

RUNNING HEAD: Computer-assisted intervention

Cluster Randomized Controlled Trial Evaluating the Effectiveness of Computer-Assisted  
Intervention Delivered by Educators for Children with Speech Sound Disorders

Sharynne McLeod

Charles Sturt University, Australia

Elise Baker

The University of Sydney, Australia

Jane McCormack

Charles Sturt University, Australia

Yvonne Wren

Bristol Speech and Language Therapy Research Unit, UK

University of Bristol, Bristol, UK

Sue Roulstone

Bristol Speech and Language Therapy Research Unit, UK

University of the West of England, UK

Kathryn Crowe

Sarah Masso

Charles Sturt University, Australia

Paul White

University of the West of England, UK

Charlotte Howland

Charles Sturt University, Australia

The University of Sydney, Australia

Correspondence:

Professor Sharynne McLeod, Ph.D.,

School of Teacher Education, Charles Sturt University

Panorama Avenue, Bathurst, NSW AUSTRALIA, 2795

Email: [smcleod@csu.edu.au](mailto:smcleod@csu.edu.au)

Telephone: +61 2 6338 4463

## Abstract

**Purpose:** To evaluate the effectiveness of computer-assisted input-based intervention for children with speech sound disorders (SSD).

**Method:** The Sound Start Study was a cluster randomized controlled trial. Seventy-nine early childhood centers were invited to participate, 45 were recruited, and 1,205 4- to 5-year-old children's parents/educators returned questionnaires. Children whose parents/educators had concerns about speech were assessed ( $n=275$ ); 132 children who were identified with phonological impairment of unknown origin underwent additional assessment. Children with SSD and no receptive language or hearing difficulties, typical non-verbal intelligence, and English as their primary language were eligible; 123 were randomized (Intervention  $n=65$ ; Control  $n=58$ ); 3 withdrew. Intervention involved Phoneme Factory Sound Sorter software administered by educators over 9 weeks; Control involved typical classroom practices.

Participants were re-assessed twice by a speech-language pathologist (SLP) blinded to the initial assessment and intervention conditions.

**Results:** For the primary outcome variable (percentage of consonants correct), the significant mean change from pre- to post-intervention for the Intervention group (mean change+6.15,  $p<.001$ ) was comparable in magnitude to the significant change for the Control group (mean change+5.43,  $p<.001$ ) with a small between groups effect size for change (Cohen's  $d=0.08$ ).

Similar results occurred for measures of emergent literacy, phonological processing, participation, and wellbeing.

**Conclusion:** Computer-assisted input-based intervention administered by educators did not result in greater improvement than typical classroom practices.

*Key words:* intervention, speech sound disorders, randomized controlled trial, Sound Start Study, speech processing, children's speech

## **Cluster Randomized Controlled Trial Evaluating the Effectiveness of Computer-Assisted Intervention for Children with Speech Sound Disorders**

There is a high prevalence of speech sound disorders (SSD) in young children (Eadie et al., 2014; Law, Boyle, Harris, Harkness, & Nye, 2000; Wren, Miller, Peters, Emond, & Roulstone, 2016) and a high number of children with SSD on speech-language pathologists' (SLPs') caseloads (e.g., Broomfield & Dodd, 2004). Accordingly, there are many interventions for SSD that are supported by research evidence (see Baker & McLeod, 2011 for a review). However, within real service contexts, the frequency and total number of intervention sessions provided can be limited by available resources (McAllister, McCormack, McLeod, & Harrison, 2011; McCormack & Verdon, 2015; Ruggero, McCabe, Ballard & Munro, 2012) and service delivery provisions (McLeod & Baker, 2014). For instance, despite evidence that children with SSD can require approximately 30 to 40 hours of intervention by SLPs to become intelligible (Hodson, 2007; Williams, 2012), children have been reported to receive an average of 6.2 hours of intervention over 12 months (Glogowska, Roulstone, Enderby, & Peters, 2000) or even less (Ruggero et al., 2012).

In an effort to close the gap between empirically-based recommendations regarding intensity of intervention, and real-world speech-language pathology service delivery constraints, some interventions for children with SSD have been trialled using alternative models of delivery in which parents or educators deliver intervention given training and guidance from an SLP (Broen & Westman, 1990; Dodd & Barker, 1990; Eiserman, McCoun, & Escobar, 1990; Lancaster, Keusch, Levin, Pring, & Martin, 2010; Ruscello, Cartwright, Haines, & Shuster, 1993). This allows intervention to be delivered more frequently, potentially meeting the needs of children. However, of the limited evidence examining parent and/or educator-delivered

interventions, the outcomes have not always been ideal. Law, Garrett and Nye (2003/2010) in a meta-analysis of interventions for expressive phonology found that intervention conditions were associated with better child progress in phonology than control conditions. However, when parent-administered interventions were removed from the analysis, the effect size increased suggesting that parent-administered interventions produced less impact on children's speech. For instance, Lancaster et al. (2010) compared SLP- and parent-delivered intervention with no intervention for 15 preschool children with SSD. The preschool children who received the SLP-delivered intervention made the most improvement. It was unclear whether this improvement was due to the difference in the intervention agent (SLP versus parent) or the difference in the intervention, as the SLPs delivered their own "eclectic" intervention involving speech production practice, auditory tasks, and phonological awareness while the parents implemented "auditory input and bombardment activities that related to their child's error patterns" (Lancaster et al., 2010, p. 187). Dodd and Barker (1990) examined parent- and teacher-delivered minimal pair intervention for 11 preschool children with SSD. Although an improvement was detected for both the parent- and teacher-implemented intervention, the improvement was described as limited for the children receiving the teacher-delivered intervention. Two barriers to implementation were identified—lack of time and intervention agent capability. Dodd and Barker (1990) noted that despite the teachers receiving approximately 24 hours of extensive training, the training did not yield competence similar to SLP-delivered intervention. One way to circumvent this barrier is to make the intervention simpler and easier to implement through the use of computer technology.

In a systematic review of computer-based interventions for children and adults with articulation and phonological disorders, Chen et al. (2016) noted that such interventions can be

effective. Unlike inexperienced intervention agents, computers can be programmed to provide accurate and consistent feedback, thus circumventing barriers to implementation noted by Dodd and Barker (1990). Ruscello et al. (1993) discovered this to be the case when they found that computer-based minimal pairs intervention delivered by a parent, was just as effective as direct speech-language pathology intervention at improving the children's speech. This study was small (12 participants), and the parents needed to attend the speech-language pathology clinic in order to access the computer program to provide the intervention. Given that the time spent training parents benefits their children, it would seem advantageous to train educators as they could expand their knowledge and skills to more children (Dodd & Barker, 1990). However, it is unknown whether computer-based intervention delivered by educators in a preschool setting would yield similar positive results to that delivered by parents. It is also unknown whether such an intervention could help prepare preschool children with SSD for future literacy instruction. This idea has appeal, because children with SSD are at risk of literacy difficulties (Anthony et al., 2011).

Models of language processing have typically identified a common pathway between speech and literacy development (Duggirala & Dodd, 1991; Kay, Lesser & Coltheart, 1992; Stackhouse & Wells, 1997) which helps explain why difficulties in these two areas commonly co-exist. What is less clear is whether differing presentations or subtypes of SSD can be identified which have distinct pathways and require different interventions. Waring and Knight (2013) identified three approaches to classification of SSD based on etiology (Shriberg et al., 2010), linguistics (Dodd, 2013) and processing skills (Stackhouse & Wells, 1997). Each approach offers alternate ways of conceptualising SSD and identifying approaches to intervention. While the approaches proposed by Shriberg et al. (2010) and Dodd (2013) offer

categorical subtypes, Stackhouse and Wells (1997) proposed a system which links a plausible underlying processing cause with speech output. Specifically, they propose three core elements: input processes (i.e., detecting and perceiving speech or seeing a written word/letter); cognitive-linguistic processes (i.e., creating, storing, and accessing lexical representations for words comprising semantic, phonological, grammatical, orthographic, and/or motor information); and output processes (i.e., producing speech or writing a word or letter). Some children with phonologically-based SSD (i.e., those who show error patterns impacting classes of sounds and/or phonotactics) are thought to have underspecified or poor quality phonological representations as a result of difficulties with input (perceptual) and cognitive-linguistic processes (e.g., Munson, Baylis, Krause, & Yim, 2010; Sutherland & Gillon, 2005). Their underspecified phonological representations are thought to manifest inaccurate speech and poor performance on tasks involving phonological awareness (Sutherland & Gillon, 2007). An intervention approach that targets input processing skills and phonological awareness could therefore result in improvements to a child's speech where the core deficit involves these areas of processing. This is indeed what Gillon (2005) discovered. Through targeting phonological error patterns via speech production practice, alongside an emergent literacy intervention targeting phoneme awareness and letter knowledge, 12 preschoolers with phonologically-based SSD became intelligible and successful readers and spellers. Gillon (2005) suggested that the phoneme awareness and letter knowledge activities might have helped the children's speech because they helped to establish "more fully specified underlying phonological representations" (p. 321). The intervention involved input and output tasks, and was delivered by an SLP or SLP student under the supervision of a qualified SLP, twice weekly in one individual session and one



small group session over an average of 25.5 hours. Three of the 12 children also needed an additional 10 to 12 hours of intervention to improve their intelligibility.

A computer-based intervention program designed to target the input processing and phonological awareness skills of children with SSD is Phoneme Factory Sound Sorter® (PFSS) (Wren & Roulstone, 2006, 2013). PFSS comprises seven interactive games that can be customized to a child's needs based on their errors in speech sound production and underlying processing skills. The games require children to listen and respond to visual and auditory stimuli, and specifically target phoneme segmentation and identification, phoneme blending, minimal pair discrimination, and rhyme detection. The PFSS software allows the user to select specific phoneme targets and goals as well as word types (nonword, real words), word position (initial, medial, final) and stimuli options (sound on/off, picture on/off), and to modify what is selected from session to session in light of a child's response to intervention—a setup known as the free configuration setting. The free configuration setting requires expertise to probe for phonological generalization, and change what is being targeted and/or played in light of a child's response to intervention. PFSS was also designed with predetermined configuration settings—known as the teacher settings. These settings comprised seven pre-set modules for children who present with common developmental phonological error patterns (i.e., stopping, final consonant deletion, velar fronting, gliding, context sensitive voicing, deaffrication, cluster reduction). These pre-set teacher settings consist of between four and seven sessions of activities. Each session contains between three and five games that have been configured to target the pattern, with each game being played multiple times across the sessions but with increasing complexity in their configuration.

Wren and Roulstone (2008) compared the effect of 12 hours of an experimental version of PFSS with conventional table-top intervention (targeting input processing and phonological awareness) and no intervention ( $n = 33$ , with 11 randomly assigned to each of three groups: computer-based PFSS, conventional table-top intervention, and no intervention). Participants received three sessions a week over 8 weeks, each session lasting 30 minutes. One session a week was delivered by the SLP with a teaching assistant observing. The teaching assistant then delivered the other two sessions each week. The free configuration settings were used rather than the teacher settings, allowing the SLP to select the most appropriate targets and stimuli for each child from session to session. Wren and Roulstone (2008) discovered that although there was not a statistically significant difference between the groups, the preschool children who received the PFSS intervention showed a trend towards more improvement compared to the children in the control group. This occurred when the children were stimutable for the target sounds prior to intervention and showed developmental rather than non-developmental error patterns. As these were the same processes targeted via the pre-set teaching settings, it was hypothesized that the same results could perhaps be achieved with the tool being delivered by non-SLPs. The effect of the intervention on the children's emergent literacy skills (i.e., phonological awareness and print knowledge), and phonological processing (i.e., ability to rapidly retrieve stored words from memory, and recall information) was not measured. The potential for intervention to indirectly address the impact of SSD on children's everyday wellbeing (i.e., activity and participation) was also not considered.

It was speculated that a larger scale study, conducted for a longer period of time and using non-SLPs, would yield clearer insight into the potential of PFSS to improve children's speech, emergent literacy, phonological processing, and wellbeing. However, for PFSS to be

implemented simply and easily, the teacher settings (rather than free configuration setting) would need to be used and children with developmental rather than non-developmental phonological error patterns selected. In light of research examining the effect of intervention intensity for children with SSD (Allen, 2013; Williams, 2012), PFSS would also need to be implemented with an adequate and feasible intensity, including dose, session frequency, and total intervention duration (Baker, 2012; Warren, Fey, & Yoder, 2007).

There is no clear answer as to what constitutes adequate and feasible intensity in intervention for children with SSD. The majority of peer reviewed published intervention research (94% of 134 studies) for children with SSD reported intervention outcomes for a pre-determined period of time rather than from referral to dismissal (Baker & McLeod, 2011). This is often because of constraints within a research context such as the time and funds available to conduct intervention research. This creates a challenge—conducting experimental research within the confines of a budget, with a feasible session dose, duration, and frequency for a period of time sufficient to demonstrate a clinically significant effect. Across a selection of peer reviewed published intervention studies involving children with SSD using pre-determine time periods and a robust RCT design, positive effects have been reported for 30-minute sessions scheduled three times weekly over 8 weeks (i.e., 24 sessions) (Allen, 2013), 12 weekly 30-minute sessions (Dodd et al., 2008), and 16 twice weekly 1 hour sessions over 8 weeks (Ruscello et al., 1993). Together, these studies suggest that twice weekly 60 minute sessions (or four 30 minute sessions each week) over approximately 8 to 12 weeks may be sufficient to demonstrate an effect in a research context.

## **Aim**

The aim of the Sound Start Study was to evaluate the effectiveness of PFSS delivered by educators (using the teacher settings) compared with typical classroom practices on the speech production, emergent literacy, phonological processing, and participation and wellbeing of 4- to 5-year-old children with phonologically-based SSD.

## METHOD

### Participant Recruitment and Eligibility

Seventy-nine early childhood education centers in Sydney, Australia were invited to be involved in the Sound Start Study between 2013 and 2015 and 45 early childhood centers were recruited<sup>1</sup>. Early childhood centers were identified by geographical location (postcode) in order to obtain a sample of children living in range of socio-economic advantage and disadvantage. Private ( $n = 7$ ), community ( $n = 9$ ), government-funded ( $n = 24$ ), and local council-funded ( $n = 5$ ) early childhood education centers were included in the sample. Directors of the early childhood education centers agreeing to participate in the study were asked to distribute screening questionnaires to parents of all 4- to 5-year-old children within their center ( $n = 1,920$ ).

The Sound Start Study was undertaken in six stages with recruitment occurring over three years (Figure 1). Eleven sites participated during more than one year. In stage 1 parent and educator screening questionnaires were returned for 1,205 4- to 5-year-old children attending one of the 45 early childhood centers. In stage 2, 275 children whose parents and/or educators were concerned about how they “talked and made speech sounds” (Glascoe, 2000) were directly assessed to screen their eligibility for the trial. In stage 3, 132 eligible children received a more comprehensive assessment. Data from stages 2 and 3 were gathered pre-intervention and served

---

<sup>1</sup> 47 sites agreed to participate but two sites were excluded due to low return rate of stage 1 questionnaires. One site did not return any questionnaires. The second site returned four questionnaires (<10% of their cohort) in which there were no concerns about children’s communication.

as baseline measures. At the end of stage 3, there were 123 children at 39 sites who were eligible for intervention in stage 4. The early childhood education centers with eligible children were randomized to the Intervention ( $n = 19$  sites, 65 children) or Control ( $n = 20$  sites, 58 children) conditions. Computerized intervention using PFSS was provided to 63 eligible children in the intervention condition. Fifty-seven children received the control condition of typical classroom practices. That is, they continued to participate in their classroom environment as usual without any input from the researchers. Three children withdrew from the study. In stage 5, immediate follow-up assessments were undertaken with 113 children. In stage 6, long-term follow-up assessments were conducted with 114 children, 6-8 weeks later. Data from stages 5 and 6 were gathered post-intervention and served as outcome measures. All direct assessments (pre- and post-intervention) were completed in a quiet room within the early childhood centers except nine stage 6 (year 1) assessments that were completed at the child's home during the school holidays due to the children's availability. This paper reports data from stages 1 to 6 (screening, baseline, intervention, and outcome), with a particular emphasis on the baseline (stages 2-3), intervention (stage 4), and outcome measures (stages 5-6). Some data regarding stage 1 are reported in McLeod et al. (2015) and stages 2 and 3 in McLeod et al. (2017).

Children's eligibility was determined across stage 1, 2, and 3. The criteria for eligibility are shown in Table 1. In summary, children were eligible for the randomized controlled trial (stage 4) if they (1) were identified as having SSD that included a phonological pattern that could be treated by the pre-set teacher settings on PFSS, (2) were not reported to have a cleft lip/palate, developmental delay, or persistent hearing loss, and (3) were found to have no difficulties with

nonverbal intelligence<sup>2</sup>, oromusculature, hearing, and receptive language during direct assessment. Children who spoke languages in addition to English were included; however, their English language skills had to be equivalent to or better than their skills in the other language(s).

### **Participants in the Randomized Controlled Trial**

Participants in the randomized controlled trial were 123 children who were eligible for intervention in stage 4 of the Sound Start Study. The participants ranged in age from 4;0 to 5;4 ( $M = 54.89$  months;  $SD = 4.2$ ) when they were assessed at stage 3 (see Table 2). There were more males ( $n = 79, 64.2\%$ ) than females ( $n = 44, 35.8\%$ ). The participants lived in a range of suburbs from the most disadvantaged (1<sup>st</sup> decile) to most advantaged (10<sup>th</sup> decile) according to the Australian Index of Relative Socioeconomic Advantage and Disadvantage (IRSAD, Australian Bureau of Statistics, ABS, 2008) that considers “people’s access to material and social resources, and their ability to participate in society” (ABS, 2008, p. 17). The mean IRSAD decile of participants was 5.97 ( $SD = 3.2$ ) and mode of 8 (25<sup>th</sup> percentile = 3, 50<sup>th</sup> percentile = 7, 75<sup>th</sup> percentile = 8). Information about family IRSAD was available for all participants. The majority of participants ( $n = 114$ ) were reported to use English at home (107 spoke English only at home and seven spoke English and an additional language at home). Nine participants spoke a language other than English at home. The parents of most ( $n = 120$ ) of the participants described their child’s English language proficiency. The majority were described by their parents as using English “very well” ( $n = 64, 52.0\%$ ), with fewer parents reporting “somewhat well” ( $n = 48, 39\%$ ), or “not very well” ( $n = 8, 6.5\%$ ). According to their parents, the participants spoke one ( $n = 100, 81.3\%$ ), two ( $n = 21, 17.1\%$ ), or three ( $n = 2, 1.6\%$ ) languages with 17 languages reported

---

<sup>2</sup> The eligibility criteria for non-verbal intelligence was adapted from year 1 (Nonverbal Index Score of >79) to year 2 (Nonverbal Index Score >70).

in total. The languages in addition to English were: Arabic, Cantonese, Filipino, Greek, Hindi, Indonesian, Korean, Malayalam, Maltese, Marathi, Portuguese, Punjabi, Spanish, Tamil, Thai and Urdu.

## **Instruments**

### **Questionnaires for parents.**

Parents were asked to complete questionnaires created by the research team at each stage of the research. The questionnaires focused on different areas including parental concern about their children's development, demographic information, languages spoken, developmental history, family history of speech and language difficulties, home literacy, home computer use, reading skills, and communication skills. The questionnaires included questions created by the researchers as well as questions from previously published scales that were included after gaining permission from the authors (e.g., Parents Evaluation of Developmental Status, PEDS, Glascoe, 2000). Typically questions were asked once; however, some key questions were repeated at different stages. For example, at each stage parents were asked to provide information regarding previous and current speech-language pathology assessment and intervention (in addition to the intervention received during the Sound Start Study).

### **Questionnaires for directors and educators at the early childhood education centers**

Directors of all participating centers completed a 22-item questionnaire that focused on different aspects of the environment, staffing and use of resources at their center. Directors were asked to describe the number of qualified staff and teaching assistants at the center each day, the organisational basis for the center, parent engagement practices, daily activities, and access, use, and attitudes towards technology at the center.

### **Instruments used for baseline and outcome measures.**

Baseline measures from stages 2 and 3 and outcome measures from stages 5 and 6 are described below (see McLeod et al., 2017 for the complete assessment protocol). The stage when assessments were administered is shown in Tables 3 and 4. Each outcome measure pertains to the individual participant level.

*Speech production instruments.*

The participants' speech production was assessed using four measures: (1) the Phonology subtest from the Diagnostic Evaluation of Articulation and Phonology (DEAP, Dodd, Hua, Crosbie, Holm, & Ozanne, 2002), (2) phonological probes created by the research team, (3) Intelligibility in Context Scale (ICS, McLeod, Harrison, & McCormack, 2012a), and (4) stimulability for each English consonant. The DEAP Phonology subtest (Dodd et al., 2002) enabled comprehensive sampling of a broad range of phonemes (consonants, vowels, and consonant clusters) in a range of contexts and syllable shapes, and provided normative data for Australian and British children. Each participant's responses to items on the DEAP were entered into Computerized Profiling (PROPH+; Long, Fey, & Channell, 2008). The percentage of consonants correct (PCC) was calculated by PROPH+ and was used as a main outcome measure. The research team developed phonological probes for each phonological pattern that could be treated using the teacher settings of PFSS (e.g., fronting, stopping). The majority of the probes consisted of 15 single words (range = 15-29) with an average of 19 opportunities to demonstrate each phonological error (range=15-39). Each word was depicted by an illustration. Participants were asked to repeat the words after the SLP. The percentage of occurrence of the phonological pattern was calculated for each phonological probe. Intelligibility was assessed using the ICS (McLeod et al., 2012a), a parent-report scale of children's intelligibility in seven different contexts. The ICS has been validated and normed for use with Australian preschool children



(McLeod, Crowe, & Shahaieian, 2015; McLeod, Harrison, & McCormack, 2012b). Consonant stimulability was assessed by asking participants to produce consonants in isolation after a model from the SLP with subsequent prompting if necessary. A score was generated regarding the number of stimuable consonants out of a possible total of 24.

### ***Emergent literacy instruments.***

Emergent literacy includes code-related skills (phonological awareness, print knowledge, and emergent writing) and oral-language skills (Whitehurst & Longian, 1998). For the purposes of this study we assessed aspects of emergent literacy that could be affected by PFSS, (i.e., phonological awareness and print knowledge). We assessed these skills using three measures: (1) a letter knowledge probe (based on a task described by Anthony et al., 2011), (2) the Preschool Word and Print Awareness measure (PWPA, Justice, Bowles, & Skibbe, 2006), and (3) the phonological awareness subtests within the Comprehensive Test of Phonological Processing – 2 (CTOPP-2; Wagner, Torgesen, & Rashotte, 2013). Children’s letter knowledge (Anthony et al., 2011) was determined by showing pairs of each capital and lower case letter within the alphabet within a random sequence. Children were asked to identify letters that they knew (e.g., Mm, Tt) and tell the SLP the name of the letter (e.g., *em*, *tee*) and the sound it made (e.g., /m/, /t/). The PWPA considered children’s knowledge of books and text by sharing a picture book and asking questions (e.g., “Show me the first letter on this page?”). A raw score out of a possible total of 17 and print-concept knowledge estimate was calculated according to instructions within Justice et al. (2006). Three subtests of the CTOPP-2 (Wagner et al., 2013) were administered and scored according to the published manual to consider participants’ phonological awareness: Elision (removing phonological segments from spoken words to form other words), Blending Words (synthesizing sounds to form words), and Sound Matching (identifying words with the same

initial and final sounds). The CTOPP-2 has been validated and normed in the US on 1,900 children.

***Phonological processing instrument.***

Individual subtests from the CTOPP-2 (Wagner et al., 2013) were used to examine participants' phonological processing and skills: the three subtests also used to measure phonological awareness (described above), Elision, Blending Words, and Sound Matching; two subtests of phonological memory, Memory for Digits (repeating numbers accurately), and Nonword Repetition (repeating nonwords accurately); and two subtests of rapid non-symbolic naming, Rapid Color Naming (rapidly naming colors) and Rapid Object Naming (rapidly naming objects).

***Children's participation and wellbeing instruments.***

Children's participation and wellbeing was considered using four measures: (1) Focus on the Outcomes of Children Under Six (FOCUS, Thomas-Stonell, Robertson, Walker, Oddson, Washington, & Rosenbaum, 2012), (2) Kiddy-Communication and Attitude Test (KiddyCAT; Vanryckeghem & Brutton, 2007), (3) Speech Participation and Activity Assessment of Children (SPAA-C, McLeod, 2004), and (4) Australian Therapy Outcome Measures for Speech Pathology (AusTOMs, Perry & Skeat, 2004). The FOCUS (Thomas-Stonell et al., 2012) enabled measurement of changes in participants' participation in communicative contexts. The 50-item parent-report measure was used to consider participants' function (speech, expressive language, pragmatics, receptive language/attention) and performance (intelligibility, expressive language, social/play, independence, coping strategies/emotions) and has been validated on 210 children in Canada (Thomas-Stonell, Oddson, Robertson, & Rosenbaum, 2010). The KiddyCAT (Vanryckeghem & Brutton, 2007) was administered to investigate participants' perceptions of

their speech ability and difficulties they had with talking. The KiddyCAT comprises 12 yes/no questions, including “Is talking hard for you?” and “Do you think that people need to help you talk?” The KiddyCAT has been standardized for children aged 3-6 years and has been used with a variety of clinical populations including children with SSD (McLeod, Harrison, McAllister, & McCormack, 2013). Ten questions from the SPAA-C (McLeod, 2004) were administered to consider participants’ feelings about talking in different communicative situations (e.g., How do you feel when you talk to the whole class?). Children were asked to color the face that best matched their response from the following options: ☺ ☹ ☹ ? O (another feeling). The total number of ☺ responses were added to calculate a score out of a possible total of 10. To date, the SPAA-C has not been normed or validated. The AusTOMs (Perry & Skeat, 2004) were completed by the SLP to describe participants’ speech impairment, activity limitation, participation restriction, and distress on a 6-point scale ranging from 5 = *no difficulty* to 0 = *profound difficulty*. Face and content validity of the AusTOMs was established by Perry and colleagues (2004) and the inter-rater and test-retest reliability of the tool was found to be “mostly satisfactory” (Morris et al., 2005, p. 344).

### **Intervention instrument: Phoneme Factory Sound Sorter**

The original British version of PFSS (Wren & Roulstone, 2006) was adapted for Australia for use in the Sound Start Study (Wren & Roulstone, 2013). The four main adaptations were: use of four Australian speakers for the auditory stimuli (cf. one British speaker in the original version), addition of cluster reduction modules in the teacher settings, replacement of culturally relevant pictures (e.g., *lolly*), and, an increase in the number of word and nonword stimuli overall. Educators delivered PFSS using the teacher setting (rather than SLPs using the free configuration setting).

## **Trial Design**

The Sound Start Study was a blinded cluster randomized controlled trial in which children's progress with the PFSS (Wren & Roulstone, 2013) completed using pre-set teacher settings with early childhood educators was compared with a group of similar size who received typical classroom practices (i.e., whatever the child would typically receive). Early childhood education centers were randomized to receive one of the two types of service delivery. The allocation of participants within clusters (sites) provided protection from contamination across trial arms and also increased the convenience of delivering the intervention within community settings.

The sites were divided and were allocated to one of two experienced speech-language pathologists (SLP1 and SLP2). These SLPs completed the pre-intervention assessments. Information about the number of children in each of their sites was sent to a statistician who randomized the sites in to trial and control conditions. The sites assessed by SLP1 were allocated to either Intervention/Control and this was only known to SLP1; SLP2 remained blinded to pre-intervention assessment results and trial condition, and vice versa. Random allocation was performed on the sites with all eligible children in a site allocated to the same trial condition (Intervention or Control). Children in the intervention arm received an individualized program based on their speech sound errors. Labels for Intervention or Control were assigned to each site based on an unrestricted realisation of a Bernoulli random variable (Binomial distribution, with parameters  $\pi = 0.5$ ,  $n = 1$ ) using the random number generator in SPSS version 20. This process ensured each site had an equiprobable chance of being the control or intervention arm. This process was independently performed for each SLP with SLP2 blinded to the list generated for SLP1 and vice versa. There were no other restrictions imposed. The post-intervention

assessments were completed by four experienced SLPs (including SLP1 and SLP2). The SLPs were blinded to the intervention condition when undertaking the post-intervention assessments (SLP1 and SLP2 did not re-assess children from their allocated pre-intervention sites).

## **Procedure**

Ethical approval was gained from Charles Sturt University (Approval number 2013/070), the NSW Department of Education and Communities State Education Research Applications Process (SERAP) (Approval number 2013267), and individual early childhood centers as required. Consent was gained from the directors, teachers, and teaching assistants at the early childhood education centers and the participants' parents/caregivers and assent was gained from the participants at each stage of the study.

### **Stage 1: Screening for eligibility**

The Sound Start Study was conducted in six stages (Figure 1). During stage 1 1,205 parents completed a 2-page screening questionnaire containing questions about their children's speech, language, and general development. Teachers subsequently completed the same screening questionnaire for children whose parents consented to participate. Children who met inclusion criteria for stage 2 ( $n = 323$ ) were invited to participate in a direct assessment of their speech and language skills.

### **Stages 2 and 3: Direct screening and comprehensive assessments**

The parents of a total of 275 participants provided consent for participation in the direct screening assessment phase of the study (see Table 1). Assessments were completed with one of two experienced SLPs in a quiet room within each participant's early childhood education center once assent had been obtained from the child. Stage 2 and stage 3 assessments took approximately 30-60 minutes each with breaks available for participants who required them.

Video and/or audio recording was completed for tasks using a Panasonic HC-V700 video camera with an external Hahnel Mk100 uni-directional microphone, and Zoom H1 audio recorder. Online broad phonetic transcription of the DEAP (Dodd et al., 2002) was completed by the assessing SLP and later checked by the same SLP for accuracy based on the audio recording. The AusTOMs measure was completed during, or immediately after, each assessment.

Parents were given a questionnaire to complete before each assessment stage and were asked to return the questionnaire to their early childhood education center. After administration of the DEAP in stage 2, a member of the research team (second author) determined which phonological probes should be administered in stage 3 to each participant based on their patterns of errors on the DEAP. For example, if a participant produced /s/ as [t] and /f/ as [p] then the stopping probe was selected. Participants' results on the probes then were used by the second author to determine the primary and secondary PFSS teacher settings to be used in the intervention phase. A primary and secondary intervention goal corresponding to a PFSS teacher setting was allocated to each participant who was eligible for stage 4 (Intervention and Control) before allocation of the participants' intervention condition.

#### **Stage 4: Intervention**

Random allocation was performed for each early childhood education center (site). The educators were asked to oversee the PFSS intervention delivered via computer four times per week for nine weeks (18 hours) in 1:1 sessions within the site. The educators were nominated by the center director based on their willingness to participate. No formal education was required for the educators to facilitate the completion of the intervention. Sites were offered financial reimbursement to partially compensate for the time spent by the educators during this research and to maintain staff:student ratios in their centers.

Intervention was individualized for each child using the preset teacher settings in PFSS based on the identified phonological error patterns present during pre-intervention assessment (stage 2). The allocation of 18 hours of intervention with PFSS was considered appropriate with respect to intervention literature (Baker, 2012; Williams, 2012) and feasible within a 10 week preschool term, as PFSS could be provided over first nine weeks of term (4 x 30 min or 2 x 60 min/week) allowing the 10th week to conduct an immediate post-treatment evaluation. The participants and educators wore headphones to listen to the PFSS program. The activities in PFSS required the participants to listen and look at stimuli and respond in various ways using the computer mouse. When participants were not proficient at using the mouse, the educators moved it to the location on the screen indicated by the participants' pointing.

The first intervention session was overseen by the SLP who undertook their stage 1-3 assessments to standardize delivery and solve any technical difficulties that arose. The same SLP monitored and videoed intervention in weeks 2-3 and 7-9 to ensure fidelity. The educators recorded the number of sessions completed (on paper). In addition, the PFSS program recorded the total number of games, total number of plays, total time spent on each game, and the score achieved (/10) for each game. Children in both conditions were allowed to attend additional speech-language pathology if organized by their parents or educators, and known attendance is noted in Figure 1.

Standard care consisted of typical classroom practice. All sites were asked to document their typical daily activities in a center questionnaire.

### **Stages 5 and 6: Immediate and long-term follow-up assessments**

The same two experienced SLPs (SLP1 and SLP2) conducted most of the immediate and long-term follow-up assessments post-intervention at stage 5 ( $n = 89$ , 78.8%) and stage 6 ( $n =$

102, 89.5%) with two other experienced SLPs completing a small proportion of the assessment in stage 5 ( $n = 24$ , 21.2%) and stage 6 ( $n = 12$ , 10.5%) in the final year of data collection. All SLPs were blinded to the information about the pre-intervention assessments (stages 1-3) and trial arm (stage 4) for the participants they assessed during stages 5 and 6. The post-intervention assessments followed a similar protocol to the pre-intervention assessments in that they took place in a quiet room in the early childhood education center that the participants attended, with the consent of parents and the assent of child participants. The assessments took approximately 60 minutes to complete. All assessment sessions were video-recorded and audio-recorded. Broad transcription of the DEAP (Dodd et al., 2002) was completed online where possible.

### **Reliability**

Inter- and intra-judge point-by-point reliability was completed for the DEAP-Phonology based on a randomly selected sample of 30 (10.9% of 275 participants) speech samples comprising 6,629 data points. Intra-judge agreement for broad phonetic transcription was 91.5% and inter-judge agreement was 90.1%. The reported level of reliability reflects “acceptable agreement” as it was >85% (Shriberg & Lof, 1991, p. 255).

### **Intervention Fidelity**

Fidelity checking was based on video samples that included at least one full PFSS game (up to 10 minutes of recorded video) of 32% of the participants ( $n = 20$ ) in the Intervention group during a session at the early childhood center with the educator. A checklist was developed to determine if the experimental tasks were completed as described. An independent observer viewed 20 videos of participants completing the intervention. Based on 264 data points, procedural fidelity for the experimental tasks was 95.5% (for further information about fidelity, see McCormack et al., 2016).



## **Data Analysis**

### **Randomization of clusters (sites).**

Prior reasoning on the design suggested on average 3 participants be allocated per cluster (early childhood education center/site). The intra-cluster correlation coefficient was conservatively estimated to be 0.1, and hence conservative design effect of 1.2. A design effect of 1.2 indicates a 20% sample size inflation is needed to maintain pre-specified power after accounting for cluster effects. In the study, a total of 19 participating sites were randomized to the intervention arm, with a mean of 3.42 participants per cluster. A total of 20 participating sites were randomized to the control arm, with a mean of 2.90 participants per cluster. The intra-class correlation coefficient for the primary outcome variable was 0.11 (95% CI ranging from 0.04 to 0.289) consistent with prior expectations. For the sample data, an analysis using individual level analyses produces the same broad conclusions as cluster-level analyses weighted by cluster size and for brevity the individual level analyses are presented in the results section.

### **Analysis of outcome variables.**

The randomized controlled trial comprised two randomized groups (Intervention, Control) with data collection at three time points (Stage 2/3, Stage 5, Stage 6) with percentage of consonants correct (PCC) as a single primary outcome variable. This design is readily amenable to analysis using Analysis of Variance (ANOVA) for a 2 by 3 within and between mixed design with randomized group as a two level between subjects factor (Group), and data collection stage as a three level repeated measures factor (Stage). This approach permits a statistical examination of the main effect for group (i.e., whether mean values averaged across the three data collection points differ between groups), a statistical examination of the main effect for Stage (i.e. whether the mean values differ between the three stages irrespective of group assignment), and the two-

way interaction between Group and Stage (i.e., are mean changes group dependent). Between groups comparisons at each stage were undertaken using the independent samples t-test, and within groups comparisons between stages were performed using the paired samples t-test.

The two-way interaction effect between Group and Stage for the primary outcome variable was the most important contrast for the design. For correlated measures ( $r \geq 0.6$ ) a minimum sample size of  $n = 46$  per group was needed for an interaction corresponding to a medium effect size with 80% power in a design without clustering. For a clustered design, with design effect of 1.2, the required minimum sample size rose to  $n = 56$  per group. To account for 10% missing values the sample size were upwardly revised to  $n = 63$  per group.

In all analyses, the effect size was quantified using partial eta squared ( $\eta_p^2$ ). In two-way and higher order ANOVAs, the partial eta squared statistic indicates the proportion of variation attributable to a factor relative to the total of the error and factor variation, and as such  $\eta_p^2$  always lies between 0 and 1. Thresholds for  $\eta_p^2$  are 0.01, 0.06, and 0.14 for small, medium and large effects respectively (Miles & Shevlin, 2001). The chi-square test of association was used to examine differences between groups for categorical data.

A missing values analysis indicated only 2.2% of outcome data was missing at stage 2 or stage 3 on those eligible for randomisation. This percentage of missing data rose to 8.9% at stage 5 but included 5.1% of randomized participants with data known to be missing completely at random (e.g., missing due to illness, or away at time of assessment) and otherwise missing with no systematic pattern. Overall percentage of missing data at stage 6 was similarly 8.9% including 5.1% with data known to be missing completely at random (e.g., due to illness, or absence at time of evaluation). Overall missing data is relatively small (<10%), is missing completely at

random (MCAR), or missing at random (MAR). Analysis under multiple imputation does not affect the conclusions drawn for this data set and for this reason the data are reported without imputation.

## **RESULTS**

An intention-to-treat analysis was carried out on the primary outcome measure (PCC) and the secondary outcome measures followed by an analysis per protocol to account for the variation (with respect to intensity of PFSS in the Intervention group), and receipt of additional SLP intervention received by each group.

### **Effectiveness of PFSS: Intention to Treat**

The 123 participants who were eligible for intervention in stage 4 of the Sound Start Study were randomized into two groups. There were 65 participants randomized to the Intervention group and 58 randomized to the Control group; however, three participants withdrew from the study (parents of two participants withdrew consent, one child did not provide assent to participate in intervention). Consequently, there were 63 children (19 sites) in the intervention arm and 57 children (20 sites) in the Control group (see Figure 1) whose results are reported. The two groups were similar on most measures, but differed on the severity of speech, and in age by one month on average (see Table 2). For percentage of consonants correct on the DEAP, the descriptive statistics suggest a higher average in those subsequently randomized to Intervention. However the standard deviation is also larger in the Control and these differences in the means and standard deviations are largely attributable to a single relatively low valued but genuine outlier in the Control group at stage 2 (consequently depressing the mean and raising the standard deviation in this group).

#### **Speech production accuracy.**

Three speech production accuracy measures were collected once before intervention (either at stage 2 or stage 3), and twice after intervention (stages 5 and 6).

*Percentage of consonants correct on the Diagnostic Evaluation of Articulation and Phonology (DEAP, Dodd et al., 2002).* Analysis using an ANOVA for a 2 by 3 mixed design indicated a statistically significant change in mean PCC on the Phonology subtest of the DEAP between the data collection stages ( $p < .001$ ,  $\eta_p^2 = .287$ ), and a statistically significant average difference between the two Groups ( $p = .009$ ,  $\eta_p^2 = .062$ ). The interaction between Stage and Group was not a significant effect ( $p < .874$ ,  $\eta_p^2 = .001$ ) (see Table 3). In the Intervention group there was a significant increase between stage 2 and stage 5 ( $p < .001$ ) and this effect was maintained at stage 6 ( $p = .458$ ). This same pattern is observed in the Control group, with a significant increase between stage 2 and stage 5 ( $p < .001$ ) and this effect was maintained at stage 6 ( $p = .078$ ). When the mean DEAP PCC was adjusted for baseline levels at stage 2, the mean DEAP PCC at stage 6 did not significantly differ between groups ( $p = 0.368$ ,  $\eta_p^2 = .007$ ).

*Primary intervention target phonological probes.* Analysis of the data using an ANOVA for a 2 by 3 mixed design indicated statistically significant changes between the Stages ( $p < .001$ ,  $\eta_p^2 = .319$ ), but no average significant difference between the two Groups ( $p = .322$ ,  $\eta_p^2 = .009$ ) nor was there a statistically significant interaction between Stage and Group ( $p < .329$ ,  $\eta_p^2 = .011$ ) (see Table 3). For the Intervention group, the mean change between stage 3 and stage 5 was statistically significant ( $p < .001$ ) as was the mean change between stage 3 and stage 6 ( $p < .001$ ) but the mean change between stage 5 and stage 6 was not statistically significant ( $p = .161$ ). The mean change in the Control group between stage 3 and stage 5 was statistically

significant ( $p < .001$ ) as was the mean change between stage 3 and stage 6 ( $p < .001$ ), and the mean change between stage 5 and stage 6 ( $p = .005$ ).

*Intelligibility in Context Scale (ICS, McLeod et al., 2012a)*. The mean score on the ICS was calculated for the participants whose parents provided data for all 7 items on the scale. Analysis using ANOVA for a 2 by 3 mixed design indicated a statistically significant change between Stages ( $p = .019$ ,  $\eta_p^2 = .094$ ) but no significant difference between Groups ( $p = .365$ ,  $\eta_p^2 = .021$ ). Additionally, the interaction between Stage and Group was not a statistically significant effect ( $p = .726$ ,  $\eta_p^2 = .008$ ) (see Table 3). In the Intervention group there was a significant change between stage 2 and stage 5 ( $p = .003$ ) and between stage 2 and stage 6 ( $p = .006$ ), but no significant change between stage 5 and stage 6. In the Control group there is a similar significant change between stage 5 and stage 6 ( $p = .009$ ) but the effect was less clear when considering stage 2 against stage 6 ( $p = .074$ ).

### **Emergent literacy skills.**

Three emergent literacy measures were collected once before intervention (stage 3), and at least once post intervention (stages 5 and/or 6) to consider outcomes over time.

*Letter knowledge (Anthony et al., 2011)*. Letter knowledge data were collected once pre-intervention (stage 3), and twice post-intervention (stages 5 and 6). Analysis using an ANOVA for a 2 by 3 mixed design indicated there were significant changes over time ( $p < .001$ ,  $\eta_p^2 = .186$ ), but the main effect for randomized groups was not a statistically significant effect ( $p = 0.110$ ,  $\eta_p^2 = .030$ ) and changes between stages were not group dependent ( $p = .190$ ,  $\eta_p^2 = .020$ ). In the Intervention group, mean values for Letter Knowledge significantly increased between stage 3 and stage 5 ( $p < .001$ ), significantly increased between stage 3 and stage 6 ( $p < .001$ ), and

significantly increased between stage 5 and stage 6 ( $p = .030$ ). In the Control group, mean values for Letter Knowledge significantly increased between stage 3 and stage 6 ( $p = .001$ ) and between stage 5 and stage 6 ( $p = .022$ ) but the effect between stage 3 and stage 5 was not significant ( $p = .129$ ).

*Preschool Word and Print Awareness (PWPA, Justice et al., 2006).* PWPA data were collected for stages 3 and 5 and the data analyzed using ANOVA for a 2 by 2 mixed design. Mean PWPA significantly increased between the two stages ( $p < .001$ ,  $\eta_p^2 = .368$ ), but there was no main effect attributable to randomized group ( $p = .858$ ,  $\eta_p^2 < .001$ ), and changes over time were not group dependent ( $p = .053$ ,  $\eta_p^2 = .035$ )

*Phonological awareness (elision, blending words and sound matching).* Phonological awareness data were collected once pre-intervention (stage 3), and twice post-intervention (stages 5 and 6). Statistically significant changes occurred between stages in elision ( $p < .001$ ,  $\eta_p^2 = .172$ ), and blending words ( $p < .001$ ,  $\eta_p^2 = .181$ ) but not in sound matching ( $p = .249$ ,  $\eta_p^2 = .014$ ). Over the three stages, the two randomized groups did not significantly differ on elision ( $p = .238$ ,  $\eta_p^2 = .014$ ), blending words ( $p = .837$ ,  $\eta_p^2 < .001$ ), or sound matching ( $p = .573$ ,  $\eta_p^2 = .003$ ), and any mean changes over time were not group dependent on elision ( $p = .208$ ,  $\eta_p^2 = .016$ ), blending words ( $p = .509$ ,  $\eta_p^2 = .007$ ) or on sound matching ( $p = .518$ ,  $\eta_p^2 = .007$ ).

### **Phonological processing skills.**

Four phonological processing measures were collected once before intervention (stage 3), and twice post intervention (stages 5 and/or 6).

*Memory for digits, nonword repetition, rapid color naming, and rapid object naming.*

Table 3 displays statistically significant increase between stages on memory for digits ( $p < .001$ ,

$\eta_p^2 = .077$ ), nonword repetition ( $p = .004$ ,  $\eta_p^2 = .057$ ), and rapid color naming ( $p = .035$ ,  $\eta_p^2 = .060$ ). However these changes are not group dependent ( $p = .624$ ,  $\eta_p^2 = .005$ ;  $p = .727$ ,  $\eta_p^2 = .003$ ;  $p = .567$ ,  $\eta_p^2 = .010$  respectively).

### **Children's participation and wellbeing.**

Four participation and wellbeing measures were collected once before intervention (stage 3), and twice post intervention (stages 5 and 6).

*Focus on the Outcomes of Children Under Six (FOCUS, Thomas-Stonell et al., 2012).*

The FOCUS total score was calculated for each participant. Analysis using an ANOVA for a 2 by 3 mixed design indicated there were no significant changes attributable to Stage ( $p = .167$ ,  $\eta_p^2 = .042$ ), Group ( $p = .606$ ,  $\eta_p^2 = .006$ ), nor for the interaction between Stage and Group ( $p = .668$ ,  $\eta_p^2 = .010$ ) (see Table 3). A post hoc analysis of means indicated that the two groups did not significantly differ in mean FOCUS total score at any stage. Mean values for the FOCUS total score did not significantly change between stages for the Intervention group or the Control group.

*Kiddy-Communication and Attitude Test (KiddyCAT, Vanryckeghem & Brutten, 2007).*

Analysis using an ANOVA for a 2 by 3 mixed design, indicated statistically significant effects for Stage ( $p < .001$ ,  $\eta_p^2 = .073$ ) but no significant effects for Group ( $p = .267$ ,  $\eta_p^2 = .012$ ), or for the interaction between Group and Stage ( $p = .292$ ,  $\eta_p^2 = .012$ ). In the Intervention group, there was a statistically significant difference between stage 5 and stage 6 ( $p = .005$ , two-sided), and between stage 3 and stage 6 ( $p < .001$ , two-sided) but with no significant change between stage 3 and stage 5 ( $p = .123$ , two-sided). In the Control group, the observed changes between stage 3

and stage 5, and stage 3 and stage 6, and stage 5 and stage 6 did not achieve statistical significance.

*Speech Participation and Activity Assessment of Children (SPAA-C, McLeod, 2004).* The number of times participants circled happy ☺ on the 10-items from the SPAA-C was determined for each stage. Analysis using an ANOVA for a 2 by 3 mixed design indicated the main effect for Stage was not statistically significant ( $p = .200$ ,  $\eta_p^2 = .016$ ), and the main effect for randomized group was not statistically significant ( $p = .080$ ,  $\eta_p^2 = .030$ ). Similarly, the Group by Stage interaction term did not achieve statistical significance ( $p = .151$ ,  $\eta_p^2 = .018$ ) (see Table 3).

*Australian Therapy Outcome Measures for Speech Pathology (AusTOMs, Perry & Skeat, 2004).* The 5-point AusTOMs scales regarding Speech Impairment, Speech Activity, Speech Participation, and Speech Distress were completed by the assessing SLP and analyzed using Pearson's chi-square statistic. Significant differences were found between the randomized groups for stage 2 and stage 5 on the Speech Activity and Speech Participation scales, but not for any other scale, or any other time. Specifically, Speech Impairment did not significantly differ between the randomized groups at stage 2 ( $p = .480$ ), nor at stage 5 ( $p = .601$ ), nor at stage 6 ( $p = .253$ ). Speech Activity did not significantly differ between the two groups at stage 2 ( $p = .716$ ), nor at stage 6 ( $p = .122$ ), but there was a significant differences at stage 5 ( $p = .048$ ). Speech Participation did not significantly differ between the two groups at stage 2 ( $p = .212$ ), nor at stage 6 ( $p = .314$ ), but there were statistically significant differences at stage 5 ( $p = .040$ ). Speech Distress did not significantly differ between randomized groups at stage 2 ( $p = .763$ ), nor at stage 5 ( $p = .921$ ), nor at stage 6 ( $p = .069$ ).

### **Typical Classroom Practices**



Of the 45 participating sites, directors of 44 sites completed and returned questionnaires (valid percentages reported) reporting their typical classroom practices. None of the sites employed an SLP or had an SLP who visited their center to provide assessment or intervention for participants.

*Emergent literacy instruction.* Children were exposed “very much” or “quite a lot” to letter name/sound activities at 18 sites (40.9%) and sound-play activities at 19 sites (43.2%). Fifteen (34.1%) sites reported that they used a specific phonological awareness program at their center. Only one site (2.3%) reported that they did not complete letter name/sound or sound-play activities at all. The directors of 34 centers (77.3%) reported that they felt children learned literacy through the use of computers.

*Daily activities and technology use.* Directors reported that they completed the following activities “very much” at their center: sitting and playing ( $n = 32$ , 72.7%), singing/stories/books ( $n = 20$ , 45.5%), individual attention in routines ( $n = 3$ , 6.8%), teaching good health practices ( $n = 15$ , 34.1%), active outdoor play ( $n = 25$ , 56.8%), and engaging in pretend play ( $n = 22$ , 50.0%). The daily activities and use of technology at each center varied with 20 sites (45.5%) reporting daily computer use and 10 sites (22.7%) reporting computer use a few times a week. Directors reported varied access to computers with 13 sites reporting that they do not have any computers dedicated for children’s use. The remaining sites had a mean of 3.29 (range = 1-12) computers dedicated for children’s use. The use of technology within the sites was also varied. Some sites reported frequent child-directed access to technology on “most days” or “every day”: four sites (9.1%) reported frequent child-led demonstrations of technology, 11 sites (25%) reported frequent child-adult interaction with technology, 15 sites (34.1%) reported frequent group technology activities, 17 (38.6%) reported frequent independent computer use by the children at

their center. The directors of 34 centers (77.3%) “agree” or “strongly agree” that computers are an essential part of learning and 41 (93.2%) “agree” or “strongly agree” that it was good to use technology to build on children’s interests. Only 14 (31.8%) directors saw facilitating internet access as a priority for the children at their center.

### **Adherence to Intervention Protocol**

Each participant in the intervention arm was requested to work on the PFSS software over 9 weeks for at least 4 sessions per week (sessions were approximately 30 minutes each indicating a total of approximately 18 hours of intervention equivalent to 128-156 separate games). Over the trial there were 39 participants of 63 (61.9%) intervention participants who received at least 70% of the intended intervention intensity. This adherence level of 70% was considered acceptable, given that few studies have reported adherence levels better than 80%, and that “positive results have often been obtained with levels around 60%” reports (Durlak & Dupre, 2008, p. 331).

Parents of participants in the Intervention and Control groups were asked to indicate whether participants received speech-language pathology intervention during the Sound Start Study. In the Intervention group, 29 did not receive additional intervention, 24 did and 10 parents did not respond. In the Control group, 30 did not receive additional intervention, 10 did and 17 parents did not respond.

By considering the adherence to protocol and participants’ additional speech-language pathology intervention, there were a total 23 participants who received an acceptable dosage of PFSS and no known additional intervention from an SLP (these participants were subsequently identified as receiving PFSS per protocol “true PFSS”) and 47 participants in the Control group who received no known additional intervention from an SLP (these participants were

subsequently identified as “true controls”). Consequently, the effectiveness of PFSS was also considered using a per protocol analysis for the 23 participants who received true PFSS and the 47 true controls.

### **Effectiveness of PFSS: Per Protocol**

Table 4 summarizes means, standard deviations and statistical significance for the per protocol analyses.

#### **Speech production accuracy**

*Percentage of consonants correct on the Diagnostic Evaluation of Articulation and Phonology (DEAP, Dodd et al., 2002).* DEAP data for true controls and true PFSS is summarized in Table 4. Analysis using an ANOVA for a 2 by 3 mixed design for these data gives the same statistical conclusions as the randomized groups analysis. That is, a statistically significant change in mean PCC on the Phonology subtest of the DEAP between the Stages ( $p < .001$ ,  $\eta_p^2 = .324$ ), and a statistically significant average difference between the two Groups ( $p = .003$ ,  $\eta_p^2 = .135$ ), but with a non-significant interaction between Stage and Group ( $p = .916$ ,  $\eta_p^2 = .001$ ).

*Primary intervention target phonological probes.* Analysis of the per protocol subset indicates statistically significant change in means between the Stages ( $p < .001$ ,  $\eta_p^2 = .283$ ), but with no average significant difference between the two Groups ( $p = .286$ ,  $\eta_p^2 = .019$ ), and any changes in means were not group dependent ( $p = .614$ ,  $\eta_p^2 = .008$ ) (see Table 4).

*Intelligibility in Context Scale (ICS, McLeod et al., 2012a).* In the per protocol analysis, average mean ICS did not significantly differ between the two Groups ( $p = .372$ ,  $\eta_p^2 = .047$ ), did

not significantly differ between the three Stages ( $p = .243$ ,  $\eta_p^2 = .008$ ), and differences between groups did not significantly change between the Stages ( $p = .996$ ,  $\eta_p^2 < .001$ ).

### **Emergent literacy skills**

Three emergent literacy measures were collected once before intervention (stage 3), and at least once post intervention (stages 5 and/or 6) to consider outcomes over time (Table 4), as noted in the intention-to-treat section.

*Letter knowledge (Anthony et al., 2011).* For the per protocol analysis, there were statistically significant changes in mean letter knowledge between Stages ( $p < .001$ ,  $\eta_p^2 = .206$ ), but the difference in means between groups did not significantly differ between Stages ( $p = .371$ ,  $\eta_p^2 = .019$ ). The two groups significantly differed at stage 3 ( $p = .043$ ), but with no significant difference at stage 5 ( $p = .128$ ,  $\eta_p^2 = .041$ ), nor at stage 6 ( $p = .440$ ,  $\eta_p^2 = .011$ ) after controlling for initial differences at stage 3.

*Preschool Word and Print Awareness (PWPA, Justice et al., 2006).* Mean PWPA significantly increased between stage 3 and stage 5 ( $p = .001$ ,  $\eta_p^2 = .159$ ), but the per protocol groups did not significantly differ on average ( $p = .483$ ,  $\eta_p^2 = .008$ ), and changes over time were not group dependent ( $p = .198$ ,  $\eta_p^2 = .027$ ).

*Phonological awareness (elision, blending words and sound matching).* Table 4 reports statistically significant changes between stages for elision ( $p < .001$ ,  $\eta_p^2 = .216$ ), and blending words ( $p < .001$ ,  $\eta_p^2 = .254$ ) but not in sound matching ( $p = .053$ ,  $\eta_p^2 = .052$ ). On average, the two per protocol groups did not significantly differ on elision ( $p = .241$ ,  $\eta_p^2 = .025$ ), blending words ( $p = .139$ ,  $\eta_p^2 = .039$ ), or sound matching ( $p = .283$ ,  $\eta_p^2 = .021$ ). The extent of the mean

difference between the two per protocol groups did not significantly change between stages for elision ( $p = .127$ ,  $\eta_p^2 = .037$ ), for blending words ( $p = .898$ ,  $\eta_p^2 = .002$ ) or sound matching ( $p = .587$ ,  $\eta_p^2 = .010$ )

### **Phonological processing skills**

Four phonological processing measures were collected once before intervention (stage 3), and twice post intervention (stages 5 and stage 6) (Table 4).

*Memory for digits, non-word repetition, rapid color naming, and rapid object naming.*

Mean memory for digits significantly increased over the three stages ( $p = .003$ ,  $\eta_p^2 = .100$ ) as shown in Table 4. However, mean changes between the stages were not significant effects for non-word repetition ( $p = .192$ ,  $\eta_p^2 = .030$ ), for rapid color naming ( $p = .065$ ,  $\eta_p^2 = .082$ ), and rapid object naming ( $p = .168$ ,  $\eta_p^2 = .041$ ). On average, the two groups did not significantly differ on any of these four measures and the extent of between groups differences was not stage dependent for memory for digits ( $p = .723$ ,  $\eta_p^2 = .006$ ), non-word repetition ( $p = .514$ ,  $\eta_p^2 = .012$ ), rapid color naming ( $p = .961$ ,  $\eta_p^2 = .001$ ), or rapid object naming ( $p = .827$ ,  $\eta_p^2 = .001$ ).

### **Children's participation and wellbeing**

Three participation and wellbeing measures were collected once before intervention (stage 3), and twice post intervention (stage 5 and stage 6) (Table 4). On average, the per protocol groups did not significantly differ on mean FOCUS ( $p = .584$ ,  $\eta_p^2 = .017$ ), nor on mean KiddyCAT ( $p = .367$ ,  $\eta_p^2 = .014$ ), nor on mean SPAA-C ( $p = .493$ ,  $\eta_p^2 = .009$ ). Mean KiddyCAT significantly decreased between stages ( $p < .001$ ,  $\eta_p^2 = .073$ ) as shown in Table 4, but the main effect of Stage did not achieve statistical significance for FOCUS ( $p = .355$ ,  $\eta_p^2 = .056$ ), nor for

SPAA-C ( $p = .128$ ,  $\eta_p^2 = .037$ ). In addition, the extent of differences between groups did not significantly change between the stages for FOCUS ( $p = .827$ ,  $\eta_p^2 = .011$ ), for KiddyCAT ( $p = .292$ ,  $\eta_p^2 = .012$ ), or for SPAA-C ( $p = .341$ ,  $\eta_p^2 = .019$ ).

## DISCUSSION

The aim of this study was to determine whether early childhood educators using the teacher settings in PFSS—a computer-based intervention targeting speech input processing and phonological awareness—could improve the speech production, emergent literacy, phonological processing, participation and wellbeing of Australian preschoolers with SSD. The experimental design was robust. Contrary to our predictions, PFSS (delivered by educators using the PFSS teacher setting) did not result in greater gains than typical classroom practice in this study. Rather, the participants in both groups showed a statistically significant improvement on most but not all measures of speech production, emergent literacy, and phonological processing from pre- to post-intervention. Most measures of wellbeing were not significant over time or between groups. These results were similar for the intention-to-treat data and per protocol data (i.e., the subgroup of 61.9% of participants in the Intervention group who received an acceptable dosage of PFSS and those in the Intervention and Control groups who did not receive any additional speech-language pathology intervention over the course of the study). The findings raise a number of issues for discussion including the meaning of statistical, clinical, and personal significance, why participants in the Intervention group (who received PFSS delivered by educators using the teacher settings) did not show greater improvement than participants in the Control group, and what could be done to address the gap between the demand and supply of speech-language pathology services for children with SSD.

## **Separating statistical significance from clinical and personal significance**

The statistical analyses revealed significant improvements in measures of speech production, emergent literacy, and some measures of phonological processing, participation and wellbeing for the participants in the Intervention and Control groups. Although this was encouraging, closer inspection of the measures suggested that the average improvement was not clinically significant, as speech, emergent literacy, and phonological processing difficulties were still apparent post-intervention. For example, the severity of the participants' SSD in the Intervention group remained mild-moderate (based on PCC), despite intervention. According to Bothe and Richardson (2011), clinical significance is only apparent when a change is sufficient to modify a clinical description or label for a presenting problem. The lack of clinical significance is also apparent when the degree of change is considered. For example, the average percentage improvement in PCC from the DEAP Phonology subtest (Dodd et al., 2002) for the participants in both groups from pre- to immediate post-intervention was approximately 6%. Given that their pre-assessment occurred in Stage 2 (6 to 8 weeks before starting intervention), that intervention was 9 weeks in duration, and that the post-assessment occurred one week later, this degree of improvement occurred over approximately 16 weeks. In a study comparing the effect of minimal versus non-minimal contrast intervention delivered weekly by SLPs in 30-minute sessions for 12 weeks, Dodd et al. (2008) reported that their preschoolers increase in PCC (using the same assessment task as the current study) was over 16%. This degree of improvement was commensurate with other studies of phonological intervention (e.g., Crosbie, Holm, & Dodd, 2005).

Measures assessing the personal experience (i.e., SPAA-C) and the impact of SSD on day-to-day life (i.e., FOCUS) suggest that the statistically significant change was not personally

significant. Personal significance is only apparent when clients report improvements that matter to them; improvements that change the way they function in day-to-day life (Bothe & Richardson, 2011). The findings from this study highlight the importance of collecting measures of impairment (e.g., PCC, % occurrence of a phonological process) in addition to personal measures of participation and wellness when interpreting the effect of intervention on children with SSD and their families.

### **Why did the intervention not work?**

PFSS was designed for children with phonologically-based SSD. In the current study, the delivery of PFSS by educators in early childhood centers did not result in significantly better speech, emergent literacy, phonological processing, or participation and wellbeing for the children in the Intervention group compared to the Control group. Five possible reasons for this are outlined below.

First, in the current study PFSS was delivered by educators using the teacher settings rather than the free configuration setting, as was used by Wren and Roulstone (2008). We chose to use the teacher settings, because it simplified the delivery of the intervention and addressed one of the primary aims of this study, to investigate the delivery of intervention by non-SLP intervention agents. It meant that the educators did not need to assess and monitor children's speech. They needed to facilitate a child's completion of activities in PFSS. However, in doing so, it is possible that the teacher settings diminished opportunities to dynamically tailor PFSS to the children's responses—selecting easier or more challenging games and/or selecting different error patterns to facilitate the children's progress from session to session.

Second, while the effectiveness of the teacher settings themselves is unknown, an alternative explanation is that delivery of intervention for children with SSD from non-SLPs is



ineffective, when SLP involvement is reduced to advising about targets and providing a program. As Dodd and Barker (1990) reported, it may be that educators may have too many barriers in their workplace setting to implement intervention for children with SSD. Many studies have included support from parents as a key element of the intervention program (Bowen & Cupples, 1999) and parents can be trained to be effective intervention agents (Sugden, Baker, Munro, & Williams, 2016), although meta-analyses of the effects of parent interventions suggest that implementation of this is by no means straightforward (Law et al., 2003/2010). The possibility that PFSS may be effective when well implemented by parents using the teacher or free configuration setting remains to be determined.

Third, not all participants in the Intervention group received the proposed intervention intensity—39 of the 63 participants (61.9%) in the Intervention group received at least 70% of the intended intervention. It is possible that under-dosing moderated the effect of PFSS. In follow-up investigations of the current study, the educators were interviewed about their experience implementing PFSS in a preschool setting (Crowe et al., 2016), and comparisons were made regarding the intensity of the intervention as recorded by the educators and the computer (McCormack et al., 2016). As part of these investigations multiple barriers and facilitators to implementation were identified including *personal factors* (e.g., child engagement with PFSS), *environmental factors* (e.g., the logistics of implementing PFSS during a busy preschool day), and *computer program factors* (e.g., program format, varied game durations). Similar to Dodd and Barker (1990), the educators found it challenging to consistently find the time to conduct the prescribed number of games each week, with the participant(s) at their center. If we are to find a solution to the gap between the demand and supply for speech-language pathology services, PFSS implemented by educators using the teacher setting does not

appear to be a viable option. Alternative or supplementary solutions need to be found. Perhaps PFSS could be effective but only when an optimal intensity of practice is adhered to, and the program is implemented at an acceptable level across intervention agents and settings (e.g., SLP, parent, educator, therapy assistant). Exactly what constitutes an acceptable level of implementation and what would be needed to ensure this level in day-to-day contexts remains to be determined. As Durlak (2015) points out, more research about implementation adherence is needed, if we are to determine the conditions under which interventions are and are not effective.

Fourth, it is possible that some children did not have input based problems underlying their surface level SSD or that input processing problems were having less impact on their SSD than output factors. Whilst individuals were assessed on the CTOPP-2 (Wagner et al., 2013), a measure of phonological processing, the results are presented for the groups and it is possible that analysis at the individual level may reveal that children with poorer performance on phonological processing measures prior to intervention made more progress with the intervention than those who showed better performance. In contrast, those with a predominantly output processing basis to their SSD would have benefited from a different approach to intervention.

Finally it is important to consider the possibility that PFSS is not effective in changing speech output, or in teaching phonological awareness in children with phonologically-based SSD. There is not enough space to discuss whether this would be the tool itself or the theoretical approach to intervention which underpins it but nevertheless, it is difficult to know whether either is the case, as other factors (e.g., use of the teacher setting, the less than optimal intervention intensity) may have masked the possible effect of PFSS on the children's speech, emergent literacy, phonological processing, and participation and wellbeing. It is also possible that PFSS works for some but not all children with phonologically-based SSD, given the range in

the amount of change in PCC from pre- to immediate post (-12.6% to +35.36%) for the participants in the Intervention group. However, it is also possible that other factors influenced the change in the participants' PCC from pre- to immediate post, given the range in the amount change (-12.5% to +29.10%) for participants in the Control group. Receipt of additional speech-language pathology services is one possible factor, given that the participant in the Intervention group who showed a change in PCC of 35.36% over approximately 16 weeks was reported to have received speech-language pathology intervention over the course of the study. This cannot be the sole factor influencing change, as the participant in the Control group who showed a change in PCC of 29.1% over the course of the intervention, was reported to have not received any additional speech-language pathology over the course of the study. Clearly, further research is needed to better understand the natural history of SSD in preschool children, and the factors associated with children who show little versus considerable improvement in speech production, emergent literacy, phonological processing, and participation and wellbeing overtime (Roulstone, Miller, Wren, & Peters, 2009).

### **Limitations**

Although this study used a robust experimental design, it is not without limitations. First, a variety of outcome measures were used to assess the effect of PFSS. Measures of the children's speech focused on speech production. Measures of speech perception were not included. It is therefore unknown if PFSS improved the quality of the children's acoustic-perceptual representations for speech. Second, PFSS was implemented by different educators from different early childhood centers. Some educators implemented the program with multiple participants; whereas, others implemented the program with one participant. Given that the effect of the number of children per educator was not controlled for, varied experience with the program may

have influenced the result. Given that individual SLPs have been reported to differentially contribute to gains in school-age children's language and literacy abilities (Farquharson, Rambyraja, Logan, Justice, & Schmitt, 2015), it is also possible that variation in early childhood educator characteristics (e.g., level of education, years of experience, age) influenced the results. Finally, not all participants received the recommended intensity. This was a difficult variable to control for, given the real-world constraints and demands on day-to-day life in a busy preschool setting. Some educators worked consistently with the participants, achieving intensity compliance rates over 90%. Some did not adhere to the prescribed intensity (Crowe et al., 2016; McCormack et al., 2016). This variation in intensity may have influenced the result. Nonetheless, a pragmatic trial is an important step in understanding how effective an intervention is likely to be in a real-world setting.

### **Future directions: Closing the gap between the demand and supply of SLPs for children with SSD**

Preschool children with SSD have the potential to become intelligible speakers and competent readers and spellers during the school years (Gillon, 2005). Although some participants in the current study improved without intervention, other participants showed relatively little improvement over time. If we are to provide all preschool children with SSD the opportunity to become intelligible before or by the time formal literacy instruction starts in the early school years and achieve literacy success, we need to identify and support those children less likely to improve without help. We also need to identify interventions suited to their needs. For some children this might be input-based interventions such as PFSS. Further research is needed to explore this possibility, controlling for the effect of intervention agent, and

intervention intensity. For some children, this intervention might involve cognitive-linguistic tasks coupled with production practice (e.g., Williams, 2012). For others, intervention might combine input and output (production) procedures (e.g., Hodson, 2007). Rvachew and Brosseau-Lapré (2015) found that whatever approach is used by an SLP, follow-up activities completed by non-SLPs (e.g., parent, educator) need to be similar to what is completed by the SLP. For instance, Rvachew and Brosseau-Lapré (2015) reported that an SLP-delivered input-based intervention comprising at least 16 different treatment procedures such as focused stimulation, chaining, delayed imitation, receptive and expressive minimal pairs was considered effective when paired with dialogic reading for home practice. It is important to note, that even though their approach was identified as an input-based approach, it still included opportunities for production practice.

Across empirical evidence reporting effective interventions for children with SSD, most involve speech production practice of carefully selected targets (with or without input-based procedures), implemented by an SLP through the use of a variety of instructional cues and feedback (Baker & McLeod, 2011; McLeod & Baker, 2017). If we are to move forward in closing the gap between supply and demand for intervention by using non-SLPs, the findings from the current study suggest that we need to focus our efforts on using effective interventions suited to children's needs, and adopt empirically-supported training strategies that facilitate faithful implementation of those interventions.

### **ACKNOWLEDGMENTS**

This research was supported by Australian Research Council Discovery grant DP130102545 and funding from the NSW Department of Education. The authors thank Felicity McKellar and the participating educators, parents, and children.

## References

- Allen, M. M. (2013). Intervention efficacy and intensity for children with speech sound disorders. *Journal of Speech, Language, and Hearing Research, 56*, 865–877.
- American Speech-Language-Hearing Association. (1985). Guidelines for identification audiometry. *ASHA, 27*, 49-52.
- Anthony, J. L., Aghara, R. G., Dunkelberger, M. J., Anthony, T. I., Williams, J. M., & Zhang, Z. (2011). What factors place children with speech sound disorder at risk for reading problems? *American Journal of Speech-Language Pathology, 20*(2), 146-160.
- Australian Bureau of Statistics (2008). *An introduction to Socio-Economic Indexes for Areas (SEIFA) 2006*. Canberra, Australia: Australian Bureau of Statistics.
- Australian Institute of Family Studies. (2007). *Growing up in Australia: The Longitudinal Study of Australian Children*. Retrieved from <http://www.aifs.gov.au/growingup/>
- Baker, E. (2012). Optimal intervention intensity in speech-language pathology: Discoveries, challenges, and uncharted territories. *International Journal of Speech-Language Pathology, 14*(5), 478-485.
- Baker, E. & McLeod (2011). Evidence-based practice for children with speech sound disorders: Part 1 narrative review. *Language, Speech, and Hearing Services in Schools, 14*, 102-139.
- Bothe, A. K. & Richardson, J. D. (2011). Statistical, practical, clinical, and personal significance: Definitions and applications in speech-language pathology. *American Journal of Speech-Language Pathology, 20*, 233-242.

- Bowen, C. & Cupples, L. (1999) Parents and children together (PACT): A collaborative approach to phonological therapy. *International Journal of Language and Communication Disorders, 34*, 35-55.
- Broen, P. A., & Westman, M. J. (1990). Project parent: A preschool speech program implemented through parents. *Journal of Speech and Hearing Disorders, 55*, 495–502.
- Broomfield, J., & Dodd, B. (2004). The nature of referred subtypes of primary speech disability. *Child Language Teaching and Therapy, 20*, 135-151.
- Chen, Y.-P. P., Johnson, C., Lalbakhsh, P., Caelli, T., Deng, G., Tay, D., . . . Morris, M. E. (2016). Systematic review of virtual speech therapists for speech disorders. *Computer Speech and Language, 37*, 98-128.
- Crosbie, S., Holm, A., & Dodd, B. (2005). Intervention for children with severe speech disorder: A comparison of two approaches. *International Journal of Language and Communication Disorders, 40*(4), 467-491.
- Crowe, K., Cumming, T., McCormack, J., Baker, E., McLeod, S., Wren, Y., Roulstone, S., & Masso, S. (2016). *Educators' perspectives on facilitating computer-assisted speech intervention in early childhood settings*. Manuscript in submission.
- Dodd, B. (2013). *Differential diagnosis and treatment of speech disordered children* (3rd ed.). Hoboken, NJ: Wiley.
- Dodd, B. & Barker, R. (1990). The efficacy of utilizing parents and teachers as agents of therapy for children with phonological disorders. *Australia Journal of Human Communication Disorders, 18*, 29-44.

- Dodd, B., Crosbie, S., McIntosh, B., Holm, A., Harvey, C., Liddy, M., . . . Rigby, H. (2008). The impact of selecting different contrasts in phonological therapy. *International Journal of Speech-Language Pathology, 10*(5), 334–345.
- Dodd, B., Hua, Z., Crosbie, S., Holm, A., & Ozanne, A. (2002). *Diagnostic Evaluation of Articulation and Phonology*. London, UK: Pearson
- Duggirala, V. & Dodd, B. (1991). A psycholinguistic assessment model for disordered phonology. *Congress for Phonetic Sciences, Aix-en-Provence, Université de Provence*, 342–345.
- Dunn, L. M., & Dunn, D. M. (2007). *The Peabody Picture Vocabulary Test - Fourth Edition (PPVT-4)*. Bloomington, MN: Pearson.
- Durlak, J. A. (2015). Studying program implementation is not easy but it is essential. *Prevention Science, 16*, 1123-1127.
- Durlak, J. A. & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*, 327-350.
- Eadie, P., Morgan, A., Ukoumunne, O. C., Ttofari Eecen, K., Wake, M., & Reilly, S. (2015). Speech sound disorder at 4 years: Prevalence, comorbidities, and predictors in a community cohort of children. *Developmental Medicine and Child Neurology, 57*(6), 578-584.
- Eiserman, W. D., McCoun, M., & Escobar, C. M. (1990). A cost-effectiveness analysis of two alternative program models for serving speech-disordered preschoolers. *Journal of Early Intervention, 14*(4), 297–317.



- Ehrler, D. J., & McGhee, R. L. (2008). *Primary Test of Nonverbal Intelligence*. Austin, TX: Pro-Ed.
- Farquharson, K., Tambyraja, S. R., Logan, J., Justice, L. M. & Schmitt, M. B. (2015). Using hierarchical linear modeling to examine how individual slps differentially contribute to children's language and literacy gains in public schools. *American Journal of Speech-Language Pathology, 24*, 504-516.
- Gillon, G. T. (2005). Facilitating phoneme awareness development in 3- and 4-year-old children with speech impairment. *Language, Speech, and Hearing Services in Schools, 36*, 308–324.
- Glascoc, F. P. (2000). *Parents' Evaluation of Developmental Status: Authorized Australian Version*. Parkville, Australia: Centre for Community Child Health.
- Glogowska, M., Roulstone, S., Enderby, P., & Peters, T. J. (2000). Randomised controlled trial of community based speech and language therapy in preschool children. *British Medical Journal, 321*, 1–5.
- Hodson, B. W. (2007). *Evaluation and enhancing children's phonological systems: Research and theory to practice*. Greenville, SC: Thinking Publications.
- Justice, L. M., Bowles, R. P., & Skibbe, L. E. (2006). Measuring preschool attainment of print-concept knowledge: A study of typical and at-risk 3- to 5-year-old children using item response theory. *Language, Speech, and Hearing Service in Schools, 37*(3), 224-235.
- Kay, J., Lesser, R., & Colheart, M. (1992). *Psycholinguistic assessments of language processing in aphasia*. Hove, UK: Lawrence Erlbaum Associates.

- Lancaster, G., Keusch, S., Levin, A., Pring, T., & Martin, S. (2010). Treating children with phonological problems: does an eclectic approach to therapy work? *International Journal of Language and Communication Disorders, 45*(2), 174–181.
- Law, J., Boyle, J., Harris, F., Harkness, A., & Nye, C. (2000). Prevalence and natural history of primary speech and language delay: Findings from a systematic review of the literature. *International Journal of Language and Communication Disorders, 35*, 165-188.
- Law, J., Garrett, Z., Nye, C. (2003/2010). Speech and language therapy interventions for children with primary speech and language delay or disorder. *Cochrane Database of Systematic Reviews*, Issue 3. Art. No.: CD004110. DOI: 10.1002/14651858.CD004110.
- Long, S. H., Fey, M. E., & Channell, R. W. (2008). *Computerized profiling* (Version MS-DOS version 9.7). Milwaukee, WI: Marquette University.
- McAllister, L., McCormack, J., McLeod, S., & Harrison, L. J. (2011). Expectations and experiences of accessing and participating in services for childhood speech impairment. *International Journal of Speech-Language Pathology, 13*, 251-267.
- McCormack, J., Baker, E., Crowe, K., Masso, S., McLeod, S., Wren, Y., & Roulstone, S. (2016). *Implementation fidelity of a computer-assisted intervention for children with speech sound disorders*. Manuscript in submission.
- McCormack, J. M., & Verdon, S. E. (2015). Mapping speech pathology services to developmentally vulnerable and at-risk communities using the Australian Early Development Census. *International Journal of Speech-Language Pathology, 17*(3), 273-286.
- McLeod, S. (2004). Speech pathologists' application of the ICF to children with speech impairment. *International Journal of Speech-Language Pathology, 6*(1), 75-81.

- McLeod, S., & Baker, E. (2014). Speech-language pathologists' practices regarding assessment, analysis, target selection, intervention, and service delivery for children with speech sound disorders. *Clinical Linguistics and Phonetics*, 28(7-8), 508-531.
- McLeod, S., & Baker, E. (2017). *Children's speech: An evidence-based approach to assessment and intervention*. Boston, MA: Pearson Education.
- McLeod, S., Crowe, K., Masso, S., Baker, E., McCormack, J., Wren, Y., Roulstone, S., & Howland, C. (2017, in press). Profile of Australian preschoolers with speech sound disorders at risk for literacy difficulties. *Australian Journal of Learning Difficulties*.
- McLeod, S., Crowe, K., McCormack, J., White, P., Wren, Y., Baker, E., Masso, S., Roulstone, S. (2015). *Preschool children's communication, motor and social development: What concerns parents and educators?* Manuscript in submission.
- McLeod, S., Crowe, K., & Shahaieian, A. (2015). Intelligibility in Context Scale: Normative and validation data for English-speaking preschoolers. *Language, Speech, and Hearing Services in Schools*, 46(3), 266-276.
- McLeod, S., Harrison, L. J., McAllister, L., & McCormack, J. (2013). Speech sound disorders in a community study of preschool children. *American Journal of Speech-Language Pathology*, 22(3), 503-522.
- McLeod, S., Harrison, L. J., & McCormack, J. (2012a). *Intelligibility in Context Scale*. Bathurst, Australia: Charles Sturt University. Retrieved from <http://www.csu.edu.au/research/multilingual-speech/ics>
- McLeod, S., Harrison, L. J. & McCormack, J. (2012b). Intelligibility in Context Scale: Validity and reliability of a subjective rating measure. *Journal of Speech, Language, and Hearing Research*, 55, 648-656.

- Miles, J. & Shevlin, M. (2001). *Applying regression and correlation: A guide for students and researchers*. London, UK: Sage.
- Morris, M., Perry, A., Unsworth, C., Skeat, J., Taylor, N., Dodd, K., . . . Duckett, S. (2005). Reliability of the Australian Therapy Outcome Measures for quantifying disability and health. *International Journal of Therapy and Rehabilitation, 12*(8), 340-346.
- Munson, B., Baylis, A. L., Krause, M. O., & Yim, D. (2010). Representation and access in phonological impairment. In C. Fougeron, B. Kühnert, M. D'Imperio, & N. Vallée (Eds.), *Laboratory phonology 10* (pp. 381–404). Berlin, Germany: Mouton de Gruyter.
- Perry, A., Morris, M., Unsworth, C., Duckett, S., Skeat, J., Dodd, K., . . . Reilly, K. (2004). Therapy outcome measures for allied health practitioners in Australia: The AusTOMs. *International Journal for Quality in Health Care, 16*(4), 285-291.
- Perry, A., & Skeat, J. (2004). *AusTOMs for speech pathology*. Melbourne, Australia: La Trobe University.
- Roulstone, S., Miller, L. L., Wren, Y., & Peters, T. J. (2009). The natural history of speech impairment of 8-year-old children in the Avon Longitudinal Study of Parents and Children: Error rates at 2 and 5 years. *International Journal of Speech-Language Pathology, 11*, 381-391.
- Ruggero, L., McCabe, P., Ballard, K. J., & Munro, N. (2012). Paediatric speech-language pathology service delivery: An exploratory survey of Australian parents. *International Journal of Speech-Language Pathology, 14*(4), 338–350.
- Ruscello, D. M., Cartwright, L. R., Haines, K. B., & Shuster, L. I. (1993). The use of different service delivery models for children with phonological disorders. *Journal of Communication Disorders, 26*, 193–203.

- Rvachew, S., & Brosseau-Lapr e, F. (2015). A randomized trial of 12 week interventions for the treatment of developmental phonological disorder in Francophone children. *American Journal of Speech-Language Pathology*, 24, 637-658.
- Shriberg, L. D., Fourakis, M., Hall, S. D., Karlsson, H. B., Lohmeier, H. L., McSweeney, J. L., . . . Wilson, D. L. (2010). Extensions to the Speech Disorders Classification System (SDCS). *Clinical Linguistics and Phonetics*, 24(10), 795-824.
- Shriberg, L. D., & Lof, G. L. (1991). Reliability studies in broad and narrow phonetic transcription. *Clinical Linguistics and Phonetics*, 5(3), 225-279.
- Stackhouse, J., & Wells, B. (1997). *Children's speech and literacy difficulties: A psycholinguistic framework*. London, UK: Whurr.
- Sugden, E., Baker, E., Munro, N., & Williams, A. L. (2016). Involvement of parents in intervention for childhood speech sound disorders: A review of the evidence. *International Journal of Language and Communication Disorders*, 51, 597–625.
- Sutherland, D., & Gillon, G. T. (2005). Assessment of phonological representations in children with speech impairment. *Language, Speech, and Hearing Services in Schools*, 36(4), 294–307.
- Sutherland, D., & Gillon, G. T. (2007). Development of phonological representations and phonological awareness in children with speech impairment. *International Journal of Language and Communication Disorders*, 42, 229-250.
- Thomas-Stonell, N. L., Oddson, B., Robertson, B., & Rosenbaum, P. L. (2010). Development of the FOCUS (Focus on the Outcomes of Communication Under Six), a communication outcome measure for preschool children. *Developmental Medicine and Child Neurology*, 52(1), 47-53.

- Thomas-Stonell, N., Robertson, B., Walker, J., Oddson, B., Washington, K., & Rosenbaum, P. (2012). *FOCUS©: Focus on the Outcomes of Communication Under Six*. Toronto, Canada: Holland Bloorview Kids Rehabilitation Hospital.
- Vanryckeghem, M., & Brutten, G. J. (2006). *KiddyCat: Communication attitude test for preschool and kindergarten children who stutter*. San Diego, CA: Plural Publishing.
- Wagner, R., Torgesen, J., & Rashotte, C. (2013). *Comprehensive Test of Phonological Processing - Second Edition (CTOPP-2)*. Austin, TX: Pro-Ed.
- Waring, R. & Knight, R. (2013) How should children with speech sound disorders be classified? A review and critical evaluation of current classification systems. *International Journal of Language and Communication Disorders*, 48, 25-40.
- Warren, S. F., Fey, M. E., & Yoder, P. J. (2007). Differential treatment intensity research: A missing link to creating optimally effective communication interventions. *Mental Retardation and Developmental Disabilities Research Reviews*, 13, 70–77.
- Whitehurst, G. J. & Lonigan, C. J. (1998). Child development and emergent literacy. *Child Development*, 69, 848-872.
- Williams, A. L. (2012). Intensity in phonological intervention: Is there a prescribed amount? *International Journal of Speech-Language Pathology*, 14(5), 456–461.
- Wren, Y., & Roulstone, S. (2006). *Phoneme Factory Sound Sorter*. London, UK: GL Assessment.
- Wren, Y., & Roulstone, S. (2008). A comparison between computer and tabletop delivery of phonology therapy. *International Journal of Speech-Language Pathology*, 10, 346-363.

Wren, Y., & Roulstone, S. (2013). *Phoneme Factory Sound Sorter* (version 2, Australian adaptation) [Computer software]. Bristol, UK: Bristol Speech and Language Therapy Research Unit.

Wren, Y., Miller, L. L., Peters, T., Emond, A., & Roulstone, S. (2016). Prevalence and predictors of persistent speech sound disorder at eight years old: Findings from a population cohort study. *Journal of Speech, Language, and Hearing Research, 59*, 647-673.

Table 1.

*Eligibility criteria from Stages 1, 2 and 3 for inclusion in stage 4 (intervention/control) of the Sound Start Study*

Stage	Skill	Criteria	Assessment tool	Informant
Stage 1: Screening for eligibility by parents/teacher	Speech	“Difficulty talking and making speech sounds” ( <i>yes or a little concerned</i> )	Parent Evaluation of Developmental Status (PEDS, Glascoe, 2000)	Parent and/or teacher
	Speech	Speech not clear to family and/or others	Parent questionnaire <sup>a</sup>	Parent and/or teacher
	Medical and developmental history	Reported no persistent hearing loss, cleft lip/palate or developmental delay	Parent questionnaire <sup>a</sup>	Parent and/or teacher
	English language use	English language skills had to be equivalent to or better than their skills in the other language	Parent questionnaire	Parent
Stage 2: Assessed for eligibility following direct screening assessment by SLP	Speech	Percentage of consonants correct (PCC) standard score $\leq 6$ (i.e., greater than one standard deviation below the mean) Had a phonological pattern that could be treated by PFSS	Diagnostic Evaluation of Articulation and Phonology – Phonology subtest (DEAP, Dodd et al., 2002)	Direct assessment
	Nonverbal intelligence	Year 1: Nonverbal Index Score $>79$ (%ile rank of $>8$ ). Year 2-3: NVI $>70$ (%ile rank of $>2$ ).	Primary Test of Nonverbal Intelligence (PTONI, Ehrler & McGhee, 2008);	Direct assessment
Stage 3: Assessed for eligibility following direct comprehensive assessment by SLP	Receptive vocabulary	Receptive vocabulary score less than 2 SD below the mean	Peabody Picture Vocabulary Test - Fourth Edition (PPVT-4, Dunn & Dunn, 2007)	Direct assessment
	Hearing	Passed hearing screening of all frequencies at 40dB during Stage 2	Pure-tone audiometry (500, 1000, 2000, 4000 Hz) at 40dB (adjusted	Direct assessment



or Stage 3

for noisy environments)  
(American Speech-Language-  
Hearing  
Association, 1985)

---

*Note.* <sup>a</sup> Question used with permission from Longitudinal Study of Australian Children (LSAC, Australian Institute of Family Studies, 2007).

Table 2.

*Characteristics of the Intervention and Control Groups in the Sound Start Study (n = 123) prior to intervention*

	Measure	Valid <i>n</i>	Intervention group	Control group	Significant difference
			<sup>a</sup> <i>M</i> (SD) <sup>b</sup> <i>n</i> (%)	<sup>a</sup> <i>M</i> (SD) <sup>b</sup> <i>n</i> (%)	
Age <sup>a</sup>	In months	123	55.3 (4.1)	56.9 (4.3)	.044
Sex <sup>b</sup>	Male	79	43 (66.2%)	36 (62.1%)	.637
	Female	44	22 (33.8%)	22 (37.9%)	
Multilingual status <sup>b</sup>	English only	100	52 (80.0%)	48 (82.8%)	.695
	Multilingual	23	13 (20.0%)	10 (17.2%)	
Socio-economic status <sup>a</sup>	IRSAD	123	6.2 (2.9)	5.8 (3.5)	.495
Speech severity <sup>a</sup>	DEAP percentage of consonants correct	123	68.15 (9.90)	63.19 (13.76)	.022
Consonant stimulability <sup>a</sup>	Number of English consonants /24	123	22.3 (1.8)	22.0 (1.9)	.338
Phonological awareness composite score <sup>a</sup>	CTOPP-2 phonological awareness composite score	122	89.0 (11.5)	87.4 (10.4)	.432
Nonverbal IQ <sup>a</sup>	PTONI raw score	123	22.4 (8.8)	22.0 (8.9)	.788

Receptive vocabulary <sup>a</sup>	PPVT raw score	123	68.1 (18.3)	71.5 (19.2)	.309
-----------------------------------	----------------	-----	-------------	-------------	------

Note. <sup>a</sup>*M* (SD), <sup>b</sup>*n* (%); IRSAD, Index of Relative Socio-Economic Disadvantage; DEAP, Diagnostic Evaluation of Articulation and Phonology (Dodd et al., 2002); CTOPP-2, Comprehensive Test of Phonological Processing – 2 (Wagner et al., 2013); PTONI, Primary Test of Nonverbal Intelligence (Ehrler & McGhee, 2008); PPVT, Peabody Picture Vocabulary Test (Dunn & Dunn, 1997).

Table 3.

*Outcomes of the Intervention and Control Groups in the Sound Start Study: Intention to Treat (Complete Cases)*

	Measure	Group	Valid <i>n</i>	Stage 2/3	Stage 5	Stage 6	Effect	<i>F</i>	<i>p</i>	$\eta_p^2$
				<i>M</i> (SD)	<i>M</i> (SD)	<i>M</i> (SD)				
<b>Speech Production</b>										
Percentage of consonants correct	DEAP	Intervention	59	69.25 (9.13)	75.40 (10.27)	76.06 (9.99)	Stage	42.36	<.001	.287
		Control	48	64.34 (11.74)	69.77 (12.97)	71.14 (12.78)	Group	6.98	.009	.062
							Interaction	0.14	.874	.001
Primary intervention target probes	Sound Start Study probes	Intervention	59	71.27 (25.7)	47.14 (38.1)	43.28 (36.0)	Stage	48.79	<.001	.319
		Control	47	73.30 (28.4)	57.41 (36.5)	48.19 (36.2)	Group	0.99	.322	.009
							Interaction	1.12	.329	.011
Intelligibility	ICS	Intervention	30	3.75 (0.36)	3.95 (0.39)	3.97 (0.45)	Stage	4.14	.019	.094
		Control	14	3.90 (0.38)	4.06 (0.48)	4.01 (0.28)	Group	0.84	.365	.021
							Interaction	0.32	.726	.008
<b>Emergent Literacy</b>										
Letter knowledge	Anthony et al. (2011)	Intervention	44	8.02 (8.58)	10.95 (9.41)	11.75 (9.40)	Stage	19.18	<.001	.186
		Control	42	5.95 (7.96)	6.90 (8.82)	9.02 (9.43)	Group	2.61	.110	.030
							Interaction	1.68	.190	.020
Word and Print awareness	PWPA	Intervention	59	6.24 (3.43)	7.59 (2.75)	-	Stage	62.31	<.001	.368
		Control	51	5.69 (3.09)	7.29 (3.16)	-	Group	0.03	.858	.000
							Interaction	3.84	.053	.035
Phonological awareness: <i>Elision</i>	CTOPP-2	Intervention	55	2.58 (4.36)	3.38 (3.83)	3.91 (4.11)	Stage	19.99	<.001	.172
		Control	43	3.02 (3.83)	4.19 (3.86)	5.37 (3.96)	Group	1.411	.238	.014
							Interaction	1.585	.208	.016
Phonological awareness: <i>Blending Words</i>	CTOPP-2	Intervention	55	4.09 (4.02)	5.33 (4.54)	6.07 (4.62)	Stage	21.44	<.001	.181
		Control	44	3.66 (3.14)	5.61 (3.40)	5.77 (3.87)	Group	0.043	.837	.000
							Interaction	0.677	.509	.007
Phonological awareness: <i>Sound Matching</i>	CTOPP-2	Intervention	56	3.88 (3.30)	4.45 (3.56)	5.77 (3.87)	Stage	1.401	.249	.014
		Control	43	4.09 (3.28)	3.93 (2.52)	4.33 (3.14)	Group	0.320	.573	.003
							Interaction	0.660	.518	.007
<b>Phonological Processing</b>										
Phonological memory: <i>Memory for Digits</i>	CTOPP-2	Intervention	55	10.44 (2.86)	11.33 (2.99)	11.60 (2.65)	Stage	7.918	<.001	.077
		Control	42	11.50 (3.19)	11.88 (3.62)	12.40 (2.58)	Group	2.316	.131	.024
							Interaction	0.473	.624	.005

Phonological memory: <i>Nonword Repetition</i>	CTOPP-2	Intervention	55	2.65 (2.78)	3.62 (3.75)	3.75 (3.64)	Stage	5.808	.004	.057
		Control	43	2.49 (3.31)	4.05 (3.63)	3.65 (3.41)	Group	0.012	.914	.000
							Interaction	0.012	.727	.003
Rapid non-symbolic naming: <i>Rapid Color Naming</i>	CTOPP-2	Intervention	34	58.29 (18.86)	54.94 (20.13)	56.76 (21.95)	Stage	3.468	.035	.060
		Control	22	58.00 (15.58)	51.14 (12.88)	52.68 (16.35)	Group	0.366	.548	.007
							Interaction	0.571	.567	.010
Rapid non-symbolic naming: <i>Rapid Object Naming</i>	CTOPP-2	Intervention	41	64.02(19.77)	63.44(21.20)	65.41(23.44)	Stage	0.964	.384	.013
		Control	34	62.38(12.61)	62.18(17.49)	64.35(18.82)	Group	0.103	.749	.001
							Interaction	0.017	.983	.000
<b>Children's Participation and Wellbeing</b>										
Parent-reported communication outcomes	FOCUS	Intervention	29	253.4 (49.3)	256.8 (52.7)	261.1 (49.6)	Stage	1.83	.167	.042
		Control	15	256.5 (38.1)	269.2 (38.4)	267.5 (49.7)	Group	0.27	.606	.006
							Interaction	0.41	.668	.010
Child reported Attitude score	KiddyCAT	Intervention	57	3.68 (2.42)	3.23 (2.61)	2.26 (2.39)	Stage	7.99	<.001	.073
		Control	46	3.76 (2.48)	3.61 (2.94)	3.15 (2.37)	Group	1.25	.267	.012
							Interaction	1.24	.292	.012
Child reported happy ☺ occurrences	SPAA-C	Intervention	56	6.95 (2.76)	6.41 (2.38)	5.90 (2.62)	Stage	1.62	.200	.016
		Control	44	5.60 (3.04)	5.87 (2.71)	5.64 (2.60)	Group	3.13	.080	.030
							Interaction	1.91	.151	.018

Note. DEAP, Diagnostic Evaluation of Articulation and Phonology – Phonology subtest (Dodd et al., 2002); ICS, Intelligibility in Context Scale (McLeod et al., 2012a); PWPA, Preschool Word and Print Awareness measure (Justice et al., 2006); CTOPP-2, Comprehensive Test of Phonological Processing – 2 (Wagner et al., 2013); FOCUS, Focus on the Outcomes of Children Under Six (Thomas-Stonell et al., 2012); KiddyCAT, Kiddy-Communication and Attitude Test (Vanryckeghem & Brutton, 2007); SPAA-C, Speech Participation and Activity Assessment of Children (McLeod, 2004).

Table 4

*Outcomes of the Intervention and Control Groups in the Sound Start Study: Per Protocol*

	Measure	Group	Valid <i>n</i>	Stage 2/3	Stage 5	Stage 6	Effect	<i>F</i>	<i>p</i>	$\eta_p^2$
				<i>M</i> (SD)	<i>M</i> (SD)	<i>M</i> (SD)				
<b>Speech Production</b>										
Percentage of consonants correct	DEAP	Intervention	22	72.87 (6.99)	78.28 (7.91)	79.29 (7.77)	Stage	28.81	<.001	.324
		Control	40	63.74 (11.96)	69.85 (13.88)	70.87 (13.30)	Group	9.36	.003	.135
							Interaction	0.09	.916	.001
Primary intervention target probes	Sound Start Study probes	Intervention	22	64.83 (27.81)	45.04 (39.01)	43.44 (37.45)	Stage	23.33	<.001	.283
		Control	39	71.83 (29.04)	57.86 (37.09)	50.85 (36.36)	Group	1.16	.286	.019
							Interaction	0.49	.614	.008
Intelligibility	ICS	Intervention	10	3.86 (0.42)	4.01 (0.32)	3.87 (0.47)	Stage	1.48	.243	.008
		Control	9	4.00 (0.43)	4.14 (0.47)	4.02 (0.31)	Group	0.84	.372	.047
							Interaction	0.00	.996	.000
<b>Emergent Literacy</b>										
Letter knowledge	Anthony et al. (2011)	Intervention	18	11.61 (9.56)	15.17 (9.54)	16.00 (9.59)	Stage	13.27	<.001	.206
		Control	35	5.74 (8.03)	7.11 (9.08)	9.31 (9.51)	Group	7.63	.008	.130
							Interaction	1.00	.371	.019
Word and Print awareness	PWPA	Intervention	22	6.50 (3.43)	7.23 (3.28)	-	Stage	11.73	.001	.159
		Control	42	5.50 (3.15)	7.12 (3.25)	-	Group	0.50	.483	.008
							Interaction	1.69	.198	.027
Phonological awareness: <i>Elision</i>	CTOPP-2	Intervention	22	2.23 (3.78)	3.32 (4.89)	3.36 (4.26)	Stage	15.16	<.001	.216
		Control	35	3.06 (4.06)	4.29 (4.00)	5.40 (4.22)	Group	1.407	.241	.025
							Interaction	2.107	.127	.037
Phonological awareness: <i>Blending Words</i>	CTOPP-2	Intervention	22	5.45 (4.70)	7.18 (5.67)	7.64 (4.94)	Stage	19.06	<.001	.254
		Control	36	3.72 (3.04)	5.64 (3.59)	6.24 (3.75)	Group	2.26	.139	.039
							Interaction	0.477	.898	.002
Phonological awareness: <i>Sound Matching</i>	CTOPP-2	Intervention	22	4.09 (3.81)	5.45 (4.34)	5.41 (5.39)	Stage	3.014	.053	.052
		Control	35	3.66 (3.06)	4.03 (2.47)	4.63 (3.14)	Group	1.176	.283	.021
							Interaction	0.536	.587	.010
<b>Phonological Processing</b>										
Phonological memory: <i>Memory for Digits</i>	CTOPP-2	Intervention	22	10.50 (3.25)	11.55 (3.10)	11.68 (2.23)	Stage	6.138	.003	.100
		Control	35	11.23 (3.18)	11.77 (2.96)	12.14 (2.55)	Group	0.451	.505	.008
							Interaction	0.325	.723	.006

Phonological memory: <i>Nonword Repetition</i>	CTOPP-2	Intervention	22	2.86 (2.95)	4.32 (4.91)	4.36 (3.81)	Stage	1.678	.192	.030
		Control	35	2.71 (3.42)	4.37 (3.73)	3.31 (3.52)	Group	0.877	.353	.016
							Interaction	0.670	.514	.012
Rapid non-symbolic naming: <i>Rapid Color Naming</i>	CTOPP-2	Intervention	15	59.47 (22.53)	52.73 (22.16)	56.00 (24.08)	Stage	2.857	.065	.082
		Control	19	57.63 (16.09)	51.79 (13.52)	53.58 (17.48)	Group	0.086	.771	.003
							Interaction	0.039	.961	.001
Rapid non-symbolic naming: <i>Rapid Object Naming</i>	CTOPP-2	Intervention	16	64.44 (25.88)	62.06 (27.53)	65.94 (29.47)	Stage	1.823	.168	.041
		Control	29	62.28 (12.79)	60.52 (16.91)	64.72 (19.46)	Group	0.072	.790	.002
							Interaction	0.048	.827	.001
<b>Children's Participation and Wellbeing</b>										
Parent-reported communication outcomes	FOCUS	Intervention	11	246.6 (49.2)	261.9 (39.0)	253.9 (36.8)	Stage	1.07	.355	.056
		Control	9	261.4 (25.7)	267.9 (47.0)	261.9 (54.9)	Group	0.31	.584	.017
							Interaction	0.19	.827	.011
Child reported Attitude score	KiddyCAT	Intervention	22	3.32 (2.15)	3.36 (2.26)	2.45 (2.35)	Stage	2.93	.057	.047
		Control	39	3.87 (2.45)	3.51 (2.86)	3.21 (2.40)	Group	0.83	.367	.014
							Interaction	0.42	.657	.007
Child reported happy ☺ occurrences	SPAA-C	Intervention	21	7.14 (2.74)	6.38 (2.46)	6.00 (2.85)	Stage	2.09	.128	.037
		Control	36	6.11 (3.04)	6.47 (2.49)	5.72 (2.53)	Group	0.48	.493	.009
							Interaction	1.09	.341	.019

Note. DEAP, Diagnostic Evaluation of Articulation and Phonology – Phonology subtest (Dodd et al., 2002); ICS, Intelligibility in Context Scale (McLeod et al., 2012a); PWPA, Preschool Word and Print Awareness measure (Justice et al., 2006); CTOPP-2, Comprehensive Test of Phonological Processing – 2 (Wagner et al., 2013); FOCUS, Focus on the Outcomes of Children Under Six (Thomas-Stonell et al., 2012); KiddyCAT, Kiddy-Communication and Attitude Test (Vanryckeghem & Brutton, 2007); SPAA-C, Speech Participation and Activity Assessment of Children (McLeod, 2004).

*Figure 1.* Participant recruitment and randomization flow diagram.