

Case Study 1: An Evidence-Based Practice Review Report

Theme: School Based Interventions for Learning

Does incorporating a Visual Phonics strategy into school literacy programs for children who are deaf or hard of hearing have a significant impact on their reading achievement?

Summary

Visual Phonics (VP) is a system of hand gestures and symbols that represent the phonemes of spoken English. It has been widely used as a strategy for literacy instruction for over 20 years, however even though it is a huge part of the literacy strategy across the US and in many schools within the UK, there appear to be comparatively few studies which have attempted to critically evaluate its efficacy. The present systematic literature review synthesised the current research, exploring the impact of using the VP strategy to develop the literacy skills of children who are deaf or hard of hearing (DHH). Five studies were reviewed in depth using Gough's (2007) weight of evidence framework: one was assigned a low rating; three were assigned a medium rating; one was assigned a high rating. A critical analysis of the research findings revealed some benefits to using VP, however, a closer look at the research designs also exposed a number of weaknesses in their methodology. Recommendations for implementation and future research are outlined and discussed.

Introduction

Visual Phonics (VP) was developed by a speech and language pathologist, Dave Krupke, in the 1970s. Krupke had identified that children who struggle to read, often benefit from literacy strategies which represent sounds in a more concrete and visible way (Montgomery, 2008). His ideas were later developed by a parent of a deaf child, in order to support children who are deaf or hard of hearing (DHH).

VP uses a multisensory approach to literacy instruction (Montgomery, 2008). Each of the 45 phonemes of English are represented by a distinct hand-shape cue and corresponding written symbol. Although VP is a visual and written system, it cannot be categorised under a communication system such as sign, speech or cueing.

Rather, it is a strategy that provides a visual, written and kinaesthetic cue to what hearing individuals hear, when they are exposed to English phonics (Musselman, 2000). VP requires 15 hours (i.e. two full days) of training to become a qualified instructor. One of the main benefits of VP is that it can easily be integrated into any traditional reading program to create a comprehensive curriculum (Smith & Wang, 2010). This makes it very time and cost effective as it does not require significant changes to be made to existing programs. The intervention can be used at any point during the school day. Trained teachers and teaching assistants are encouraged to use the gestures consistently across lessons, be it during small group work, individual work or at the whole class level. The intervention should not be confined to literacy lessons, but instead to all subjects within the curriculum to maximise student exposure.

There are other phonics interventions available, which have adopted the use of a gestural component to support the child's memory and storage of English sounds.

The Jolly Phonics method uses a synthetic approach that teaches children a series of actions which represent each of the 42 letter sounds. Although the use of actions can make the sounds more salient to the child (e.g. pretending to be a train by moving arm forward and backward whilst saying 'ch' 'ch'), the actions provide little or no information with regards to the place of articulation (Campbell, Torr, & Cologon, 2014). In contrast, the VP method uses hand shapes that are kinaesthetically similar to the method in which the sound is produced; the shape of the tongue, mouth and throat. For example, the /p/ sound is produced by using a flat hand facing out at the chin then moving forward and opening quickly (i.e. to represent a puff of air blowing out of the lips). The hand cue will represent the phoneme sound, rather than the orthographic representation, allowing for a differentiation between the hard /c/ in "cool" (i.e. hand at side of mouth with fingers in a C-like shape, moving forward) and the soft /c/ in "cereal" (i.e. index finger moving slightly out from lips). As a result, deaf pupils like hearing pupils, learn that /c/ can have two sounds. Similarly, the written symbols are visual representations of the position and/or movement of the hand to create the cue. An important factor to consider when deaf children are learning to read is the fact that the English language has a finite set of speaking and spelling rules (Woolsey, Satterfield & Roberson, 2006). According to Nussbaum, Waddy-Smith and Wilson (2004) there are 396 ways to spell all the words in English. The advantage of teaching a coding system such as VP, is that it provides the deaf child with rules of phoneme combinations which they can use to eliminate some impossible combinations of sounds (Narr, 2006).

Theoretical Basis

DHH student's literacy achievements tend to lag behind their hearing peers (Smith & Wang, 2010), which has led researchers to investigate these abilities and propose numerous theoretical frameworks to explain discrepancies in performance. There appears to be a clear divide between 1) theorists who believe that the process of reading acquisition in the DHH population is radically *different* from that of the hearing population, so that practices which support phonemic awareness are meaningless and inappropriate (Musselman, 2000) and 2) theorists who believe that the literacy skills between these two populations is very *similar* so that DHH children follow the same trajectory and sequence of skills when learning to read as that of hearing children (i.e. in terms of their ability to make rhyme judgements, identify individual sounds, blend sounds and decode unfamiliar words). Therefore, under the similarity assumption, DHH children can be taught to read in traditional ways (e.g. King & Quigley, 1985; 1998; 2001)).

The VP strategy adopts the latter view, supporting the *qualitative similarity hypothesis* (Paul, 2008), which assumes that the reading development of DHH is qualitatively similar to that of hearing peers, but quantitatively delayed (Wang, Trezek, Luckner & Paul, 2008). There is a huge body of research to support this position. For example, studies carried out by Hammill (2004) indicated that there is a link between phonological awareness and reading skills in hearing children and a number of more recent studies suggest that this relationship also exists in the development of accomplished deaf readers (e.g. Dyer, MacSweeney, Szczerbinski, Green & Campbell, 2003). Indeed, many studies have identified that the ability to manipulate phonological information is the key skill which distinguishes poor or average DHH readers from proficient ones (e.g. Perfetti & Sandak, 2000). The VP

strategy acknowledges the importance of the ability to manipulate phonological information in literacy development and provides children who are DHH with accessible mental images of phonemes. As a result, the DHH child is no longer relying solely on them listening to the sounds. Such a phonological representation of English, directly assists the students in manipulating the phonemes and acquiring phonemic awareness and phonics skills (Schirmer & McGough, 2005). As children who are DHH will struggle to access phonological representations of the sounds, the VP program takes a novel approach to literacy intervention, by adopting a multimodal representation of phonemes that can be used to support their access to the phonological components of the English language.

Rationale

Being able to read and write has become an enormously important skill, particularly in today's society where young people rely heavily on written text to communicate (e.g. via phone messages, email and social media). Working with individuals with literacy difficulties is an integral part of the Educational Psychologist's (EP) role (Meece & Miller, 2001). It is therefore pertinent that these professionals know the implications and potential benefits/pitfalls of implementing an array of literacy programmes that are available to DHH children in schools and children's centres. EPs can then advise these institutions on effective strategies which have a sound evidence base in psychological research.

As an intervention strategy, VP appears to show great potential, so much so that it has been widely recommended (i.e. particularly within the US) and is generally viewed by teachers as a robust strategy for raising achievement scores of DHH students (Luetke-Stahlman & Nielson, 2003). This appears to be the case, despite

there being very little empirical evidence to support these views. In fact, many researchers have warned against advising this strategy as they argue that VP has “virtually no research base” (e.g. Woolsey, Satterfield & Roberson, 2006). Teachers’ “belief” or professional intuition with regards to the efficacy of VP cannot be sufficient when selecting an educational programme as the implications of suggesting such programmes with little evidence base could significantly impede student progress. In light of these findings, the present paper has carried out a systematic review of the literature in an attempt to shed light on the following question: ***“Does incorporating the VP strategy into school literacy programs for students who are DHH have a significant impact on their level of reading achievement?”***

Critical Review of the Evidence Base

Literature Search

Initial searches were carried out in January 2016 using the electronic databases PsychInfo, ERIC and PubMed. Search items used to retrieve the target studies are listed in Table 1.

Table 1: Search terms applied to databases.

Databases	Search terms			
PsychINFO	Literacy	Visual Phonics	Deaf	Children
ERIC	Reading	Visuals	Hard of Hearing	School
PubMed	Intervention	Gesture		Students

Note: A broad search was carried out, using not only the search term “Visual Phonics” but also terms such as “gesture” and “visuals”. The rationale behind using broader search terms was that many studies do not specifically mention VP in their titles, despite this strategy being an integral part of the intervention. This is due to the fact that VP is rarely used on its own, but rather in conjunction with wider literacy interventions.

Inclusion and exclusion criteria

Figure 1 provides a breakdown of the number of studies found within each search engine. The initial search across the databases generated a large number, equalling 279 studies. Two hundred and sixty four studies were immediately excluded due to a lack of relevance to the current research question. Of the 61 remaining studies, 40 were excluded for not meeting the inclusion criteria listed in Table 2 and seven were excluded as duplicates. A total of 14 articles were read in full text, of which nine were excluded (Appendix A). A total of five studies were then used in the final review question (see Table 2). Details of the studies are listed in Table 4.

Figure 1: Flowchart depicting the screening process of articles retrieved from the database search:

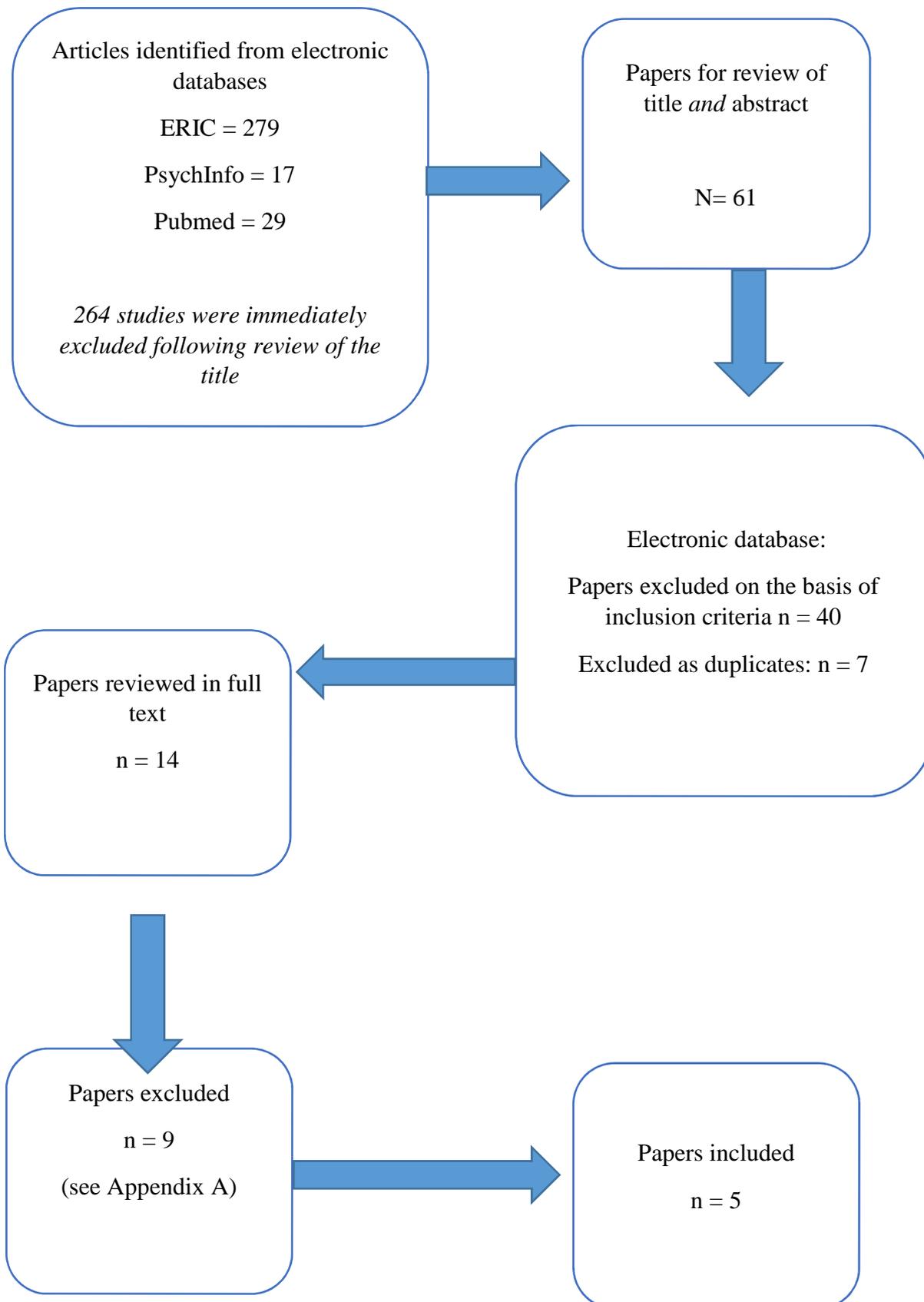


Table 2: Inclusion and exclusion criteria

	Inclusion Criteria	Exclusion Criteria	Rationale
1. Type of Publication	<p>a) The study must be a peer reviewed-journal or been submitted for peer review.</p> <p>b) The study contains primary empirical data.</p>	<p>a) The study is not peer reviewed or study has not been submitted for review.</p> <p>b) The study does not collect primary empirical data (e.g. a review paper).</p>	<p>The purpose of a peer review is for reviewers to assess the quality of the study. The research will therefore have undergone a certain level of scrutiny ensuring that it meets the required standards.</p> <p>By using empirical data, it is possible to explore and review the impact of Visual Phonics</p>
2. Language and Setting	<p>a) Study must be written in English. There are no restrictions as to where the data was collected. Non-UK studies were also included.</p>	<p>a) The study is written in a language other than English.</p>	<p>Saves the reviewer time and resources.</p>
3. Participants	<p>a) Participants must be children who are deaf or hard of hearing.</p>	<p>a) Children who are not deaf or hard of hearing</p>	<p>Children who are deaf or who have a significant hearing loss are the population of interest for this paper.</p>
4. Intervention	<p>a) The study implements the Visual Phonics Programme</p>	<p>a) The study does not employ the Visual Phonics Programme</p>	<p>The purpose of the present article is to review the effectiveness the Visual Phonics Programme</p>

	Inclusion Criteria	Exclusion Criteria	Rationale
	b) The study focusses primarily on literacy difficulties.	b) The study focusses primarily on other measures of achievement (e.g. vocabulary development).	Enhancing literacy skills is the primary focus of the present paper.
	c) Intervention must be school-based	c) Intervention is based at home or any place other than school	School-based interventions are the focus of the present review.
5. Type of Design	a) The study must report outcome data, pre and post intervention. Alternatively the study must employ a group design which reports between group outcomes	a) Outcome data is only collected after the intervention has been completed and/or the study does not employ a group design.	Having both pre and post measures or a comparison between groups allows the review to explore the impact of the literacy intervention over time.

Table 3: List of final five studies included in the literature review:

Included studies

Tucci, S. and Easterbrooks, S. (2013). A Syllable Segmentation, Letter-Sound, and Initial-Sound Intervention With Students Who Are Deaf or Hard of Hearing and Use Sign Language. *The Journal of Special Education*, 48(4), 279-289.

Narr, R.F. (2008) Phonological Awareness and Decoding in Deaf/Hard-of-Hearing Students Who Use Visual Phonics. *Journal of Deaf Studies and Deaf Education* 13 (3), 406-416

Trezek, B., Wang, Y., Woods, D., Gampp, T. and Paul, P. (2007). Using Visual Phonics to Supplement Beginning Reading Instruction for Students Who Are Deaf or Hard of Hearing. *Journal of Deaf Studies and Deaf Education*, 12(3), 373-384.

Trezek, B. & Wang. Y. (2005). Implications of Utilizing a Phonics-Based Reading Curriculum With Children Who Are Deaf or Hard of Hearing. *Journal of Deaf Studies and Deaf Education*, 11(2), 202-213.

Trezek, B. & Malmgren (2005). The Efficacy of Utilizing a Phonics Treatment Package with Middle School Deaf and Hard-of-Hearing Students. *Journal of Deaf Studies and Deaf Education*, 10(3), 256-271.

Table 4: Detailed summary of each of the five studies which appear within the present paper

Study ID	Author(s) and aim (relevant to review)	Participants, Range of Hearing Loss and Method of Communication	Design	Intervention Type	Duration of Intervention	Measures	Key findings
1	Tucci and Easterbrooks (2015) To investigate syllable segmentation, identification of letter-sound correspondence and initial sound identification in DHH who are receiving Foundations curriculum supplemented by Visual Phonics	3 kindergarten students who are DHH 103dB and 110dB respectively for each child Total Communication, ASL, Pidgin, Simultaneous Communication	Single-case designs- multiple baseline comparisons	Foundations curriculum supplemented by Visual Phonics.	One hour per day, four days per week, for 25 weeks.	Performance on syllable segmentation, letter-sound correspondence and initial sound identification	Increased performance in letter-sound correspondence and initial-sound identification when Foundations curriculum was supplemented by Visual Phonics

Table 4: Detailed summary of each of the five studies which appear within the present paper

Study ID	Author(s) and aim (relevant to review)	Participants, Range of Hearing Loss and Method of Communication	Design	Intervention Type	Duration of Intervention	Measures	Key findings
2	Narr (2008) To examine the relationship between performance on literacy tasks and length of time using Visual Phonics in a cohort of DHH children.	10 DHH kindergarten through Grade 3 students 68dB- 118dB unaided puretone average in better ear Combination of ASL and Sign Supported English	Correlational design	Teachers used an "eclectic" approach to literacy instruction (e.g. integrated resources from programmes such as Reading A-Z and use of levelled readers by Houghton-Mifflin, Rigby and Wright Group. Supplemented by Visual Phonics.	Intervention ranged from one year to three years	Picture rhyme task and decoding task (i.e. based on Dyer et al., 2003 and Sterne & Gotswami, 2000), used to assess phonological awareness and decoding respectively	Students were able to use phonological information to make rhyme judgements and decode words- no significant interaction between reading ability and length of time using Visual Phonics instruction

Table 4: Detailed summary of each of the five studies which appear within the present paper

Study ID	Author(s) and aim (relevant to review)	Participants, Range of Hearing Loss and Method of Communication	Design	Intervention Type	Duration of Intervention	Measures	Key findings
3	Trezek, Wang, Woods, Gampp and Paul (2007)	20 DHH kindergarten and 1st grade students 70dB - 112dB unaided pure-tone average in better ear. Study also included five children with cochlear implants. Total Communication Approach	Pre/post quasi experimental design	LACES, phonics-based reading curriculum supplemented by Visual Phonics	One school year	Scores on 1) a post-test reading achievement measure and 2) Dominie Reading and Writing Assessment Portfolio	Improvements in beginning reading skills

Table 4: Detailed summary of each of the five studies which appear within the present paper

Study ID	Author(s) and aim (relevant to review)	Participants, Range of Hearing Loss and Method of Communication	Design	Intervention Type	Duration of Intervention	Measures	Key findings
4	Trezek and Wang (2006). To evaluate the improvements made by deaf and hard of hearing kindergarten and first-grade students when given 1 year of instruction from a phonics-based reading curriculum supplemented by Visual Phonics	13 DHH kindergarten and 1st grade students 82dB pure tone average in better ear. Study also included two children with cochlear implants. Total Communication Approach	Pre/post quasi experimental design	LACES,phonics-based reading curriculum supplemented by Visual Phonics	One school year	Four subtests of the WIAT-II (i.e. word reading, non-word decoding, reading comprehension)	Increased performance on reading comprehension, non-word reading and word reading

Table 4: Detailed summary of each of the five studies which appear within the present paper

Study ID	Author(s) and aim (relevant to review)	Participants, Range of Hearing Loss and Method of Communication	Design	Intervention Type	Duration of Intervention	Measures	Key findings
5	Trezek and Malgrem (2005). To evaluate the efficacy of implementing a phonics treatment package which includes Visual Phonics to a group of middle-school-aged students	22 DHH middle school students 8dB-110dB pure tone average in better ear. Study also included one child with auditory neuropathy. Method unknown	Pre/post quasi experimental design	<i>Decoding A</i> in combination with Visual Phonics, Baldi, pictorial glossary and one of three reading curriculums (i.e. Reading Milestones, Pair-It books or Wildcat Series)	Eight weeks	Scores on: 1) a curriculum -based post-test and 2) a generalisation test created specifically for the study	Increased performance on grapheme-phoneme identification both in isolation and within words. Better performance on word decoding tasks. These results did not appear to be affected by degree of hearing loss

Critical appraisal for quality and relevance

Evaluation

The quality and relevance of the five studies included within the present paper were appraised using the Weight of Evidence (WoE) approach (Gough, 2007). Using this framework it was possible to review four aspects of the evidence in a very systematic way: WoE A appraised the methodological quality of the research by rating the way in which the study was executed; WoE B appraised the methodological relevance of the study; WoE C appraised the appropriateness of the study aims in relation to the present research question; WoE D provided an overall rating by averaging the weights for each factor. For overall ratings for each of the studies, see Table 5. For further information on individual appraisals, see Appendix C.

In order to evaluate WoE A (methodological quality), the execution of the studies was judged using criteria set out by either Horner et al., (2005) (i.e. for the single case studies or small sample sizes) or Gersten et al., (2005) (i.e. for the quasi-experimental pre/post design studies). Complete coding protocols are listed in Appendix B. Protocols for judging WoE B and WoE C were designed by the reviewer and developed specifically for the present paper (see Appendix C).

Table 5: Overview of the WoE ratings for each of the five studies.

	Methodological quality (WoE A)	Methodological relevance (WoE B)	Relevance to the present research question (WoE C)	Overall weighting (WoE D)
Tucci and Easterbrooks (2015)	Medium (2)	Medium (2)	Medium (2)	Medium (2)
Narr (2008)	Low (1)	Low (1)	Medium (2)	Low (1)
Trezek, Wang, Woods, Gampp & Paul (2007)	High (3)	Medium (2)	High (3)	High (3)
Trezek and Wang (2006)	Medium (2)	Medium (2)	High (3)	Medium (2)
Trezek and Malgrem (2005)	Medium (2)	High (3)	Medium (2)	Medium (2)

Participants

The total number of participants within the five studies was 68 (i.e. ranging from sample sizes between 3-22 participants). Only two of the three participants in the Tucci and Easterbrooks (2013) study were included in the present paper, as the third participant was not exposed to VP. Potential generalizability may be limited due to the small number of participants within each of the studies. The age of the participants was between 5 years; 0 months and 9 years; 10 months. The studies recruited participants from a range of ethnic backgrounds, namely Caucasian, Asian, Hispanic and African American. The gender of participants was not listed in all of the studies and can therefore not be reported. The level of hearing difficulties varied both within and across studies ranging from “slight” to “profound” (i.e. 30dB- 120dB).

Fourteen out of the 68 participants had a cochlear implant and 1 had a diagnosis of auditory neuropathy. The method of deaf communication used also varied within and across studies. These included the Total Communication, Oral/Aural, Sign Supported English and American Sign Language (ASL) approaches (see Table 4). Descriptions of how participants were selected and exact measures of the level of hearing difficulty were reported in four out of the five studies (i.e. Tucci & Easterbrooks (2015), Trezek et. al., (2007), Trezek & Wang (2006) and Trezek and Malgrem (2005)) which contributed to higher WoE A ratings.

Intervention

Teacher Training

All of the studies employed experienced DHH teachers (i.e. between 2 and 23 years of work experience), trained in VP to carry out the intervention, with the exception of Tucci and Easterbrooks (2015) in which the primary author participated in the implementation of the programme, alongside the key teacher. In effect, all of the studies received either a high or medium rating for WoE C, which highlights the need for experienced teachers when working with students within the deaf community.

Teachers who had not already been trained in VP received between a total of 8 and 16 hours of training and all of the studies (i.e. with the exception of Narr (2008) who collected data post-intervention) also regularly monitored the implementation throughout the intervention process. The importance of regular monitoring by the researchers is stated in WoE C, which emphasises the importance of fidelity; by ensuring that participants across groups are receiving equal VP input and that instructors have ample opportunities to ask questions if necessary. Trezek and Wang (2006) took a very systematic approach to monitoring, using regular focus

groups, classroom observations and refresher sessions. This led to a WoE C rating of “high” for this particular study.

Intervention Programmes

Each of the five studies used a range of literacy intervention programmes in conjunction with the VP strategy. Some of the interventions targeted the basic foundation skills of literacy such as phonemic awareness and phonics skills (e.g. Trezek & Malmgren, 2005) whereas others focussed on higher level skills such as fluency, vocabulary and comprehension (e.g. Trezek & Wang, 2006).

Some of the studies used one main programme in addition to VP (e.g. Trezek et al., 2007), whereas others used a more “eclectic approach”, by incorporating numerous other programmes (e.g. Narr, 2006). For example, Trezek and Malmgren (2005) employed five different interventions, including a computer software programme.

Studies which limited the number of additional strategies were rated more highly in WoE C as it meant that changes in outcome measures could be more confidently attributed to VP, due to a reduction in confounding variables. A number of the programmes (e.g. such as LACES literacy programme in Trezek et al., (2007)) and assessments (e.g. Generalisation Test in Trezek and Malmgren, 2005) were either developed by professionals working within the school districts or by the researchers themselves. From an educational perspective there are clear benefits to developing and tailoring programmes to meet the needs of the individual children (i.e. particularly when we consider how varied the individual needs of the DHH population can be). However, from a research perspective, such programmes are difficult to monitor and control as one cannot obtain reliability and validity measures.

Length of Intervention

The amount of time participants were exposed to VP varied greatly across studies. The least intensive programmes with regards to time allocation were Trezek and Wang, (2006), which was carried out over eight weeks. The study by Narr (2008) assessed students who had undergone VP instruction for between one and three years. However, this study did not employ a pre-post test design and the researchers had no control of the independent variable. This contributed to a low WoE A rating, which stresses the importance of having measures at numerous data points (i.e. including baseline measures) and evidence of systematic manipulation of the independent variable. Ultimately, researchers found no relationship between performance on reading ability and length of time in literacy instruction with VP.

Design

All five of the papers used a quantitative approach to research. Multiple baseline/single case designs (Tucci et al., 2013), pre-test post-test designs (Trezek & Malmgren 2005; Trezek & Wang, 2006; Trezek et al., 2007) and correlational designs were used in the present review. The WoE B was designed to reflect the suitability of the methodology used within each of the studies. However, there are a number of important elements of the study designs which require further discussion. For example, two out of the three pre-test post-test studies (i.e. Trezek & Wang, 2006; Trezek et al., 2007) did not employ an active control group. This makes it very difficult to ascertain whether it was the literacy programme, the use of VP system or a combination of both that had a positive effect on the students' literacy skills. Furthermore, the sample size of each of the pre-post design studies was small, with the largest sample size reaching 22 participants (i.e. Trezek & Malmgren, 2005). According to Cohen (1992) the required group size to pick up large effect sizes (ES)

(i.e. of 0.8) is 26 participants (when comparing means at the .05 significance level). None of the studies in this review met this criteria. Therefore one must be cautious when interpreting results as the studies lack power. Studies which lack power are likely to produce non-significant results, even when ES may be large. Inadequate sample sizes undermine ES results and compromise generalizability. One must also consider that the administration of the pre-and post-test measures was the responsibility of the teachers in two out of the three studies (i.e. Trezek & Wang, 2006; Trezek et al., 2007). This could potentially have caused bias in administering and scoring of the test measures, the importance of which is reflected in the WoE A criteria. Administration of pre-post measures within the Trezek and Malmgren (2005) study was carried out by the first researcher. Although this may reduce bias in the ways mentioned above, the researcher was aware of the group assignment posing further threats to the internal validity of the study.

Finally, all of the studies were carried out using convenience sampling within specialist schools for deaf pupils within the US. Although random assignment of participants is viewed as methodologically superior to convenience sampling (Kratochwill, 2003), DHH children are often allocated to specific teachers to meet Individual Education Plan (IEP) targets. This makes random sampling and assignment both logistically and pragmatically difficult. This was taken into account when devising the methodological WoE criteria and studies were not penalised for assigning participants to conditions, based on teacher suitability. Trezek and Malmgren (2005) took measures to increase the internal validity of the study by using a coin toss to determine the assignment of the participants to either the experimental or the control group, which was reflected in the higher rating on WoE B.

Measures

The outcome measures varied greatly across studies; from standardised assessments (e.g. Trezek & Wang (2006) used the WIATII and Trezek et al., (2007) used the Reading and Writing Assessment Portfolio) and informal assessments, developed by the researchers (e.g. Tucci & Easterbrooks (2013) used PowerPoint slides to assess syllable segmentation and rhyme judgement). The WoE C criteria stress the importance of using a range of outcome measures when evaluating progress in literacy skills. This is due to the complexity of literacy development, which requires a number of prerequisite skills to develop effectively (e.g. Strong & Stuckless, 1995). As mentioned above, Trezek et. al., (2007) and Trezek and Wang (2006) used a number of formal assessments, covering a range of skills to assess student progress. These included phonemic awareness, rhyme judgment, syllable segmentation, writing, spelling and whole word recognition; to name a few. As a result, these studies were awarded “high” ratings for WoE C. It is important to note that although methods for assessing participants at pre and post-test phases were consistent across each of the studies, responses to assessment questions were provided differently depending on the students preferred mode of communication. For example in the study by Trezek et al., (2007), students who were oral/aural communicators used a combination of speech and VP to respond to test items, whereas students who were simultaneous communicators responded by using a combination of speech, sign language and VP to respond. This meant that the standardised assessments (i.e. such as the Dominic Reading and Writing Assessment Portfolio (DeFord, 2001) had to be adapted to suit the students’ needs, which will ultimately have had detrimental effect on the validity of the results. Trezek

and Wang (2006) explicitly compare participant scores to that of the national population average for students who are DHH, which provided the study which contributed to the high WOE C rating.

Findings

Due to the variation in research methods and designs, effect sizes (ES) needed to be calculated differently for each of the studies (i.e. either using Cohen's *d* (Cohen 1988), Standard Mean Difference (SMD) (Morris, 2007) or Percent all non-overlapping data (PAND) (Parker et al., 2011). In order to make comparisons between the studies, ES were provided with a label of "small", "medium" or "large", depending on the rating criteria listed in Appendix D.

ES for each of the studies are listed in Table 6. The majority of ES outcomes in the studies were large. This suggests that VP has a high efficacy in developing literacy skills in the DHH population. This was particularly evident in studies which measured early literacy skills. For example, Tucci and Easterbrooks found medium to high ES for performance on syllable segmentation tasks and Narr (2008) reported high ES for tasks requiring knowledge of phonological awareness and rhyme onset. However, it must be noted that there were no pre-test scores available in the latter study (i.e. the ES were calculated through the relationship between the length of time in literacy instruction and performance on phonological and decoding tasks). Therefore scores must be interpreted and compared with caution, as they cannot be directly compared to studies which have used a control group. The significance of having a control group is reflected in WoE B. Due to a lack of control group, Narr (2008) received a rating of "low". Higher order literacy skills were demonstrated by Trezek and Wang

(2006), who found high ES for in word reading, non-word reading and reading comprehension tasks. The study carried out by Trezek and Malgrem (2005) revealed significantly higher mean scores on post-tests than their matched controls, which is reflected in the high ES. Researchers reported scores within the treatment group which showed z scores that were nearly three standard deviations higher than the controls. Furthermore, participants appeared to make gains in word decoding skills, regardless of their level of hearing impairment. This appears to suggest that VP benefits not only students with a range of hearing difficulties but also children with a range of hearing devices (i.e. including cochlear implants and hearing aids).

Teacher reports collected from three of the five studies (i.e. Trezek et. al., (2007), Trezek and Wang (2006) and Trezek and Malgrem (2005)) appear to suggest that the majority of administrators found the intervention simple to use, engaging to students and easy to integrate into their existing literacy curriculum. Whether or not teacher's perceptions of VP interventions were positive in the other two studies is unclear as no information was reported. Studies which discussed the social importance and practicality of the intervention received higher WoE A ratings.

Furthermore, all five of the studies have demonstrated that VP can benefit the child's ability to retain new information, the speed at which the child is able to make sound-print connections and the child's ability to identify rhyme without accessing auditory information. The VP method has been demonstrated to take advantage of muscle memory and allow DHH children to *see* and *feel* the phonological concepts which underpin the foundations literacy development; that is phonological awareness.

Table 6: Effect sizes for data from each of the five studies

Authors	Outcome measure	Method of effect size calculation	Effect size
Tucci and Easterbrooks (2015).	<i>Initial Sound Identification</i>	Calculated using (Perla, Provost & Murray, 2011)	<i>PAND = 50%</i> (Josephine) <i>PAND = 86%</i> (Ben)
	<i>Letter Sound Correspondence</i>		<i>PAND = 55.6%</i> (Josephine) <i>PAND = 55%</i> (Ben)
	<i>Syllable Segmentation</i>		<i>PAND = 77.8%</i> (Josephine) <i>PAND = 75%</i> (Ben)
Narr (2008)	<i>Phonological Awareness Test</i>	Extracted from the paper	<i>d = 1.84</i>
	<i>Decoding test</i>	Extracted from the paper	<i>d = 1.60</i>
Trezek, Wang, Woods, Gampp and Paul (2007)	<i>Sentence Writing Phoneme</i>	Extracted from the paper	<i>d = 1.37</i>
	<i>Sentence Writing Spelling</i>	Extracted from the paper	<i>d = 1.49</i>

Table 6: Effect sizes for data from each of the five studies

Authors	Outcome measure	Method of effect size calculation	Effect size
Trezek and Wang (2006)	<i>Phonemic Awareness Segmentation</i>	Extracted from the paper	$d = 1.21$
	<i>Phonemic Awareness Deletion</i>	Extracted from the paper	$d = 1.47$
	<i>Phonics Onsets</i>	Extracted from the paper	$d = 1.30$
	<i>Phonics Rimes</i>	Extracted from the paper	$d = 1.17$
	<i>Word Reading</i>	Extracted from the paper	$d = -1.60$
	<i>Pseudoword Decoding</i>	Calculated using Becker (1988)	$SMD = 1.64$
	<i>Reading Comprehension</i>	Calculated using Becker (1988)	$SMD = 1.31$
Trezek and Malgrem (2005)	<i>Curriculum based pre and post test</i>	Calculated using Morris (2007)	$SMD = 2.54$

Conclusions

There appears to be a good case for implementing VP as a tool to promote the acquisition of grapheme-to-phoneme correspondence in children who are DHH. The present review aimed to evaluate in depth, the effectiveness of VP with this unique population. The large variation in overall WoE ratings across the five studies (i.e. one high, three medium and one low) reflects the complexity of the DHH population and the lack of controlled variables. Children who are DHH will differ in their method of communication and in their approach to literacy learning. Furthermore, teasing out the direct effects of VP on literacy skills can be difficult as it is always delivered as part of a larger treatment package.

Tucci et al., (2014) suggest that the weak evidence base for DHH literacy interventions can be explained through 1) a low incidence of DHH training programmes 2) a low existence of DHH teachers and 3) a limited number of researchers interested in the classroom interventions of children who are DHH. More work needs to be done to address these issues and increase access to alternative strategies and effective curricula materials.

In summary, the following review provides promising support for the use of VP with students who are DHH. Further research is required to evaluate whether the evidence base is sufficient to suggest using such strategies throughout schools in the UK.

Recommendations

Practical Recommendations

The present review highlights the ease with which VP can be integrated into any literacy curriculum. Nevertheless, there are a number of practical issues which need

to be taken into account before adopting this approach. For example, although instructors only need to undertake 15 hours of direct VP instruction by a certified trainer to become qualified, the intervention requires a number of hours of practice before the gestures become fluid and engrained into memory. Careful consideration must therefore be taken, when selecting instructors (e.g. Are they motivated to participate in the intervention? Do they play a lead role in literacy teaching? Are they willing to take the time to integrate VP into the existing curriculum?). One must also consider that VP will be most effective, when the students have a high level of exposure to the gestures (e.g. Waddy-Smith & Wilson, 2003). It is therefore recommended that a number of staff undergo the training (i.e. and not solely the class teacher), in order to increase consistency for the students. Finally, research suggests that parents play a huge role in developing their child's literacy skills (e.g. Schirmer & McGough, 2005). It is recommended that schools involve parents and siblings as much as possible in the process (e.g. through home school communication books and parent-teacher meetings) in order to increase the chances of skills being generalised.

Recommendations for Future Research

In order to ascertain a deeper understanding of the critical components of VP intervention, there are a number of recommendations to be made for future research in this area. First, one would need to control for the method of communication, as this varied greatly between participants in the present studies. Some participants used gesture to communicate (e.g. ASL) where as others used the aural/oral method. We know that the syntax of ASL varies greatly to that of written English (e.g. Van Staden, 2013) and will therefore have an effect on how DHH children process

and produce written information. Follow up measures would also need to be taken and collected. Out of the five studies described in the present paper, none reported taking any follow up measures. Although Trazek and Malgrem (2005) included a generalisation task, assessing the lasting effects of VP intervention would provide stronger support for its implementation. It is also recommended that the authors control more closely for researcher bias. Three out of the five studies reviewed in the present paper were co-written by the same author (i.e. Beverley Trezek). This could affect the validity of the research findings. However this is testament to how few studies on VP exist in the current literature and how much still needs to be done. The benefits of VP are noteworthy, however, the majority of the research into VP has been carried out in the US. Further research is required to assess whether VP would be as effective in the UK; delivered in conjunction with a British literacy curriculum. It is also important to note that parental involvement was not considered in any of the five studies. The deaf community are reported to have strong views when it comes to choosing a method of communication for their children (e.g. Strong & Stuckless, 1995). Reporting and monitoring these differences may provide more insight into some of the barriers and facilitators of integrating literacy strategies into the school curriculum.

Teachers within two of the studies reported that the VP strategy allowed them to teach portions of their existing phonics curriculum that they previously struggled to address. It is therefore recommended that further research be carried out in this area, as it would provide information with regards to the development of VP.

Finally, one must consider the significance of evidence based practice when implementing an intervention. Survey studies, which have interviewed teachers and staff working with deaf students have shown that phonology instruction is rarely

deemed important enough to be integrated into their phonology instruction (e.g. Coley & Bockmiller, 1980; LaSasso & Mobley, 1997). Leybaert (1993) also argued that the reading problems of DHH children arise for precisely this reason: the education system's failure to address the phonological components of reading instruction. This highlights the importance of providing schools with a strong evidence base for implementing different interventions. Studies should always consider the practical implications of their findings and relay this information to the teachers working with the DHH population.

References

- Campbell, S., Torr, J. & Cologon, K. (2014). Pre-packaging Preschool Literacy: what drives early childhood teachers to use commercially produced phonics programs in prior to school settings. *Contemporary Issues in Early Childhood*, 15(1),40.
- Cihon, T. M. Gardner, R., Morrison, D., & Paul, P. (2008). Using visual phonics as a strategic intervention to increase literacy behaviours for kindergarten participants at-risk for reading failure. *Journal of Early and Intensive Behaviour and Intervention*, 5, 138-155.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155 – 159.
- Coley, J., & Bockmiller, P. (1980). Teaching reading to the deaf: An examination of teacher preparedness and practices. *American Annals of the Deaf*, 125, 909-915.

Dyer, A., MacSweeney, M., Szczerbinski, M., Green, L., & Campbell, R. (2003). Predictors of Reading Delay in Deaf Adolescents: The Relative Contributions of Rapid Automated Naming Speed and Phonological Awareness and Decoding. *Journal of Deaf Studies and Deaf Education, 8*(3), 215-229.

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

Gough, D. (2007). Weight of evidence: a framework for the appraisal of the quality and relevance of evidence. *Research Papers in Education, 22*(2), 213-228.

Hammill, D. (2002). What we know about correlates of reading. *Exceptional Children, 70* (4), 453-468.

Horner, R. H., Carr, E. G., Halle, J., Mcgee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children, 71*(2), 165-179.

Kratochwill, T. R. (2003). *Task force on evidence-based practice interventions in school psychology*. Retrieved on January 25th 2014 from:
www.switzerland.k12.in.us/Rti/RTI%20Resources%20Galore_Updated%206.28.09/SSP%20EBI%20Manual.pdf

King, C. M., & Quigley, S.P. (1985). *Reading and deafness*. San Diego: College Hill Press.

LaSasso, C., & Mobley R. T. (1997). National Survey of Reading Instruction for Deaf or Hard-of –Hearing Students in the U.S. *The Volta Review*, 99(1), 31-59.

Leybaert, J. (1993). Reading in the deaf: The roles of phonological codes. In M. Marschark & M. D. Clark (Eds.), *Psychological Perspectives on Deafness* (269-309). Hillside, NJ: Lawrence Erlbaum Associates.

Luetke-Stahlman, B. & Nielson J. (2003). The Contribution of Phonological Awareness and Receptive and Expressive English to the Reading Ability of Deaf Students with Varying Degrees of Exposure to Accurate English. *Journal of Deaf Studies and Deaf Education*, 8(4), 464-484.

Meece, J. & Miller, S. (2001). A Longitudinal Analysis of Elementary School Students' Achievement Goals in Literacy Activities. *Contemporary Educational Psychology*, 26(4), 454-480.

Montgomery, J. (2008). Dave Krupke: What Exactly is Visual Phonics? *Communication Disorders Quarterly*, 29(3),177-182.

Morris, S. (2007). Estimating Effect Sizes From Pretest-Posttest-Control Group Designs. *Organizational Research Methods*, 11(2), 364-386.

Musselman, C.(2000). How do children who can't hear learn to read an alphabetic script? A review of the literature on reading and deafness. *Journal of Deaf Studies and Deaf Education*, 5(1), 9-31.

Narr, R. & Cawthon, S. (2010). The "Wh" Questions of Visual Phonics: What, Who, Where, When, and Why. *Journal of Deaf Studies and Deaf Education*, 16(1), 66-78.

Nussbaum, D., Waddy-Smith, B., & Wilson, V. (2004). See the sound: Visual Phonics. *Clinicians in the Classroom*, 13(50), 15-29.

Parker, R., Vannest, K. and Davis, J. (2011). Effect Size in Single-Case Research: A Review of Nine Nonoverlap Techniques. *Behavior Modification*, 35(4),303-322.

Perfetti, C. A., & Sandak, R. (2000). Reading optimally builds on spoken language. Implications for deaf readers. *Journal of Deaf Studies and Deaf Education*, 5, 32-50.

Quigley, S.P., & King, C.M. (1985). *Reading Milestones*. Austin, TX: PRO-ED, Inc

Rhoades, E. (2006). Research Outcomes of Auditory-Verbal Intervention: Is the Approach Justified? *Deafness & Education International*, 8(3),125-143.

Schirmer, B. R., & McGough, S. M. (2005). Teaching reading to children who are deaf: Do the conclusions of the national reading panel apply? *Review of Educational Research*, 75(1), 83-117.

Smith, A., & Wang, Y. (2010). The Impact of Visual Phonics on the Phonological Awareness and Speech Production of a Student Who Is Deaf: A Case Study. *American Annals of the Deaf*, 155(2),124-130.

Steck-Vaughn Company, (1998). *Pair-it books*. Toronto, CAN: Harcourt Canada Ltd.

Strong, M. and Stuckless, R. (1995). A Review of Bilingual/Bicultural Programs for Deaf Children in North America. *American Annals of the Deaf*, 140(2), 83-94.

Torgesen, J. and Mathes, P. (2000). *A basic guide to understanding, assessing, and teaching phonological awareness*. Austin, Tex.: Pro-Ed.

Tucci, S., Trussell, J., & Easterbrooks, S. (2014). A Review of the Evidence on Strategies for Teaching Children Who Are DHH Grapheme-Phoneme Correspondence. *Communication Disorders Quarterly*, 35(4),191-203.

Van Staden, A. (2013). An Evaluation of an Intervention Using Sign Language and Multi-Sensory Coding to Support Word Learning and Reading Comprehension of Deaf Signing Children. *Child Language Teaching and Therapy*, 29 (3), 305-318

Waddy-Smith, B., & Wilson, V. (2003). See that sound! Visual phonics helps deaf and hard of hearing students develop reading skills. *Odessey*, 5, 14-17.

Woolsey, M.L, Satterfield, S.T. & Roberson, L. (2006). Visual Phonics: an English Code Buster? *Journal of American Annals of the Deaf*, 4, 452-151.

Wang, Y., Trezek, B., Luckner, J.L., & Paul, P. (2008). The role of phonology and phonologically related skills in reading instruction for students who are deaf and hard of hearing. *American Annals of the Deaf*, 153(4), 456-481.

Wright Group (2001). *Wildcat series*. New York: McGraw Hill.

Appendices

Appendix A: Studies excluded by full text:

Excluded paper	Rationale for exclusion
----------------	-------------------------

Excluded paper	Rationale for exclusion
EBeal-Alvarez, J. S & Easterbrooks, Susan R. (2013). Increasing Children's ASL Classifier Production: A Multicomponent Intervention. <i>American Annals of the Deaf</i> , 158 (3), 311-333.	Exclusion criteria 4 b): Main focus is not literacy
Berent, G. P., Kelly, R.S., Schmitz, K.L., Kenney, P. (2009). Visual Input Enhancement via Essay Coding Results in Deaf Learners' Long-Term Retention of Improved English Grammatical Knowledge. <i>Journal of Deaf Studies and Deaf Education</i> , 14 (2), 190-204.	Exclusion criteria 4 a): Use of Visual Phonics
Dostal, H. M. & Wolbers, K.A. (2014). Developing Language and Writing Skills of Deaf and Hard of Hearing Students: A Simultaneous Approach. <i>Literacy Research and Instruction</i> , 53 (3), 245-268.	Exclusion criteria 4 a): Use of Visual Phonics
Hoffman, M. & Wang, Y. (2010). The Use of Graphic Representations of Sign Language in Leveled Texts to Support Deaf Readers. <i>American Annals of the Deaf</i> , 155 (2),131-136.	Exclusion criteria 4 a): Use of Visual Phonics
Liu, H., Liu, C. and Andrews, J. (2014). Literacy and Deaf Students in Taiwan: Issues, Practices and Directions for Future Research: Part I. <i>Deafness & Education International</i> , 16(1),2-22.	Exclusion criteria 4 a): Use of Visual Phonics
Liu, H., Andrews, J. and Liu, C. (2014). Literacy and Deaf Students in Taiwan: Issues, Practices and Directions for Future Research: Part II. <i>Deafness & Education International</i> , 16(1), 23-36.	Exclusion criteria 4 a): Use of Visual Phonics
Stoefen-Fisher, J. and Lee, M. (1989). The Effectiveness of the Graphic Representation of Signs in Developing Word Identification Skills for Hearing Impaired Beginning Readers. <i>The Journal of Special Education</i> , 23(2),151-167.	Exclusion criteria 4 a): Use of Visual Phonics
Tetzchner, S. Roger, O.S, Lilleeng, M.K. (1997). Literacy Intervention for a Deaf Child with Severe Reading Disorder. <i>Journal of Literacy Research</i> , 29 (1), 25-46.	Exclusion criteria 4 a): Use of Visual Phonics

Excluded paper	Rationale for exclusion
van Staden, A. (2013). An Evaluation of an Intervention Using Sign Language and Multi-Sensory Coding to Support Word Learning and Reading Comprehension of Deaf Signing Children. <i>Child Language Teaching and Therapy</i> , 29 (3), 305-318	Exclusion criteria 4 a): Use of Visual Phonics

Appendix B: Coding Protocol for Methodological Quality (WoE A)

The five selected studies were evaluated using Gough's (2007) Weight of Evidence (WoE) Framework and reviewed using the following two protocols (i.e. depending on their research design):

- ❖ *Single case design*: Horner, R. H., Carr, E. G., Halle, J., Mcgee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children*, 71(2), 165-179.
- ❖ *Quasi-experiment*: Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. S. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Exceptional children*, 71(2), 149-164.

Study ID Number 1
Case

Research Design: Single

Date: 10/01/16

Name of Coder: Maike Feuerstein

Name of paper: Tucci, S. and Easterbrooks, S. (2013). A Syllable Segmentation, Letter-Sound, and Initial-Sound Intervention With Students Who Are Deaf or Hard of Hearing and Use Sign Language. *The Journal of Special Education*, 48(4), pp.279-289.

Horner et al., (2005). The Use of Single-Subject Research to Identify Evidence-Based Practice in Special Education

Quality Indicators Within Single-Subject Research

Article Reference:

Description of Participants and Setting

Participants are described with sufficient detail to allow others to select individuals with similar characteristics; (e.g., age, gender, disability, diagnosis).

Yes – Age, gender and level of hearing loss information was provided.

- No
- N/A
- Unknown/Unable to Code

The process for selecting participants is described with operational precision.

- Yes
- No** - The authors only state that “three children participated in the study”
- N/A
- Unknown/Unable to Code

Critical features of the physical setting are described with sufficient precision to allow replication.

- Yes
- No** - Not enough information was provided with regards to what kinds of intervention the participants had already received for literacy.
- N/A
- Unknown/Unable to Code

Overall Rating of Evidence: 3 2 1 0

Dependent Variable

Dependent variables are described with operational precision.

- Yes** – Syllable segmentation, letter-to –sound correspondence, initial – sound identification were the dependent variables. Specific examples provided.
- No
- N/A
- Unknown/Unable to Code

Each dependent variable is measured with a procedure that generates a quantifiable index.

- Yes** - Means are reported.
- No
- N/A
- Unknown/Unable to Code

Measurement of the dependent variable is valid and described with replicable precision.

- Yes** - Percentage of correct responses on syllable segmentation tasks, letter-sound tasks and initial sound tasks.
- No
- N/A

Unknown/Unable to Code

Dependent variables are measured repeatedly over time.

Yes -Data points were collected in terms of correct responses at regular intervals over the course of the intervention.

No

N/A

Unknown/Unable to Code

Data are collected on the reliability or inter-observer agreement associated with each dependent variable, and IOA levels meet minimal standards

Yes – Interobserver agreement was calculated by a second trained researcher – observer for 19 sessions using point-by-point formula.

No

N/A

Unknown/Unable to Code

Overall Rating of Evidence: 3 2 1 0

Independent Variable

Independent variable is described with replicable precision.

Yes - Yes- all the information is available for replication.

No

N/A

Unknown/Unable to Code

Independent variable is systematically manipulated and under the control of the experimenter.

Yes - Foundations combined with VP

No

N/A

Unknown/Unable to Code

Overt measurement of the fidelity of implementation for the independent variable is highly desirable.

Yes

No

N/A

Unknown/Unable to Code

Overall Rating of Evidence: 3 2 1 0

Baseline

The majority of single-subject research studies will include a baseline phase that provides repeated measurement of a dependent variable and establishes a pattern of responding that can be used to predict the pattern of future performance, if introduction or manipulation of the independent variable did not occur.

- Yes
- No
- N/A
- Unknown/Unable to Code

Baseline conditions are described with replicable precision.

- Yes
- No
- N/A
- Unknown/Unable to Code

Overall Rating of Evidence: 3 2 1 0

Experimental Control/internal Validity

The design provides at least three demonstrations of experimental effect at three different points in time.

- Yes – At every session.
- No
- N/A
- Unknown/Unable to Code

The design controls for common threats to internal validity (e.g., permits elimination of rival hypotheses).

- Yes
- No- Insufficient descriptors provided.
- N/A
- Unknown/Unable to Code

The results document a pattern that demonstrates experimental control.

- Yes - Each participant acted as their own baseline control.
- No
- N/A
- Unknown/Unable to Code

Overall Rating of Evidence: 3 2 1 0

External Validity

Experimental effects are replicated across participants, settings, or materials to establish external validity.

- Yes - Accurately described methods.**
- No
- N/A
- Unknown/Unable to Code

Overall Rating of Evidence: 3 2 1 0

Social Validity

The dependent variable is socially important.

- Yes - The research was aimed at determining whether or not VP in combination with Foundations was an effective literacy strategy for children who are deaf or hard of hearing.**
- No
- N/A
- Unknown/Unable to Code

The magnitude of change in the dependent variable resulting from the intervention is socially important.

- Yes**
- No
- N/A
- Unknown/Unable to Code

Implementation of the independent variable is practical and cost effective

- Yes
- No
- N/A
- Unknown/Unable to Code**

Social validity is enhanced by implementation of the independent variable over extended time periods, by typical intervention agents, in typical physical and social contexts.

- Yes
- No
- N/A
- Unknown/Unable to Code**

Overall Rating of Evidence: 3 2 1 0

Average WoE A across the 7 judgement areas:

Sum of X / N = 13/7 = 1.9

X = individual quality rating for each judgement area

N = number of judgement areas

Overall Rating of Evidence: 1.9 (MEDIUM rating)

LOW = 0-1.4

MEDIUM = 1.5-2.4

HIGH = 2.5 – 3.0

Study ID Number 3

Research Design: Quasi-experimental

Date: 10/01/16

Name of Coder: Maike Feuerstein

Full Study Reference: Trezek, B., Wang, Y., Woods, D., Gampp, T. and Paul, P. (2007). Using Visual Phonics to Supplement Beginning Reading Instruction for Students Who Are Deaf or Hard of Hearing. *Journal of Deaf Studies and Deaf Education*, 12(3), pp.373-384.

Intervention Name (description of study): Effectiveness of using Visual Phonics alongside Direct Instruction reading programmes and a separate phonic-based reading curriculum for children who are deaf or hard of hearing.

Research design: pre/post quasi experimental

Type of Publication: Journal

Gersten et al., (2005). *Quality Indicators for Group Experimental and Quasi-Experimental Research in Special Education*

Essential Quality Indicators

Describing Participants

Was sufficient information provided to determine/confirm whether the participants demonstrated the disability(ies) or difficulties presented?

Yes – Degree of hearing loss and age of onset information were included.

No

N/A

Unknown/Unable to Code

Were appropriate procedures used to increase the likelihood that relevant characteristics of participants in the sample were comparable across conditions?

Yes

No

N/A- No comparison group available.

Unknown/Unable to Code

Was sufficient information given characterizing the interventionists or teachers provided? Did it indicate whether they were comparable across conditions?

Yes –Teachers were chosen on the basis of degrees and certificates earned in teaching and number of years working in the field.

- No
- N/A
- Unknown/Unable to Code

Implementation of the Intervention and Description of Comparison Conditions

Was the intervention clearly described and specified?

Yes – ***Interventions were clearly described (e.g. Visual Phonics and LACES programme) and time frames were noted.***

- No
- N/A
- Unknown/Unable to Code

Was the fidelity of implementation described and assessed?

Yes – ***Facilitators were trained over a 11 month period and focus groups were held to review teacher progress. Authors also visited teachers in the classroom to observe sessions and monitor implementation.***

- No
- N/A
- Unknown/Unable to Code

Was the nature of services provided in comparison conditions described?

- Yes
- No
- N/A**
- Unknown/Unable to Code

Outcome Measures

Were multiple measures used to provide an appropriate balance between measures closely aligned with the intervention and measures of generalised performance?

Yes – ***Multiple (relevant) subtests of the Dominic Reading and Writing Assessment Portfolio (DeFord, 2001) were used.***

- No
- N/A
- Unknown/Unable to Code

Were outcomes for capturing the intervention's effect measured at the appropriate times?

Yes – ***Student progress was tested at the beginning of the academic year 2003 and then again as a post test at the beginning of 2004.***

- No
- N/A
- Unknown/Unable to Code

Data Analysis

Were the data analysis techniques appropriately linked to key research questions and hypotheses? Were they appropriately linked to the unit of analysis in the study?

Yes – *Paired samples t tests were used to calculate the difference between pre test and post test raw scores on the Dominie subtests. Effect sizes, cohen's d were used to calculate the difference scores (pretest-post test) of all participants across subtests.*

- No
- N/A
- Unknown/Unable to Code

Did the research report include not only inferential statistics but also effect size calculations?

Yes -*Effect sizes provided (Cohen's d)*

- No
- N/A
- Unknown/Unable to Code

Desirable Quality Indicators

Was data available on attrition rates among intervention samples? Was severe overall attrition documented? If so, is attrition comparable across samples? Is overall attrition less than 30%?

Yes – *20 students finally participated.*

- No
- N/A
- Unknown/Unable to Code

Did the study provide not only internal consistency reliability but also test-retest reliability and interrater reliability (when appropriate) for outcome measures? Were data collectors and/or scorers blind to study conditions and equally (un)familiar to examinees across study conditions?

- Yes
- No**
- N/A
- Unknown/Unable to Code

Were outcomes for capturing the intervention's effect measured beyond an immediate posttest?

- Yes
- No**
- N/A
- Unknown/Unable to Code

Was evidence of the criterion-related validity and construct validity of the measures provided?

- Yes
- No**

- N/A
- Unknown/Unable to Code

Did the research team assess not only surface features of fidelity implementation (e.g. number of minutes allocated to the intervention or teacher/interventionist following procedures specified), but also examine quality of implementation?

Yes – *training conducted over 1 month period and contained four parts: initial training in visual phonics, lesion demonstration, lesson observation and three 30 minute focus groups. Initial Visual Phonics training was carried out in a one 2-hour session after school. Teachers provided with a resource guide and again a focus group was held one week following training. 45 minute observations were carried out in the classrooms by researchers.*

- No
- N/A
- Unknown/Unable to Code

Was any documentation of the nature of instruction or series provided in comparison conditions?

- Yes
- No
- N/A**
- Unknown/Unable to Code

Did the research report include actual audio or videotape excerpts that capture the nature of the intervention?

- Yes
- No**
- N/A
- Unknown/Unable to Code

Were results presented in a clear, coherent fashion?

Yes – *Tables and figures were used to present data from participants pre and post assessment.*

- No
- N/A
- Unknown/Unable to Code

Overall Rating of Evidence: 3 2 1 0

This study was given an overall rating of three because four of the essential criteria were met and at least three of the desirable criteria were met.

Appendix C: Weight of Evidence (WoE A) criteria

Ratings for Quasi-Experimental Designs – Methodological Quality (WoE A)

Gersten et al., (2005) have devised a set of dichotomous criteria to rate the quality of experimental and quasi-experimental studies. The authors have proposed a set of quality indicators which can be used to define the quality of research as either “acceptable” or “high”. These definitions of acceptable and high quality are tentative, and have been adapted for the present review in such a way as to produce criteria for high, medium and low weighting for the 3 quasi-experimental studies (i.e. studies 3, 4 and 5. The studies were rated as follows:

For a HIGH rating of 3 , the study must meet:	For a MEDIUM rating of 2 , the study must meet:	For a LOW rating of 1 , the study must meet:
At least <u>three</u> essential criteria indicators	At least <u>two</u> essential criteria indicators	At least <u>two</u> essential criteria indicators
At least <u>three</u> desirable criteria indicators	At least <u>three</u> desirable criteria indicators	At least <u>one</u> of the desirable criteria indicators

Note: an unknown/unable to code score is coded as = 0.

Ratings for Single Case Designs – Methodological Quality (WoE A)

Horner et al.,(2005) devised a set of criteria for single case designs, which has been used to guide the ratings for WoE A of studies 1 and 2. Horner described a set of professional standards (i.e. which relate to the measurement of the independent and dependent variable, as well as the social validity level of experimental), which are crucial when determining whether a study is of “high” quality research. Studies 1 and 2 were rated using the following criteria:

(Note: “Unknown” responses are recorded as = 0)

a) Participant and Setting Description

	For a HIGH weighting of 3 , the study must have:	For a MEDIUM weighting of 2 , the study must at least have:	For a LOW weighting of 1 , the study must <u>have 1/2 criteria below</u> :
Criteria			
1. A clear definition of the context in which the intervention was used	√		
2. A clear description of participants (including level of hearing difficulty)	√	√	√
3. A clear description of how participants were selected.	√	√	√

b. Dependent Variable

	For a HIGH weighting of 3 , the study must have:	For a MEDIUM weighting of 2 , the study must have:	For a LOW weighting of 1 , the study must <u>have at least 2/3 criteria below</u> :
Criteria			
1. Dependent variables are operationally defined to allow valid, consistent assessment and facilitate replication.		√	

2. A quantifiable index for each dependent variable is provided.	√	√	√
3. Measurement of the dependent variable is valid and described with replicable precision.	√		
4. Dependent variables are measured at a number of points to allow identification of a performance pattern.	√	√	√
5. Inter-rater reliability is at 80% +	√	√	√

c. Independent Variable (IV)

Criteria	For a HIGH weighting of 3 , the study must have:	For a MEDIUM weighting of 2 , the study must have:	For a LOW weighting of 1 , the study must <u>have at least 1/2 criteria below:</u>
1. A precise description of the IV is given.	√	√	√
2. Evidence that the IV has been systematically manipulated.	√	√	√
3. Documentation that the practice was implemented with fidelity.	√		

d. Baseline

Criteria	For a HIGH weighting of 3 , the study must have:	For a MEDIUM weighting of 2 , the study must have 1/2 of	For a LOW weighting of 1 , the study does not meet either of the criteria.

the criteria
below:

1. Provides a baseline measurement.	√	√
2. Baseline measurement is described with replicable precision.	√	√

d. Experimental Control

Criteria	For a HIGH weighting of 3 , the study must have:	For a MEDIUM weighting of 2 , the study must have:	For a LOW weighting of 1 , the study must <u>have at least 1/2 criteria below:</u>
1. 3+ baseline points are collected from the data.	√	√	√
2. Evidence for controlling the effects of extraneous variables is provided.	√		
3. Demonstration of experimental effects are recorded 3+ times.	√	√	√

e. External Validity

Criteria	For a HIGH weighting of 3 , the study must have:	For a MEDIUM weighting of 2 , the study must have:	For a LOW weighting of 1 , the study will not meet the criteria
External validity is established by ensuring that the experimental effects can be replicated across participants, settings, or materials.	√	n/a	

f. Social Validity

Criteria	For a HIGH weighting of 3 , the study must fulfil 4 criteria:	For a MEDIUM weighting of 2 , the study must fulfil at least the following two criteria:	For a LOW weighting of 1 , the study must fulfil one of the following criteria:
1. The DV is socially important.	√	√	√
2. The magnitude of change in the DV resulting from the intervention is socially important.	√	√	√
3. Implementation of the IV is practical and cost effective.	√		

4. IV is implemented by typical change agents, in typical social contexts. ✓

Scores on each of the criteria were averaged for each study. The study was then awarded a **HIGH** score if the average was between **0- 1.4**, a **MEDIUM** score if the average was between **1.5 – 2.4** and a **LOW** score if the average was between **2.5- 3.0**.

Methodological Relevance (WoE B)

The WoE B considers the methodological relevance of each of the 5 studies. The aim is to determine to what extent the research designs within each study were suitably selected to measure the effectiveness of the Visual Phonics (VP) programme for children who are deaf or hard of hearing (DHH).

Criteria	For a HIGH weighting for relevance to the review question, the study must have the following:	For a MEDIUM weighting for the relevance to the review question, the study must have:	For a LOW weighting for the relevance to the review question, the study must have at least one of the following:
The study must have a control group	✓		
Outcome data are collected pre and post intervention	✓	✓	✓

The programme is delivered by staff who are experienced in working with the DHH population and who are trained in VP	√	√	√
Efforts were made to monitor the implementation of the intervention	√	√	
VP is delivered in conjunction with no more than one other intervention package	√	√	
Outcome measures cover a wide range of skills involved in proficient reading	√	√	√

Overall Weight of Evidence (WoE D)

Each of the 5 studies were analysed and weighted using the criteria listed above. Each study was given a weighting of High (3), Medium (2) or Low (1) for WoE A, B and C. The scores were then averaged to provide an overall weight of evidence for WoE A B and C:

Overall Weight of Evidence	Average Scores
High	>2.5
Medium	1.5-2.4
Low	<1.4

Appendix D: Effect Size Interpretation Criteria

Table 6: Distributions and descriptors for Cohen's *d* and SMD

<i>Type of effect size</i>	Small	Medium	Large
Cohen's <i>d</i> and SMD value:	0.2	0.5	0.8

Table 7: Percentile distributions and descriptors for PAND (Parker et al., 2011)

<i>Type of effect size</i>	No effect	Small	Moderate	Medium	Large
-----------------------------------	------------------	--------------	-----------------	---------------	--------------

Doctorate in Educational and Child Psychology

PAND	0.2	0.38	0.64	0.86	1
Percentile Rank	10th	25th	50th	75th	90th
