

Implementation fidelity of a computer-assisted intervention for children with speech sound disorders

Abstract

Background: Implementation fidelity refers to the degree to which an intervention or program adheres to its original design. This paper examines implementation fidelity in the Sound Start Study, a clustered randomised controlled trial of computer-assisted support for children with speech sound disorders (SSD).

Method: 63 children with SSD in 19 early childhood centres received computer-assisted support (Phoneme Factory Sound Sorter [PFSS] – Australian version; Wren & Roulstone, 2013). Educators facilitated the delivery of pre-set games in PFSS targeting phonological error patterns identified by a speech-language pathologist. Implementation data were gathered via: (1) the computer software, which recorded when and how much intervention was completed over 9 weeks; (2) educators' records of practice sessions; and (3) scoring of fidelity (intervention procedure, competence and quality of delivery) from videos of intervention sessions.

Result: Less than one third of children received the prescribed number of days of intervention, while approximately one half participated in the prescribed number of intervention plays. Computer data differed from educators' data for total number of days and plays in which children participated; the degree of match was lower as data became more specific. Fidelity to intervention procedures, competency and quality of delivery was high.

Conclusion: Fidelity may impact intervention outcomes and so needs to be measured in intervention research; however, the way in which implementation fidelity is measured may impact on data.

Introduction

Evidence-based practice (EBP) is core to the provision of health services throughout the world, and speech-language pathology is no exception (e.g. American Speech-Language-Hearing Association, 2005; Speech Pathology Australia, 2011). According to Carroll and colleagues (2007), “evidence-based practice assumes that an intervention is being implemented in full accordance with its published details” (p. 2). According to this reasoning, selecting and implementing a particular intervention simply because empirical research exists to support the approach would not equate to evidence-based practice, unless it was implemented as directed. Furthermore, adaptation, based on clinical experience or client preferences, would not be appropriate. However, adaptations can sometimes lead to improved outcomes for clients (Durpak & DuPre, 2008) and other forms of evidence still have a place in clinical decision-making. Indeed, individualisation is recognised as an important component of intervention (Roth & Worthington, 2015). For this reason, Dollaghan (2007) suggested that engaging in EBP requires clinicians to consider and integrate multiple forms of evidence in clinical practice: empirical research, clinical experience, and information from clients. However, at times, tension exists between these forms of evidence (Odom, 2009). For instance, empirical research might exist to support the use of a particular intervention approach, but clinicians need to adapt the approach in the clinical setting due to issues including time, resourcing, or client characteristics (Roulstone, Wren, Bakapoulou, & Lindsay, 2012). So how do we determine the components of an intervention approach that should not be modified, and how might adaptations influence the effectiveness of the intervention we deliver? Evaluations of the implementation fidelity of interventions might assist.

Implementation fidelity

Implementation fidelity refers to the degree to which an intervention approach is

implemented in accordance with its published details, and as intended by its developers (Dusenbury, Brannigan, Falco, & Hansen, 2003). In intervention research, evaluation of implementation fidelity provides insight into the contribution of the intervention to outcomes obtained, and reduces the likelihood of researchers drawing false conclusions about an intervention's effectiveness (Carroll et al., 2007). When unexpected results are obtained, evaluations of implementation fidelity can assist in determining whether the intervention itself was ineffective or whether the quality of implementation had an impact on the intervention outcomes (Carroll et al., 2007). Evaluations of implementation fidelity can also provide insights into the aspects of the intervention that are most necessary for the intervention to be effective, the aspects of the intervention that are most difficult to translate from research into clinical settings, and the training and support strategies that those implementing the intervention require in order to do so effectively. That is, implementation fidelity can assist in the development of effective, practical, sustainable and clinically-relevant interventions.

Currently, in the field of speech sound disorders (SSD), there is a growing body of empirical research, which supports the provision of intervention for children. Law and colleagues (2013) conducted a meta-analysis of 25 randomised controlled trials (RCTs) to examine the outcomes of interventions for a range of speech and language difficulties. While they concluded that the evidence for interventions targeting some communication difficulties was variable, they found the evidence for interventions targeting SSD showed that they were effective. The RCTs examined within the Law et al. (2013) meta-analysis covered a range of different approaches. Additional approaches have been documented within a narrative review and found to be effective (Baker & McLeod, 2011). However, few of these studies have reported implementation fidelity. In other health fields, a lack of research examining implementation fidelity has also been noted (Brietenstein et al., 2010). Consequently, we

know that these interventions have the potential to be effective, but the necessary components are less clear, as are the impacts of contextual adaptations.

Elements of implementation fidelity

One barrier to evaluating implementation fidelity is the variation in how the construct is defined and measured. Carroll et al. (2007) conducted a literature review of studies (primarily from 2002-2007) examining implementation fidelity in order to determine key elements within this construct, and then proposed a framework to illustrate the relationships between them. A description of each of these elements is given in Table I.

[Insert Table I here]

In their framework, Carroll et al. (2007) proposed that measurement of implementation fidelity was essentially the measurement of *adherence*, which they proposed as an overarching term to include the faithful delivery of intervention content, and the faithful delivery of the intervention at the prescribed intensity including: frequency, duration and coverage (dose). Based on Warren, Fey, and Yoder's (2007) conceptualization of intervention intensity, *frequency* can refer to the number of times a particular dose or session is provided per unit of time (e.g., 30 minutes x twice weekly; 100 trials x twice weekly), *duration* as the time period of a session (e.g., 30 minutes), and/or the time period over which intervention is conducted (e.g., 10 weeks), and *coverage* as different aspects of dose or amount of intervention completed. Coverage includes *session dose*--the number of teaching episodes in a session (e.g., 50 trials) and *cumulative dose*--the total number of teaching episodes completed over the total period of intervention (e.g., 1000 trials) and/or the total amount of time spent on intervention (e.g., 10 hours). Warren et al. (2007) suggested that cumulative intervention intensity can be calculated via the product of session dose x session frequency x total intervention duration in time (e.g., 50 trials x 2 week x 10 weeks = 1000 trials). Carroll et al. proposed that the degree of adherence can be moderated by a series of other factors

including the quality of delivery. Finally, they noted that an analysis of outcomes could help to identify the components of the intervention required for the intervention to be effective (i.e. program differentiation). Similarly, Brietenstein and colleagues (2010) recognised adherence as a key element of implementation fidelity, alongside competence. They described adherence as the degree to which the behaviours of those implementing the intervention conformed to the intervention protocol; and competence as the skillfulness of those people in intervention delivery (including communication skills, technical abilities, and responsiveness to the needs of participants). Thus, Carroll’s concept of “quality delivery” and Brietenstein’s concept of “competence” may be similar.

Brietenstein et al. (2010) suggested that adherence could be measured by examining the quantity or presence of prescribed behaviours, through self-report or observations (live or via video/audio recordings). However, they cautioned that the content of the fidelity instrument was important in order that it “capture behaviours and processes that are congruent with the underlying theoretical framework and reflective of the core components of the intervention” (Brietenstein et al., 2010, p. 7).

The purpose of this paper is to report on the implementation fidelity of the computer-assisted intervention program, Phoneme Factory Sound Sorter (PFSS – Australian version; Wren & Roulstone, 2013), delivered to children with SSD in the Sound Start Study, a clustered randomised control trial.

Phoneme Factory Sound Sorter—Australian version (PFSS) (Wren & Roulstone, 2013)

PFSS is a computer-assisted program developed to target input processing skills in children with SSD of unknown origin. The design of the program was guided by the psycholinguistic model (Stackhouse & Wells, 1997), which recognises that speech output errors can be the result of underlying difficulties with input processing. Consequently, the program aims to strengthen input processing skills (including auditory perception and

phonological representations) via speech processing games, rather than targeting output skills directly.

Within PFSS, there are seven interactive game types, each of which can be customised to a child's specific needs based on their speech sound errors. Each game type targets a different aspect of input processing such as rhyme awareness (i.e. listening to spoken words and identifying whether they rhyme), phoneme detection (i.e. listening to spoken words and identifying the sounds within the word), phoneme blending (i.e. listening to a series of sounds and identifying the word they combine to produce) and minimal pair discrimination (i.e. distinguishing between two spoken words which differ in one sound only such as *tea* and *key*). For example, in the "Pair and Pick" game, picture pairs (e.g., *key*, *tea*) are presented on animated bubbles on the computer screen. The child then hears one of the words (e.g., *key* [ki]) via the computer speakers. The child is instructed to pick the picture that best matches the spoken word. An animated character provides the child with feedback about the response accuracy (correct/incorrect). All of the games require children to complete 10 trials, and once completed, the program stores children's performance data. If the full 10 trials are not completed, no data is recorded on the computer running the program.

PFSS has two intervention settings: the free configuration setting, and the teacher setting. The free configuration setting allows the user to determine how the program will be used from session to session. The teacher setting comprises a series of pre-set modules, each designed to target a specific phonological error pattern identified in a child's speech. Each module contains four to seven levels, with each level comprising three to five games. A level could be repeated or played multiple times before starting the next level when the child is ready or at regular agreed intervals in time. The levels within a module represent an increasingly level of difficulty such that children start at an easy level and progress through to more challenging levels either as they improve or over time. In this way, the pre-set settings

can follow a time-based or a performance-based criterion. Figure 1 includes a screen shot of the four games that comprise the level 1 for the gliding module in the teacher setting.

Wren and Roulstone (2008) examined the effectiveness of PFSS compared to a traditional table-top approach and no therapy using a randomised controlled trial design with 33 children. In their study, the intervention was delivered three times a week. One of the sessions was conducted by an SLP. The other two sessions were conducted by an assistant, who observed the weekly SLP session. Wren and Roulstone (2008) used the free-configuration setting, tailoring the processes targeted and the type and number of games played from session to session, in light of a child's performance. They found that although the children's speech production skills were not statistically significantly between groups after intervention, the children who received intervention showed signs of greater improvement compared with the children in the control group. Given these promising results, PFSS was modified for the Australian context and investigated in a larger community-based RCT. For this investigation, PFSS was delivered by educators using the pre-set teacher setting. This service delivery option was considered in an effort to identify a solution that could help address the gap between the demand and supply for SLP services for children in Australia (McAllister, McCormack, McLeod, & Harrison, 2011; Ruggero, McCabe, Ballard, & Munro, 2012).

Research aims

Implementation fidelity is essential to examine the relationships between outcomes, adherence, and barriers and facilitators to implementation in clinical research. In this paper, adherence to the Sound Start Study implementation protocol is examined, focusing on *coverage* (dose) with respect to the total number of days on which PFSS was played and the total number of games played over the total period of intervention. Adherence was measured in two ways to examine the impact of measurement tools on the adherence rates obtained.

Furthermore, the quality of the delivery (or competence in delivering the intervention) was evaluated to examine the moderating influence of this construct on adherence. Consequently, the aims of the current research were to determine:

1. adherence to the prescribed coverage of PFSS intervention in terms of cumulative dose, measured in time (days) and total number of games played (plays);
2. the impact of measurement tools on the evaluation of adherence, specifically comparing measurement by the computer and educators; and
3. the quality of delivery by considering procedural fidelity.

Method

Context of the study

The Sound Start Study was designed to explore the effectiveness of PFSS – Australian version (Wren & Roulstone, 2013) in supporting the speech and emergent literacy skills of Australian preschool children with SSD, when delivered by educators, using the pre-set teacher settings. The study was a blinded clustered randomised controlled trial in which the performance of children with SSD who received the PFSS intervention was compared with a group of children with SSD who did not. Children were randomly allocated to the intervention/control arm of the study based on the early childhood education centres they attended (i.e. centres were randomised to receive the program, or not). Educators (teachers and/or teaching assistants) facilitated PFSS with children in their centres using intervention targets prescribed by speech-language pathologists (SLP) based on assessment outcomes. The effectiveness of the PFSS intervention is discussed elsewhere (see McLeod et al., 2016); however, to summarize, PFSS intervention administered by educators did not result in greater improvement than typical classroom practices. In the current paper, the adherence to the intervention protocol was examined via comparison of three data sets: (1) the computer software, which provided evidence of the number of days and games played by each child

each week for the entire period of intervention, (2) educators' records of number of intervention days, sessions, and games played each week on a hard-copy recording sheet; and (3) SLPs' fidelity scoring from videos of the intervention sessions.

Participants

The Sound Start Study was conducted over 3 years. Early childhood centres across Sydney were approached by the research team and invited to participate in the study. Centres were chosen to represent a broad range of socioeconomic regions. A total of 19 early childhood centres were involved in the implementation of PFSS across the three years; three of these sites participated in more than one year (i.e. 16 unique settings). The settings were New South Wales Department of Education and Communities preschools ($n = 10$), community preschools ($n = 2$), local council preschools ($n = 2$), a preparatory program in an independent private school ($n = 1$), and a privately owned long day care centre ($n = 1$). Centres had between 1 and 13 children participating in the PFSS program ($M = 3.9$).

The PFSS intervention was the fourth of six stages in the Sound Start Study. Thus, children who received the intervention had already progressed through stages 1-3. In stage 1, children were identified by parents/early childhood teachers with communication concerns via a written questionnaire. In stage 2, they participated in a speech assessment with a SLP and some were diagnosed with SSD. In stage 3, those diagnosed with SSD participated in further assessments to identify the nature of their SSD. Those with a SSD which could not be attributed to a structural or genetic cause, and appeared primarily phonological in nature (i.e. one or more phonological patterns were present) were assigned to the control or intervention arm of the study based on their centre. Phonological impairment was diagnosed on the basis of their performance on the Diagnostic Evaluation of Articulation and Phonology (Dodd, et al., 2002), and study-specific phonological patterns probes were administered to provide further information about each child's most pervasive phonological patterns. Further

information about the research design is provided in McLeod et al. (2016).

Across the three years, 65 children commenced the intervention, but during the intervention phase, 2 withdrew. Thus, data is reported for 63 participants. The participants ranged in age from 4;1 to 5;5 ($M = 55.4$ months; $SD = 4.2$) when they were assessed. There were more males ($n = 41$, 65.1%) than females ($n = 22$, 34.9%). The participants lived in a range of metropolitan suburbs from the most disadvantaged (1st decile) to most advantaged (10th decile) according to the Australian Index of Relative Socioeconomic Advantage and Disadvantage (IRSAD, ABS, 2008). The mean IRSAD decile of participants was 6.2 ($SD = 2.8$). The majority of participants ($n = 51$, 81.0%) only spoke English only at home, ten spoke English and an additional language at home, and two spoke English and two additional languages at home. The additional languages were Arabic, Filipino, Greek, Hindi, Korean, Malayalam, Marathi, Punjabi, Spanish and Thai.

Intervention: Phoneme Factory Sound Sorter (PFSS) program

Each child who was assigned to the intervention arm of the trial was allocated specific pre-set modules using the teacher setting in PFSS (Wren & Roulstone, 2013). The specific module(s) selected for each child targeted the phonological error patterns with the highest percentage occurrence in the child's speech. The patterns were identified by the second author (SLP) following an analysis of each child's speech samples from the DEAP Phonology Assessment (Dodd et al., 2002), and the study-specific phonological probes. When two or more patterns had the same percentage occurrence, the pattern with the earlier age of disappearance developmentally was prioritised (e.g. 2-element cluster reduction was prioritised over 3-element cluster reduction).

Intervention protocol

An intervention protocol was developed, prescribing the amount of intervention to be provided in the intervention arm of the trial. The protocol was based on previous intervention

research with children with SSD which has indicated 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 (Allen, 2013; Dodd et al., 2008; Ruscello et al., 1993). It was also influenced by the practicalities of a busy preschool schedule and that most children attend preschool two or three days per week rather than every weekday. Given these findings, the Sound Start Study protocol stipulated that the PFSS program should be facilitated by an educator at the child's early childhood centre over a 9 week period. Each week, the child was to receive four sessions of their PFSS program. A session was defined as the completion of the three to five games comprising a level of a module. Given the attendance schedule of most children, this typically meant two sessions on one day, and two sessions on another day of that same week (e.g. one session in the morning and one in the afternoon twice a week). In this way, if a child was assigned a pre-set module containing six levels, the child would have to complete the three to five games comprising level one, twice on two days in a week, equivalent to four sessions per week (see Figure 1). The child would start a new level the following week until all six levels in the pre-set module had been completed. Given that the total intervention duration was 9 weeks, children completed as many pre-set modules as possible over 9 weeks. If a child was part-way through the levels in a pre-set module by the 9th week, the child stopped at that level rather than completing the module.

Each session was anticipated to last for approximately 15-20 minutes. The protocol also stipulated that the educator select the same level of a module for each of the four sessions across two days in one week regardless of the child's performance, then progress to the next level the following week, again regardless of the child's performance. That is, the intervention had a time-based criterion for progress, rather than a performance-based criterion. Changing the level each week increased the complexity of the games that the child

completed (e.g. more complex words, more contrasts, less visual support).

In order to ensure the intervention was implemented consistently across sites, the protocol described the roles and responsibilities of the educators who were facilitating the children's participation in the intervention, and the research team members who were the points of contact for these staff. Responsibilities of the educators included: completing training in the program, monitoring the children's participation, changing the PFSS program level each week, noting attendance/participation in the program and participating in interviews post-intervention. The responsibilities of the research team included: providing the training to educators, identifying the appropriate pre-set module for children (i.e. the phonological error pattern that would be targeted), resolving technical or implementation issues as these arose, and visiting the preschools to check that records were being maintained.

Procedure

Once randomisation of centres had occurred, those in the intervention arm were provided with lap-tops on which the PFSS program was downloaded. Educators who would facilitate PFSS sessions were nominated by each centre director, typically based on their willingness to participate. The nominated educators were provided with a copy of the intervention protocol and given initial training in the features of the PFSS program by one of the research SLPs. The same SLP attended the first intervention session at the early childhood centre to ensure the intervention was facilitated consistently by educators across sites and to resolve any difficulties that arose. Intervention was expected to continue for 9 weeks, with the educator required to record the details of intervention for each child receiving the PFSS program at the centre. The research SLP monitored and videoed intervention in weeks 2-3 and 7-9 for later fidelity checks; 30 (47.6%) children were videoed once in week 2 or 3 and once in week 7, 8 or 9, 27 (42.86%) were videoed once in week 2-3 or week 7-9, and six (9.5%) children were not videoed in weeks 2-3 or weeks 7-9.

Measures

Two data sets enabled an evaluation of the adherence to the prescribed intervention protocol within the Sound Start Study: the data stored within the PFSS computer program regarding the children's participation throughout the intervention, and data recorded simultaneously by the educators on a paper-based weekly summary sheet (see Figure 1). The computer-based data comprised details of the number of days on which intervention was completed each week, the number of games within the level for the week (between 3 and 5), and the number of plays (i.e., the recommended 4 sessions x 3-5 games = 12-20 plays). The educator's data comprised the dates, times and intervention sessions undertaken by a particular child, as well as the number of games played during each session, each week.

In order to examine quality of delivery, the research SLP viewed 20 (32%) of the videos children completing the intervention that included at least one full PFSS game (up to 10 minutes of recorded video), at the conclusion of the study. A 12-item checklist was developed (see Appendix) to determine if the intervention was completed as described in the intervention protocol.

[Insert Figure 1 here]

Analysis

Data from both the computer program and the educator summaries were entered into Microsoft Excel. In order to examine adherence to the prescribed intervention coverage in days (time), the total number of days that each child was reported to have received intervention over 9 weeks was calculated for both data sets. The number of children who received 18 or more days of intervention (the prescribed amount) was then calculated to determine the proportion who received the prescribed amount. The total mean number of days and range was identified for the sample. In order to examine adherence to the prescribed *coverage* in plays, the total number of plays that each child was reported to have had was

calculated for both data sets. The number of children who had 108 or more plays (the minimum amount, based on 3 games played 4 times per week for 9 weeks) was then calculated to determine the proportion who received the minimum prescribed amount. The mean number of plays and range was also identified for the sample.

In order to examine the impact of measurement tools on the evaluation of adherence, point by point agreement was determined for the number of days and number of plays recorded by the computer and the educators for each child. Data points could only be compared when data existed across both data sets. The number of exact matches was calculated for each child with a complete data set for each week of intervention. The number of children with complete data sets was different each week, as reflected in the results.

Finally, in order to examine quality of delivery, procedural fidelity was checked across the 20 videos. Each video was checked against 12 criteria (yes/no items), resulting in 246 data points with which to explore the degree of match.

Results

The results will be examined in three ways. Firstly, adherence to the protocol is examined by presenting the prescribed coverage in days and plays and comparing this with the data recorded by the educators and the computer program. Secondly, the degree of match between the days and plays recorded by the computer and that recorded by the educators is presented, to examine the consistency of the data collected, and the impact of the measurement tool on the results. Finally, the quality of delivery (procedural fidelity) is reported to examine the moderating influence that it might have had on the results.

Adherence to the protocol

Cumulative dose: total intervention days

The total prescribed and reported number of days that children received intervention is presented in Table II. According to the weekly summaries completed by educators,

children received an average of 14.56 days of intervention; however, this ranged from 2-28 (data were missing for 2 children). That is, some children received many more days of intervention than prescribed and some received much less. When the proportion of children who received 18 or more days of intervention was calculated, to determine the number who received intervention on the prescribed number of days, only 27.41% were found to have done so. While the figures from the computer-based data differed slightly to those recorded by the educators, the trends were still the same. According to the computer-based data, children received an average of 14.69 days of intervention, but this ranged from 1-28, and only 23.4% were recorded to have received the prescribed amount.

[Insert Table II here]

Cumulative dose: total intervention plays

The prescribed number of plays (shown in Table II) was based on a calculation of the prescribed number of sessions ($n=4$) multiplied by the prescribed number of intervention weeks ($n=9$) multiplied by the number of games to be played each session (which ranged from $n=3-5$ depending on the pre-set module that children were completing). Thus, the prescribed number of plays ranged from a minimum of 108 (3 games each session each week) to 180. According to the weekly summaries completed by educators, children participated in an average of 91.88 plays; however, this ranged from 7-155 (data were missing for 2 children). When the proportion of children who participated in 108 or more plays was calculated, to determine the number who participated in the minimum prescribed number, only half were found to have done so. The figures from the computer-based data were slightly better than those recorded by the educators; however, the trends were still the same. According to the computer-based data, children participated in an average of 101.16 plays, but this ranged from 4-160. Just over half of the children (56.25%) were recorded to have received the prescribed amount.

Consistency / Validity of measurement

The degree of match between the days and plays recorded by the computer and by the educators is presented in Table III. The number of available data points varied each week, as data could only be matched when both the educators and computer had recorded information for the same child. The reasons for missing data varied, but included child absences, technical issues resulting in computer data not being saved and educators forgetting to complete the summaries or return them to the researchers. The total number of data points (i.e. children) that could be examined for degree of match each week is given in Table III. The number ranged from 59 (week 1) to 21 (week 9), and resulted in a total of 423 points across all weeks. For each data point, exact matches were determined (i.e. when the number of days/plays recorded by the computer and the educator were identical). These are presented in Table III also.

The results indicated that there was consistently a difference between the educator and computer-recorded data for both the total intervention days and plays. The degree of exact match ranged from 64.29% to 90% for days, and from 25% to 45.10% for plays, indicating that the degree of difference increased as the data sets became more specific (i.e. the number of plays each week compared to the number of days).

[Insert Table III here]

Procedural Fidelity

Procedure fidelity examined (1) adherence to the protocol, (2) educator's competence to select an appropriate environment to conduct the intervention, *attempt* to keep a child on tasks, and *attempt to solve* practical problems as they arose (e.g. computer monitor freezing, headphones use), and (3) educator's quality of implementation by *successfully* keeping a child on task via verbal and non-verbal remarks, *successfully solving* practical problems and responding appropriately to the child's questions and comments during the task to ensure the

session was completed (see Appendix for the checklist template). Procedural fidelity for the experimental tasks across adherence to the protocol, educator's competence, and implementation quality was high at 91.9% based on 246 data points.

Discussion

Evaluating the implementation fidelity within the Sound Start Study revealed a lack of adherence to the prescribed coverage (dose) but high procedural fidelity or quality of delivery. Potential factors or “moderators” (Carroll et al., 2007) associated with the poor adherence to coverage were explored in interviews with the educators following the Sound Start Study (see Crowe et al., 2016). Three overarching factors that impacted implementation of PFSS were identified by the educators: personal factors (child, peers, educators), environmental factors (policy and philosophical, physical, logistics) and PFSS factors (format, games, game duration). In the current paper, we discuss the issue of adherence to the prescribed coverage, in terms of measurement and potential impact on outcomes, and we expand upon the discussion of barriers to implementation in order to identify future directions for intervention research.

Issues of adherence

Both sets of data (educator-reported and computer-reported) differed from the prescribed dose in the intervention protocol. Examination of the total days of intervention over 9 weeks revealed some children received intervention on more days than prescribed; however, the majority received less. Similarly, when total plays was examined, only half participated in the prescribed number of plays. This finding has implications for our interpretation of the findings from the Sound Start Study. However, it also has broader implications for the way in which we design and research interventions, the way in which we adapt interventions, and the way in which we provide training and support to those who will facilitate the delivery of those interventions. Each of these issues will be addressed in turn.

Within the Sound Start Study, the majority of children who received PFSS intervention received less than the prescribed dose in terms of number of intervention days and plays—an implementation issue which may have impacted on their outcomes. In recent years, the importance of intervention amount (or dose) in studies of intervention effectiveness has been highlighted (Allen, 2013; Glogowska et al., 2000; Williams, 2012). For instance, Glogowska and colleagues conducted a community-based RCT in which children were assigned to an intervention group or a “watchful waiting” (control) group and outcomes were compared after a 12 month period. The children in the intervention group received an average of 6.2 hours of intervention during that time period, and their outcomes were not significantly different to the control group at the end of that time. It was subsequently argued that 6 hours of intervention was insufficient to result in significant change for children with speech/language difficulties (Law & Conti-Ramsden, 2000). Williams (2012) and Allen (2013) examined dose required for the multiple oppositions approach to effect change and concluded that a minimum of 30 sessions (and 50 trials) was necessary, and gains were greater when the frequency of the intervention was more intense. From these studies it is clear that an “effective” intervention is only effective when delivered at an optimal intensity – this includes an optimal dose during a session, an optimal frequency of sessions, and an optimal total number of sessions or overall duration.

In the current study, educators were unable to implement the necessary dosage of intervention due to a range of factors (Crowe et al., 2016). While children who received intervention typically made gains in their speech skills post-intervention, their improvement was not significantly different from the gains made by children who did not receive intervention (McLeod et al., 2016). However, further examination of the results, revealed that children who received the amount of intervention prescribed in the protocol did not have significantly different outcomes to those who received less. So, the issue may be a broader

one of determining the optimal intervention intensity considering, dose, session frequency, and duration (including session duration and total period of time over which intervention is delivered) than an issue of poor adherence to a prescribed dose influenced by a time-based criterion. Furthermore, many continued to present with SSD that required ongoing support and intervention, beyond the completion of this project. Further research is required to determine whether the provision of intervention for a longer period of time would assist the children's speech development and whether a greater session dose (e.g. more than the protocol-defined dose) delivered in more frequent sessions and/or delivered by SLPs rather than educators would have been more effective, or indeed whether a different intervention approach involving speech production practice would yield better outcomes than PFSS.

Issues of intervention design and research

Results from this study reinforce the importance of measuring implementation fidelity in order to explore the impact on outcomes. In this study implementation fidelity in terms of coverage (dose) was poor; however, implementation fidelity did not appear to impact on treatment effectiveness. Neither children who received intervention according to the protocol nor those who received less improved significantly more than children in the control group. This raises further questions about the effectiveness of the intervention program, but also the adequacy of the prescribed dosage and the facilitation of the intervention agents. Thus, the implementation fidelity data from this study can guide future research efforts.

Intervention research has not traditionally reported fidelity information for many reasons including the time and financial challenges associated with gathering the implementation data. However, another reason for the limited reporting of implementation fidelity may be because the researchers who design the intervention are often the same as those implementing it within a research study and thus believe that they will remain true to the protocols they have devised. However, this cannot be guaranteed. Researchers need

objective evaluation to check that they are doing what they intend, and believe, they are doing. In addition, the growing facilitation of speech-language pathology intervention by those other than a SLP (i.e. parents or teaching staff) means this information is vital to gather in order to understand the potential results (Sugden, Baker, Munro, & Williams, 2016).

In proposing the need for implementation fidelity to be routinely measured and reported, we do recognise a need for appropriate measurement tools and procedures to be available to do so. In the current study, the two data-sets used to determine adherence (computer-reported and educator summaries) differed from the protocol, and from each other, for both the total number of days and intervention plays (trials) in which children participated. The degree of match between the computer and educators' data became less as the data became more specific (i.e. the match was better for days than for plays; 77.54% compared to 38.06%). This reflects two issues: (1) that the protocol was not followed, and (2) that one (or both) of the fidelity measures was inaccurate.

Poor adherence to intervention protocols may be due to a range of factors, some of which are discussed in the following section on *adaptation*. In community-based research, such as the Sound Start Study, gathering adherence data can assist us to determine the capacity of children, staff and services to participate in a stipulated program of intervention, facilitators and barriers to participation, and modifications that may need to be made if programs are to be effective in those contexts.

However, we need to ensure that the fidelity measures will provide accurate data. Fidelity measures may be inaccurate due to human error, which may in turn be due to issues of time (for gathering and recording data) or understanding (knowing what data to collect or how to record it). This reveals the need to consider the type of data it might be most reasonable to have people record/collect as evidence of fidelity. For instance, given the reliance of Australian SLPs on parent involvement in intervention for children with SSD

(McLeod & Baker, 2014), it raises the importance of examining the fidelity of parent involvement, including whether parents are able to adhere to a recommended schedule of practice sessions and the dosage during practice with their children, between SLP sessions.

Issues of adaptation

Researchers have also noted the impact of organizational, instructional, and client barriers and facilitators on adherence rates (e.g. availability of staff/resources to support the intervention, alignment of the intervention with organisation/community goals and philosophies, training and support, sustainability, time, intervention impact) (Brietenstein et al., 2010; Carroll et al., 2007). For instance, in the current study, when children were absent from the centre for a day/week, they did not receive the intervention for that week. As the intervention was provided over a 9-week period, missed sessions could not be made up later, and so the total amount of intervention that those children received was less than that prescribed in the protocol. Furthermore, there were times when more than one factor acted as a barrier for facilitation. For example, the physical environment of the early childhood centre and the requirement of centres to maintain strict staffing ratios acted as a combined barrier to implementation in the current study. Barriers and facilitators to the implementation of PFSS are discussed further in Crowe et al. (2016).

If those implementing interventions are unable to adhere to an intervention protocol, there is a need to gather practice-based evidence data showing the effect of the intervention in the way it was implemented in an everyday setting.

Issues of training and support to intervention agents

In the current study, we undertook measurement of quality of delivery through observations of a sample of videos showing intervention taking place. This enabled a check of whether the intervention was being facilitated as prescribed in the protocol and outlined in training sessions. While the results showed high levels of fidelity to the intervention

procedures, the adherence data showed a much lower level of fidelity to coverage (dose). The training and support provided to the educators delivering the intervention focused around the logistics of the intervention, such as navigating the PFSS software and protocols for how much intervention should be delivered and when. It may therefore be important to also provide educators with training to develop their understanding of the concepts of dose and why achieving the prescribed dose is important to children's outcomes. Thus, this suggests a need to support therapy "agents" (e.g. educators or parents) to ensure they are able to provide intervention as recommended, in both the mode but also the dose, and to work with them to develop an understanding of why fidelity is important.

Limitations

A number of limitations need to be taken into consideration to understand the outcomes of this study. Some of these limitations relate to the type of data collected. For instance, the PFSS program only recorded completed games, so if children commenced but did not complete all 10 trials within a game, data were not recorded. Thus, the total number of plays recorded for each child (i.e. number of games across the week) might not fully reflect the number of intervention trials (or attempts) they had. In contrast, the educators may have over- or under-estimated plays as a result of completing the summary some time after the child had completed the sessions. If summaries were completed retrospectively, the data may have lacked accuracy. Another limitation related to missing data. In a small number of cases, the software failed to save the child's performance data at the end of the session and so no results were recorded for those days. At other times, educators did not complete session summaries. In these instances, the data from the computer and the educator could not be matched, and instead had to be classed as "missing".

Within the current study, procedural fidelity based on the 32% sample was high; however, this was only checked at one period of time. Brietenstein et al. (2010) suggested

that “ongoing assessments of fidelity may capture issues related to practitioners’ drift, contextual issues that may influence the implementation and receipt of the intervention, identifying adaptations of the intervention, and provide important information for supervising and training practitioners” (p. 7). In future studies, the collection of additional contextual data, and fidelity checks across multiple time points, might provide further insight into the way in which the educator or the environmental variation influenced the implementation and outcomes. Furthermore, the collection of information about how intervention agents are selected to facilitate the intervention, and their perceptions of training and support needs could be gathered in order to ensure that they were best prepared to deliver the support in line with the protocol requirements.

Future Directions

There is a real need for intervention research to document and rationalise dose decisions, particularly in the field of phonological impairment. Some work has been done regarding dose manipulation in intervention studies with children with childhood apraxia of speech (Thomas, 2014) where a randomised controlled trial was conducted (Murray et al., 2015) and then dose and delivery mode was manipulated in later studies. At present, there is limited understanding as to the optimal intervention intensity required (including session duration, frequency and dose, and cumulative dose over time) in order for the PFSS to be effective. The intensity prescribed in the current study was guided by intervention research based on the amount required in order to see change with other approaches, and it is clear that the majority of children in the current study did not receive this amount. Thus, it is possible that lack of adherence to the protocol, and lack of sufficient dose, may have impacted on their outcomes. Further research is needed to explore this. Barber et al. (2007) suggested a need for researchers to explore the role of implementation fidelity in their analyses of intervention effectiveness in order to identify acceptable levels of adherence and competence for

establishing and maintaining intervention effects. This may assist in determining the degree to which interventions can be adapted to suit contextual needs, without losing their effectiveness (Brietenstein, 2010). Similarly, the adaptation of intervention for different contexts requires a clear understanding of program differentiation, or the aspects of intervention that are most responsible for change (Carroll et al., 2007). By determining the most important components to retain, and those that are non-essential, we might develop protocols for implementation that are better able to be implemented in non-research, and potentially non-clinical, settings.

Once an optimum intensity has been identified for an intervention, further adaptation of the intervention may be required for implementation in an education setting. For example, the intervention protocol may need to be adapted to align with the experience of the intervention agents and specific client factors (Dollaghan, 2007). Or, additional training of intervention agents may be required to address instructional barriers prior to implementation. Finally, the feasibility of an intervention protocol within an organizational structure needs to be considered to identify and minimize any barriers prior to implementation.

Conclusion

There is an established body of evidence indicating that intervention for SSD is effective, and a range of approaches have empirical research to support their use. However, there is increasing recognition in the implementation research literature that components of intervention often need to be adapted for everyday implementation (Meyers et al., 2012), due to organisational, instructional and client barriers and facilitators (Durpak & DuPre, 2008). Thus, tensions can exist between implementing an intervention exactly as designed and modifying the intervention to suit contextual needs (Odom, 2009). In the current study, the tension was centred around the challenge of adhering to the coverage or dose (including days and plays) stipulated in the implementation protocol (and based on prior research), when

delivering the intervention in an early childhood centre, with multiple children attending on different numbers/times of days, and with a timetable of other tasks to be completed each day as well. The result was that not all children received the amount of intervention recommended, but the impact of this on their outcomes is not yet clear.

This tension of implementing an empirically-based intervention in a real world setting sits at the heart of what it means to “do” evidence-based practice. While we do not have the answer to eliminating the tension, it may be mediated by considering the range of evidence that we utilise in order to build our understanding of the effectiveness of particular interventions. That is, implementing an intervention exactly as designed is drawing on external evidence, but undertaking and evaluating contextual modifications creates internal evidence, and the combination of both are important. Indeed, Durlak and DuPre (2008) have noted that “fidelity and adaptation frequently co-occur and each can be important to outcomes” but that “most researchers have considered program adaptation as an implementation failure (i.e. a failure to achieve fidelity) and have not assessed its possible contribution to outcomes” (p. 341).

Baker (2012) proposed there is a need to find practical solutions when disparities exist between empirically-based recommendations for children with SSD and the limitations in the workplace. Implementation fidelity can assist us to examine the impact of barriers and facilitators on adherence rates, and on outcomes, in order to design interventions that are suited to clinical contexts without losing their rigor or effectiveness.

References

- Allen, M. M. (2013). Intervention efficacy and intensity for children with speech sound disorder. *Journal of Speech, Language, and Hearing Research*, 56(3), 865-77.
- American Speech-Language-Hearing Association. (2005). *Evidence-based practice in*

- communication disorders* [Position Statement]. Available from www.asha.org/policy.
- Australian Bureau of Statistics. (2011, June). *Information paper: Measures of socioeconomic status*. Canberra: author.
- Baker, E. (2012). Optimal intervention intensity in speech-language pathology: Discoveries, challenges, and uncharted territories. *International Journal of Speech-Language Pathology*, 14(5), 478-85.
- Baker, E. (2010). The experience of discharging children from phonological intervention. *International Journal of Speech-Language Pathology*, 12(4), 325-328.
- Baker, E. & McLeod, S. (2011). Evidence-based practice for children with speech sound disorders: Part 1 narrative review. *Language, Speech, and Hearing Services in Schools*, 42, 102-139.
- Breitenstein, S. M., Gross, D., Garvey, C., Hill, C., Fogg, L., Resnick, B. (2010). Implementation fidelity in community-based interventions. *Research in Nursing and Health*, 33(2), 164–173.
- Carrey, N. J., Curran, J. A., Greene, R., Nolan, A., McLuckie, A., (2014). Embedding mental health interventions in early childhood education systems for at-risk preschoolers: An evidence to policy realist review. *Systematic Reviews*, 3(84).
- Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implementation Science*, 2(40).
doi:10.1186/1748-5908-2-40
- Chen, Y.-P. P., Johnson, C., Lalbakhsh, P. et al. (2016). Systematic review of virtual speech therapists for speech disorders. *Computer Speech and Language*, 37, 98-128.
- 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., 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.
- Dollaghan, C. (2007). *The handbook for evidence-based practice in communication disorders*. Baltimore, MD: Paul H. Brookes.
- 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(3-4), 327-350.
- Dusenbury, L., Brannigan, R., Falco, M. & Hansen, W. B. (2003). A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. *Health Education Research*, 18, 237-256.
- Forman, S. G., Olin, S. S., Eaton Hoagwood, K., Crowe, M., & Saka, N. (2009). Evidence-based interventions in schools: Developers' views of implementation barriers and facilitators. *School Mental Health*, 1, 26–36.
- Glogowska, M., Roulstone, S., Peters, T. J., & Enderby, P. (2000). Randomised controlled trial of community-based speech and language therapy for pre-school children. *British Medical Journal*, 321, 923-926.
- Law, J., Garrett, Z., & Nye, C. (2013). 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.
- Law, J., & Conti-Ramsden, G. (2000). Treating children with speech and language impairments: Six hours of therapy is not enough. *British Medical Journal*, 321, 908-909.
- 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.
- 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., McCormack, J., Wren, Y., Roulstone, S., Crowe, K., Masso, S., White, P., & Howland, C. (2016). *Randomized controlled trial evaluating the effectiveness of computer-assisted intervention delivered by educators for children with speech sound disorders*. Manuscript in submission.
- Meyers, D. C., Katz, J., Chien, V., Wandersman, A., Scaccia, J. P., & Wright, A. (2012). Practical implementation science: Developing and piloting the quality implementation tool. *American Journal of Community Psychology*, 50(3-4), 481-96.
- Mihalic S. (2004). The importance of implementation fidelity. *Emotional and Behavioral Disorders in Youth*, 4, 83-86, 99-105.
- Murray, E., McCabe, P., & Ballard, K. (2015). A randomized controlled trial for children with childhood apraxia of speech comparing rapid syllable transition treatment and the Nuffield Dyspraxia Programme-Third Edition. *Journal of Speech, Language, and Hearing Research*, 58(3), 669-686.
- Odom, S. L. (2009). The tie that binds: Evidence-based practice, implementation science, and outcomes for children. *Topics in Early Childhood Special Education*, 29(1), 53-61.
- Roth, F., & Worthington, C. K. (2015). *Treatment resource manual for speech-language pathology*. Clifton Park, NY: Delmar, Cengage Learning.
- Roulstone, S. E., Wren, Y. E., Bakapoulou, I. & Lindsay, G. (2012). Exploring educational and SLT interventions for children with SLCN. *Child Language Teaching and Therapy*, 28(3), 325-341.

- 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.
- 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.
- Thomas, D., McCabe, P., & Ballard, K. (2014). Rapid Syllable Transitions (ReST) treatment for childhood apraxia of speech: The effect of lower dose frequency. *Journal of Communication Disorders*, 51, 29-42.
- 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. (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 (2nd ed.)*. Bristol, UK: Bristol Speech and Language Therapy Research Unit.

Table I

Elements of implementation fidelity proposed by Carroll et al. (2007) applied to the Sound Start Study*

Element	Description	Application/ Measurement in the Sound Start Study
Adherence		
Adherence to content of an intervention	The content of an intervention is delivered as it was designed or researched (Mihalic, 2004)	Computer data and educator summary
Adherence to intensity: frequency, duration and coverage	Participants receive the prescribed intervention intensity (e.g., frequency of sessions, dosage per session, and total amount of intervention per unit of time) prescribed by its designers.	Computer data and educator summary
Moderators*		
Quality of delivery	The manner in which an intervention agent delivers a program (Mihalic, 2004).	Video observations (fidelity checks)
Participant responsiveness	The level of engagement or responsiveness of participants to the intervention. It involves judgments by participants or recipients about the outcomes and relevance of an intervention.	Interviews (see Crowe et al., 2016)
Intervention complexity	The degree to which the complexity of an intervention acts as a barrier to its adoption.	Interviews (see Crowe et al., 2016)
Facilitation strategies	The strategies put in place to optimise the level of fidelity achieved. Such strategies may include the provision of manuals, guidelines, training, monitoring and feedback, capacity building, and incentives.	Interviews (see Crowe et al., 2016)
Analysis of components of intervention		
Program differentiation	The identification of intervention components that make a difference to outcomes and those that may be redundant.	

*The focus of the current paper is on issues of adherence, and quality of delivery. The role of the listed moderators is explored in detail in Crowe et al. (2016).

Table II

Prescribed and reported total days and plays of intervention over 9 weeks

Prescribed coverage	Educator-reported coverage (n=62)		Computer-reported coverage (n=64)	
	No. (%) who received ≥ 18 days	Mean (SD) range	No. (%) who received ≥ 18 days	Mean (SD) range
Days: 18	17 (27.41%)	14.56 (5.47) 2-28	15 (23.4%)	14.69 (5.42) 1-28
Plays: 144 (108-180)*	No. (%) who received ≥ 108 plays	Mean (SD) range	No. (%) who received ≥ 108 plays	Mean (SD) range
	31 (50.00%)	94.84 (39.99) 7-155	36 (56.25%)	101.16 (41.10) 4-160

*The range in plays was influenced by the inherent variation in the number of games (three to five) across levels, per pre-set module across participants.

Table III

Weekly degree of match between computer-and educator-reported coverage: days of intervention and plays per week

Week	Number of data points	Days	Plays
		Number (%)	Number (%)
		of match	of match
1	59	43 (72.88%)	21 (35.59%)
2	56	36 (64.29%)	19 (33.93%)
3	57	42 (73.68%)	22 (38.60%)
4	51	39 (76.47%)	23 (45.10%)
5	47	40 (85.11%)	20 (42.55%)
6	48	41 (85.42%)	17 (35.42%)
7	44	36 (81.82%)	22 (50.00%)
8	40	36 (90.00%)	10 (25.00%)
9	21	15 (71.43%)	7 (33.33%)
Total	423	328 (77.54%)	161 (38.06%)

Appendix A. Sound Start Study Fidelity Checklist for Phoneme Factor Sound Sorter (PFSS) Implementation

View 10 minute recording per child that includes playing one whole game on PFSS				
Participant number: _____		Participant video number and week of Rx: _____		
Date of fidelity check: _____		Fidelity scorer's name: _____		
Descriptor	Yes	No	N/A	Not apparent from video
A. Adherence to protocol				
Equipment				
1. Computer is in working order with PFSS software installed				
2. Wireless mouse and x2 headphones are connected to the computer				
3. Only one child is sitting at the computer (+educator)				
Phoneme Factory Sound Sorter Program				
4. The child is sitting on a chair, wearing headphones, looking at the computer, and playing the games.				
5. The child completes 10 trials within the game (i.e. required dose per game) before starting another game.				
6. a) The child navigates during the game using the mouse and/or				
b) the educator provides assistance in using the computer mouse as needed.				
B. Competence (Educator)				
7. The educator selects an appropriate environment for the games to be completed (e.g. seated in an area away from distractions). (i.e. were there avoidable distractions?)				
8. The educator ATTEMPTS TO KEEP the child on task during the PFSS session (during set-up and during games) via re-adjusting the child's seat, re-positing the child's headphones, using verbal remarks and non-verbal behaviours (e.g. pointing, facial expressions) that attempt to prompt the child to listen to and respond to questions during PFSS games. (cf. question 10)				
9. The educator IDENTIFIES practical problems related to implementation of the intervention (e.g. issues with headphone use, computer problems such as the computer monitor freezing) (cf. question 11)				
C. Quality of implementation (Educator)				
10. The educator SUCCESSFULLY KEEPS a child on task during the PFSS session via verbal remarks and non-verbal behaviours so that games are completed (cf. question 8)				
11. The educator SOLVES practical problems related to implementation of the intervention (e.g. successfully modifying headphones so that child is compliant and wears headphones, solves computer problems so that games continue, and session is completed) (cf. question 9)				
12. The educator appropriately answers child's questions and responds to child's extraneous comments during PFSS games, so that the child continues games.				
TOTAL				= /12 (yes)