2010 SREE Conference Abstract Template


Abstract Preparation and Submission Procedures

Save this document to your computer. Fill in each of the sections below with the relevant information about your proposed poster, paper, or symposium. Make sure to save the document again when completed. When ready, submit your abstract at http://www.sree.org/conferences/2010/submissions/

The template consists of the following sections: title page, abstract body, and appendices (references and tables and figures). Figures and tables included as part of submission should be referred to parenthetically—“(please insert figure 1 here).” The body section of your abstract should be no longer than 5 pages (single spaced, using the Times New Roman 12-point font that has been set for this document). The title page and appendices do not count toward this 5-page limit.

Insert references in appendix A of this document. Insert tables and graphics in appendix B. Do not insert them into the body of the abstract.

For questions, or for help with abstract preparation or submission, contact us at inquiries@sree.org, or 847-467-4001

Title:
The Effects of Vocabulary Intervention on Young Children’s Word Learning: A Meta-Analysis

Author(s):
Loren M. Marulis, The University of Michigan
Susan B. Neuman, The University of Michigan
Abstract Body

Limit 5 pages single spaced.

Background/context:
Compelling research studies converge on the important of vocabulary knowledge as the foundation for literacy achievement (Dickinson & Neuman, 2006; Neuman & Dickinson, 2001). Children who come to school with well-developed vocabulary skills—a foundation for reading acquisition—are likely to become successful in learning to read. Studies of economically disadvantaged children (Hart & Risley, 2003; Neuman & Celano, 2006), however, indicate wide disparities in oral language and vocabulary knowledge from their more economically advantaged peers. Specifically, linguistically disadvantaged children know about 5,000 words on entry into first grade compared to linguistically advantaged at 20,000 (Snow, Burns & Griffin, 1998). These differences are particularly disturbing considering that vocabulary size is an effective predictor of reading comprehension not only in the elementary grades, but high school as well (Chall, Jacob, & Baldwin, 1990; Cunningham & Stanovich, 1997; Scarborough, 2001).

If we are to close language and vocabulary gaps, interventions must work toward improving and potentially accelerating economically disadvantaged children’s language skills prior to kindergarten entry. This meta-analysis examines to what extent vocabulary interventions stimulate word learning in the early years.

Purpose / objective / research question / focus of study:
Description of what the research focused on and why.

Specifically, our meta-analysis addresses the following questions: 1) Are vocabulary interventions an effective method for teaching word learning to young children? And, 2) What are the characteristics of vocabulary interventions that significantly increase word learning for young children?

Setting:
Description of where the research took place.
*N/A (meta-analysis)

Population / Participants / Subjects:
Description of participants in the study: who (or what) how many, key features (or characteristics).
*N/A (meta-analysis)

Intervention / Program / Practice:
Description of the intervention, program or practice, including details of administration and duration.
*N/A (meta-analysis)

Research Design:
Description of research design (e.g., qualitative case study, quasi-experimental design, secondary analysis, analytic essay, randomized field trial).
*See below (meta-analytic review)

Data Collection and Analysis:
Description of the methods for collecting and analyzing data.

Study Eligibility Criteria

In order to be included in our meta-analysis, studies had to meet the following criteria:

1. The study needed to include:
• An intervention or specific program, or training designed to increase word learning
• A dependent variable that measures word learning (e.g. expressive/receptive vocabulary)
  The measure could be standardized (e.g., the Picture Peabody Vocabulary Test (PPVT) or
  author created (e.g., Coyne, McCoach, & Kapp's Receptive definitions, Expressive
  definitions and Context assessments, Coyne et al, 2007).
• Use an experimental or quasi-experimental design; a randomized controlled trial, a
  pretest-intervention-posttest design, a control group that received no training, or a post-
  intervention comparison between pre-existing groups (e.g., two Kindergarten
  classrooms).
• Study participants were aged 0-9.0 (approximately birth to the middle of 3rd grade) who
  were free of developmental or neurological impairments such as Down's syndrome,
  William's Syndrome or Cerebral Palsy.
• Training was conducted in English

Study Search and Retrieval

We searched the following electronic databases for published and unpublished studies:
PsycINFO, ISI Web of Science, Education Abstracts ProQuest Dissertations & Theses and
Educational Resources Information Center (ERIC CSA & OCLC FirstSearch) through
September 22, 2008 using the following search term: Word learning OR Vocabulary AND
intervention OR instruction OR training OR building OR experience OR learning OR
development OR teaching. This search yielded 53,754 citations. In order to maintain and code
our library of citations, we imported them into the Bibliographic software Endnote. Using
preliminary coding, 3548 citations were deemed potentially relevant and subsequently retrieved
and read in full. In addition to the electronic search, we contacted experts and authors in the
field for any published and unpublished data (their own or that of colleagues) and relevant
references. We received a 29% response rate, generating 33 manuscripts. Therefore, through
both electronic search, manual search and author communication, we attained a total of 3581
papers.

Inclusion Coding

Four University graduate students received extensive training in both general meta-analysis
coding procedures and procedures specific to our vocabulary meta-analysis. After sufficient
training was completed, the four coders read 10 studies together and discussed whether each
should be included based on our five criteria. All disagreements were resolved through
discussion until 100% agreement was reached. Subsequently, a training set of 50 studies was
coded separately by all four coders. The level of agreement reached between the four raters on
their inclusion determination (Fleiss' Kappa = .96) falls well within the “almost perfect
agreement” range (Landis & Koch , 1977). Subsequently, each coder individually coded the
remainder of the studies. One hundred and four papers met all five criteria.

Study Characteristics/ Potential Moderators

In order to address our second question regarding the key elements of successful vocabulary
interventions, we coded all characteristics of each study and intervention that we believed, based
on past research and theory (e.g., NPR 2000, Mol et al, 2008; Elleman et al, 2009) would
influence the effect sizes.

Due to the large number of variables and importance of accuracy, training for this process
was conducted for 4-5 hours per week over a span of 6 weeks by the first author and involved
tutorials on research design, variable coding and practical coding techniques. In addition, the first author created a coding sheet with accompanying coding manual. At the conclusion of the training, all four coders coded 5 papers in full together with extensive discussion and revision of the coding manual and sheet in accordance to the sample studies. Next, the coders coded 5 more papers independently and met to compare and discuss. Fleiss' kappa was calculated for the four coders at .67, which, though it falls within the "Substantial agreement" range was not as high as is needed to allow for proper use of moderator analysis. Lipsey & Wilson 2001 recommend an agreement level of at least .81 and it was clear that more revisions to the coding sheet were needed based on the studies coded. All coding disagreements were resolved through discussion and reference to the tutorial materials from training. The coders independently coded an additional 35 papers (over 60 studies; 150 effect sizes) and achieved an agreement level within the “almost perfect agreement” range (k=.89).

Effect Size Calculation

We entered the data (both reported in the manuscript and obtained from authors) into the software program Comprehensive Meta-Analysis (CMA; Borenstein, Hedges, Higgins, & Rothstein, 2008). CMA allows for multiple types of reported data, including means and standard deviations, F statistic data, categorical data, odds ratios and so on. All effect sizes were expressed as Hedges’ g, a slightly more conservative derivative of Cohen’s d. Hedges’ g includes a correction for biases due to sample size, weighting each effect size by the standard error of the effect size so that less precise estimates are given less weight in the analyses. One overall combined Hedges’ g was calculated for each study and delayed post test effect sizes (post tests given one week or more after intervention) were separately analyzed. We utilized a random effects model in order to generalize our results beyond the studies included in our meta-analysis. In addition, due to the variance in vocabulary interventions, we had reason to suspect that the variability would not be limited to sampling error, in which case a random effects model is recommended (Lipsey & Wilson, 2001). We used the Q statistic to examine heterogeneity.

Publication Bias

We empirically examined our publication bias using the classic fail-safe N test, which indicated that we would need to be missing 61,226 studies in order to potentially invalidate our significant effect size results.

Findings / Results:
Description of main findings with specific details.

Our 103 papers (33 unpublished) comprised of 107 total studies. These studies yielded 345 effect sizes which produced a mean effect size of 0.94 with a standard error of 0.05. The 95% confidence intervals produce a lower limit of .84 and an upper limit of 1.0, p<.0001. The overall effect size for delayed post-tests did not differ significantly or in magnitude from immediate post tests (g=1.1, SE=.14, 95% CI=.8 to 1.3). This significant overall effect size of all vocabulary interventions using all vocabulary measurements is considered both educationally relevant and large according to Cohen’s conventions (Cohen, 1988). (please inert table 1 here).

As we suspected, our sample was heterogeneous (Q=1261.42, df=106, p<.0001). Thus, we examined study characteristics to look for homogeneity and influences on effect size.

Intervention Characteristics

An important moderating variable related to the independent variable was who conducted the
vocabulary intervention. Interventions given by day care teachers were significantly less successful (received significantly lower gains by the children) \( g = .11, \ SE = .1 \) than when conducted by either an experimenter \( g = 1.0, \ SE = .12 \), a school teacher \( g = .89, \ SE = .07 \), or a parent \( g = .73, \ SE = .13 \), \( p < .0001 \). Though there were magnitude differences favoring the experimenters, there were no significant differences found among the other interventionists.

Group size (individually administered interventions vs. small groups of 5 or less children vs. large groups of 6 or more children) did not significantly alter the effect size in either magnitude or significance. Children in all group sizes benefitted equally.

We also coded for the type of word learning intervention (whether the words/concepts were taught in an explicit/extended manner such as detailed definitions and examples given during a storybook reading and discussion or an implicit/incidental manner such as a storybook intervention). Children made significantly higher gains with interventions that used an explicit method \( g = 1.0, \ SE = .07 \) or a combination of implicit and explicit methods \( g = 1.12, \ SE = .08 \) than those that employed an implicit method \( g = .67, \ SE = .04 \), \( p < .0001 \).

Subject level characteristics

Children who were determined by the coders to be at-risk in any way received similar gains to children who were not. Within the category of at-risk, the only variable that was found to be a moderator was income level. Low SES children obtained a significantly lower effect size \( g = .7, \ SE = .09 \) than all other children \( g = 1.4, \ SE = .17 \), \( p < .001 \).

Standardized Measures vs. Author-Created Measures

Another meaningful difference we found in effects sizes was produced by the usage of standardized vs. author-created measures. In general, gains obtained on the standardized assessments were significantly lower \( g = .68, \ SE = .042 \) than those on author-created measures \( g = 1.1, \ SE = .045 \), \( p < .0001 \).

The National Reading Panel (2000) concluded that standardized vocabulary tests did not seem to be sufficiently sensitive to vocabulary changes to be used as dependent measures in vocabulary intervention studies and predicted an underestimate of effect sizes when combining standardized tests with author-created tests. Thus, in order to investigate vocabulary interventions in a more nuanced and appropriate manner as indicated by our findings and as recommended by the NRP, we separated our analyses into the interventions that measured vocabulary growth through standardized measures and interventions that measured vocabulary growth through author-created measures. All subsequent analyses are based on this separation.

SES

Within interventions that utilized author-created measures, there was no difference between gains of low income children and non-low income children. However, within standardized measures, low income children gained significantly less \( g = 0.54, \ SE = .07 \) than non-low income children \( g = .88, \ SE = .1 \), \( p < .05 \).

Interventionists

The vocabulary interventions that were measured with author-created tests were more sensitive to the implementer. School teachers \( g = 1.19, \ SE = .16 \) and experimenters \( g = 1.15, \ SE = .12 \) conducting the interventions were associated with significantly higher gains in children's vocabulary learning than parents \( g = 0.36, \ SE = 1.0 \), \( p < .0001 \).

For vocabulary interventions measured by standardized tests, there were slight magnitude differences but no significant difference in vocabulary growth no matter the intervener. The resultant sample sizes of day care interventionists were too small to run moderator analyses once the studies were separated by author-created and standardized vocabulary measures.
Conclusions:
Description of conclusions and recommendations based on findings and overall study.

These results provide strong support for the overall effectiveness of vocabulary interventions in preschool through 9 year old children. In addition, due to the inclusion of 33 unpublished studies with significantly lower effect sizes (g=0.66, SE=0.038) than published studies (g=0.95, SE=0.045), p<0.0001, it’s likely that our overall effect size is on the conservative end.

The significantly higher gains obtained by school teachers, experimenters and parents as compared to the day care teachers solicits further research into effective training programs for day care teachers who spend large amounts of time with young children during important times for vocabulary growth.

It makes sense that teachers and experimenters who have intervention experience and background as well as have vested interest in the results would tend to produce higher learning gains than parents, who likely know their children the best but may lack training or understanding in learning interventions. This, however, was not evident until we partitioned our sample into the fundamentally different approaches to measuring vocabulary growth. It may be the case that the standardized measures were too obtuse to detect nuanced growth and conceptual changes in vocabulary learning in general and thus were less tractable across interventionists. It may also follow that teachers and experimenters' expertise (and resultant higher vocabulary gains for their students) were more readily detected through the use of the more sensitive and targeted to authentic learning author-created measures.

It is important to note that while these findings are encouraging to learning interventions, relying solely on these targeted tests may overinflate the true or transferable understanding of the children. It may be best to interpret vocabulary learning effect sizes in tandem, being cognizant of the differences between the types of tests and their ability to report on learning growth. If we were to apply this method, we would find that, overall vocabulary interventions are highly effective when measured by either author-created or standardized measures. Furthermore, even though the standardized measures may not be as sensitive or able to measure nuanced and complex vocabulary acquisition and are associated with significantly lower growth than author-created measures, children are still able to make moderate to large gains. This could reflect the most conservative end of the spectrum of vocabulary acquisition in young children with the growth on standardized measures reflecting the other end in which young children are able to make very large vocabulary. It may be that one factor in the achievement gaps of low income children vs. their middle-class peers is related to measurement. We can only speak to the effect within young children's vocabulary learning but it would serve educational researchers to investigate this effect within other disciplines and grade levels.

Appendices
Not included in page count.
Appendix A. References
References are to be in APA version 6 format.

References

*Indicates meta-analyzed studies


*Ammon, P. R., & Ammon, M. S. (1971). Effects of training black preschool children in
vocabulary versus sentence construction. *Journal of Educational Psychology, 62*(5), 421-26


vocabulary outcomes. Charleston, WV: AEL.


*Notari-Syverson, A., & et al. (1996). *Facilitating language and literacy development in preschool children: To each according to their needs*. Retrieved from ERIC database. (ED395692)


Appendix B. Tables and Figures
Not included in page count.