Huo and Wang (2017) reviewed 15 studies investigating the effects of phonologically based instruction on ELLs. They found moderate effect sizes for phonemic awareness ($d = 0.62$) and non-word reading ($d = 0.55$) and a small effect size for word reading ($d = 0.33$). It should be noted that the studies reviewed were diverse, including data from various locations, such as Hong Kong, Malaysia, India, and Japan. The effect size of NCSs was not reported independently and was instead included in the overall analysis. In addition, Murphy Odo (2021) conducted a review of 45 studies on EPA and phonics instruction for ELLs and found no significant differences in word reading effects between ELLs with a logographic background ($g = 0.43$) and those with an alphabetic one ($g = 0.39$). The effect size of pseudoword reading was larger for ELLs with a logographic background ($g = 0.82$) than for those with an alphabetic one ($g = 0.52$). These studies demonstrated the efficacy of phonological training in enhancing the EPA and reading skills of ELLs, but they only included studies with direct or explicit training strategies. Also, no meta-analyses have yet synthesized evidence concerning the effectiveness of NCSs' EPA and reading training. The present study aims to cover a broader variety of research related to improving EPA and reading ability for NCSs beyond direct training approaches. Such an analysis will help us understand the nature of English phonologically based training within a logographic language context and the relative distinctness and effectiveness of this training compared to those in other native language contexts.

To comprehensively assess the training effects of NCSs' EPA and reading ability, it is crucial to examine phoneme, syllable, and rhyme awareness separately in addition to overall EPA ability. Pan et al. (2019) conducted a meta-analysis to investigate the relationship between EPA and reading success (at the word, sentence, and text levels) for NCSs. They found that English phoneme awareness was the most robust correlate of individual differences in NCSs' reading competence, followed by English onset-rime awareness and syllable awareness. What is more, Tao et al. (2007) reported that English onset-rime awareness in NCS elementary students uniquely predicted English word reading and pseudoword reading, while phoneme awareness did not. Other research has shown that English phoneme awareness, syllable awareness, and rime awareness are significant independent predictors for NCSs' word reading development (S. S. Yeung & Chan, 2013; S. S. Yeung & Ganotice, 2014), suggesting their distinct roles in NCSs' word reading skills. Therefore, a thorough understanding of the effects of training on NCSs' EPA and reading ability necessitates the evaluation of all aspects of EPA.

Additionally, it is essential to evaluate real word and pseudoword reading as separate outcome measures since they require different cognitive processes and carry distinct instructional implications. Pseudowords lack meaning (Z. Zhang & Peng, 2022), and reading pseudowords involves additional phonological demands that sight words do not have (McWeeny et al., 2022). By separating these two measures, researchers can assess different aspects of reading ability that may relate to diverse skills or strategies. Recent review studies on ELLs have consistently found that phonologically based training programs have a greater impact on the ability to read pseudowords than real words (Huo & Wang, 2017; Murphy Odo, 2021). These findings imply that participants who receive EPA and phonics training may experience greater improvements in decoding skills, which are essential for deciphering unfamiliar words.

TRAINING STRATEGIES AND TRAINING PROGRAMS IN NATIVE CHINESE SPEAKERS

A wide range of training programs in improving EPA and reading skills have been carried out within NCS populations, with a distinction made here between those that either directly or indirectly teach EPA.

Within direct EPA training strategies, EPA is taught either alone or in combination with other skills, such as phonics and vocabulary. Studies conducted for NCSs have categorized four types of direct instructional programs. The first type of program is known as the "EPA only" program, which explicitly teaches an individual's sound awareness at the syllable, onset-rime/rhyme, or phoneme level of a word solely through oral instruction without any print exposure. To give an example, researchers have directly trained EPA alone during instruction for NCSs (H. Y. Wang, 2012; Z. M. Xu, 2013). EPA components, such as syllable, rhyme, and phoneme awareness, can be practiced via the explicit training processes of blending, segmentation, or deletion. The second type of program is referred to as "EPA+phonics," which involves both EPA training and instruction in letter–sound associations. For example, 12-year-old students in primary school received instruction about single and multiple phonographs for several weeks, followed by training on multiple processing skills of phonemic awareness (Xiong, 2014). The primary distinction between the first two programs lies in whether the EPA training is conducted with or without exposure to written letters. Previous meta-analyses of a series of English-language studies revealed that programs combining phonological training with written letters or words might be more effective in word reading than purely phonetic programs (Bus & Van Ijzendoorn, 1999; Ehri et al., 2001).

This suggests that children need additional support for linking phonological processes to reading. However, the majority of these studies focused on readers with alphabetic contexts; whether this conclusion can be applied to NCSs across age groups is less known. The third type is the "EPA and/or phonics+semantics" program, wherein the instruction of lexical semantic knowledge is integrated with either EPA training or a combination of EPA and phonics skills. For instance, Chinese kindergarteners were first taught vocabulary knowledge in meaningful contexts, followed by the instruction of the EPA skills of the trained vocabulary (S. S. S. Yeung et al., 2013). Recent research has demonstrated that English vocabulary is a unique predictor of NCS children's English word reading (Y. Liu et al., 2017). However, the lack of sound–meaning correspondence may hinder NCSs' enhancement from phonics training. Castles et al. (2018) have argued that in alphabetic systems, the phonemes of the language are represented by letters or groups of letters. If a child learns to decode that symbol-to-sound relationship, then that child will have the ability to translate printed words into spoken language, thereby accessing information about meaning. Nonetheless, even if NCS children manage to decode new words with the letter–sound correspondence rules through phonics instruction, the connections of the sound, form, and meaning of words cannot be activated. As a result, the benefits of phonics pedagogy for NCSs may be limited. Therefore, the inclusion of lexical semantic knowledge may be particularly pivotal for NCSs' reading development. The fourth type is referred to as the "EPA and/or phonics+other contents" program, which integrates EPA and/or phonics training with other practical lessons, such as the International Phonetic Alphabet, Chinese characters, Chinese phonological awareness, or reading comprehension. For instance, Siu et al. (2018) investigated whether metalinguistic and working memory training affected Chinese children's reading gains. In an 8-week training program, both EPA and phonics knowledge of Chinese and English were taught. The present study aims to investigate whether programs combining EPA training with written letters or words may be more effective in word reading than purely EPA programs and whether training involving lexical semantic knowledge leads to greater effects compared with other types of training programs.

Within indirect EPA training strategies, EPA (as an outcome measure) was not explicitly or directly taught. Rather, alternative training programs were implemented, such as viewing cartoon videos, receiving vocabulary instruction, engaging in parent–child reading, or taking part in interactive dialog activities (Chow et al., 2021; Chow et al., 2010; Y. Li, 2012; S. S. Yeung et al., 2020; C. Y. Zhang et al., 2010). Indirect training strategies have also been shown as effective in improving NCSs' EPA and reading skills (Y. Li, 2012; S. S. Yeung et al., 2020; C. Y. Zhang et al., 2010). For example, Zhang et al. (2010) simply arranged for Chinese kindergarteners with a mean age of 66 months to watch English-language Disney cartoons for 20 min a day, 5 days a week, over an 18-week period. As a result of this intervention, the children's EPA ability (assessed by the authors as an outcome measure) was significantly enhanced. Other researchers implemented vocabulary instruction in a story-based context among Hong Kong kindergarteners and found that while EPA was not a direct instructional target, the participants gained greatly not only in their receptive and expressive general vocabulary knowledge but also in their phonemic awareness skills (S. S. Yeung et al., 2020).

The effectiveness of direct and indirect training programs has been compared by scholars. Sermier Dessemontet et al. (2019) conducted a meta-analysis on phonics instruction for individuals with intellectual disabilities and reported that direct training yielded stronger effect sizes than indirect training. But this meta-analysis did not report the participants' native language in the studies analyzed. Hence, it is unclear whether direct and indirect training strategies impact NCSs' EPA training outcome. Investigating this issue is of great significance since most EPA/phonics and reading programs targeting NESs or other native language contexts are explicit and have been found effective (Bus & Van Ijzendoorn, 1999; Ehri et al., 2001; Rehfeld et al., 2022). Indirect training programs may also offer great benefits to NCSs by providing them with the necessary contextual knowledge. If the indirect strategy is found to be effective, it is recommended that educators place greater emphasis on utilizing such strategies in future English teaching for NCSs.

ADDITIONAL FACTORS INFLUENCING TRAINING EFFICACY

Language proficiency is a crucial factor in evaluating the language skills of ELLs (Goetze & Driver, 2022; Mohsen, 2022). In this study, we aim to evaluate language proficiency by using "sample characteristics" and "grade" as moderating variables to examine their impact on EPA and reading training for NCSs. Previous training programs for NCSs have primarily focused on the typically developing population (Y. Li, 2012; Sun et al., 2015). Only a few studies have investigated special populations, including individuals with dyslexia, at-risk readers, and individuals with learning disabilities impacting English language learning (Tam & Leung, 2019; Xiong, 2014; S. S. Yeung & Savage, 2020). Examining the difference in training effects between a typically developing population and the special population of NCSs is helpful for educators in adopting more targeted teaching.

The effects of phonologically based training programs differ widely in terms of grade. Meta-analyses, primarily for children with alphabetic contexts, consistently indicate that the effects are larger when training programs are implemented early (Bus & Van Ijzendoorn, 1999; Ehri et al., 2001). This is likely because the window of benefit for phonological training is shorter for English-speaking children who typically learn to read English at a young age. Another meta-analysis for ELLs extended the sample range of EPA and phonics training to middle schools and demonstrated that the effect sizes of word reading for middle school students are larger than those in the elementary period (Murphy Odo, 2021). However, the influence of training effects on ELLs with a logographic language background is unclear. For NCSs, who may be exposed to English at varying ages, the benefits of phonological training may manifest a slightly longer time window of efficacy. Moreover, NCSs' EPA skills were found to be significantly lower than those of NESs across age groups (Pan & Gai, 2013). Hence, it appears necessary to examine the training effects of NCSs in different age groups.

Instructional intensity is another critical factor related to training effects. The American National Reading Panel's meta-analysis reported that when phonemic awareness instruction lasts between 5 and 18 hours in total for children, it yields the best effect (Ehri et al., 2001). Murphy Odo (2021) found that the more time devoted to ELLs phonics instruction, the larger the effect of training on English word reading. And the largest mean effect size of pseudoword reading was observed when ELLs were taught for between 1000 and 2000 min. Thus, when assessing additional factors related to training effects, it is essential to consider the intensity of training, including the number of weeks and the distribution of sessions.

Finally, to investigate potential publication bias, we will examine the impact of peer-reviewed articles and unpublished articles on overall EPA and overall reading competence.

THE PRESENT RESEARCH

This study aims to examine the training effects on EPA and reading in NCSs across age groups. The following research questions are addressed.

1. What is the impact of EPA and reading training programs on the overall development of EPA, including its components (syllable awareness, rhyme awareness, and phoneme awareness), as well as overall reading, including its components (word reading and pseudoword reading), among NCSs?
2. To what degree do training strategies (direct and indirect) and type of program (EPA only, EPA+phonics, EPA and/or phonics+semantics, and EPA and/or phonics+other contents) result in different effect sizes on NCSs' EPA and word reading?
3. How do sample characteristics, grade, and training intensity influence effect sizes on NCSs' EPA and word reading?

METHOD

Literature search

This meta-analysis was conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Liberati et al., 2009; see Figure 1). In April 2022, an electronic search of four databases was conducted: Scopus, PsycInfo, Web of Science, and China National Knowledge Infrastructure (CNKI).

We used different search strategies in three English databases and one Chinese database. In the three English databases, we chose three subjects ("phonological awareness," "Chinese,"

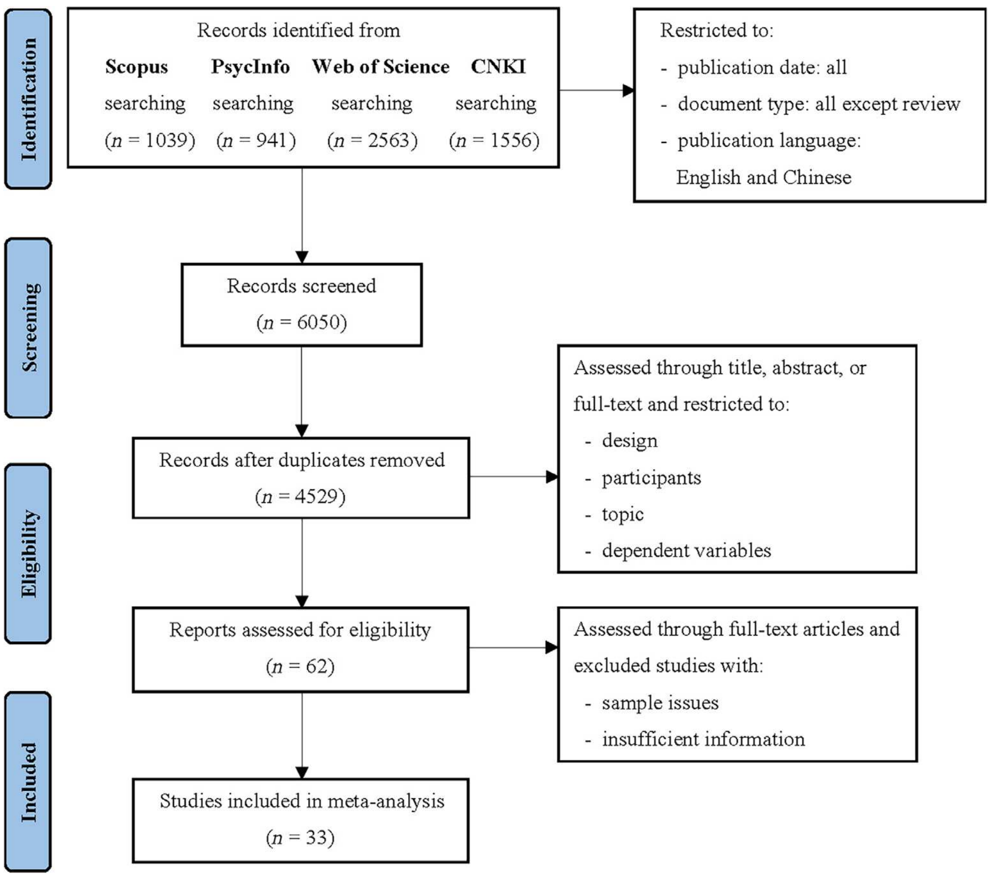

FIGURE 1 PRISMA diagram of study selection process.

and "training"), thus creating separate search strings. We also located their alternatives to have a comprehensive result. For the first subject, "phonological awareness," we searched *phonological sensitivity, phonological processing, phonological recoding, phonological skills, phonem* awareness, metalinguistic awareness, phonic*, letter-sound, grapheme-phoneme,* and *phoneme-grapheme*. We combined them with the Boolean operator OR. Similarly, for the second subject, "Chinese," we searched *China, Mandarin*, Cantonese, Hong Kong, Taiwan*, L2, second language, language learner*, English learner*, ESL, EFL, ELL,* and *bilingual**. For the third subject, "training," we searched *interven*, teach*, instruct*, pedagog*, treat*, remedia*, program*, strateg** and *effect**. As a last step, the three separate searches were conducted on the title, abstract, keywords, or subject in each database and combined with the Boolean operator AND. We limited publication languages to English and Chinese. Articles from peer-reviewed journals, book chapters, reports, non-peer-reviewed journals, dissertations, and conference proceedings were also collected. This yielded 1039 articles in Scopus, 941 articles in PsycInfo, and 2563 articles in Web of Science.

We conducted two steps in the Chinese database, CNKI, based on "subject" searching. First, in the advanced search of CNKI, we restricted to Chinese literature with "journal type." We separately searched 21 subject words or phrases in Chinese (*英语语音意识English phonological awareness, 语音意识训练phonological training, 语音意识干预phonological awareness intervention, 语音意识教学phonological awareness instruction, 音位意识phoneme awareness, 首音意识onset awareness, 韵脚意识rime awareness, 押韵意识rhyme awareness, 英语韵律English prosody, 韵律意识prosody awareness, 重音意识stress awareness, 语调意识intonation awareness, 节奏意识rhythm awareness, 元语言意识metalinguistic awareness, 语音敏感性phonological sensitivity, 英语语音训练English phonetic training, 拼读法phonics, 英语阅读训练English reading training, 英语阅读干预English reading intervention, 英语拼写训练English spelling training*, and *英语读写训练English literacy training*). Then, we searched English and Chinese articles of doctoral dissertations, master's theses, proceedings, and newspapers. In total, this yielded 1556 articles. The publication date was not restricted in any of the four databases because there is generally a dearth of training literature in this area.

Inclusion and exclusion criteria

To be included, studies had to meet the following criteria: (a) *Design*. The experimental designs were restricted to randomized control trials and quasi-experimental designs. Case studies and observational studies were excluded. (b) *Participants*. Participants should be NCSs who learn Chinese as their first language and English as a second language. (c) *Topic*. The study was included if the topic was related to the intent to improve EPA, with or without phonics, within NCSs. (d) *Dependent variables*. EPA (or its components) was a must as a dependent variable for each study. Other dependent variables could include real word reading or pseudoword reading.

The title, abstract, or full text (when necessary) was independently screened by two authors for each article. The agreement between researchers was determined using an inter-rater reliability analysis ($\kappa = .92$). The researchers resolved all discrepancies through discussion. Sixty-two articles met the inclusion criteria. Two researchers independently read them. Another inter-rater reliability analysis was performed to determine researcher agreement on excluded articles ($\kappa = .90$).

The following circumstances and study features resulted in exclusion: (a) *Sample issues*. NCSs were not the only target population, and their results were not separately analyzed (five studies). The experimental and the control groups were not at the same baseline at the start of the training (five studies), for example, the experimental group involved poor English readers while the control group involved typically developing readers. Parts of the samples were duplicated (one study). (b) *Insufficient information*. Studies lacked key descriptions of training procedures (one study), pretests, or posttests (17 studies). For instance, in some studies, authors did not report what the pretests included or how the tests were administered and assessed. Studies not reporting pretest scores on any of the outcomes were also excluded.

Finally, thirty-three articles were retained for data extraction and coding. Six studies used a randomized control trial design, and 27 used a quasi-experimental design. Seventeen were peer-reviewed articles, and 16 were unpublished papers.

Data extraction and effect size calculation

We conducted seven independent meta-analyses. We first coded effect sizes of overall results for EPA and reading outcomes separately. Then, we coded effect sizes for the results of syllable awareness, rhyme awareness, phoneme awareness, word reading, and pseudoword reading outcomes, respectively. For the overall effect size in EPA, the results of the measures selected in each article could be a composite score of EPA or any of the following components, including syllable awareness, rhyme awareness, phoneme awareness, or any of the processing skills in phoneme awareness, such as phoneme deletion, phoneme blending, or phoneme segmentation (see Appendix A). Due to the limited studies reporting onset-rime awareness and rhyme awareness as outcomes, they were combined into one category—rhyme awareness. The overall reading effect size, as determined by the measures chosen in each article, could be word reading or pseudoword reading. Several studies included more than one measure of EPA or reading. In the first two independent meta-analyses, we averaged the effect sizes to produce one "g effect size" for each sample (shown as "combined" in Appendix A). Two articles had more than one comparison; we adopted the mean of these comparisons (Chow et al., 2010; Siu et al., 2018). Moreover, Cheung (1999) and C. Y. Zhang et al. (2010) reported the scores of participants from different grades separately, so we kept all of them and coded each as an independent sample. Chow et al. (2021) reported participants' scores from two conditions: the group with low vocabulary knowledge and the group with high vocabulary knowledge. In each condition, participants were assigned to the experimental group or the control group. We kept these two conditions and coded each as an independent sample. Consequently, we got 37 independent samples from 33 articles with EPA outcomes and 17 independent samples from 16 articles with reading outcomes. Similarly, we used the same procedure to separately calculate the combined effect sizes for syllable awareness, rhyme awareness, phoneme awareness, word reading, and pseudoword reading.

We used the Comprehensive Meta-Analysis Version 3.3.070 software to analyze the data. The effect size of this meta-analysis was the standardized mean difference (Hedges's *g*) between the experimental and control conditions on the posttest. We chose the effect size of Hedges's *g* over Cohen's *d* because it corrected for small sample sizes (Borenstein et al., 2009). According to Cohen (1988), values of 0.8, 0.5, and 0.2 represented large, medium, and small effect sizes, respectively. The raw means and standard deviations of the posttests were preferred for calculating the effect sizes. However, we used inferential statistics to compute the effect size for the studies that did not report means and standard deviations (Card, 2015). These procedures included the use of *t*-statistics and *p*-values. We adopted the random-effects model to calculate the combined effect sizes and reported results based on that. The random-effects model allowed for between-study variance beyond sampling error (Borenstein et al., 2009). The Q-statistic was used to evaluate the heterogeneity of the combined effect sizes.

Coding of potential moderators

To evaluate the impact of outcome heterogeneity on the meta-analysis results, subgroup analyses were performed to compare contrasts based on categorical moderator variables (e.g., typical or special populations) across all meta-analyses. To ensure adequate statistical power, categorical moderator variables with at least three contrasts were included in this meta-analysis due to the limited number of available papers.

Coding was completed independently by two authors using a structured coding sheet. Two authors independently coded studies for characteristics reflecting potential moderator variables (see Appendix A). They coded for 10 categorical variables that were used as moderators. Inter-rater reliability ranged from 91% to 100% for various categorizations. Finally, all disagreements were resolved.

The following elements were coded. (1) *Sample characteristics*. This moderator aimed to examine the training effects on typical and special populations. (2) *Training strategy*. The studies were divided into two more general groups, namely direct training studies and indirect training studies. (3) *Type of program*. We only categorized the studies that used direct training strategies in terms of the focus of the training. Studies with indirect training strategies will be compared with those using direct training strategies in other sections. Hence the types of training programs include "EPA only," "EPA+phonics," "EPA and/or phonics+semantics," and "EPA and/or phonics +other contents." (4) *Grade*. "Grade" was categorized into five types: pre-elementary (including preschool and kindergarten), early elementary, upper elementary, middle school, and university. (5) *Number of weeks*. Two types were "≤12" and ">12." This moderator aimed to examine whether the length of training time was a potential influence on outcomes. (6) *Number of sessions*. This moderator was categorized into ">60" sessions, "30–60" sessions, "≤30" sessions, and "unclear," which aimed to see whether the intensity of training had an impact on outcomes. (7) *Publication status*. Two types of articles were "peer-reviewed" and "unpublished."

Publication bias assessments

Several approaches were applied separately to the overall EPA and overall reading samples to determine whether publication bias existed. The asymmetrical distribution of the studies on the funnel plot implied the possibility of publication bias. The Begg and Mazumdar rank correlation test was a direct statistical analog of the funnel graph approach (Begg & Mazumdar, 1994). Egger's regression method tested for a linear association between the training effect and its standard error (Egger et al., 1997). The fail-safe procedure helped estimate how many missing studies were needed to nullify a statistically significant effect (Borenstein et al., 2010). With random-effects models, the "trim and fill" method was used to examine the possible effects of studies that were probably missed because of publication bias (Duval & Tweedie, 2000).

RESULTS

Overall effects

Forest plots and funnel plots for overall EPA and overall reading are shown in Figures 2–5.

We examined the overall effect size across 37 independent samples from 33 articles in EPA ($n = 3137$) and 17 independent samples from 16 articles in reading ($n = 1506$). Results from the random-effects model analysis showed a mean estimated $g = 0.651$ in EPA, 95% confidence interval (CI) [0.509, 0.793], $z = 8.984$, $p < .001$, and $g = 0.498$ in reading, 95% CI [0.269, 0.727], $z = 4.261$, $p < .001$. Homogeneity test was significant in EPA, $Q\ (36) = 129.309$, $p < .001$, with $I^2 = 72.160$, and in reading, $Q\ (16) = 71.933$, $p < .001$, with $I^2 = 77.757$. This indicated that effect sizes had a large amount of heterogeneity, thus adding further justification for the utilization of random-effects models.

The funnel plot revealed a roughly symmetrical distribution of EPA outcomes, with no apparent outliers. The Begg and Mazumdar rank correlation test showed no evidence of publication bias, $p = .205$. In addition, the fail-safe number was 2632, which implied a robust effect. There were no indications of publication bias for reading outcomes in the funnel plot. The fail-safe number was 310, which suggested a strong

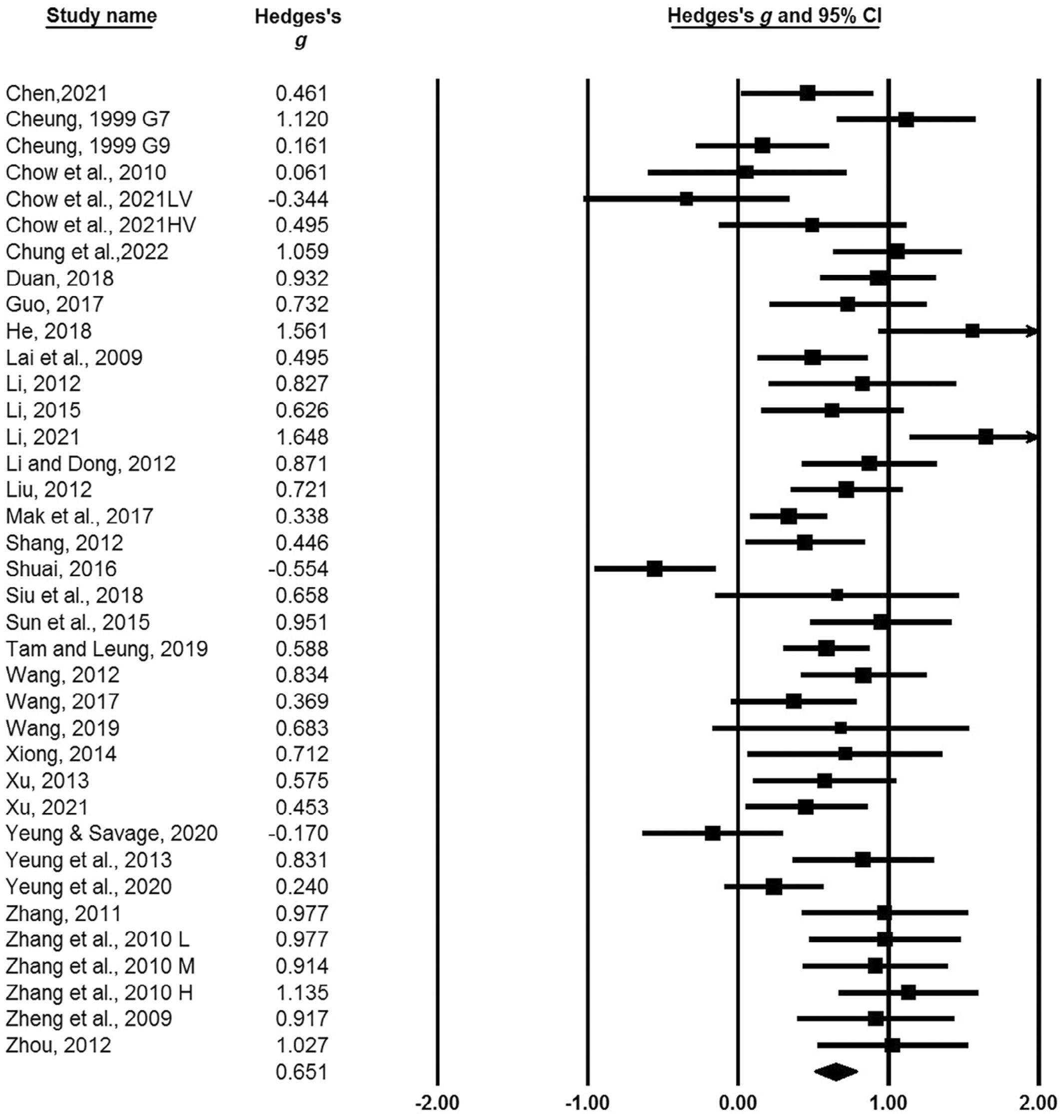

FIGURE 2 Overall English phonological awareness (EPA) forest plot.

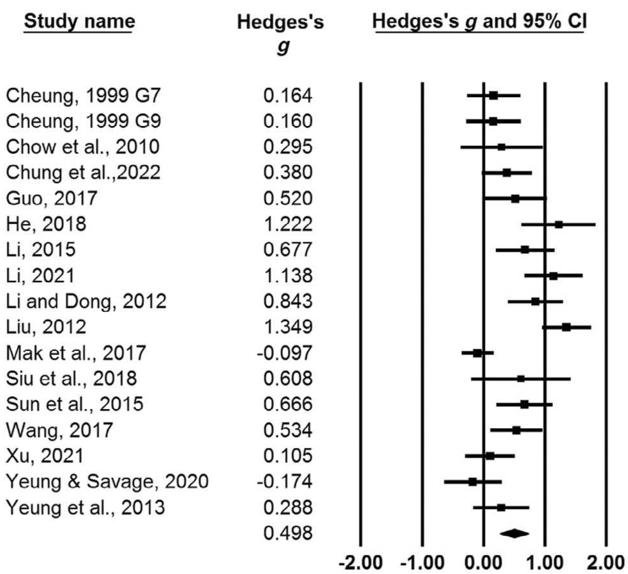

FIGURE 3 Overall reading forest plot.

impact. The Begg and Mazumdar rank correlation test also showed no publication bias, $p = .232$.

Effects of different components

For different outcome measures, there were no outliers according to the standard of Takacs and Kassai (2019). An overview of the effects on the overall EPA, overall reading, and different components is shown in Table 1. The random-effects model analysis showed that training had a moderate effect on syllable awareness ($g = 0.468$, $p = .015$), a large effect on rhyme awareness ($g = 0.948$, $p < .001$), and a medium to large effect on phoneme awareness ($g = 0.736$, $p < .001$). Results also showed a small impact on word reading ($g = 0.297$, $p = .009$) and a moderate effect on pseudoword reading ($g = 0.586$, $p = .008$).

The results suggested no publication bias for syllable awareness, rhyme awareness, and pseudoword reading (Begg and Mazumdar rank correlation test, $ps > .536$; Egger's regression

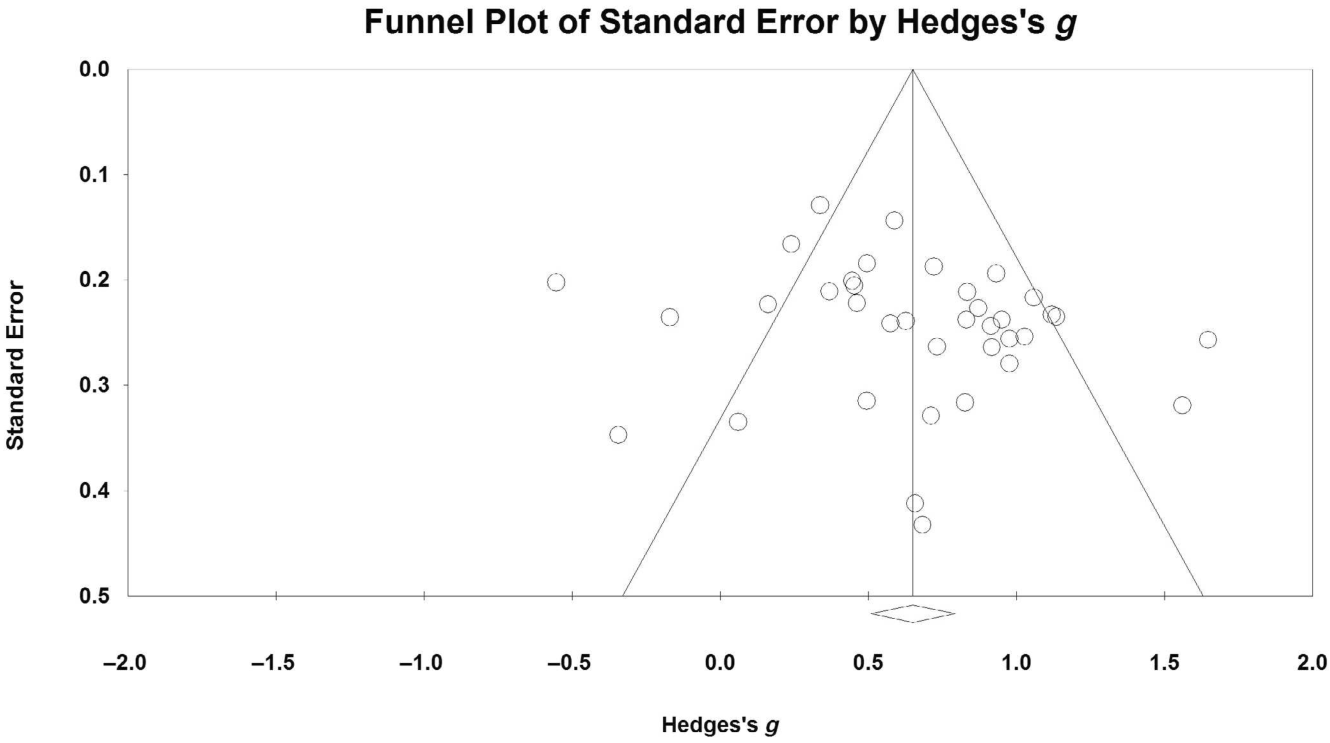

FIGURE 4 Overall English phonological awareness (EPA) funnel plot.

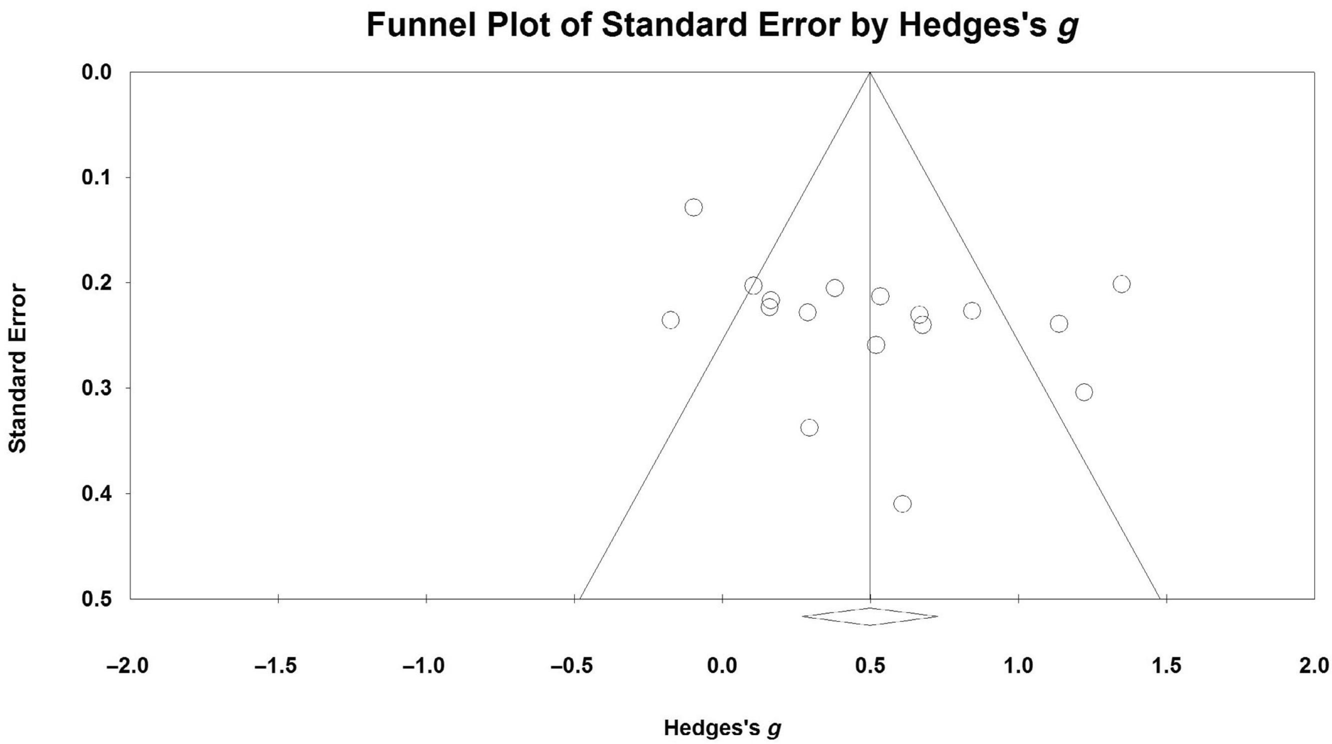

FIGURE 5 Overall reading funnel plot.

test, *ps* > .327). The funnel plot presented a sign of asymmetric distribution for phoneme awareness. The Duval and Tweedie trim and fill analysis procedure estimated that five hypothetical studies are missing. However, the adjusted effect size was still significant (estimate: 0.552, 95% CI [0.349, 0.755]). Moreover, the funnel plot presented a sign of asymmetric distribution for word reading; however, the trim and fill analysis revealed that with the adjustment, the estimate of the unbiased average effect size would be $g = 0.208$. It was very close to the average effect size found in our meta-analysis, $g = 0.297$. It suggested that publication bias was not a significant issue within this meta-analysis.

TABLE 1 Overview of the effects

Outcome	*k*	*g*	95% CI	*p*	*Q*	I^2
EPA	37	0.651	[0.509, 0.793]			
Syllable awareness	8	0.468	[0.091, 0.846]	<.05	38.768***	81.944
Rhyme awareness	7	0.948	[0.640, 1.256]	<.001	16.753*	64.186
Phoneme awareness	18	0.736	[0.546, 0.925]	<.001	63.229***	73.113
Reading	17	0.498	[0.269, 0.727]			
Word reading	13	0.297	[0.076, 0.518]	<.01	39.400***	69.543
Pseudoword reading	7	0.586	[0.155, 1.016]	<.01	47.605***	87.396

Note: k = the number of the effect sizes. EPA, English phonological awareness.
$^*p < .05$; $^{***}p < .001$.

TABLE 2 Mixed-effects categorical moderator analysis (EPA)

Moderator/level	*k*	Sample size	*g*	95% CI	*Q*	*p*
Main effect	37	3137	0.651	[0.509, 0.793]		
Training strategy					0.316	0.574
Direct	29	2575	0.672	[0.513, 0.831]		
Indirect	8	562	0.564	[0.223, 0.905]		
Type of program					6.485	0.090
EPA only	7	620	0.682	[0.433, 0.930]		
EPA+phonics	11	810	0.564	[0.218, 0.909]		
EPA and/or phonics+semantics	5	348	1.030	[0.715, 1.345]		
EPA and/or phonics+other contents	6	799	0.563	[0.363, 0.763]		
Sample characteristics					1.202	0.273
Special	3	311	0.368	[−0.162, 0.899]		
Typical	34	2826	0.676	[0.527, 0.826]		
Grade					10.797	0.029
Pre-elementary	9	728	0.779	[0.525, 1.033]		
Early elementary	11	889	0.458	[0.234, 0.681]		
Upper elementary	8	596	0.969	[0.710, 1.228]		
Middle school	6	650	0.435	[−0.012, 0.882]		
University	3	274	0.584	[0.337, 0.831]		
Number of weeks					3.718	0.054
>12	19	1820	0.779	[0.568, 0.989]		
≤12	18	1317	0.508	[0.330, 0.685]		
Number of sessions					9.957	0.019
>60	7	594	0.950	[0.720, 1.180]		
≤30	16	1324	0.619	[0.435, 0.803]		
30–60	8	809	0.740	[0.435, 1.045]		
Unclear	6	410	0.215	[−0.228, 0.658]		
Publication status					0.667	0.414
Peer-reviewed	21	1892	0.599	[0.430, 0.767]		
Unpublished	16	1245	0.724	[0.474, 0.975]		

Note: k = the number of the effect sizes. EPA, English phonological awareness.

Moderator analyses

Using mixed-effects categorical moderator analysis, we calculated effect sizes and 95% CIs for each level of all seven categorical moderators. Results are displayed in Tables 2 and 3, along with all between-group heterogeneity tests. For EPA as an outcome, the type of program had a marginally significant effect on the effect size, Q (3) = 6.485, p = .09. All four types of training programs—"EPA only," "EPA+phonics," "EPA and/or phonics+semantics" and "EPA and/or phonics+other contents"—showed significant training effects (ps < .01). However, the "EPA and/or phonics+semantics" program resulted in the largest trend of NCSs' improvement in EPA (g = 1.030, p < .001). Post-hoc analysis revealed that the "EPA and/or phonics+semantics" program had a significantly larger effect size than the "EPA only" program (p = .069). The grade level had a significant effect on the effect size, Q (4) = 10.797, p = .029, with middle school showing a marginally significant training effect (p = .057) and all the other grades showing significant training effects (ps < .001). Additionally, post-hoc analysis showed that the average effect size in upper elementary was significantly larger than that in early elementary and middle school (ps < .05). Furthermore, the number of sessions was a significant moderator, Q (3) = 9.957, p = .019. Except for the "unclear" type, all the other types showed significant training effects (ps < .001). Post-hoc analysis revealed that training with more than 60 sessions produced a significantly larger effect size than training with fewer than 30 sessions (p = .039). In addition, the number of weeks had a marginally significant effect on the effect size, Q (1) = 3.718, p = .054. Training programs with more than 12 weeks yielded a larger average effect size than those with fewer than 12 weeks (g = 0.779, p < .001).

For reading as an outcome, after excluding the university category, which involved only one study, grade significantly affected the effect size, Q (3) = 11.885, p = .008. Post-hoc analysis demonstrated that upper elementary showed a significantly larger effect size than those in early elementary and middle school (ps < .05). Pre-elementary and upper elementary had significant training effects (ps <0.05), and middle school showed a marginally significant training effect (p = .086). Moreover, publication status was a significant moderator, Q (1) = 11.770, p < .01. Both peer-reviewed and unpublished articles indicated significant training effects (ps < .01), but the latter led to a larger average effect size (g = 0.905, p < .001).

DISCUSSION

This meta-analysis investigated the effectiveness of direct and indirect training programs on NCSs' EPA and reading outcomes. The results revealed the efficacy of training in improving EPA and reading for typically developing Chinese students. Many correlational studies have reported that EPA is a strong predictor of NCSs' later English reading development (B. L. Li et al., 2011; Lin et al., 2017). Although correlational findings provide valuable insights, they are insufficient to determine a causal relationship between variables. In this study, we aimed

TABLE 3 Mixed-effects categorical moderator analysis (Reading)

Moderator/level	k	Sample size	g	95% CI	Q	p
Main effect	17	1506	0.651	[0.509, 0.793]		
Type of program					2.973	0.396
EPA only	5	424	0.379	[0.083, 0.676]		
EPA+phonics	5	403	0.717	[0.175, 1.258]		
EPA and/or phonics+semantics	3	216	0.647	[0.138, 1.157]		
EPA and/or phonics+other contents	3	395	0.206	[−0.217, 0.628]		
Grade					11.885	0.008
Pre-elementary	3	243	0.779	[0.525, 1.033]		
Early elementary	5	537	0.266	[−0.111, 0.643]		
Upper elementary	5	382	0.988	[0.664, 1.313]		
Middle school	3	248	0.386	[−0.054, 0.827]		
Number of weeks					2.278	0.131
>12	7	730	0.725	[0.257, 1.193]		
≤12	10	776	0.334	[0.137, 0.531]		
Number of sessions					0.100	0.751
≤30	10	822	0.457	[0.174, 0.741]		
30–60	5	538	0.550	[0.053, 1.046]		
Publication status					11.770	0.001
Peer-reviewed	11	1034	0.262	[0.064, 0.460]		
Unpublished	6	472	0.905	[0.595, 1.214]		

Note: k = the number of the effect sizes. EPA, English phonological awareness.

to extend beyond correlational evidence by synthesizing relevant intervention studies to assess the causal influence of EPA in NCSs who are learning to read. The synthesis of existing literature may provide stronger support for the causal role of EPA in the development of reading skills.

Native Chinese speakers' English phonological awareness and reading skills improved significantly

To our knowledge, the present meta-analysis is the first study to synthesize previous research on the effects of EPA and reading training for ELLs in a logographic context. The analysis included studies published both in English and Chinese to better capture information about EPA and reading training for NCSs. NCSs face additional challenges in learning EPA and reading skills as their writing systems are not based on an alphabet but rather on logographic characters or symbols. Results from the present study can lead to more targeted and effective support for English language education for NCSs.

EPA and reading training exerted a moderate effect on syllable awareness, but the effect size was smaller compared to that of rhyme and phoneme awareness. This is likely because NCSs already have relatively strong syllable awareness, so they make relatively minor improvements in their ability to perceive English syllables. Chinese is spoken with syllable timing, and Chinese speakers tend to use syllabic segmentation when reading English words (Deng et al., 2019). Consequently, NCSs may reach the ceiling early on the syllable awareness task (McBride-Chang & Kail, 2002). Wei (2017) reported that the accuracy of English syllable awareness (79.8%) of 10-year-old Chinese children was significantly higher than that of phoneme awareness (70.9%) and onset-rime awareness (57.7%), suggesting that NCSs' performance in syllable awareness is relatively good. Similarly, McBride-Chang et al. (2004) found no significant difference in syllable awareness between first graders in Hong Kong and age-matched Canadian children, and the elementary children from mainland China even demonstrated superior performance in a syllable deletion task compared to age-matched English speakers.

Furthermore, EPA and reading training exerted the most substantial impact on improving NCSs' rhyme awareness. There are two potential reasons. First, the Chinese learning experience might have contributed to the development of onset-rime awareness in NCSs, given the prevalence of phonosemantic characters in Chinese writing, where the whole character typically rhymes with its sound component (Tao et al., 2005). Moreover, English has a strong tradition of rhyming verse (Bradley & Bryant, 1983; McKie, 1997), while most Chinese poems and song lyrics obey tail rhyme to some extent (X. Wang et al., 2011). As a result, NCSs may possess a tendency to transfer their acuity in discerning rhyme patterns of Chinese characters to their sensitivity towards English rhyme awareness and word reading. This is supported by Tao et al. (2007), who found that Chinese onset-rime awareness accounted for unique variances in English word reading, while Chinese tone awareness influenced nonword decoding, and their effects on English word reading were mediated by English onset-rime awareness. Second, rhyme awareness might be easier to acquire than other phonological units. As Ziegler and Goswami (2005) posit in the grain size theory, people's phonological awareness typically progresses from larger to smaller units. These findings have important implications for the design of EPA and reading training for NCSs. Particularly, it is crucial to increase their awareness of English rhyme patterns and improve their word reading abilities.

In addition, EPA and reading training programs yielded a moderate to large effect on phoneme awareness. This is likely due to NCSs' underdeveloped skills in phoneme awareness, which make them particularly responsive to such training. Additional challenges in learning a second language are likely to occur when the phonology and orthography of the target language are very contrastive to the native language (Ellis, 1999). English words consist of individual phonemes that can be manipulated to form new words (McBride-Chang & Kail, 2002), so phoneme awareness may be particularly vital for English language acquisition. Conversely, phonemes are not explicitly represented in the Chinese system (Branum-Martin et al., 2012), and thus NCSs do not typically require phoneme-level awareness to read Chinese (Shu et al., 2008). This difference could pose challenges for NCSs in developing their English phoneme awareness abilities. What is more, the findings can also be supported by grain size theory, which suggests small grain size units of processing, such as single letters and phonemes, are favored, even when large unit information is available (Ziegler & Goswami, 2005).

Noticeably, although NCSs' phoneme awareness can be greatly improved through training, the effect size was smaller than that in NESs, as reported in the meta-analyses of Bus and Van Ijzendoorn (1999) and Ehri et al. (2001). For NESs, phonemic awareness has a much stronger predictive power than the syllable and rhyme factor for later reading (Høien et al., 1995). The smaller effect of instruction on NCSs compared to NESs raises questions regarding the optimal intensity and mode of delivery of PA instruction for NCSs. It should be noted that compared to training for NESs, which has involved a variety of "phonemic awareness only" training methods (Castle et al., 1994; Ehri et al., 2001), few EPA training studies for NCSs focused only on the phoneme level of words—perhaps a more intense focus on this advanced level of phonological awareness might yield greater effects for some learners.

In terms of understanding the struggles of NCSs in acquiring EPA, it is worth considering phonemic aspects of English that might provide a particular challenge for NCSs. Some English phonemes, such as the consonants /θ/ and /ð/ and vowels /æ/ and /ɜː/, are uncategorized in the Chinese language, making them more difficult to discriminate against and manipulate for NCSs learning English after early childhood. This is consistent with the Perceptual Assimilation Model, which underscores how people's native language may influence their perception of non-native languages (Best, 1995). A resultant question is whether targeted phonetic category training could effectively improve NCSs' EPA ability. A meta-analysis reviewing 25 years of perception training studies conducted by Sakai and Moorman (2018) revealed that second language learners made significant progress in their phonetic production abilities after receiving phonetic perception training. Moreover, studies on NCSs have also shown that training on sensitivity to the English /l/−/d/ contrast, perception, and production of /θ/ and /ð/ is effective (Flege, 1989; Y. F. Li & Dong, 2009, 2011). Considering the lower gains in EPA and reading for NCSs within EPA-based training programs, we, therefore, suggest that a combination of EPA-based training and English phonetic categories training would offer a fruitful future direction of study. The findings may have implications for educators and instructors when designing appropriate NCS language curricula with a focus on subcomponents of EPA.

The present study also found that the training effect on NCSs' pseudoword reading was larger than that on word reading. This finding is in line with previous meta-analyses that focused on ELLs (Huo & Wang, 2017; Murphy Odo, 2021). Pseudoword reading requires strong decoding skills to read unfamiliar, nonsense words, and EPA-related training programs may provide NCSs with more chances to improve their decoding abilities. In contrast, each real word has its lexical semantic knowledge. With insufficient oral vocabulary acquired in daily life, NCSs may have limited progress in word reading during a training program.

Programs integrating lexical semantic knowledge produced the largest trend in enhancing English phonological awareness

Both direct and indirect training strategies facilitated NCSs' EPA development. Although this conclusion is different from the meta-analytic conclusion of Sermier Dessemontet et al. (2019), which found that direct strategy training had a larger effect size, this may imply that substantial English language environment exposure benefiting from the indirect strategy is particularly important for NCSs. Thus, whilst parents and schools may be keen to continue with direct EPA training activities, and their value is undeniable, this analysis demonstrates that indirect activities may also be beneficial. As Ellis (2015) claims, language processing is very sensitive to usage frequency. Usage contains a wealth of latent linguistic structures, and learners' apprehension of these structures is achieved mainly through indirect learning. Traditionally, indirect EPA and reading training programs have received limited attention within the field. In the future, educators can explore the value of generally increased English exposure, such as watching English television or plays, listening to books being read aloud, or encouraging English discussion and debate.

The purpose of this study was to evaluate the effectiveness of four direct training programs, namely "EPA only," "EPA+phonics," "EPA and/or phonics+semantics," and "EPA and/or phonics+other contents," in improving NCSs' EPA and reading competence. The results indicated that all four programs were effective in enhancing NCSs' EPA skills. However, the "EPA and/or phonics+semantics" program displayed the largest tendency in enhancing NCSs' EPA competence, and it produced a significantly larger effect size than the "EPA only" program, highlighting the significance of oral vocabulary instruction. This finding is in line with the lexical restructuring model, which shows that one's early word representations are more holistic with smaller vocabulary sizes (Walley & Metsala, 1990). As a child's vocabulary grows, these holistic representations are gradually restructured so that smaller segments of sound, like rimes and phonemes, are ultimately represented (Goswami, 2002). This might suggest that a combination of EPA/phonics and vocabulary-rich instruction is recommended for early NCSs. Nevertheless, further investigation requires examining whether vocabulary should be taught before phonics instruction or in a more integrated manner. Increased phonics training intensity for NCSs, compared to NESs, may also enable more rapid progress in English reading; however, this requires further empirical study.

Additionally, the type of training program used did not significantly affect NCSs' reading performance. This finding contradicts earlier research by Bus and Van Ijzendoorn (1999) and Ehri et al. (2001), which reported that the combination of EPA training with letter instruction yielded better word reading outcomes than phonetic training alone. The present meta-analysis indicates that both EPA training and its combination with letter–sound knowledge are effective approaches to improving NCSs' reading ability.

English phonological awareness and reading training were most effective for native Chinese speakers in upper elementary grades

In terms of assessing language proficiency, on the one hand, it has been observed that training significantly enhances NCSs' EPA and reading development across age groups. This finding aligns with our hypothesis, positing that the advantages of phonological training for NCSs may exhibit an extended time window of efficacy. Moreover, the findings indicate that Chinese students in the upper elementary grades are particularly inclined to benefit most from EPA and reading training. On the other hand, the characteristics of the sample population do not seem to play a significant role in NCSs' EPA and reading training, as both typical and special NCSs can effectively enhance their skills.

Compared to NCSs in early elementary, those in the upper elementary stage exhibited significantly larger effect sizes on EPA and reading. This is likely because higher-grade students have made greater progress in pinyin (the official romanization system for standard Mandarin Chinese) and character recognition. As a result, their increasing proficiency in Chinese can be transferred to the development of their EPA and reading competence. Tao et al. (2007) found that the elementary NCSs' Chinese onset-rime awareness accounted for unique variances in English word reading and that Chinese tone awareness accounted for unique variances in English pseudoword reading. Other studies have also revealed that Chinese upper elementary students exhibited higher levels of English reading, EPA, and

Chinese phonological awareness than those in the early elementary stage (B. L. Li et al., 2011; F. Xu & Dong, 2005).

On the other hand, Chinese middle school students did not benefit as much from EPA and reading training compared to upper elementary school students. This is probably because instructional strategies employed in middle school classrooms may provide fewer opportunities for NCSs to practice their decoding skills. Rao (1996) states that Chinese English teachers mainly employ rote-memory strategies rather than phonics when teaching English vocabulary. Consequently, NCSs may depend less on their EPA skills while reading English words. As Pan et al. (2019) reported in a meta-analysis, the correlation between EPA and reading ability (at the word, sentence, and text levels) was lowest in middle school compared to other age groups, including pre-elementary, elementary, and university stages.

Native Chinese speakers require more intensive training to enhance English phonological awareness and reading skills

In this study, we found that while training weeks and sessions were not significant moderators for reading as an outcome, Chinese students who got more intensive training showed bigger advances in EPA ability. It suggests that increasing the intensity of training may be an effective strategy for improving EPA ability in NCSs. Moreover, for EPA, publication status was not a significant moderator. However, for reading, the effect sizes of the unpublished papers were larger than those of the peer-reviewed ones. This result was somewhat surprising. This is likely because experimenters conducted less rigorous design and execution compared to studies published in academic journals. Another potential reason for this phenomenon is that unpublished articles may not have undergone rigorous scrutiny during the peer-review process, resulting in fewer quality checks.

LIMITATIONS AND IMPLICATIONS

In this meta-analysis, we primarily examined EPA and reading-related components; other outcomes related to second language learning, such as spelling, vocabulary, and reading comprehension, were not examined. Moreover, in several studies, the experimental group chose an additional time for training, but there was no active control group that received training in different skills at the same time. This meant training time for the experimental group and the control group was not the same, which might affect the accuracy of the evaluation. Furthermore, the long-term effects of EPA training were not examined because most studies did not administer follow-up tests.

In interpreting the results of this study, other limiting factors warrant consideration. One factor is the quantity of the sample available, which may affect the analysis of potential moderators. In cases where subgroup numbers were insufficient, relevant moderators were not examined. For instance, we did not evaluate the indirect strategy's influence on reading because there was only one subgroup for indirect strategy use. Another limitation is the discrepancy between unpublished and peer-reviewed studies. Contrary to the main theoretical effects of the file drawer problem (Hunter & Schmidt, 1990), unpublished studies showed larger effect sizes for EPA than peer-reviewed studies. This divergence may be attributable to an overestimation of effect size in low-quality studies. To increase the generalizability of results among the NCS population in this domain, more peer-reviewed and high-quality research is needed. While publication bias analysis indicates little impact on results, the possibility for publication bias to inflate the training effects from unpublished studies should not be ignored. Hence caution should still be exercised when interpreting certain effect sizes.

China has the largest number of English learners in the world, and while this review demonstrates some effectiveness in EPA training in NCSs, the immediate impact on word reading is more modest, a finding that warrants further attention. The specific combination of EPA and phonics instruction is one area for additional study. Approaches to English phonics instruction can vary greatly, including combinations of synthetic phonics, analytic phonics, analogy-based phonics, phonics through spelling, embedded phonics, and onset-rime phonics instruction (Armbruster, 2010). While the effectiveness of various approaches to teaching children to read has been compared in NES samples (Comaskey et al., 2009; Johnston et al., 2012), few previous studies have directly compared different phonics approaches in NCS samples.

CONCLUSION

This meta-analysis extends our understanding of the effectiveness of EPA and reading training for ELLs in a logographic context. The results showed that EPA-based training programs were moderately effective at promoting typically developing NCSs' EPA and reading development. Training in rhyme awareness was particularly critical, as shown by the large effect size in our analysis. We also identified that NCSs across age groups benefited from EPA training. However, participants from the upper elementary stage produced significantly larger effect sizes on EPA and reading than those from early elementary and middle school. Additionally, both direct and indirect training strategies promoted NCSs' EPA and reading ability. And the "EPA and/or phonics+semantics" program produced the largest trend in improving NCSs' EPA competence, and it demonstrated a significantly larger training effect than the "EPA only" program. Finally, more intensive training showed a larger effect on EPA than less intensive training.

Currently, there is increasing interest in phonics programs among English teachers in China, and this paper suggests that EPA, which is often neglected, should be considered an integral part of China's English curriculum standards and classrooms. It is hoped that this article will help stimulate further research in this field, such as investigating the longer-term effects of EPA-based training on EPA and reading.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflicts of interest.

ORCID

Ying Jiang https://orcid.org/0000-0002-3805-0201
Xiaosong Gai https://orcid.org/0000-0001-6836-5442
Meryem S. Üstün-Yavuz https://orcid.org/0000-0002-0117-9044
Jenny M. Thomson https://orcid.org/0000-0002-4293-4851

REFERENCES

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

Armbruster, B. B. (2010). *Put reading first: The research building blocks for teaching children to read: Kindergarten through grade 3*. Diane Publishing.

Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics*, *50*(4), 1088–1101. https://doi.org/10.2307/2533446

Best, C. T. (1995). A direct realist view of cross-language speech perception: Speech perception and linguistic experience. In W. Strange (Ed.), *Cross-language speech perception* (pp. 171–206). York Press.

Bialystok, E., Majumder, S., & Martin, M. M. (2003). Developing phonological awareness: Is there a bilingual advantage? *Applied Psycholinguistics*, *24*(1), 27–44. https://doi.org/10.1017/S014271640300002X

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). When does it make sense to perform a meta-analysis? In M. Borenstein, L. V. Hedges, J. P. T. Higgins, & H. R. Rothstein (Eds.), *Introduction to meta-analysis* (pp. 357–364). Wiley.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods*, *1*(2), 97–111. https://doi.org/10.1002/jrsm.12

Bradley, L., & Bryant, P. E. (1983). Categorizing sounds and learning to read: A causal connection. *Nature*, *301*(5899), 419–421. https://doi.org/10.1038/301419a0

Branum-Martin, L., Tao, S., Garnaat, S., Bunta, F., & Francis, D. J. (2012). Meta-analysis of bilingual phonological awareness: Language, age, and psycholinguistic grain size. *Journal of Educational Psychology*, *104*(4), 932–944. https://doi.org/10.1037/a0027755

Bus, A. G., & Van Ijzendoorn, M. H. (1999). Phonological awareness and early reading: A meta-analysis of experimental training studies. *Journal of Educational Psychology*, *91*(3), 403–414. https://doi.org/10.1037/0022-0663.91.3.403

Card, N. A. (2015). *Applied meta-analysis for social science research*. Guilford Publications.

Castle, J. M., Riach, J., & Nicholson, T. (1994). Getting off to a better start in reading and spelling: The effects of phonemic awareness instruction within a whole language program. *Journal of Educational Psychology*, *86*(3), 350–359. https://doi.org/10.1037/0022-0663.86.3.350

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

*Chen, J. Y. (2021). *Research on the influence of phonetic awareness training on the word reading and spelling ability in junior class of vocational colleges*. [Unpublished master's thesis]. Minnan Normal University.

*Cheung, H. (1999). Improving phonological awareness and word reading in a later learned alphabetic script. *Cognition*, *70*(1), 1–26. https://doi.org/10.1016/S0010-0277(98)00074-2

*Chow, B. W. Y., Hui, A. N. N., Li, Z., & Dong, Y. (2021). Dialogic teaching in English-as-a-second-language classroom: Its effects on first graders with different levels of vocabulary knowledge. *Language Teaching Research*., Advance online. 136216882098139. https://doi.org/10.1177/1362168820981399

*Chow, B. W. Y., McBride-Chang, C., & Cheung, H. (2010). Parent-child reading in English as a second language: Effects on language and literacy development of Chinese kindergarteners. *Journal of Research in Reading*, *33*(3), 284–301. https://doi.org/10.1111/j.1467-9817.2009.01414.x

Chung, K. K. H., & Lam, C. B. (2020). Cognitive-linguistic skills underlying word reading and spelling difficulties in Chinese adolescents with dyslexia. *Journal of Learning Disabilities*, *53*(1), 48–59. https://doi.org/10.1177/0022219419882648

*Chung, K. K. H., Li, X., Lam, C. Y., Lam, C. B., Fung, W. K., & Lai, P. Y. (2022). The effects of bilingual reading program on Chinese children from low socioeconomic status families. *Early Education and Development*, *33*(2), 204–218. https://doi.org/10.1080/10409289.2021.1891803

Cohen, J. (1988). Statistical power analysis for the behavioural science. *Technometrics*, *31*(4), 499–500. https://doi.org/10.1080/00401706.1989.10488618

Comaskey, E. M., Savage, R. S., & Abrami, P. (2009). A randomised efficacy study of web-based synthetic and analytic programmes among disadvantaged urban kindergarten children. *Journal of Research in Reading*, *32*(1), 92–108. https://doi.org/10.1111/j.1467-9817.2008.01383.x

Deng, Q., Choi, W., & Tong, X. (2019). Bidirectional cross-linguistic association of phonological skills and reading comprehension: Evidence from Hong Kong Chinese-English bilingual readers. *Journal of Learning Disabilities*, *52*(4), 299–311. https://doi.org/10.1177/0022219419842914

*Duan, Y. L. (2018). *An empirical study on the application of phonics instruction in the cultivation of English phonological awareness of primary school students*. [Unpublished master's thesis]. Hunan Normal University.

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*(2), 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, *315*(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629

Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly*, *36*(3), 250–287. https://doi.org/10.1598/rrq.36.3.2

Ellis, N. C. (2015). Implicit and explicit language learning. In *Implicit and explicit learning: Their dynamic interface and complexity*. John Benjamins.

Ellis, R. (1999). *The study of second language acquisition*. Shanghai Foreign Language Education Press.

Flege, J. E. (1989). Chinese subjects' perception of the word-final English /1/–/d/ contrast: Performance before and after training. *Journal of the Acoustical Society of America*, *86*(5), 1684–1697. https://doi.org/10.1121/1.398599

Goetze, J., & Driver, M. (2022). Is learning really just believing? A meta-analysis of self-efficacy and achievement in SLA. *Studies in Second Language Learning and Teaching*, *12*(2), 233–259. https://doi.org/10.14746/ssllt.2022.12.2.4

Goswami, U. (2002). In the beginning was the rhyme? A reflection on Hulme, hatcher, Nation, Brown, Adams, and Stuart (2002). *Journal of Experimental Child Psychology*, *82*(1), 47–57. https://doi.org/10.1006/jecp.2002.2673

Guan, Y. J., Li, Y. F., & Dong, Q. (2006). 学前儿童英语语音意识、字母知识及口语词汇之间关系的研究 [the interrelations among preschooler's phonological awareness, letter knowledge and oral vocabulary]. *Psychological Development and Education*, *22*(2), 7–11. https://doi.org/10.16187/j.cnki.issn1001-4918.2006.02.002

*Guo, F. (2017). *An experimental study on the effectiveness of phonics instruction on primary school pupil's English vocabulary acquisition*. [Unpublished master's thesis]. Guangzhou University.

*He, S. Q. (2018). *An empirical study of the impact of phonics on the reading and writing of English vocabulary in primary schools*. [Unpublished master's thesis]. Shanghai International Studies University.

Høien, T., Lundberg, I., Stanovich, K. E., & Bjaalid, I. K. (1995). Components of phonological awareness. *Reading and Writing*, *7*(2), 171–188. https://doi.org/10.1007/BF01027184

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings*. Sage Publications, Inc.

Huo, S., & Wang, S. (2017). The effectiveness of phonological-based instruction in English as a foreign language students at primary school level: A research synthesis. *Frontiers in Education*, *2*(15), 1–13. https://doi.org/10.3389/feduc.2017.00015

Johnston, R. S., McGeown, S., & Watson, J. E. (2012). Long-term effects of synthetic versus analytic phonics teaching on the reading and spelling ability of 10 year old boys and girls. *Reading and Writing*, *25*(6), 1365–1384. https://doi.org/10.1007/s11145-011-9323-x

Jongejan, W., Verhoeven, L., & Siegel, L. S. (2007). Predictors of reading and spelling abilities in first- and second-language learners. *Journal of Educational Psychology*, *99*(4), 835–851. https://doi.org/10.1037/0022-0663.99.4.835

*Lai, Y., Tsai, H., & Yu, P. (2009). A multimedia English learning system using HMMs to improve phonemic awareness for English learning. *International Forum of Educational Technology & Society*, *12*(3), 266–281.

Li, B. L., Tao, S., & Dong, Q. (2011). 英语语音意识在汉语儿童英语单词阅读及拼写中的作用 [the roles of phonological awareness in English word reading and spelling among mandarin-speaking children]. *Psychological Development and Education*, *4*, 388–393. https://doi.org/10.16187/j.cnki.issn1001-4918.2011.04.002

Li, G. Z., & Dong, Y. P. (2012). The effects of phonological awareness training on English language learners' word reading skills. *Journal of PLA University of Foreign Languages*, *1*, 59–63.

*Li, H. D. (2015). *An experimental study of the impact of THRASS (a systematic phonics learning program) on elementary school students in China*. [Unpublished master's thesis]. Hangzhou Normal University.

*Li, Y. (2012). *The research of child ESL learners' English phonological awareness under the English cartoon context teaching*. [Unpublished master's thesis]. Liaoning Normal University.

*Li, Y. F. (2021). *The application of phonics in English vocabulary teaching for primary school students*. [Unpublished master's thesis]. Xinan University.

Li, Y. F., & Dong, Q. (2009). 汉语母语儿童和成人视听双通道英语语音知觉的训练研究 [the training study of Chinese children and adults' audiovisual English phonetic perception]. *Psychological Development and Education*, *9*(1), 72–78. https://doi.org/10.16187/j.cnki.issn1001-4918.2009.01.009

Li, Y. F., & Dong, Q. (2011). 不同语言知觉训练方式对汉语儿童和成人英语语音产出的作用 [the effect of the different phonetic perception training on Chinese children and adults' English phonetic production]. *Chinese Journal of Applied Psychology*, *17*(4), 325–330. https://doi.org/10.3969/j.issn.1006-6020.2011.04.006

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *Journal of Clinical Epidemiology*, *62*(10), 1–34. https://doi.org/10.1136/bmj.b2700

Lin, D., Liu, Y., Sun, H., Wong, R. K. S., & Yeung, S. S. S. (2017). The pathway to English word reading in Chinese ESL children: The role of spelling. *Reading and Writing*, *30*(1), 87–103. https://doi.org/10.1007/s11145-016-9664-6

*Liu, X. H. (2012). *An empirical study on the application of phonics in English vocabulary teaching in primary schools*. [Unpublished master's thesis]. Guangzhou University.

Liu, Y., Yeung, S. S. S., Lin, D., & Wong, R. K. S. (2017). English expressive vocabulary growth and its unique role in predicting English word reading: *A longitudinal study involving Hong Kong Chinese ESL children*. *Contemporary Educational Psychology*, *49*, 195–202. https://doi.org/10.1016/j.cedpsych.2017.02.001

*Mak, B. S. Y., Cheung, A. C. K., Guo, X., Abrami, P. C., & Wade, A. (2017). Examining the impact of the ABRACADABRA (ABRA) web-based literacy program on primary school students in Hong Kong. *Education and Information Technologies*, *22*(6), 2671–2691. https://doi.org/10.1007/s10639-017-9620-3

McBride-Chang, C., Bialystok, E., Chong, K. K. Y., & Li, Y. (2004). Levels of phonological awareness in three cultures. *Journal of Experimental Child Psychology*, *89*(2), 93–111. https://doi.org/10.1016/j.jecp.2004.05.001

McBride-Chang, C., & Kail, R. V. (2002). Cross-cultural similarities in the predictors of reading acquisition. *Child Development*, *73*(5), 1392–1407. https://doi.org/10.1111/1467-8624.00479

McKie, M. (1997). The origins and early development of rhyme in English verse. *The Modern Language Review*, *92*, 817–831. https://doi.org/10.2307/3734202

McWeeny, S., Choi, S., Choe, J., LaTourrette, A., Roberts, M. Y., & Norton, E. S. (2022). Rapid automatized naming (RAN) as a kindergarten predictor of future reading in English: A systematic review and meta-analysis. *Reading Research Quarterly*, *57*(4), 1187–1211. https://doi.org/10.1002/rrq.467

Mohsen, M. A. (2022). Computer-mediated corrective feedback to improve L2 writing skills: A meta-analysis. *Journal of Educational Computing Research*, *60*(5), 1253–1276. https://doi.org/10.1177/07356331211064066

Murphy Odo, D. (2021). A meta-analysis of the effect of phonological awareness and/or phonics instruction on word and pseudo word reading of English as an L2. *SAGE Open*, *11*(4), 215824402110591. https://doi.org/10.1177/21582440211059168

Pan, Y., Deng, X. P., Lin, Z. J., & Liu, Y. (2019). 中国学生英语语音意识与英语阅读成绩关系的元分析 [a meta-analysis of the relationship between Chinese students' English phonological awareness and reading achievements]. *Studies of Psychology and Behavior*, *17*(2), 11. https://doi.org/10.3969/j.issn.1672-0628.2019.02.012o

Pan, Y., & Gai, X. S. (2013). 汉语母语与英语母语学生英语语音意识发展趋势的比较研究 [a comparison of the developing trends of English phonological awareness between mandarin-speaking students and English-speaking students]. *Studies of Psychology and Behavior*, *11*(1), 73–77. https://doi.org/10.3969/j.issn.1672-0628.2013.01.012

Parrila, R., Kirby, J. R., Mcquarrie, L., & Rate, A. (2004). Articulation rate, naming speed, verbal short-term memory, and phonological awareness: Longitudinal predictors of early reading development? *Scientific Studies of Reading*, *8*(1), 3–26. https://doi.org/10.1207/s1532799xssr0801

Rao, Z. (1996). Reconciling communicative approaches to the teaching of English with traditional Chinese methods. *Research in the Teaching of English*, *30*, 458–471.

Rehfeld, D. M., Kirkpatrick, M., O'Guinn, N., & Renbarger, R. (2022). A meta-analysis of phonemic awareness instruction provided to children suspected of having a reading disability. *Language, Speech, and Hearing Services in Schools*, *53*(4), 1117–1201. https://doi.org/10.1044/2022_lshss-21-00160

Sakai, M., & Moorman, C. (2018). Can perception training improve the production of second language phonemes? A meta-analytic review of 25 years of perception training research. *Applied PsychoLinguistics*, *39*(1), 187–224. https://doi.org/10.1017/S0142716417000418

Sermier Dessemontet, R., Martinet, C., de Chambrier, A. F., Martini-Willemin, B. M., & Audrin, C. (2019). A meta-analysis on the effectiveness of phonics instruction for teaching decoding skills to students with intellectual disability. *Educational Research Review*, *26*(1), 52–70. https://doi.org/10.1016/j.edurev.2019.01.001

*Shang, L. N. (2012). *The influence of English phonological awareness training on high school students' listening comprehension*. [Unpublished master's thesis]. Liaoning Normal University.

Shu, H., Peng, H., & McBride-Chang, C. (2008). Phonological awareness in young Chinese children. *Developmental Science*, *11*(1), 171–181. https://doi.org/10.1111/j.1467-7687.2007.00654.x

*Shuai, J. L. (2016). *An empirical study on effect of phonological awareness training on spelling ability of first-year students in rural junior middle school*. [Unpublished master's thesis]. Gannan Normal University.

*Siu, T. S. C., McBride, C., Tse, C. S., Tong, X., & Maurer, U. (2018). Evaluating the effects of metalinguistic and working memory training on reading fluency in Chinese and English: A randomized controlled trial. *Frontiers in Psychology*, *9*(12), 1–13. https://doi.org/10.3389/fpsyg.2018.02510

*Sun, B., Zhu, B., Chen, J., & Zhou, H. (2015). The effects of phonological awareness training on the reading performance by child EFL learners in China. *Chinese Journal of Applied Linguistics*, *38*(3), 339–354. https://doi.org/10.1515/cjal-2015-0021

Takacs, Z. K., & Kassai, R. (2019). The efficacy of different interventions to foster children's executive function skills: A series of meta-analyses. *Psychological Bulletin*, *145*(7), 653–697. https://doi.org/10.1037/bul0000195

*Tam, I. O. L., & Leung, C. (2019). Evaluation of the effectiveness of a literacy intervention programme on enhancing learning outcomes for secondary students with dyslexia in Hong Kong. *Dyslexia*, *25*(3), 296–317. https://doi.org/10.1002/dys.1626

Tao, S., Feng, Y. J., & Li, W. (2007). 语音意识的不同成分在汉语儿童英语阅读学习中的作用 [the roles of different components of phonological awareness in English reading among mandarin-speaking children]. *Psychological Development and Education*, *23*(2), 82–92. https://doi.org/10.16187/j.cnki.issn1001-4918.2007.02.014

Tao, S., Huang, X. M., & Li, W. (2005). 儿童汉英双语语音意识:跨语言一致性、差异与迁移 [mandarin-speaking children's phonological awareness: Similarities differences and transference between Chinese and English]. *Journal of Beijing Normal University (Social Sciences)*, *3*, 47–54. https://doi.org/10.3969/j.issn.1002-0209.2005.03.006

Vadasy, P. F., & Sanders, E. A. (2011). Efficacy of supplemental phonics-based instruction for low-skilled first graders: How language minority status and pretest characteristics moderate treatment response. *Scientific Studies of Reading*, *15*(6), 471–497. https://doi.org/10.1080/10888438.2010.501091

Walley, A. C., & Metsala, J. L. (1990). The growth of lexical constraints on spoken word recognition. *Perception & Psychophysics*, *47*(3), 267–280. https://doi.org/10.3758/BF03205001

*Wang, C. W. (2019). Can raising phonological awareness enhance the English reading comprehension of Taiwanese primary school-aged children? *Asian EFL Journal*, *23*, 4–27.

*Wang, H. Y. (2012). *A correlational study on the explicit instruction of English phonological awareness and listening competence development in college EFL context*. [Unpublished master's thesis]. Yanshan University.

*Wang, M. X. (2017). *The application of English picture books in phonics on first graders in primary school*. [Unpublished master's thesis]. Shenzhen University.

Wang, X., Chen, X., Yang, D., & Wu, Y. (2011). Music emotion classification of Chinese songs based on lyrics using tf*idf and rhyme. *Proceedings of the 12th International Society for Music Information Retrieval Conference*, 765–770. https://doi.org/10.5281/zenodo.1416458

Wei, X. B. (2017). 语音及韵律意识在以汉语为母语的儿童英语单词认读中的作用 [the roles of phonological awareness and prosodic sensitivity in EFL Chinese children's word-level reading]. *Foreign Language Education in China (Quarterly)*, *10*(1), 47–55.

*Xiong, W. (2014). *The phonological awareness and spelling intervention study on primary students with learning disabilities in English*. [Unpublished master's thesis]. Hunan Normal University.

Xu, F., & Dong, Q. (2005). 汉语儿童汉语与英语语音意识发展的关系 [research on the relationship between development of Chinese and English phonological awareness in primary school]. *Psychological Development and Education*, *21*(1), 31–35. https://doi.org/10.3969/j.issn.1001-4918.2005.01.006

*Xu, Y. (2021). 大学生结构化英语语音意识干预的有效性 [effectiveness of structured phonological awareness intervention on Chinese college students]. *Foreign Language Learning Theory and Practice*, *1*, 101–108.

*Xu, Z. M. (2013). *The effect of English phonological awareness training on children's Chinese phonological awareness and early reading abilities*. [Unpublished master's thesis]. Shandong University.

Yeong, S. H. M., Fletcher, J., & Bayliss, D. M. (2014). Importance of phonological and orthographic skills for English reading and spelling: A comparison of English monolingual and mandarin-English bilingual children. *Journal of Educational Psychology*, *106*(4), 1107–1121. https://doi.org/10.1037/a0036927

Yeung, S. S., & Chan, C. K. K. (2013). Phonological awareness and oral language proficiency in learning to read English among Chinese kindergarten children in Hong Kong. *British Journal of Educational Psychology*, *83*(4), 550–568. https://doi.org/10.1111/j.2044-8279.2012.02082.x

Yeung, S. S., & Ganotice, F. A. (2014). The role of phonological awareness in biliteracy acquisition among Hong Kong Chinese kindergarteners who learn English-as-a-second language (ESL). *Asia-Pacific Education Researcher*, *23*(3), 333–343. https://doi.org/10.1007/s40299-013-0108-7

*Yeung, S. S., Ng, M., Qiao, S., & Tsang, A. (2020). Effects of explicit L2 vocabulary instruction on developing kindergarten children's target and general vocabulary and phonological awareness. *Reading and Writing*, *33*(3), 671–689. https://doi.org/10.1007/s11145-019-09982-3

*Yeung, S. S., & Savage, R. (2020). Teaching grapheme–phoneme correspondences using a direct mapping approach for at-risk second language learners: A randomized controlled trial. *Journal of Learning Disabilities*, *53*(2), 131–144. https://doi.org/10.1177/0022219419894563

*Yeung, S. S. S., Siegel, L. S., & Chan, C. K. K. (2013). Effects of a phonological awareness program on English reading and spelling among Hong Kong Chinese ESL children. *Reading and Writing*, *26*(5), 681–704. https://doi.org/10.1007/s11145-012-9383-6

*Zhang, C. Y., Sang, B., Zhou, A. B., & Zhang, Y. M. (2010). 动画片语境学习对幼儿英语语音意识的促进 [effects of learning in the cartoon context on the development of phonological awareness in kindergarteners]. *Psychological Science*, *33*(5), 1042–1046. https://doi.org/10.16719/j.cnki.1671-6981.2010.05.058

*Zhang, Q. (2011). *A study on the effect of phonological awareness training on rural children's English vocabulary learning*. [Unpublished master's thesis]. Northwest A&F University.

Zhang, Z., & Peng, P. (2022). Reading real words versus pseudowords: A meta-analysis of research in developmental dyslexia. *Developmental Psychology*, *58*(6), 1035–1050. https://doi.org/10.1037/dev0001340

*Zheng, X. B., Wang, Z. K., Liu, D. M., Xu, J., Li, W. L., & Meng, X. Z. (2009). 语音训练对幼儿英语语音意识和字母知识的促进 [phonological instruction promotes the phonological awareness and knowledge of letters of children in kindergarten]. *Psychological Development and Education*, *1*, 66–71. https://doi.org/10.16187/j.cnki.issn1001-4918.2009.01.018

*Zhou, X. (2012). *The effect of English phonological awareness training on children's Chinese phonological awareness and early reading abilities*. [Unpublished master's thesis]. Shandong University.

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, *131*(1), 3–29. https://doi.org/10.1037/0033-2909.131.1.3

How to cite this article: Jiang, Y., Gai, X., Üstün-Yavuz, M. S., Zhang, M., & Thomson, J. M. (2023). A meta-analysis of training effects on English phonological awareness and reading in native Chinese speakers. *PsyCh Journal*, 1–19. https://doi.org/10.1002/pchj.675

APPENDIX A

A.1 | OVERVIEW OF STUDIES INCLUDED IN THE META-ANALYSIS

Article	Sample size	Sample characteristics	Grade	Type of program	Training strategy	Number of weeks	Number of sessions	Publication status	Outcome measure	Effect size (g)
Chen, 2021	82	Typical	University	EPA and/or phonics +other contents	Direct	>12	30–60	Unpublished	EPA (1 measure: syllable+phoneme +onset-rime)	0.461
Cheung, 1999										
G7	84	Typical	Middle school	EPA only	Direct	≤12	≤30	Peer-reviewed	EPA (1 measure: phoneme deletion)	1.120
									Reading (1 measure: word reading)	0.164
G9	79	Typical	Middle school	EPA only	Direct	≤12	≤30	Peer-reviewed	EPA (1 measure: phoneme deletion)	0.161
									Reading (1 measure: word reading)	0.160
Chow et al., 2010	34	Typical	Kindergarten	/	Indirect	≤12	≤30	Peer-reviewed	EPA (1 measure: syllable+phoneme)	0.061
									Reading (1 measure: word reading)	0.295
Chow et al., 2021										
Low level	32	Typical	Early elementary	/	Indirect	≤12	Unclear	Peer-reviewed	EPA (1 measure: syllable awareness)	−0.344
High level	40	Typical	Early elementary	/	Indirect	≤12	Unclear	Peer-reviewed	EPA (1 measure: syllable awareness)	0.495
Chung et al., 2022	99	Typical	Kindergarten	EPA and/or phonics +other contents	Direct	>12	≤30	Peer-reviewed	EPA (1 measure: syllable awareness)	1.059
									Reading (1 measure: word reading)	0.380
Duan, 2018	120	Typical	Upper elementary	EPA+phonics	Direct	>12	30–60	Unpublished	EPA (3 measures-combined: phoneme awareness, syllable awareness, onset-rime awareness)	0.932
Guo, 2017	60	Typical	Upper elementary	EPA and/or phonics +semantics	Direct	>12	≤30	Unpublished	EPA (1 measure: syllable+phoneme +rhyme)	0.732
									Reading (1 measure: word reading)	0.520
He, 2018	50	Typical	Upper elementary	EPA+phonics	Direct	>12	>60	Unpublished	EPA (1 measure: syllable+phoneme +onset-rime)	1.561
									Reading (1 measure: word reading +pseudoword reading)	1.222
Lai et al., 2009	120	Typical	Early elementary	EPA and/or phonics +other contents	Direct	≤12	≤30	Peer-reviewed	EPA (1 measure: phoneme awareness)	0.495
Li, 2012	42	Typical	Kindergarten	/	Indirect	>12	Unclear	Unpublished	EPA (1 measure: syllable+rhyme +onset-rime)	0.827
Li, 2015	72	Typical	Upper elementary	EPA+phonics	Direct	>12	30–60	Unpublished	EPA (1 measure: phoneme segmentation)	0.626
									Reading (1 measure: pseudoword reading)	0.677

(Continues)

Article	Sample size	Sample characteristics	Grade	Type of program	Training strategy	Number of weeks	Number of sessions	Publication status	Outcome measure	Effect size (*g*)
Li, 2021	80	Typical	Upper elementary	EPA and/or phonics +semantics	Direct	>12	30–60	Unpublished	EPA (1 measure: phoneme awareness)	1.648
									Reading (1 measure: word reading)	1.138
Li & Dong, 2012	8	Typical	Middle school	EPA only	Direct	≤12	≤30	Peer-reviewed	EPA (2 measures-combined: phoneme deletion, phoneme segmentation)	0.871
									Reading (2 measures-combined: word reading, pseudoword reading)	0.843
Liu, 2012	120	Typical	Upper elementary	EPA+phonics	Direct	>12	≤30	Unpublished	EPA (2 measures-combined: syllable awareness, phoneme awareness)	0.721
									Reading (1 measure: pseudoword reading)	1.349
Mak et al., 2017	249	Typical	Early elementary	EPA and/or phonics +other contents	Direct	>12	30–60	Peer-reviewed	EPA (2 measures-combined: phoneme segmentation, initial sound fluency)	0.338
									Reading (2 measures-combined: word reading, pseudoword reading)	−0.097
Shang, 2012	100	Typical	Middle school	EPA only	Direct	>12	Unclear	Unpublished	EPA (1 measure: syllable+phoneme +onset-rime)	0.446
Shuai, 2016	100	Typical	Middle school	EPA+phonics	Direct	>12	Unclear	Unpublished	EPA (1 measure: syllable+phoneme +onset-rime)	−0.554
Siu et al., 2018	25	Typical	Early elementary	EPA and/or phonics +other contents	Direct	≤12	30–60	Peer-reviewed	EPA (1 measure: syllable+phoneme)	0.658
									Reading (2 measures-combined: word reading-consistent, word reading-inconsistent)	0.608
Sun et al., 2015	80	Typical	Early elementary	EPA only	Direct	≤12	≤30	Peer-reviewed	EPA (6 measures-combined: rhyme detection, syllable synthesis, syllable segmentation, phoneme synthesis, phoneme segmentation, initial phoneme deletion)	0.951
									Reading (2 measures-combined: non word recognition, non-word reading)	0.666
Tam & Leung, 2019	202	Special	Middle school	EPA and/or phonics +other contents	Direct	>12	>60	Peer-reviewed	EPA (2 measures-combined: initial sound, blending phonemes)	0.588
Wang, 2012	96	Typical	University	EPA only	Direct	>12	≤30	Unpublished	EPA (1 measure: syllable+phoneme +onset-rime)	0.834
Wang, 2017	90	Typical	Early elementary	EPA+phonics	Direct	≤12	30–60	Unpublished	EPA (1 measure: phoneme+rhyme)	0.369
									Reading (1 measure: word reading)	0.534
Wang, 2019	21	Typical	Early elementary	EPA+phonics	Direct	≤12	≤30	Unpublished	EPA (1 measure: syllable+phoneme +rhyme)	0.683

Article	Sample size	Sample characteristics	Grade	Type of program	Training strategy	Number of weeks	Number of sessions	Publication status	Outcome measure	Effect size (*g*)
Xiong, 2014	38	Special	Upper elementary	EPA+phonics	Direct	≤12	≤30	Unpublished	EPA (4 measures-combined: rhyme awareness, initial phoneme recognition, phoneme detection, phoneme counting)	0.712
Xu, 2013	70	Typical	Early elementary	EPA only	Direct	≤12	≤30	Unpublished	EPA (1 measure: phoneme+rhyme +onset-rime)	0.575
Xu, 2021	96	Typical	University	EPA only	Direct	≤12	Unclear	Peer-reviewed	EPA (1 measure: syllable+phoneme +onset-rime)	0.453
									Reading (1 measure: word reading)	0.105
Yeung & Savage, 2020	71	Special	Early elementary	EPA+phonics	Direct	≤12	≤30	Peer-reviewed	EPA (1 measure: phoneme elision)	−0.170
									Reading (2 measures-combined: word reading, word attack)	−0.174
Yeung et al., 2013	76	Typical	Kindergarten	EPA and/or phonics +semantics	Direct	≤12	≤30	Peer-reviewed	EPA (5 measures-combined: syllable deletion, rhyme deletion, rhyme generation, phoneme identification, phoneme deletion)	0.831
									Reading (2 measures-combined: word reading, pseudoword reading)	0.288
Yeung et al., 2020	157	Typical	Preschool	/	Indirect	≤12	≤30	Peer-reviewed	EPA (2 measures-combined: syllable deletion, phoneme identification)	0.240
Zhang, 2011	56	Typical	Upper elementary	EPA+phonics	Direct	>12	>60	Unpublished	EPA (1 measure: syllable+phoneme +onset-rime)	0.977
Zhang et al., 2010										
Low grad	67	Typical	Preschool	/	Indirect	>12	>60	Peer-reviewed	EPA (1 measure: phoneme awareness)	0.977
Middle grade	73	Typical	Preschool	/	Indirect	>12	>60	Peer-reviewed	EPA (1 measure: phoneme awareness)	0.914
High grade	83	Typical	Kindergarten	/	Indirect	>12	>60	Unpublished	EPA (1 measure: phoneme awareness)	1.135
Zheng et al., 2009	63	Typical	Preschool	EPA and/or phonics +semantics	Direct	≤12	>60	Peer-reviewed	EPA (3 measures-combined: syllable synthesis, rhyme detection, phoneme detection)	0.917
Zhou, 2012	69	Typical	Early elementary	EPA and/or phonics +semantics	Direct	>12	30–60	Unpublished	EPA (1 measure: phoneme+onset-rime)	1.027