

**Critical Review:**  
**Does Classroom-Based (“Push-in”) Language Intervention Provide Comparable Treatment Results to Traditional (“Pull-out”) Intervention in Children with Language Impairments?\***

Cox, T.  
M.Cl.Sc (S-LP) Candidate  
School of Communication Sciences and Disorders, Western University

### ***Introduction***

Two service delivery models are used in treating school-age children with speech and language impairments are: (a) traditional pull-out intervention, wherein services are delivered outside of the classroom, and (b) classroom-based intervention, wherein services are delivered within the classroom, either directly by the clinician, or through consultation and collaboration with the classroom teachers (Cirrin et al., 2010). In recent decades, trends in legislation and the education system have promoted more inclusive perspectives and practices with regard to special-needs in the classroom. This trend extends to speech and language services in school settings, with a substantial increase in the use of classroom-based interventions reported (ASHA, 2002; Beck & Dennis, 1997; Cirrin et al., 2010; Law et al., 2002). The economic benefits (Boyle et al., 2007) and therapeutic efficacy of classroom-based intervention have been reported (Cirrin, et al., 2010; Farber & Klein, 1999; Roller et al., 1992). However, concerns have been raised about the universal application of the classroom-based model (Law et al., 2002), and the significant obstacles facing clinicians who attempt to implement it (Beck & Dennis, 1997; Elksnin & Capilouto, 1994). In spite of this, reportedly little research has focused on the relative treatment gains of classroom-based versus pull-out intervention (Alberta Health & Wellness [AHW], 2004; Cirrin et al., 2010). Therefore, given the increase in classroom-based service delivery in schools, the efficacy of this approach as compared to the pull-out approach must be explored.

### ***Objectives***

To critically evaluate existing literature comparing the treatment results of classroom-based and pull-out models of speech and language service delivery in school settings

### ***Methods***

#### **Search Strategy**

PubMed, ProQuest, and Scopus online databases were searched using the following terms: ((service delivery

models) OR (intervention)) AND ((speech-language) OR (language impairment)) AND (school). Additional relevant studies were gleaned from reference lists of the searched articles.

#### **Selection Criteria**

Articles selected for inclusion were required to examine or compare the effects of classroom-based and pull-out models of speech-language intervention as applied to children with language impairments.

#### **Data Collection**

The search yielded the following types of articles: randomized controlled trial (2), quasi-experimental (2).

### ***Results***

#### **Randomized Controlled Trials**

In a randomized clinical trial (RTC) of 39 inner-city African-American preschool children with formally identified Specific Language Impairment (SLI), Valdez and Montgomery (1997) compared the treatment outcomes of two intervention conditions: *Inclusion Group* intervention provided in the classroom with classmates, and *Pull-Out Group* intervention provided outside of the classroom in small groups. A randomized block design was used, with stratification by language impairment severity (mild, moderate, severe), as determined by a standardized pre-treatment language assessment (Clinical Evaluation of Language Fundamentals – Preschool, or CELF-P). Subjects were randomly assigned to one of the four participating Head Start Centres. Two of the centres received the experimental “nclusion group intervention, and two received the comparison Pull-Out Group” intervention. Two speech-language pathologists (S-LPs) conducted the interventions, with each clinician conducting one of each intervention type to avoid the possible effects of the clinician’s model preference.

All groups received intervention from the S-LP in weekly 90-minute sessions, for 6 months (36 hours total). The intervention for all groups consisted of “concept development”. Inclusion Groups received intervention in their classroom with 10-15 of their

\* This paper was created as a required assignment for the CSD9639 Evidence Based Practice for Clinicians course at Western. While it has been evaluated by course instructors for elements of accuracy and style, it has not undergone formal peer-review.

typically developing peers. Pull-Out Groups received intervention in a separate room; Pull-Out Group size was unspecified.

In this study, intervention outcomes were measured by post-treatment CELF-P2 scores. To avoid clinician bias during testing, the SLPs who administered pre- and post-test language assessments were not responsible for intervention. Descriptive statistics were used to compare outcome differences between the two treatment groups. No significant clinical differences were found in total language, receptive language, or expressive language scores between the Inclusion Group and Pull-Out Group. However, the Pull-Out Group showed greater gains in mean receptive language scores than the Inclusion Group. Comparison of means for total language scores revealed that one of the Pull-Out Groups scored higher than the other three groups. Children with severe deficits in the Inclusion Group showed greater improvement in standard scores than those in the Pull-Out Group.

In their discussion, the authors conclude that their findings suggest that “children who are specifically language impaired may be expected to show the same progress in an inclusion setting or a pull-out setting” (p. xxx). However, this conclusion appears to be made by comparing post-treatment scores, rather than post-treatment gains, of the two treatment groups. Comparable gains were found between the treatment groups’ Total and Expressive Language Scores, but considerably larger gains in receptive language score was made by the Pull-Out Group. It is unclear how the authors reached such a conclusion, as clinically relevant findings were described but statistical values were not reported. There are additional reasons to interpret the findings of this study with caution. Most notable is the lack of a detailed procedure description. Beyond the vague description of “concept development,” no details about the intervention were provided, such as the concepts targeted for development, how target concepts were determined, or the specific techniques used to develop them.

Treatment groups were poorly defined by the authors, which may have led to possible confounds in the experiments design. According to the authors, one Pull-Out teacher was “highly motivated... following up with language-based activities in the classroom” (p. xxx). This suggests that Pull-Out Groups received some classroom-based intervention in addition to Pull-Out intervention. More evidence of a design confound can be found in the discussion, where the authors mention techniques “used by the clinician and emulated by the Head Start teachers,” such as remodeling, recasting, and expansion. This suggests that Pull-Out teachers

received consultation that resulted in additional indirect language stimulation for the Pull-Out Groups. Finally, the size of groups receiving Pull-Out intervention was not reported. This information is important in determining the face validity of the study, because if intervention sessions were conducted with all 10 of the subjects from the classroom, it may not be representative of what most clinician would consider to be the traditional pull-out model of speech/language therapy.

Despite the weaknesses discussed above, this study had considerable strengths, including the use of randomization, baseline equivalencies, a lack of attrition effects, and blinding of clinicians conducting pre- and post-treatment language assessments. As such, this study provides suggestive evidence that Inclusion and Pull-Out models of speech and language intervention yield comparable treatment outcomes.

Wilcox, Kouri, and Caswell (1991) compared treatment gains of preschoolers with language impairments randomly assigned to two conditions: (a) classroom-based intervention, and (b) pull-out intervention conducted one-on-one, outside of the classroom ( $n = 20$ ). Interventions employed the *interactive modeling technique* to increase the productive use of target vocabulary. Bi-weekly intervention sessions took place over 12 weeks for a total of 18 hours.

Data recorded during intervention sessions included the number of target word models, spontaneous imitations, spontaneous productions, and attempted productions. In addition, 30 minute language samples were taken 3 times at home during the study to measure generalization. Data analysis revealed no significant difference in target vocabulary growth between groups. However, children in the classroom-based condition were significantly more likely to demonstrate their target words in the home setting.

This study utilized appropriate statistical methods, randomization, subject selection, test reliability, and intervention procedures. Overall, the study provides compelling evidence that classroom-based therapy is as effective as pull-out therapy in improving vocabulary, and more effective in generalizing vocabulary gains.

#### Quasi-Experimental Studies

Bland and Prelock (1996) compared the treatment gains of 7 matched pairs ( $n = 14$ ) of elementary students with communication disorders assigned to two conditions: (a) classroom-based intervention, and (b) pull-out intervention conducted outside of the classroom. Communication disorders were determined from

baseline scores on the Clinical Evaluation of Language Fundamentals – Revised (CELF-R) (Wiig, Secord, & Semel 1987). A range of communication problems were present in the condition groups, including verbal fluency, semantics, expressive organization, pragmatics, word finding, syntax, and oral/written expression. Subjects were matched by age within a 6-month range, and, where possible, by gender and type of communication impairment.

Classroom-based intervention employed the Language in the Classroom (LINC) program, which was previously developed by the one of the study authors (Prelock, Miller, & Reed, 1995). In this intervention, an SLP, an SLP student, and a LINC-trained teacher collaborated to plan and deliver LINC activities, which incorporated curricular materials and goals. No details on specific therapeutic techniques were reported in the study, but a lesson plan included in the appendix indicates that the modeling, prompting, and cuing of target vocabulary was used with some of the subjects. Curricular and subject-specific communication goals were targeted in this intervention. Intervention was provided on a weekly basis in 30-45 minute sessions. Pull-out intervention was conducted once or twice a week in 30-45 minute sessions. No details on the specific therapeutic techniques used during these sessions were provided. In both conditions, intervention was provided to the matched pairs for 1 to 3 years, throughout the school year.

Audio recorded language samples of 100-200 utterances made by subjects in conversation with the SLP were obtained twice a year (spring and fall). Systematic Analysis of Language Transcripts (SALT) was applied to orthographic transcripts of language samples. The SALT analysis produced the following measures: Word/Morpheme Summary, Distributional Summary, and Frequency/Percent of Utterance Types.

Analysis of data was limited to each year of sampling, as subjects reportedly left and entered the study, and switched condition groups between each year in the 3-year timeframe of the study. Between group comparison of results revealed that class-room based subjects produced significantly more intelligible/complete utterances during time period six ( $U = .0250$ ). No other significant differences were found between groups for the first 5 time periods.

Overall, a lack of methodological detail made validity difficult to determine in this study. It was unclear what specific therapeutic techniques were employed in each of the intervention models applied. Also, in relation to the SALT transcription procedures, no attempt to establish intra- or inter-transcriber reliability was

mentioned. Additionally, potential confounds were found in the treatment conditions of this study. More specifically, the authors noted that some subjects moved between condition groups from year to year. This may have led to a confound in the treatment condition as it is unclear whether it was possible for subjects to be able to apply strategies or skills learned in one condition to their activities in the following condition. Lastly, the number of subjects participating in the study within a given year was not reported, making it difficult to interpret the reportedly significant findings of time period 6. In consideration of the weaknesses outlined above, it was concluded that this study provided equivocal evidence that classroom-based therapy is as effective as pull-out therapy.

Throneburg et al. (2000) compared the treatment gains of school-aged subjects with speech or language impairments in 12 classrooms randomly assigned to 3 conditions ( $n = 63$ ).

Two classroom-based conditions were used: collaborative, and independent. In the collaborative condition, the teacher and S-LP planned and delivered the intervention together. In the independent condition, only the S-LP was responsible for intervention. Both classroom-based conditions involved 12 weekly 40-minute language activities that targeted 60 curricular vocabulary and subject-specific speech and language goals (600 minutes total). Subjects in these conditions also received minimal pull-out therapy for their specific speech and language goals (approx. 60 minutes total) when it was deemed necessary by their SLP. The pull-out condition involved 12 weekly 50 minute small group therapy sessions with the same goals and materials as the other conditions, but conducted in a separate room (600 minutes total).

Pre- and post-test data were obtained via a vocabulary test designed and administered by the authors and a team of trained S-LP students. The test was designed to be sensitive to different levels of vocabulary comprehension and use. A graded point system was applied to the following levels of subject responses: (a) correct verbal definitions, (b) correct use of the word in a sentence, (c) recognitions of the word's meaning from two choices. Sub-levels of response criteria were also used in the scoring procedures. Two examiners scored the vocabulary tests from audio recordings. Pearson's product moment correlations were used to establish intra- and inter-judge reliability (.99 and .97, respectively). The authors noted that 13% of the pre-test recordings were discarded due to poor quality or incomplete recording. In these cases, within-session scores were used in the following data analysis.

Results revealed a significant difference in test gains between the three condition groups (ANOVA  $p = 0.045$ ). A Duncan post-hoc analysis revealed that the collaborative group test gains were significantly higher than those of the independent and pull-out groups.

A critical review of this study found several inherent weaknesses that may have undermined the results. First, the lack of assessor blinding may have led to examiner bias during pre- and post-intervention vocabulary testing, which involved a degree of subjective interpretation of responses. Further, any bias present may have been magnified in the 13% of pre-test scores taken within the session by one examiner. Second, treatment confounds may have influenced the results of this study. Subjects in the classroom-based group received a minimum of 60 minutes of pull-out intervention in addition to their classroom-based intervention. Third, the proportion of articulation and language disorders present in each condition group was not controlled for, and no attempts to examine impairment-based differences were reported. The collaborative group had a smaller proportion of language-disordered subjects (42%) when compared to the independent and pull-out conditions (73% and 67%, respectively). Further, given that vocabulary is a language skill, and that vocabulary was the only measure of test gains obtained, the inclusion of subjects with solely articulation-based impairments raises questions about the content-validity of this study's design. In consideration of these significant weaknesses, it was concluded that this study provides equivocal evidence that classroom-based therapy is as effective as pull-out therapy.

### *Discussion*

Overall, this critical review has provided suggestive evidence that classroom-based intervention provides treatment results that are comparable to those of traditional pull-out intervention. However, the suggestive evidence was only found in 2 of the 4 articles, and is thus quite limited in scope.

Two studies examined the relative efficacy of the pull-out and classroom-based models in producing vocabulary gains. Only one of these studies provided a high level of evidence (Wilcox, Kouri, & Caswell, 1991). This same study also produced compelling evidence for superior generalization of vocabulary from the classroom-based model (Wilcox, Kouri, & Caswell, 1991). However, as this evidence is limited to one study, replication of these results is required to confirm this effect. Unfortunately, the second study that used vocabulary growth as a measure of treatment gains contained design flaws that severely compromised its findings (Throneburg et al., 2000). Most detrimental,

was the fact that subjects in this study had both speech and language impairments, while the intervention and measuring procedures targeted solely a language skill (vocabulary). Therefore, since there was no apparent attempt to group impairment types during data analysis, the findings of Throneburg et al. cannot be used to answer the research question of the present study.

Broader measures of treatment gains were used in 2 of the 4 studies. One such study produced suggestive evidence that classroom-based intervention can produce comparable CELF-P score gains to pull-out intervention (Valdez & Montgomery, 1997). While no significant difference was found between groups in this randomized block study, the stratification of subject by impairment severity (mild, moderate, severe) revealed an interesting pattern in treatment gains. From the figures provided, classroom-intervention appeared more successful in improving scores of subjects with severe impairments, while pull-out intervention appeared to produce greater overall movement of severely and moderately impaired subjects into the mild range of impairment. Unfortunately, no attempt to analyze the significance of these differences was reported.

Taken as a whole, three significant weaknesses were present in the majority of the studies reviewed: (a) confounds in treatment conditions, (b) a lack of subject impairment profiles, and (c) a lack of specific detail in the intervention procedures.

First, treatment confounds were present in several forms: (a) pull-out groups receiving some degree of classroom-based intervention (Valdez & Montgomery, 1997), (b) classroom-based groups receiving some degree of pull-out intervention (Throneburg et al., 2000), and (c) movement of subjects between conditions between testing periods (Bland & Prelock, 1996). While I recognize the limitations presented by the ethical and logistical issues of conducting research with impaired children in a school setting, it was feasible to control for these confounds in all cases, either through improved experimental design, or data analysis procedures.

Second, the lack of detailed subject-impairment profiles produces an inherent weakness in the findings of these studies. Namely, it is possible that one of the intervention models was more effective with a specific type of impairment, but without specific impairment profiles, such effects cannot be measured or analyzed. Ultimately, this would likely impact the clinical implications of findings

Lastly, the lack of detail in intervention procedures in 3 of the 4 studies makes it difficult to determine the

validity of these procedures. For example, the type of specific therapeutic techniques may have differed between conditions. Alternatively, environmental factors (e.g., visual/auditory distraction) may have impacted the efficacy of a therapeutic technique. As such, this lack of detail impedes the critical analysis of these studies, and undermines the quality of evidence they have produced.

### **Conclusion**

Despite the limited scope and amount of high quality evidence, this critical review found suggestive evidence that classroom-based intervention produces treatment gains comparable to pull-out intervention in children with language impairments.

In light of the weaknesses outlined in the present research, and the overall lack of research addressing this question, additional research is necessary. Future research on this topic should focus on the following:

- The relative treatment gains of subjects with grouped impairment profiles in classroom-based and pull-out conditions.
- Targeting and measuring a variety of specific language skills (e.g., Brown's morpheme use, syntax, phonological awareness skills, etc.) in the comparison of the two models.

### **Clinical Implications**

As previously mentioned, the evidence found in this review was limited in both scope and quantity; therefore implications should be interpreted accordingly. When targeting vocabulary in language intervention, the compelling evidence for comparable treatment gains and superior generalization of classroom-based intervention indicate that this approach as the recommended model for evidence-based practice. However, this review has not found sufficient evidence for recommending or discouraging the use of classroom-based intervention over pull-out intervention when targeting other language skills in therapy. Given the reported benefits of both service delivery models (Cirrin et al., 2010), and the current lack of research on this topic, it is reasonable to allow factors other than relative treatment gains to inform the selection of intervention approaches, namely, factors such as time and budget constraints (ASHA, 2002; Boyle et al., 2007), feasibility of successful implementation (Beck & Dennis, 1991; Elksnin & Capilouto, 1994; Law et al., 2002), and theoretically-sound, experience-based clinical judgment.

### **References**

- Alberta Provincial Government, Departments of Children's Services, Community Development, Health and Wellness, and Learning. (2004). *Review of Speech-Language Services for Children and Youth in Alberta: A Literature Review*.
- American Speech-Language-Hearing Association. (2002). A workload analysis approach for establishing speech-language caseload standards in the schools: Guidelines and technical report. *ASHA Leader*, 8(7), 21-30.
- Beck, A., & Dennis, M. (1997). Speech-language pathologists' and teachers' perceptions of classroom-based interventions. *Language, Speech, & Hearing Services in Schools*, 28(2), 146-152.
- Bland, L., & Prelock, P. (1995). Effects of collaboration on language performance. *Journal of Children's Communication Development*, 17(2), 31-38.
- Boyle, J., McCartney, E., Forbes, J., & O'Hare, A. (2007). A randomized controlled trial and economic evaluation of direct versus indirect and individual versus group modes of speech and language therapy for children with primary language impairment. *Health Technology Assessment*, 11(25), 1-158.
- Cirrin, F., Tracy, L., Schooling, N., Nelson, S., Diehl, P., Flynn, M., Staskowski, T., & Adamczyk, F. (2010). Evidence-based systematic review: Effects of different service delivery models on communication outcomes for elementary school-age children. *Language, Speech & Hearing Services in Schools*, 3(07), 233-264
- Elksnin, L., & Capilouto, G. (1994). Speech-language pathologists' perceptions of integrated service delivery in school settings. *Language, Speech, and Hearing Services in Schools*, 25(1), 258-267.
- Farber, J., & Klein, E. (1999). Classroom-based assessment of a collaborative intervention program with kindergarten and first-grade students. *Language, Speech, and Hearing Services in Schools*, 30(1), 83-91.
- Law, J., Garrett, Z. and Nye, C. (2004). Speech and Language Therapy Interventions for Children With Primary Speech and Language Delay or Disorder. *Journal of Speech, Language, and Hearing Research*, 47(4), 924-943.
- Law, J., Lindsay, G., Peacey, N., Gascoigne, M., Soloff, N., Radford, J., & Band, S. (2002). Consultation as a model for providing speech and language therapy in schools: A panacea or one step too far? *Child Language Teaching and Therapy*, 18(2), 145-163.
- Prelock, P., Miller, B., & Reed, N. (1995). Collaborative partnerships in a language in the classroom program. *Language, Speech, and Hearing Services in Schools*, 26, 286-292.

- Roller, E., Rodriguez, T., Warner, J., & Lindahl, P. (1992). Implementing collaborative consultation: Integration of self-contained children with severe speech-language needs into the regular education classroom. *Language, Speech, and Hearing Services in Schools, 23*, 365–366.
- Throneburg, R., Calvert, L., Sturm, J., Paramboukas, A., & Paul, P. (2000). A comparison of service delivery models: Effects on curricular vocabulary skills in the school setting. *American Journal of Speech-Language Pathology, 9*(1), 10–20.
- Valdez, F., & Montgomery, J. (1997). Outcomes from two treatment approaches for children with communication disorders in Head Start. *Journal of Children's Communication Development, 18*, 65–71.
- Wiig, E., Secord, W., & Semel, E. (1991). *Clinical evaluation of language fundamentals - preschool*. Psychological Corporation: New York.
- Wilcox, M., Kouri, T., & Caswell, S. (1991). Early language intervention: A comparison of classroom and individual treatment. *American Journal of Speech-Language Pathology, 1*(1), 49–62.