Effects of Stuttering Treatment: A Systematic Review of Single-Subject Experimental Design Studies

Kimberly A. Murza, PhD, CCC-SLP
Martine Vanryckeghem, PhD, CCC-SLP
Chad Nye, PhD
Anu Subramanian, PhD, CCC-SLP
Acknowledgments: The authors wish to acknowledge Melissa Hamilton and Megan McMenamy for their work in the initial literature search and coding procedures. The authors also wish to thank Dr. Oliver Wendt for his consultation during the analysis.

Structured Abstract

Clinical Question: For individuals who stutter (P), which behavioral treatment approaches (I, C) are most effective in promoting fluent speech (O)?

Method: Systematic Review and Meta-Analysis

Study Sources: CINAHL, CINAHL Plus (with Full Text), ERIC, MEDLINE, PsycINFO, and ProQuest Dissertations & Theses Full Text

The authors also conducted an ancestry search from all studies meeting the inclusion criteria previously described.

Search Terms: stutt* OR stamm* AND treat* OR interven* OR therap* AND single subject OR multiple baseline

Number of Included Studies: 19

Primary Results: A total of 19 studies representing 74 participants met the inclusion criteria and yielded an average mean baseline reduction (MBLR) of stuttering behavior by 67%. An overall improvement of nonstuttered speech behavior represented by the percentage of non-overlapping data points (PND) resulted in a 49% improvement representing 19 participants. Analyses suggest that a fluency shaping program or assertiveness training program, in one form or another lead to fluency improvement. The question remains if participant characteristics (i.e., age, stuttering severity, level of social anxiety, using coping behaviors, presence of negative speech-associated cognition, among others), are variables to account for when considering a treatment approach.

Conclusions: These data are consistent with previous meta-analysis and systematic review stuttering intervention group results indicating that, for the vast majority of the participants, treatment has a positive effect on their speech and support for a range of treatment procedures exists. Further research is needed to determine which participant characteristics are important to consider when determining the best treatment approach for an individual.
Effects of Stuttering Treatment: A Systematic Review of Single-Subject Experimental Design Studies

Kimberly A. Murza, PhD, CCC-SLP
Martine Vanryckeghem, PhD, CCC-SLP
Chad Nye, PhD
Anu Subramanian, PhD, CCC-SLP

Clinical Scenario

Malia has been a speech-language pathologist for 20 years and has worked in a variety of settings with individuals of all ages. Five years ago, Malia joined a private practice specializing in fluency. She had always enjoyed working with children and adults with fluency disorders and was excited to further specialize in the area. Over the past five years, Malia has continued to develop her knowledge and skills while working with individuals with fluency disorders but has recognized that what seems to work for one client does not necessarily work for another. Although this is not unique to fluency, Malia is interested in understanding whether certain behavioral intervention approaches are more appropriate for certain clients. In accordance with the American Speech-Language-Hearing Association’s position statement on evidence-based practice (2005), Malia works to stay current on the research literature in fluency and takes time to understand her clients’ unique needs and preferences. She set out to review the state of the research specific to behavioral stuttering interventions for her clients from preschool through adult to determine if certain approaches were more effective than others for certain types of clients.

Background Information

Treatment for stuttering has essentially centered on its modification, or the shaping of fluency. For some, stuttering is a multidimensional disorder, and as a result, its intervention includes more than the behavioral dimension to encompass the “person” who stutters and his or her negative emotional reaction to aspects of communication, negative speech-related attitude, and the use of concomitant behaviors. For others, focusing on the dysfluency itself is the core and often only dimension of treatment. Aside from the behavioral and cognitive-behavioral therapies for stuttering and integrated interventions, other approaches may include counseling, using an assistive device, and pharmaceutical treatments, among others.

Several specific therapeutic methods exist that are considered types of behavioral stuttering approaches. Stuttering modification techniques are aimed at producing less dysfluent speech by managing or modulating the stuttering, sometimes using classical conditioning (Adams, 1980; Bloodstein & Ratner, 2008; Brutten, 1975; Guitar, 2014; Van Riper, 1973, 1982; Webster, 1979). Management techniques may involve systematic desensitization (Brutten & Shoemaker, 1967, 1969; Walton & Mather, 1963; Wolpe, 1958), and counter conditioning and deconditioning of previously learned behaviors through a gradual stepwise approach relative to speech units and stress level. In addition to some of the aforementioned approaches, Van Riper (1973, 1982), suggested using cancelation, pull-out, and preparatory set, allowing the person who stutters (PWS) to “stutter more fluently.” Complementary to modifying the stuttering, the affective and cognitive dimensions surrounding the disorder are given serious consideration during intervention. Following this approach, negative emotional reaction and negative speech-related cognition are important components in the treatment of the PWS (Beilby, Byrnes, & Yaruss, 2012; Brutten & Vanryckeghem, 2003a, 2003b, 2007; Gregory, Gregory, Campbell, & Hill, 2003; Guitar, 2014; Langevin, Kully, Teshima, Hagler, Narasimha Prasad, 2010; Menzies, Onslow, Packman, & O’Brien, 2009; Vanryckeghem & Brutten, 2011; Williams & Dugan, 2002; Yaruss & Quesal, 2006).

Fluency shaping treatments, using operant conditioning principles, and stimulus-response paradigm techniques, focus on the actual speech behaviors exhibited by the PWS with the goal of establishing fluent speech. Fluency shaping approaches include the contingent application of aversive stimuli, such as shock or time out, to increase fluency (Flanagan, Goldiamond, & Azrin, 1958; Haroldson,
Effects of Stuttering Treatment: A Systematic Review of Single-Subject Experimental Design Studies

Martin, & Starr, 1968; Martin & Siegel, 1966; Shames & Sherrick, 1963); positive contingency applications, such as listener feedback (Ingham, Martin, Haroldson, Onslow, & Leney, 1985; Martin, Haroldson, & Triden, 1984); or choral reading (Stager, Denman, & Ludlow, 1997). Other methods have included the use of phonation modification intervals during prolonged speech (Ingham, 1990; Wingate, 1976), awareness training (Goldiamond, 1965; La Croix, 1973), and habit reversal (de Kinkelder & Boelens, 1998; Woods, Miltenberger, & Lumley, 1996), among others (Costello Ingham, 1999; Ryan, 1986).

To treat stuttering, integrated or hybrid approaches that involve using more than one class of treatment have gained impetus. These integrated approaches might use fluency enhancing techniques, anxiety reducing procedures, cognitive restructuring, and include introducing assistive devices (Baumeister, Caspar, & Herziger, 2003; Gregory et al., 2003; Guitar, 2014; Langevin et al., 2006; Ramig & Bennett, 1997).

Since the 1970s, several efforts have been made to provide research data summaries regarding stuttering treatment’s effect for both children and adults. Primarily, these reviews were undertaken for one of three purposes: (1) to describe the extent of the research conducted (Ingham & Lewis, 1978; Thomas & Howell, 2001), (2) to attempt to draw conclusions about the effectiveness of stuttering interventions (Nye et al., 2013), or (3) to document critical variables when evaluating stuttering treatment efficacy (Craig, 1998; Curlee & Yairi, 1997).

In 2013, Nye et al. conducted a systematic review and meta-analysis of randomized control trials and quasi-experimental design studies of behavioral stuttering interventions for children and adolescents who stutter. Their inclusion criteria yielded nine studies representing a total of 327 treated participants. Results showed the overall effect of behavioral interventions was a reduction in percentage of stuttered syllables or words by approximately one standard deviation compared to control participants. However, the authors noted several limitations; the first was the limited number of included studies, which was notable considering that stuttering research has been conducted for more than 75 years. Another limitation was the lack of adequate reporting of both intervention and participant characteristics. Without specific detail on variables such as dosage, stuttering severity, or treatment procedures, it is impossible to determine which approach is best for which type of individual.

Clinical Question

Malia was interested in determining whether certain behavioral stuttering approaches were more effective than others for certain types of clients. To guide her search for external evidence, Malia developed her clinical question based on the PICO format (patient/problem, intervention, comparison, outcome; Dollaghan, 2007; Sackett, Straus, Richardson, Rosenberg, & Haynes, 2000): For individuals who stutter (P), which behavioral treatment approaches (I, C) are most effective in promoting fluent speech (O)?

Search for the Evidence

Inclusion Criteria

Studies were evaluated at two levels in the search process: title and abstract screening and a full text review of the study. Because the Nye et al. (2013) meta-analysis included only group design research, Malia was interested in reviewing single-subject design research.

Title and abstract inclusion criteria. All titles and abstracts advanced to the next stage of the screening if they met at least one of the following inclusion criteria:

1. Treatment appeared to be a behavioral intervention for stuttering (e.g., stuttering modification, stuttering management techniques, counter conditioning or deconditioning approaches, fluency shaping)
2. Appeared to have used a single case design that was either ABAB, multiple baseline, alternating treatment, or changing criterion

Full text inclusion criteria. Full text was obtained for all studies meeting the inclusion criteria at the title/abstract inclusion level. The following criteria were met by all studies included in the final analysis.

1. Single-subject experimental design as described above
2. Behavioral fluency intervention
3. Outcome measures related to stuttered speech (e.g., stuttered words, syllables)
4. Participants who stuttered of any age

Information Retrieval

Six electronic databases were selected and systematically searched for relevant literature for this review. These databases included CINAHL, CINAHL Plus with Full Text,
Effects of Stuttering Treatment: A Systematic Review of Single-Subject Experimental Design Studies

Data extraction

All included studies were double coded using a coding form and an accompanying codebook defining all terms. Each study was coded by two independent coders; any conflicts were resolved to consensus by discussion between the coders or a third colleague.

Calculating and Synthesizing Effect Size

Calculating the treatment effects were based on changes in the measurement of outcome data from baseline through treatment phases. Quantifying these treatment effects was accomplished for each study using both or one of the following meta-analytic statistical procedures: (1) percentage of non-overlapping data (PND) or (2) mean baseline reduction (MBLR) (see Kazdin, 1982; Mastropieri & Scruggs, 1985; Scruggs & Mastropieri, 1998; Scruggs, Mastropieri, Forness, & Kavale, 1988).

To evaluate the treatment effects of particular participant characteristics, both PND and MBLR were calculated where appropriate. That is, the PND was calculated for those studies that measured the increase of either total syllables per minute and/or total words per minute, and MBLR was calculated for the decrease in the percent of syllables stuttered or percent words stuttered.

The PND and MBLR statistics provide a metric of treatment impact in which the larger the percentage, the more effective the treatment. The PND treatment interpretation scale uses 90% or greater to mean the intervention is considered to be very effective, 70–89% reflects a moderate level of treatment effect, 50–69% suggests a mildly effective or questionable treatment effect, and less than 50% is said to be ineffective (Scruggs & Mastropieri, 1998). The MBLR statistic is used when the outcome measured is a reduction of the observed behaviors (e.g., disfluent behaviors) that reflects a positive change in behavior. This is achieved by calculating the amount of change between the last three baseline data points and the last three treatment data points and then converting to a percentage.

There is no scale equivalent available for the MBLR values like described for the PND standards. Because the MBLR offers an improvement measurement (as does the PND), albeit through measuring the decrease in a targeted behavior, applying the very conservative PND scale to the MBLR value was used to express the treatment effect’s magnitude to guard against a potential inflation of treatment effects.

Evaluating the Evidence

Search Results

The initial search of the electronic databases retrieved a total of 183 results across the following databases: CINAHL = 16; CINAHL Plus with Full Text = 20; ERIC = 12; MEDLINE = 57; PsycINFO = 63; ProQuest Dissertations & Theses Full Text = 15. Stage 1 criteria were applied to these citations (i.e., treatment appeared to be a behavioral intervention for stuttering and the design appeared to be either ABAB, multiple baseline, alternating treatment, or changing criterion). A study title might advance to Stage 2 for consideration if the title or abstract was unclear regarding any one of these criteria. Stage 2 criteria were then applied (i.e., design as described in Stage 1, behavioral fluency intervention, outcome measures related to stuttered speech, participants who stuttered at any age). Figure 1 shows the number of studies retrieved and eliminated at each stage of the decision process.

Methodological Quality Assessment and Effect Size Findings

The methodological quality assessment was accomplished using a coding format developed by Schlosser and Wendt (2008) based on a taxonomy designed by Simeonsson and Bailey (1991). The quality appraisal includes assessing the research design and implementation, inter-observer agreement, and treatment integrity. The quality appraisal results can be found in Table 1. The data presented in Table 2 are a summary of each individual’s findings included in this study.
Findings

The results of this investigation were based on the examination of available data for 19 multiple baseline, single-subject experimental design studies. The overall effectiveness of the combined treatments achieved a moderate to highly effective result that included an intensive treatment program for adolescents and adults involving the introduction of fluency shaping procedures such as DAF and CAFET, in addition to a relapse management program using Power2, during which time, attention was given to the affective and cognitive variables surrounding stuttering (Blood, 1995a, 1995b).

Some of the interventions involving awareness training were determined to be moderately to highly effective; this included the three experiments by Gagnon and Ladouceur (1992). Experiments one and two were identical in that they used awareness training (identification of stuttering modeled by the therapist), regulated breathing, and easy speech. The only difference between the two experiments was the children’s age (ages 10–11 versus ages 6–7). In experiment three (ages 7–11), an additional component of parental participation (i.e., using booster and group sessions) was part of the interventional procedure. The combined elements of this approach led to a high effect size, which culminated when including other’s support (i.e., parent and group participation, booster sessions) was added to the original procedure.

Regulated breathing, as described in the Williamson, Epstein, and Coburn (1981) study on an adult male and a simplified form of regulated breathing accompanied by awareness training, competing response, and social support (parents) in the study of Wagaman, Miltenberger, and Arndorfer (1993) with eight children (ages 6–10) were also among the treatment approaches that received a moderate treatment effectiveness score. To a lesser extent, the combination of awareness training, regulated breathing, and discussion and modeling of diaphragmatic breathing, as reported by Miltenberger, Wagaman, and Arndorfer (1996), was effective in reducing stuttering in two young adults. The remaining two interventions that received a moderate to high MBLR score included using a response contingent punisher (verbal slow down stimulus) with two preschool/kindergarten children (Reed & Godden, 1977) and a syllable-timed speech approach (uttering each syllable to a rhythmic beat) used with three preschoolers (Trajkovski et al., 2009).

Treatment procedures that proved to be moderately to highly effective in the previously cited studies, did not necessarily lead to the same results in other studies. Some studies related to awareness training and regulated breathing (see Table 1) showed only a mild effect (Caron & Ladouceur, 1989; Ladouceur, Caron, & Caron, 1989) or were ineffective (Elliott, Miltenberger, Rapp, Long, & McDonald, 1998). Self-modeling as an intervention strategy resulted in a mild effectiveness rating in the Bray (1997) study with elementary school children and was shown to be ineffective in the Harasym, Langevin, and Kully (2015) intervention with adults who stutter, as well as in the Northrup (2012) investigation with elementary school children. Only one other study (LaSalle, 2015) revealed mild effectiveness when the clinician modeled slow speech rate to preschoolers. The remaining studies involving implementation of the Lidcombe program without direct parent involvement (Jang, 2008) and assertiveness training with adults (Schloss, Freeman, Smith, & Espin, 1987) failed to reach a level of effectiveness.

As described, several of the interventions involved more than one strategy within a study, making it difficult to draw conclusions about which specific treatment subcomponent, if any, provided the major force in the obtained outcome. Also, in some studies the participants had been subjected to other treatment strategies before the start of the experiment. This was the case in the Wagaman et al. (1993) study; five of the eight participants had received a different treatment. The time frame when this occurred was not indicated. Variables such as these might create confounding results, and without specific treatment history, it is impossible to assess the real impact of the present treatment protocol.

Another factor that hampers inference from the available data is the participant characteristic variations, such as stuttering severity. As indicated in the results, people with a more severe stutter generally showed more speech improvement compared to the individuals with a milder form of stuttering. Logically, of course, a person with a more severe stutter has more room for improved speech in terms of reduction of percent syllables stuttered or words stuttered, compared to a person with a mild stutter. More surprising were the results related to age; specifically, the fact that the outcome measures of adolescents were higher than those of preschoolers and adults. This is indeed unexpected because receiving treatment as close as possible to the disorder’s onset typically has a greater impact than later treatment after the disorder has taken on a more complex structure (Manning,...
& DiLollo, 2018). In addition, adolescents are typically an age group that is less likely to be interested in therapy and less motivated and devoted to put time and effort into clinical practice. However, as mentioned in the results section, given the few number of adolescent participants in the intervention studies, drawing major conclusions from this analysis would be premature.

Another issue related to the treatment effect’s evaluation leads to the typical appraisal of intervention success in terms of fluency improvement (percent syllables stuttered or percent words stuttered). However, the question remains: What does a reduction in stuttering mean for the individual who stutters? Does the person who stutters evaluate his or her speech as improved when the percent syllables or words stuttered is reduced below the 5%, 2%, or 3% boundary (Bothe, Davidow, Bramlett, & Ingham, 2006; Craig, 2010)? Or, is the experience and belief of success related to the simple reduction in stuttering crucial and, if so, is a decrease of 50%, 30%, or 25% stuttering meaningful for the individual? It might well be that the mere modification of his or her stuttering in terms of reducing the number of syllable reiterations, the length of a prolongation, or a decrease in the use of concomitant behaviors is the determining factor of personal success and improvement. No universal data are available to the clinician or the client to estimate successful treatment and outcome in terms of these subjective criteria.

Fifteen of the 19 studies included follow-up measurement ranging from eight weeks up to one year posttreatment. As indicated by the almost identical effect sizes for immediate posttreatment and follow-up data for the majority of the studies, the fluency gains were maintained. An exception to this observation was evident in the Gagnon and Ladouceur (1992) second intervention, where the two participants did not maintain a decrease in stuttering. Similarly, only one of the two participants in the Miltenberger et al. (1996) study maintained the posttreatment fluency level, and the Northrup (2012) study showed mixed results. It goes without saying that the extent to which treatment outcomes are generalized and maintained is one of the major goals of intervention. Stuttering treatments should always have the intention to move on beyond the boundaries of the secure clinical setting. Carryover and ultimate preservation of an acceptable level of fluency are a condition sine qua non for treatment to be considered successful.

The current data are consistent with previous meta-analysis and systematic review stuttering intervention group results indicating that, for the vast majority of the participants, treatment has a positive effect on their speech (Andrews, Guitar, & Howie, 1980; Herder, Howard, Nye, & Vanryckeghem, 2006; Nye et al., 2013), and support for a range of treatment procedures exists. As alluded to earlier in this paper and in other publications (Nye et al., 2013), the question remains if participant characteristics, in terms of age, stuttering severity, level of anxiety, using coping behaviors, and presence of negative speech-associated cognition, among others, are variables to take into account when considering a treatment approach.

The Evidence-Based Decision

Malia undertook this review to answer the following clinical question: For individuals who stutter (P) which behavioral treatment approaches (I, C) are most effective in promoting fluent speech (O)? Malia’s overall findings were consistent with previous meta-analysis and systematic review stuttering intervention group results indicating that, for the vast majority of the participants, treatment has a positive effect on their speech (Andrews et al., 1980; Herder et al., 2006; Nye et al., 2013), and support for a range of treatment procedures exists.

Malia was also encouraged to learn about certain programs that are particularly efficacious. However, it is clear that further research needs to be conducted to determine whether there are certain participant characteristics that predict better success with a given treatment approach. As is often the case when conducting research, Malia has started asking different questions about stuttering treatment. She wonders whether a key issue to consider in the study of treatment efficacy is whether a reduction in overt stuttering is considered the major outcome for a person who stutters? Are the typically established outcome data, such as less than 2% syllables stuttered, meaningful to the individual who stutters? Or, would the person who stutters self-report in terms of impact of speech improvement to daily functioning be an important factor to consider when evaluating an intervention’s effectiveness? The evidence-based decision-making process involves integrating external and internal evidence. Malia understands this and decided to spend more time talking with her clients of all ages about the outcomes that would be meaningful in their lives at the beginning of treatment. She looks forward to reading new research that addresses the shortcomings of her review.
Authors’ Note

*Kim Murza, PhD, CCC-SLP, is a speech-language pathologist and associate professor at the University of Northern Colorado specializing in language and literacy and autism spectrum disorder. Her research interests include social communication and engagement intervention, inference generation, systematic review and meta-analysis, disciplinary literacy, and high-quality professional learning.

Martine Vanryckeghem received her PhD from Southern Illinois University and is currently a professor at the University of Central Florida. From 1990 until 2000, Dr. Vanryckeghem served as managing editor of the *Journal of Fluency Disorders*. She is a board certified fluency specialist and an ASHA fellow.

Chad Nye is professor emeritus, University of Central Florida. He has authored systematic reviews for Cochrane Collaboration, Campbell Collaboration, and several journal publications of systematic reviews and meta-analysis using both group and single-subject designs.

Anu Subramanian, PhD, CCC-SLP, is a clinical associate professor in the Department of Communication Sciences and Disorders at the University of Iowa. She supervises graduate students in pediatric clinical placements, including early intervention. Her interests are in clinical education, early intervention, and stuttering.

*Corresponding author contact information

Kimberly A. Murza, PhD, CCC-SLP
Associate Professor, Audiology & Speech Language Sciences
University of Northern Colorado
Campus Box 140
Gunter Hall 1510
Greeley, CO 80639
Kimberly.Murza@unco.edu
Phone: 970-351-1084
Fax: 970-351-2974

References

*Studies included in the systematic review and meta-analysis.


Figure 1. Search Flowchart Showing the Number of Retrieved Citations and Exclusions at Each Decision Stage
### Table 1. Quality Appraisal

<table>
<thead>
<tr>
<th>Study</th>
<th>Appraisal (mean MBLR)</th>
<th>Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blood, 1995a</td>
<td>Suggestive (88%)</td>
<td>Maintenance, but no generalization. Clear treatment description and integrity of implementation data. Inter-rater reliability completed across all phases and results &gt; 80% agreement. Completed on 15% of sample.</td>
</tr>
<tr>
<td>Blood, 1995b</td>
<td>Inconclusive (88%)</td>
<td>Maintenance, but no generalization. Clear treatment description but no integrity of implementation data. No inter-rater reliability checks reported.</td>
</tr>
<tr>
<td>Bray, 1997</td>
<td>Conclusive (88%)</td>
<td>Maintenance and generalization completed. Clear treatment description and integrity of implementation data. Inter-rater reliability checks on 100% of data across phases, results &gt; 80% agreement.</td>
</tr>
<tr>
<td>Caron &amp; Ladouceur, 1989</td>
<td>Suggestive (63%)</td>
<td>Maintenance, but no generalization. Clear treatment description but no integrity of implementation data. Inter-rater reliability checks completed on only 10% of the sample, results &gt; 80% agreement.</td>
</tr>
<tr>
<td>Elliott, Miltenberger, Rapp, Long, &amp; McDonald, 1998</td>
<td>Suggestive (48%)</td>
<td>Generalization, but no maintenance. Clear treatment description and integrity of implementation data. Inter-rater reliability completed across all phases on 33% of sample, all results not &gt; 80% agreement.</td>
</tr>
<tr>
<td>Gagnon &amp; Ladouceur, 1992 (exp 1)</td>
<td>Suggestive (83%)</td>
<td>Maintenance, but no generalization. Clear treatment description and integrity of implementation data. Inter-rater reliability completed across all phases and results &gt; 80% agreement. Completed on 10% of sample.</td>
</tr>
<tr>
<td>Gagnon &amp; Ladouceur, 1992 (exp 2)</td>
<td>Suggestive (87%)</td>
<td>Maintenance, but no generalization. Clear treatment description and integrity of implementation data. Inter-rater reliability completed across all phases and results &gt; 80% agreement. Completed on 10% of sample.</td>
</tr>
<tr>
<td>Gagnon &amp; Ladouceur, 1992 (exp 3)</td>
<td>Suggestive (90%)</td>
<td>Maintenance, but no generalization. Clear treatment description and integrity of implementation data. Inter-rater reliability completed across all phases and results not all at &gt; 80% agreement. Completed on 10% of sample.</td>
</tr>
<tr>
<td>Harasym, Lanevin, &amp; Kulik, 2015</td>
<td>Preponderant (2.3%)</td>
<td>No generalization or maintenance. Clear treatment description and integrity of implementation data. Inter-rater reliability completed across all phases and results &gt; 80% agreement. Completed on 20% of sample.</td>
</tr>
<tr>
<td>Jang, 2008</td>
<td>Suggestive (38%)</td>
<td>No generalization or maintenance. Clear treatment description and integrity of implementation data. Inter-rater reliability completed across all phases and results not all at &gt; 80% agreement. Completed on 30% of sample.</td>
</tr>
<tr>
<td>Ladouceur, Caron, &amp; Caron, 1989</td>
<td>Suggestive (65%)</td>
<td>Maintenance, but no generalization. Clear treatment description but no integrity of implementation data. Inter-rater reliability completed across all phases and results &gt; 80% agreement. Completed on 100% of sample.</td>
</tr>
<tr>
<td>LaSalk, 2015</td>
<td>Conclusive (53%)</td>
<td>ABAB with at least three data points in each phase. Clear treatment description and integrity of implementation data. Inter-rater reliability completed across all phases and results &gt; 80% agreement. Completed on 20% of sample.</td>
</tr>
<tr>
<td>Miltenberger, Wagaman, &amp; Arndorfer, 1996</td>
<td>Preponderant (72%)</td>
<td>Generalization and maintenance but only two participants in this multiple baseline across participants design. Clear treatment description and integrity of implementation data. Inter-rater reliability completed across all phases and results &gt; 80% agreement. Completed on 25% of sample.</td>
</tr>
<tr>
<td>Northrup, 2012</td>
<td>Preponderant (29%)</td>
<td>Maintenance, but no generalization. Clear treatment description and integrity of implementation data. Inter-rater reliability completed across all phases and results &gt; 80% agreement. Completed on 100% of sample.</td>
</tr>
<tr>
<td>Reed &amp; Godden, 1977</td>
<td>Suggestive (86%)</td>
<td>Generalization, but no maintenance. Only two participants in this multiple baseline across participants design. Clear treatment description but no integrity of implementation data. Inter-rater reliability completed across all phases and results &gt; 80% agreement. Completed on 100% of sample.</td>
</tr>
<tr>
<td>Schloss, Freeman, Smith, &amp; Espin, 1987</td>
<td>Preponderant (35%)</td>
<td>Generalization and maintenance. Clear treatment description and integrity of implementation data. Inter-rater reliability completed across all phases and results &gt; 80% agreement. Completed on 33% of sample.</td>
</tr>
<tr>
<td>Trajkovski, Andrews, Onslow, Packman, O’Brien, &amp; Menzies, 2009</td>
<td>Suggestive (92%)</td>
<td>No generalization or maintenance. Clear treatment description but no integrity of implementation data. Inter-rater reliability completed across all phases and results &gt; 80% agreement. Completed on 20% of sample.</td>
</tr>
<tr>
<td>Wagaman, Miltenberger, &amp; Arndorfer, 1993</td>
<td>Conclusive (84%)</td>
<td>Generalization and maintenance. Clear treatment description and integrity of implementation data. Inter-rater reliability completed across all phases and results &gt; 80% agreement. Completed on 25% of sample.</td>
</tr>
<tr>
<td>Williamson, Epstein, &amp; Coburn, 1981</td>
<td>Suggestive (87%)</td>
<td>No generalization or maintenance. Clear treatment description but no integrity of implementation data. Inter-rater reliability completed across all phases and results &gt; 80% agreement. Completed on 100% of sample.</td>
</tr>
<tr>
<td>Study</td>
<td>Participant characteristics</td>
<td>Intervention and outcome characteristics</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----------------------------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td>Blood, 1995a</td>
<td>CA: 14:2, M 15:4, M 15:8, M</td>
<td>Intervention approach (intervention category): Power 2 Sutter-free fluency shaping program (fluency shaping)</td>
</tr>
<tr>
<td>Blood, 1995b</td>
<td>CA: 21, M 20, M 25, M</td>
<td>Intervention approach (intervention category): Power 2 Sutter-free fluency shaping program, CAFET (fluency shaping)</td>
</tr>
<tr>
<td>Bray, 1997</td>
<td>CA: 9, M 13, M 11, M 8, F</td>
<td>Intervention approach (intervention category): Self-modeling (modeling)</td>
</tr>
<tr>
<td>Caron &amp; Ladouceur, 1989</td>
<td>4 subjects ages 6–9 Gender NG</td>
<td>Intervention approach (intervention category): Awareness training, regulated breathing (assertiveness training)</td>
</tr>
<tr>
<td>Elliott, Miltenberger, Rapp, Long, &amp; McDonald, 1998</td>
<td>CA: 11:7, M 8:8, M 7:9, M 9:2, M 5:9, M</td>
<td>Intervention approach (intervention category): Awareness training, regulated breathing/competing response, parental social support (assertiveness training)</td>
</tr>
<tr>
<td>Gagnon &amp; Ladouceur, 1992 (exp 1)</td>
<td>4 males, 10–11 yrs.</td>
<td>Intervention approach (intervention category): Awareness training, easy speech group practice, regulated breathing (assertiveness training)</td>
</tr>
<tr>
<td>Gagnon &amp; Ladouceur, 1992 (exp 2)</td>
<td>4 males, 6–7 yrs.</td>
<td>Intervention approach (intervention category): Awareness training, easy speech group practice, regulated breathing (assertiveness training)</td>
</tr>
<tr>
<td>Study</td>
<td>Participant characteristics</td>
<td>Intervention approach (intervention category)</td>
</tr>
<tr>
<td>-------</td>
<td>-----------------------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>CA</td>
<td>3 males, 7–11 yrs.</td>
<td>Awareness training, group practice, regulated breathing, relapse prevention</td>
</tr>
<tr>
<td></td>
<td>3 mild, 3 moderate, 3 severe</td>
<td>Awareness training, regulated breathing, cognitive restructuring, relapse prevention</td>
</tr>
<tr>
<td></td>
<td>7 males, 15–37 yrs.</td>
<td>Slow rate speaking, demand capacity model (modeled slow speech)</td>
</tr>
<tr>
<td></td>
<td>3 mild, 3 moderate, 3 severe</td>
<td>Awareness training, regulated breathing, competing response (assertiveness training)</td>
</tr>
<tr>
<td>Study</td>
<td>Participant characteristics</td>
<td>Intervention and outcome characteristics</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----------------------------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td>Reed &amp; Godden, 1977</td>
<td>5:10, M 3:9, F</td>
<td>Response contingent punisher (fluency shaping)</td>
</tr>
<tr>
<td>Schloss, Freeman, Smith, &amp; Espin, 1987</td>
<td>23, M 21, M 20, M</td>
<td>Modeling, behavioral rehearsal, feedback, social reinforcement, homework (assertiveness training)</td>
</tr>
<tr>
<td>Wagaman, Miltenberger, &amp; Arndorfer, 1993</td>
<td>6 males, 2 females; 6–10 yrs.</td>
<td>Awareness training, regulated breathing/ competing response, social support (assertiveness training)</td>
</tr>
<tr>
<td>Williamson, Epstein, &amp; Coburn, 1981</td>
<td>39, M 3:11, M 3:8, M</td>
<td>Regulated breathing (assertiveness training)</td>
</tr>
</tbody>
</table>

*MBLR* = Mean between-case improvement ratio; *PND* = Proportion of normal duration.