

Investigating the relationship between support for vocabulary development in Jumpstart classrooms
and children's vocabulary development

Jumpstart Vocabulary Study Final Report

Julie Dwyer, Ph.D.

Boston University

August 28, 2012

Acknowledgements

I would like to thank Jumpstart for awarding me the funds to make this study possible. In addition, I would like to thank the Jumpstart Corps Members, classroom teachers, parents, and children who generously participated in this study. Finally, I would also like to thank my research team for all of their dedication and hard work throughout the year.

In the past twenty years, a great deal of research has documented the relationship between early vocabulary knowledge and successful reading comprehension and achievement (Cunningham & Stanovich, 1997; National Reading Panel, 2000; Scarborough, 1998; Senechal, Oulette, & Rodney, 2006). Unfortunately, there are large differences in early vocabulary knowledge and growth trajectory between children from different socioeconomic backgrounds (Chall, Jacobs, & Baldwin, 1990; Denton, West, & Walston, 2003; Hart & Risley, 1995), which undoubtedly contributes to the unacceptable achievement gaps we see between children from advantaged and disadvantaged backgrounds in our nation's schools (Jencks & Phillips, 1998). Considering these issues, it is heartening that a body of research demonstrates that vocabulary instruction can exert an influence on vocabulary knowledge and later reading ability for young, preliterate children (Beck & McKeown, 2007; Biemiller & Boote, 2006; Mol, Bus, & de Jong, 2009; Hargrave & Senechal, 2000; Lonigan & Whitehurst, 1998; National Early Literacy Panel, 2009; National Reading Panel, 2000; Snow, Burns, & Griffin, 1998; Wasik, Bond, & Hindman, 2006).

Given these findings, one might expect to see a great deal of attention paid to vocabulary learning and teaching with our nation's youngest children. However, there is little evidence that this issue is treated with urgency in the education of young children. Research has demonstrated that there is a paucity of intentional, rich, explicit vocabulary instruction in preschool or primary grade curricular materials or classrooms (Beck, McCaslin, & McKeown, 1980; Blachowicz & Fisher, 2000; National Reading Panel, 2000; Neuman & Dwyer, 2009).

This lack of attention to early vocabulary learning and teaching in preschool and primary grade classrooms and instructional materials is surprising, given what we know about early vocabulary differences, the importance of vocabulary for later school success, and the power of instruction to influence vocabulary knowledge. If we are to narrow vocabulary gaps between children from different socioeconomic backgrounds prior to kindergarten entry, it is critical that we

continue to study the efficacy of vocabulary instructional practices and interventions that can be implemented in seamlessly in early childhood classrooms.

The Potential of the Jumpstart Program

Jumpstart is supplementary preschool program that has the potential to positively influence children's vocabulary development. It is a nationally implemented program designed to work toward ensuring that every child enters formal schooling prepared to succeed. To do this, Jumpstart trains adults to work with low-income preschool children in early childhood centers to develop early language, literacy, social, and emotional competencies using developmentally appropriate and active practices. Jumpstart team leaders and corps members work in low-income area preschools twice weekly for two hours implementing research-based, nationally-adopted session plans. Each visit includes a combination of whole- and small-group activities focusing on many aspects of early language and literacy development, including vocabulary development (Jumpstart for Young Children, 2010; www.jumpstart.org).

Because we know there are such large discrepancies between children from different socioeconomic backgrounds in vocabulary knowledge, and vocabulary knowledge is such a key aspect of later reading achievement, this study focuses on the unique capacity of Jumpstart to foster vocabulary development in low-income children. Efforts by Jumpstart to increase general vocabulary knowledge include providing corps members with a list target vocabulary for each session, suggestions for instructional strategies to teach these target words, and reminders to corps members to use both generally rich vocabulary and target vocabulary frequently during the sessions. This innovative, research-based program has the potential to narrow vocabulary gaps that exist between children prior to school entry and provide all children with an equal chance for school success.

This study is designed to investigate the extent and influence of support for vocabulary development happening during Jumpstart classroom sessions compared to the normal course of business in Jumpstart classrooms and typical comparison preschools. Research questions addressed in this study are as follows:

- 1) Do preschool children in Jumpstart classrooms acquire more receptive, expressive, and Jumpstart target vocabulary during the year than children in comparison classrooms?
- 2) Is the extent of support for vocabulary development different in Jumpstart sessions, Jumpstart preschool classrooms during non-Jumpstart time, and comparison preschool classrooms during the usual course of business?
- 3) Is there a relationship between the extent of support for vocabulary development found in Jumpstart and comparison classrooms and growth in children's receptive, expressive, and Jumpstart target vocabulary size?
 - 3a) What is the relationship between the support for vocabulary development provided by regular classroom lead teachers and children's growth in receptive, expressive, and Jumpstart target vocabulary size?
 - 3b) What is the relationship between the support for vocabulary development by Jumpstart Team Leaders and Corps Members and children's growth in receptive, expressive, and Jumpstart target vocabulary size?

Method

Sample

Preschool Sites. The sample for this study was drawn from four individual preschool organizations in five sites (the treatment group has three sites and six classrooms and the comparison group has two sites with four classrooms). The Jumpstart site at this university partners with three participating city preschools, all serving low-income children. The Jumpstart program guidelines for qualifying early education partners state that from 75-100% of children served by potential program partners must qualify as low-income, defined as living at or below state or federal low-income thresholds (personal communication, Jumpstart site manager at a large northeastern university, 2/17/2011). The comparison sites in this study were both Head Start sites. To qualify for Head Start, children must be between three and five years old and live in a family with an income that meets federal guidelines for a particular proportion below the poverty line.

Teachers. In this study, only one “lead”¹ teacher in each classroom was the subject of classroom observations (described in later sections), with a total of 10 teachers identified as “lead” teachers in their classroom who are the focus of classroom observations (6 lead teachers in the treatment group and 4 lead teachers in the comparison group). Table 1 outlines descriptive information about these lead teachers (it should be noted that two of the lead teachers in the treatment group failed to return their questionnaires despite multiple efforts from the research team).

¹ “Lead” teacher is in quotation marks because there were a few cases where the teacher who served as the lead teacher for this study was not technically considered the lead teacher at that site. There were two classrooms where there were two teachers who said that they were truly co-teachers in the classroom. In both of those cases, teachers were asked which individual would like to be the focus of classroom observations and that individual was considered the “lead” teacher for the purposes of this study. In one other classroom the permanent lead teacher was on sick leave during the first observation, so the assistant teacher was the focus of the classroom observation. Thereafter, that teacher was considered the “lead” teacher in that classroom. Anecdotal observation suggested that this individual did interact with children instructionally almost as much as the technical lead teacher.

Table 1

Demographics of Teachers in Jumpstart and Comparison Groups

	Jumpstart Teachers (N = 4)	Comparison Teachers (N = 4)
Female	100%	75%
Average Age Range		
20 or under	25%	25%
21-25 years old	50%	0%
26-30 years old	25%	50%
31-39 years old	0%	25%
Race/Ethnicity		
Black/African	75%	0%
Asian/Pacific Islander	0%	0%
White	25%	0%
Hispanic/Latino/a	0%	50%
Other	0%	25%
Mixed Race	25%	0%
Language Background		
English is first language	100%	25%
Spanish is first language	0%	50%
Bilingual	0%	25%
Years of Experience Teaching PreK as Lead		
Less than 1 to 2 Years	67%	50%
3-6 Years	0%	0%
7-9 Years	0%	50%
Certifications		
Child Development Associates (CDA)	0%	50%
Early Childhood Education Certificate	100%	0%
CDA and Early Childhood Certificate	0%	50%
Highest Level of Education		
High School/GED	0%	0%
Some College	25%	0%
Associate's Degree	50%	75%
Bachelor's Degree	25%	25%

As can be seen in Table 1, teachers in the treatment group tend to be younger, have less experience teaching preschool, and are more likely to speak English as a first language than teachers in the comparison group. All of the teachers in the treatment group have earned an Early Childhood Education certificate only, whereas the teachers in the comparison group have either earned a CDA or earned both an Early Childhood Education Certificate and a CDA. Finally, teachers in the treatment group are somewhat less highly educated than teachers in the comparison group. Although there are some observable differences between these groups (the sample was too small to successfully conduct statistical tests), most seem to favor the comparison group.

Children. There were a total of 161 children in all classrooms whose parents consented for them to participate in the study (of 191 total possible children who were enrolled in these preschool classrooms). Over the course of the study, 28 children “moved”, which means the child either left the preschool site or changed classrooms within the site during the school year². In addition, there were three children for whom we did not have any teacher report to indicate that the child moved from the study, but the child was missing a great deal of data. For all of the reasons cited above, those 28 children were removed from all analyses³.

After removing the children who “moved” from the sample, there were 133 children in the sample (treatment n = 77; comparison n = 56). Descriptive information about these children is depicted in Table 2. Children in the Jumpstart treatment group were slightly younger and more likely to be male than children in the comparison group. In addition, there were many more children in the treatment group that spoke English as their primary language (62% in the treatment group compared to 43% in the comparison group). Seventy-two percent of children in the treatment group and none of the children in the comparison group had ever experienced the Jumpstart program.

Table 2

Demographics of Children in the Jumpstart and Comparison Groups

	Jumpstart Treatment Group (N = 77)	Comparison Group (N = 56)
Mean Age in Months at Pretest	47	50*
Female	51%	57%
Language Background		
Primarily English	62%	43%
Bilingual in English and Another Language	37%	48%
Primarily Spanish	0%	5%
Primarily Another Language	1%	5%
Did not have previous experience with Jumpstart	72%	100%

² Children in both situations were considered to have moved, as in either situation children were not experiencing the same classroom environment throughout the year. In the case of children in the treatment group who changed classrooms within their site, they were also exposed to two different Jumpstart teams during the year.

³ Missing data analyses indicated that the children who “moved” did not have statistically different vocabulary pretest scores than children that did not move. This is one indication that movers and non-movers are a similar population.

Unfortunately, despite many efforts to obtain additional background information on children in this study, only 49% of parents/caregivers returned the background questionnaire that provided this information. Approximately 55% of parents in the treatment condition and 41% in the comparison group returned the questionnaire. Although it is unconventional to do so, for the purposes of this report the demographic information gleaned from this background questionnaire (information on the race/ethnicity of the child, parental education level, and family income) is presented separately from demographic information collected by the research team from children or their teachers (presented in Table 2). This information can be found in Table 3.

Table 3

Demographic Information on Children From Parent Questionnaires

	Jumpstart Treatment Group (N =42)	Comparison Group (N = 23)
Race/Ethnicity		
Black/African	20% ^a	14%
Asian/Pacific Islander	0%	9%
White	24%	5%
Hispanic/Latino/a	20%	32%
Other	7%	23%
Mixed Race	29%	18%
Parent Education Level		
High School/GED	17%	57%
Some College	22%	29%
Associate's Degree	7%	5%
Bachelor's Degree	27%	10%
Master's Degree	12%	0%
Doctorate	15%	0%
Yearly Family Income		
\$25,000 or under	33%	90%
\$26-30,000	10%	5%
\$31-40,000	8%	5%
\$41-50,000	0%	0%
\$51-60,000	0%	0%
\$61-70,000	3%	0%
Over \$70,000	46%	0%

^a All percentages in this table should be read as the proportion of children within that treatment condition who have that characteristic. For example, this number should be interpreted as follows: 20% of the children in the treatment group are of Black/African descent.

Within the sample of children whose parents/caregivers returned these questionnaires (and seen in Table 3), there are significant differences between the treatment and comparison group in

race/ethnicity ($\chi^2=12$, $p \leq .05$), parent education level ($\chi^2=15$, $p \leq .01$), and family income ($\chi^2=18$, $p \leq .001$). In terms of race/ethnicity, though the treatment and comparison groups had similar proportions of children of Black/African descent (20% vs. 14%), there were fewer White children and a larger proportion of children of Hispanic/Latino/a descent in the comparison group. With respect to parental education level, there were stark discrepancies between the two groups—almost 91% of children in the comparison group had parents/caregivers with an associate’s degree or less, whereas 54% of children in the treatment group had parents/caregivers with a bachelor’s degree or higher. Differences between the two groups in family income are equally concerning. Ninety percent of children in the comparison group live in homes that had a total yearly income of less than \$25,000. In contrast, almost half (46%) of children in the treatment group live in homes that had a total yearly income of \$61,000 or more.

Because the information reported in the previous paragraph (and in Table 3) is available for less than half of the children in the sample, it can hardly be considered representative of the entire sample in this study. However, it is important to present this information in this report for two reasons. First, it may be important information for Jumpstart to consider—these data suggest that the children Jumpstart is working with (in this site at least) may be a significantly different and more advantaged population than Jumpstart intends. Secondly, it is important for the rigor of the study to look carefully at differences between groups on critical demographic factors that are known to be associated with children’s learning and achievement⁴. Demographic differences between

⁴ To address this point, missing data analyses were conducted to determine if there were significant differences on pre- and posttest vocabulary outcomes between children whose parents returned the questionnaire and those who didn’t. This was done because it is possible that the parents/caregivers who returned the questionnaires may be qualitatively different than parents/caregivers who didn’t return the questionnaire. Since we don’t have any demographic data on children or parents for whom we don’t have a completed questionnaire, the best data we have to investigate this issue is child vocabulary outcomes—which may serve as a sort of proxy for other demographic differences. Importantly, there were no significant mean differences in any pre- or posttest vocabulary outcomes between children whose parents returned the questionnaire and children whose parents did not return the questionnaire. Though far from conclusive, this suggests that children whose caregivers didn’t return the questionnaire *may not* be that different from children whose caregivers did return the questionnaire. This indicates that the data in the questionnaires, though clearly incomplete, *may* be representative of true demographic differences between the treatment and comparison groups.

groups is worrisome for assuming parity between groups in this study. Data collected from the questionnaire, when taken at face value and considered representative of the sample writ large, suggest that the children in the Jumpstart treatment group may be a qualitatively different population than children in the comparison group, which makes it somewhat difficult to compare their learning (see Appendix A for an analysis of pretest and posttest differences between children based on parental education level, family income, and race/ethnicity; this information is placed in an appendix rather than the body of the paper because these variables will not be used in multivariate analyses due to the large amount of missing data—however, these data may be useful to Jumpstart in understanding how these important factors are related to children’s outcomes).

Jumpstart Corps Members. Thirty-six Jumpstart corps members signed consent forms agreeing to participate in the study and filled out background questionnaires. Of these 36 individuals, six were Team Leaders in the Jumpstart program. Each of these individuals was observed and audio taped twice during the year during classroom observations (described more fully in the following pages). Of the remaining program participants who agreed to participate, one member from each Jumpstart team was selected as an “average” corps member who was in their first year as a Corps Member. The Jumpstart Program Coordinator at the PI’s university selected this “average” corps member based on her initial site visits during the first months of the Jumpstart program and on the recommendation of each team leader. This average corps member completed a background questionnaire, was observed during two rounds of classroom observations, and was audio taped during these observations.

Table 4

Demographic Information on Jumpstart Corps Members

	Jumpstart Corps Members (N = 36)
Female	87%
Mean Age in Years	20
Race/Ethnicity	
Black/African	3%
Asian/Pacific Islander	14%
White	65%
Hispanic/Latino/a	8%
Other	0%
Mixed Race	8%
Language Background	
English is first language	68%
Bilingual	30%
School Where Corps Member Enrolled	
School of Education	14%
College of Arts and Sciences	32%
College of General Studies	14%
College of Health and Human Sciences	11%
School of Management	5%
School of Communication	5%
Dual Enrollments	14%
Year in College	
Freshman	27%
Sophomore	22%
Junior	32%
Senior	14%
Beyond Senior Year	3%
Courses Taken	
Corps Member Has Taken Early Childhood Courses	49%
Corps Member Has Taken Literacy Courses	41%
Years of Experience in Jumpstart	
This was first year	73%
This was second year	16%
This was third year	8%

All of the 36 Boston University undergraduate and graduate students participating in the Jumpstart program and who agreed to participate in this study completed background questionnaires. As Table 4 indicates, Jumpstart Corps Members (JCM) at this university are overwhelmingly female and average 20-years-old. Sixty-five percent self-identify as White, 14% as Asian/Pacific Islander, 3% as Black/African, 8% as Hispanic/Latino/a, and the remainder of respondents as another or mixed race. Sixty-eight percent speak English as their first language and the remainder of students reported that they are bilingual with English and another language.

Educationally, JCMs range across years of college and the school in which they are enrolled—interestingly, only 14% of JCMs were enrolled in the School of Education. Almost half of JCMs have taken at least one course focusing on some aspect of early childhood and about 41% have taken a course that focuses on some aspect of literacy learning or teaching. Almost three-quarters of JCMs were participating in Jumpstart for the first time this year.

Child Measures

General Receptive Vocabulary. To measure general receptive vocabulary, the Peabody Picture Vocabulary Test, Fourth Edition was individually administered to all children in both the Jumpstart treatment and comparison groups at pretest and posttest. Children were shown a series of cards with four pictures on each card: the target and three distractors. The children were told to point to the picture that best depicts the target word. As a widely-used, respected, valid, and reliable assessment, this measure serves as a standardized and norm-referenced measure of children's general receptive vocabulary knowledge, taking age into account. Form A was administered at pretest and Form B was administered at posttest. Across test forms and ages, median reliability of this measure was above .90 (Dunn & Dunn, 2007). In preparation for multivariate analyses, the pre- and posttest versions of this measure were standardized to have a mean of 0 and a standard deviation of 1.

General Expressive Vocabulary. To measure general expressive vocabulary at pretest and posttest, all children in both groups were individually administered the Expressive Vocabulary Test, Second Edition (EVT-2)(Williams, 2007). Children were shown individual pictures on cards and asked to name the item depicted on the card. This assessment is standardized, norm-referenced, and takes age into account. In addition, it has been designed to be administered easily with the PPVT-4. Form A was administered at pretest and Form B was administered at posttest. Reliability of this

measure is above .90 (Williams, 2007). As with the PPVT, this measure was standardized in preparation for multivariate analyses to have a mean of 0 and a standard deviation of 1.

Jumpstart Target Word Assessment (JTWA). This assessment is a receptive vocabulary assessment similar to the PPVT that measures children's knowledge of a sampling of Jumpstart target words. Children were shown a series of 50 cards with four pictures on each card: the target item and three distractors. Children were told "Point to ___(target word)___" and the assessor recorded the child's choice.

To develop this measure, words that Jumpstart targets in weekly session plans were identified. These target words were taken from one of two different areas in the session plans: a) words identified from the trade books that Jumpstart Corps members read with children, and b) words that are identified as germane to "Let's Find Out About It" or "Science" activities in Jumpstart session plans. There were 309 total target words (or phrases) that adhered to these criteria and were identified in Jumpstart session plans. Of this number, 201 were from trade books, 89 were from "Let's Find Out About It", and 19 from Science (see Table 5 for a complete list of target words from each category). Of these words, 26 appear more than one time in session plans as target words (in other words, 26 of these words are supposed to be reviewed at least one additional time beyond the initial exposure) (see Table 5 for a list of words that are identified as target words more than one time, identified by an underline).

Table 5

Jumpstart Target Vocabulary Words and Those Selected for Target Word Assessment

Activity Setting For the Focus on Jumpstart Target Vocabulary Words	Vocabulary Words
Trade Books (Total n = 201; Target Word Assessment n = 38)	<u>Address</u> *, alligator, ambled ** <i>, apron</i> , artist, as usual, balance , banged, bank, barnyard, <u>bees</u> , blackberries , blur, brother, burping, bursting with pride, bustling, butterfly, candles , <u>cellar</u> , chill, chop, <u>clothespins</u> , clover , <u>collar</u> , collection, <u>colorful</u> , corner, cozy, crack, cradle , crawfish , crib, crisp, crocodile, cupboard, darts, dashes, dawn, delicatessen, delicious, diamonds , disappear, disgusting, dough, downstream, draped , dreams, droplets, drumming, ducklings, echoed, edges, enormous, <u>envelope</u> , everyone, everywhere, except, <u>faded</u> , <u>feather</u> , felled, fetch, filthy , flapped, gasp, <u>gentle/y</u> , goose, grasshopper , hare , harmonica , hatch, hedge, hooray, hopscotch, howling, hurtles, invite, island, jam, jerk, jig , <u>kite</u> , kitten, knead , lap, large appliances, lawn, leaps , leash, lick, lightning , lullaby, magician, marbles, mare , mash, merganser, merry-go-round, moaned, molt, monster, monumental, mouse, mozzarella, neighbors, nibble, nudges, ordinary, pail, pajamas, parrot, peeking , perfume, photograph, pig, piglet, pinwheel, pizza slicer, pollen, proud, puppy, raccoon , races, rainbows, raindrops, rascal, <u>reflection</u> , <u>ripe</u> , rummage, sail, saleslady, saucer, scent, screamed, <u>sea horse</u> , seal, seashells, seeds, <u>shadow</u> , shovels, shrub , signs, sip, slammed, sleeve, slurps, smash, smudge, snarled, soaked , soared , sobs, sour, spoiled, squares, stain, <u>stamp</u> , starfishes, <u>stems</u> , storm, stray, <u>stretched</u> , striped , sturdy, sundown, supper, swamp, swishing, tears, teenager , thorn , thunder, thundercloud, tight , tilted , tiptoed, together, topping, traffic , trapped, ugly, under, unfolded, unison, unlatched, vegetables , vines, waddled, wails, warblers , water lilies, watering cans, wheelbarrow , whispered , windowsill, wing, wish, wondering , worms, wrecked, yawned
“Let’s Find Out About It” (Total n = 89; Target Word Assessment n = 10)	ads, air, airplane, airtight, aisle, balloon, birth, birth announcement, birth date, blossom, breakable, bright, carry, cashier , clothesline, clumps, color names, cool, country, crease, <u>damp</u> , <u>dark/er</u> , decoration, dragon, <u>dry</u> , dull, evaporate, expand, expose, fan , flag, flashlight, float, flower names, fold , fruit, fruit & vegetable names, glass, groceries, grocery store, hair, heavier, image, imagination, imagine, inches, inflated , ingredients, initials, invitation, length , letter , lid, light, <u>lighter</u> , loose, mammal , measure, message, mirror , moist, mud, neck, nest , newborn, nurse, nutrients, opaque, particles, pillow slip (pillowcase), postcard, predators, rich, roots, sandy, shiny, skin, smooth, soil, stationery, steps, sticks, supermarket, through, translucent, transparent , twigs, weight, windy
“Science” (Total n = 19; Target Word Assessment n = 2)	<u>absorb</u> , absorbent , dried grass, drop , eyedropper, mix, mold, overlap, plant, release, repel, seed coat, shape, spread, squeeze, straw, <u>waterproof</u> , weave, wheat

*Underlined words are words that are identified as a target words more than once in Jumpstart session plans

** **Bolded and italicized words** are words that appear in the Target Word Assessment used in this study

The Jumpstart target vocabulary assessment is comprised of a sampling of 50 of the 309 identified Jumpstart target words (see Table 5 for the 50 words that were included in the final Target Word Assessment, denoted in bold). Words were chosen based on the following criteria:

- 1) **Two or More Per Session Plan:** An attempt was made to include at least two words from each session plan to represent the words covered in each plan. Due to difficulties depicting

some words, there are two sessions that have only one word represented on the Jumpstart Target Vocabulary Assessment. In addition, there are more than two words included in the assessment from several session plans.

2) Difficulty Level: All 309 words were coded for their difficulty level according to Andrew Biemiller's (2010) Words Worth Teaching List. This list categorizes words into 6 categories: a) E: Words that are easy and are generally known by most children at the end of second grade without any direct teaching, b) T2: Words that are a high priority to be taught to children by the end of grade 2 because they may not be learned by all children if they are not taught directly, c) L2: Words that are a low priority to be taught to children by the end of grade 2 because they are somewhat less common words, d) T6: Words that are a high priority to be taught to children by the end of grade 6, e) L6: Words that are a low priority to be taught to children by the end of grade 6, and f) D: Words that are so difficult that they are rarely known by children by the end of grade 6. In addition, words were coded as NA if it was not found on Biemiller's (2010) list. Table 6 depicts what proportion of the 309 Jumpstart target words fell into each difficulty level. In an attempt to represent in this assessment the general difficulty of the total list of Jumpstart target words, assessment words were chosen in part to mirror these proportions.

Table 6

Proportion of Target Words of Each of Biemiller’s (2010) Difficulty Levels

	E	T2	L2	T6	L6	D	NA
Proportion of Total Jumpstart Target Vocabulary Words (n = 309)	36%	16%	3%	6%	2%	4%	34%
Proportion of Words in the Target Word Assessment (n=50)	30%	20%	2%	8%	4%	18%	18%

3) **Part of Speech:** All 309 Jumpstart target words were assigned a part of speech (e.g. the target word *grasshopper* is a noun). Although it is more difficult to depict verbs on this type of assessment, every effort was made to include a substantial number of verbs in this assessment. The final target word assessment that was administered included 32 nouns, 8 verbs, and 10 adjectives.

In addition to choosing target words, in a receptive vocabulary measure it is equally important to carefully choose three distractors words and associated stimulus pictures to be depicted with each target word. The distractors words were chosen based on the following criteria:

1) **Semantic Similarity:** For each target vocabulary word chosen to be included in the target word assessment, a semantic category was assigned. Then three distractor words were chosen for their inclusion in that semantic category, their semantic similarity to the target word, and their semantic similarity to one another. For example, for the target vocabulary word “knead” the semantic category was “actions you do while cooking” and the distractors were “mix”, “sift”, and “grate”.

2) **Difficulty Level:** Distractor vocabulary words were also chosen based on their difficulty level. After target words were chosen and semantic categories assigned for each target word, the three distractors that were to be depicted on each picture plate were chosen such that one distractor word was in Biemiller’s (2010) E category (words that are easy, or likely known

by children at grade 2 without instruction), one distractor was in the T2 or L2 categories (high or low priority to be taught in second grade, respectively), and one distractor was in the T6, L6, or D categories (high priority to be taught by 6th grade, low priority to be taught by 6th grade, or difficult even for 6th graders, respectively). Choosing distractors based on difficulty level increased the chances that there would be one distractor that the child is more likely to know (E), one that they may or may not know (T2 or L2), and one that they are unlikely to know (T6, L6, or D).

To choose stimulus pictures for both target words and distractors, a trained research assistant and the PI searched the internet for free-source color drawings that clearly and accurately depicted each target word. Every effort was made to ensure that none of the individual pictures on a picture plate were more salient than the others in terms of color, shape, texture, or interest.

Before the study began, a pilot version of this assessment containing 60 items was piloted in two ways. Once stimulus pictures for both target and distractor items were chosen, these pictures were sent to 17 college educated adults. These adults were asked to provide one word that they thought best described each stimulus picture. This rationale for this procedure was to determine if mature language users identified the stimulus pictures as the word that it was intended to depict. Adult responses to each picture were coded as “exact match”, “derivative of word”, “phrase that includes word or derivative of the word”, “synonymous with word”, or “other word”. Stimulus pictures that were had a majority of adult responses for that item coded as “exact match”, “derivative of word”, “phrase that includes word or derivative of the word”, or “synonymous with word” were left unchanged. Stimulus pictures that had a majority of responses coded as “other word” were either discarded or, if appropriate, were replaced with more clear pictures. This resulted in a list of 60 target words and corresponding distractor items.

After changes were made to the assessment based on the adult pilot data, a second round of piloting was conducted with a sample of 21 three-, four-, and five-year-old children attending an area Head Start program that had not experienced the Jumpstart program. Children were shown each of the 60 picture plates. Each plate included the target item and three distractors. Children were told “Point to (target word)”. Children’s exact responses were recorded and also coded as accurate or inaccurate. Each item was analyzed for the proportion of children that correctly identified the item and the proportion of children that chose each of the three distractors.

Since the goal of this assessment was to document learning that occurs over the course of children’s exposure to the Jumpstart program (and natural learning that might occur in comparison classrooms), it was important to create an assessment that had as few items as possible that many children already knew before the Jumpstart intervention began. Therefore, items that were correctly identified by 65% or more of the children in the pilot sample were removed from the assessment. In addition, items where 10% or less of the sample of children correctly identified the item were carefully scrutinized for the source of the difficulty. If it was decided that the item was adequately depicted but just unknown by this sample of children, the item was kept in the assessment. If the source of children’s difficulty with the item seemed to be that it was somewhat poorly depicted, then the picture was modified or changed where possible. If it was not possible to find an acceptable alternative picture, the item was removed from the assessment. Finally, if there were discernible patterns in the distractors that children seemed to be drawn to choosing, these distractors were carefully analyzed, and in many cases changed for another picture or word.

This second and final round of piloting resulted in a 50 item target word assessment that was administered to all children participating in the study at both pretest and posttest. These words are presented in Table 5. Though all 50 items were administered at each time point to each child, post-hoc analyses suggested that some of the items did not work well for various reasons. Those items

were subsequently removed from the measure, resulting in a 33 item measure⁵. Alpha reliability of this measure averaged across pretest and posttest administrations was adequate (Cronbach's $\alpha = .73$). In addition, this measure appears to have concurrent validity with other widely-used and well respected measures of children's vocabulary knowledge—at pretest it was significantly correlated with pretest PPVT ($r = .65, p \leq .01$) and pretest EVT ($r = .61, p \leq .01$); at posttest it was correlated with posttest PPVT ($r = .64, p \leq .01$) and posttest EVT ($r = .62, p \leq .01$). A total score was calculated for each child by summing the number of items he or she correctly identified. Scores were then standardized such that they had a mean of 0 and a standard deviation of 1.

Classroom Observation

Observations of Support for Vocabulary Development. To investigate the extent and nature of support for vocabulary development in Jumpstart and non-Jumpstart preschools, two two-hour long observations were conducted in each Jumpstart treatment classroom during Jumpstart sessions, in Jumpstart classrooms during non-Jumpstart time, and in comparison classrooms. During observation in Jumpstart classrooms during non-Jumpstart time and in comparison classroom during the normal course of business, trained observers observed, coded, and audio-taped the lead adult for two hours of the preschool day, focusing on documenting all observed instances of vocabulary instruction or support.

Observers were trained to document each instance of support for vocabulary development that was observed on an electronic protocol sheet. Instances of support for vocabulary development that they were to document included the following: 1) the teacher/Jumpstart Corps member provides

⁵ All items where children scored at (25% of children got the item right) or below chance (less than 25% of children got the item right) at pretest or posttest were removed; this included *length, transparent, crawfish, amble, peek, knead, mare, warbler, draped, hare, tilted, shrub, mammal, and cellar*. Additionally, *inflated* and *jig* were removed because anecdotal observation suggested that a majority of children were drawn to the target stimulus pictures because they were interesting pictures, not because they knew the correct answer. Finally, *absorbent* was removed because the item had strange pictures that appeared to confuse children.

children with the meaning of a target vocabulary word (e.g. the adult says, “It’s *cloudy*. That means it’s going to rain”), 2) the teacher/Jumpstart Corps member helps child/children to determine the meaning of a target vocabulary word (e.g. the adult has a back and forth exchange with children about the meaning of the word *endangered*), 3) the teacher/Jumpstart Corps member labels an object (e.g. the teacher says “We have two *eyes*” and touches her eye), and 4) the teacher/Jumpstart Corps member uses gestures to help child understand the meaning of a word (e.g. the teacher says “He was *shivering*” as she mimics shivering). In addition, each observer coded the entire observation for changes in the activity setting (e.g. small group, whole group, individual) and content focus (e.g. science).

Interrater Reliability of Classroom Observations. Interrater reliability (IRR) between observers was calculated in two rounds. The first round of IRR was conducted near the beginning of the first round of classroom observations, immediately following an initial paired visit that observers did as an extension of the training (described in subsequent sections regarding training). For the first round of IRR, observers conducted observations in pairs with another trained observer during their second observation. Each of these individuals conducted the observation independently and then met with their partner to come to a consensus on instances of support for vocabulary. The second round of IRR was conducted during the second round of classroom observations, where each observer again went on one paired visit with another observer.

Between the two rounds of observations, there were 6 paired observations completed that could be used to determine interrater reliability. There were 32 total observations conducted in the study, so IRR could be established for about 19% of total observations. The remainder of the observations were done independently by each observer.

To calculate IRR, a research assistant carefully analyzed the similarities and differences between the coding of the pair of research assistants for each paired visit. Each instance of support

for vocabulary development that was recorded by either member of the pair was documented and it was determined whether each member of the pair both recorded and agreed on that instance. The total number of agreements between a pair was then divided by the total number of instances recorded by both observers to get a percent agreement.

Unfortunately, the first round of calculating interrater reliability for these observations was quite low (~ 28% agreement). Because of time scheduling issues with completing classroom observations, the amount of time it took observers to “clean” their protocols, and the time it took to calculate interrater reliability, several independent observations had been completed by the time these reliability issues were discovered. After carefully analyzing where IRR broke down, it became clear that observers were not reliably coding instances where teachers labeled items or used gestures. Some observers were highly meticulous in recording every instance where a teacher referred to a referent and others were less observant. In addition, gestures were often hard to interpret, leading to failures in reliability.

To address this issue, IRR analyses were conducted again with paired observations, this time only focusing on observers’ documentation of instances of support for vocabulary when teachers actually defined a word for children. For example, the following counted as an instance where a teacher provided a meaning of a word for children: “The author is the person who wrote the book”. Author is the vocabulary word and the teacher clearly defined it for children. On the other hand “Look, it’s a giraffe (teacher points to a picture of a giraffe)” no longer was counted as an instance because the teacher only gestured at and labeled the item but did not define it in any way.

Recalculations were done by calculating, for each member of a paired observation, the total number of instances where the teacher defined a word for children. Next, the total number of times both observers recorded the same instances of defining a word was calculated. Then the total number of instances of defining a word across the pair was calculated. Finally, their percent

agreement was calculated by dividing the number of agreements with the number of total instances between the two observers. When only looking at instances of support where the teacher defined the word, IRR was much higher—when averaged across paired observations it was 75% agreement. This was an acceptable proportion of agreement that suggested reliability across observers in coding this type of support for vocabulary development. The remainder of classrooms observations were then recoded such that only instances of support for vocabulary that involved the teacher explicitly defining the word were used and counted.

There are several factors that support the decision to only look at instances of defining words for children. First, clearly interrater reliability analyses suggested that observers were only reliably recording instances of support for vocabulary development that involved defining words for children. Although there are many other ways to help children learn vocabulary, it is beyond the scope of this study to include other types of vocabulary support. Second, providing definitions to children has been shown to contribute to increased vocabulary outcomes (Beck & McKeown, 2007; Biemiller & Boote, 2006; Stahl & Fairbanks, 1986), so it is an important aspect of classroom instruction to include.

Adding further justification for a focus on defining are results from an analysis of suggestions for instruction found in Jumpstart session plans. By identifying the different types of support for vocabulary found in the session plans, it was clear what types of support for vocabulary development the Jumpstart program hopes to provide for children. To conduct this investigation, the PI and a research assistant went through all 20 Jumpstart session plans, pulled out all instances of support for vocabulary development, and coded them by type of support provided. We found many different types of support for vocabulary that Jumpstart suggests that Corps Members use, including defining a word, labeling, using gestures, use in a sentence, among others (Appendix B includes a complete list of the strategies that we found in our analysis). We also conducted analyses that

allowed us to determine what vocabulary instructional moves were most prevalent in Jumpstart's session plans. To do this, we counted the total number of instances of each type of instruction across session plans and then divided by the total number of instances of support for vocabulary across session plans. We found that providing definitions was one of the most prevalent instructional strategies suggested by Jumpstart, accounting for 21% of all instances (Appendix C includes a table that illustrates the percent of total instructional moves found in the session plans that could be characterized as each particular type of instructional move). Taken together, it seems that focusing on instances of support for vocabulary where Corps Members are defining words for children is empirically sound and in line with the goals of Jumpstart.

Preparing Observation Data for Analysis. It was our intention to conduct two observations in each classroom setting (Jumpstart sessions, Jumpstart classrooms during non-Jumpstart time, and non-Jumpstart classrooms during business-as-usual) that lasted two hours each. However, there were several factors that made this difficult to do. First, it was relatively common for teachers to take the entire class of children out of the room during regular classroom time. They often went to another part of the building or outside for gross motor time. For logistical reasons, observers did not follow teachers or children out of the classroom. Therefore, during the time the teachers and children were out of the room, nothing was observed or recorded. This did not happen in all classes, but it happened in enough that it was important to account for this missing data. The average length that lead classroom teachers were in the classroom during observations, being observed and recorded, was 100 minutes (SD = 23 minutes). Secondly, although Jumpstart sessions are slated to be 2 hours long, that was not true in our observations. Rather, average Jumpstart observation across classrooms and time periods was just over an hour and a half (mean = 93 minutes, SD = 15 minutes).

Because observations both within and across teachers and Jumpstart Corps members varied in length, it was important to account for these differences when calculating total instances of support for vocabulary. It would not be fair to compare the total number of instances of support provided when one individual was observed for 2 full hours and another for only an hour and 10 minutes. To make observations with different time lengths comparable, the total average number of instances of support per 60 minutes was calculated for each observation. Then, for each individual lead teacher or Jumpstart corps member, an average across the two observations was calculated. The resulted in an estimate of the average “density” of support for vocabulary per hour that children in a particular setting experience.

Other Measures

Background Questionnaires. Jumpstart corps members, parents/caregivers of participating children, and lead preschool teachers who consented to participate in the study each were asked to complete a questionnaire. These questionnaires were designed such that where possible, they include questions that are similar across participant type (e.g. gender, opinions about the importance of different cognitive, language, and socio-emotional domains) However, there were also questions that were unique to each type of participant (e.g. preschool teachers were asked about their experience teaching preschool, but obviously parents were not asked this question). These questionnaires provide important demographic information. The goal was to use this information in multivariate analyses to predict teachers/Jumpstart Corps member practice and/or children’s vocabulary growth and/or to statistically control for any pre-existing differences between the treatment and comparison teachers or children. However, due to limited response rates (only 49% of parents/caregivers returned these questionnaires), including any variables from the parent/caregiver questionnaire severely limits sample size and therefore cannot be used in multivariate analyses. This

information was used, however, to provide some suggestive descriptive information about the characteristics of the treatment and comparison group (included in previous sections of this report).

Other Child Demographic Information. In addition to filling out questionnaires about their own background, teachers provided information about children's previous experience with Jumpstart (no experience with Jumpstart coded as 0, previous experience with Jumpstart coded as 1), children's language background (primarily English coded as 1, primarily another language or bilingual coded as 0), and children's gender (female coded as 1, male coded as 0). In addition, because children could not all be tested on the same day at pretest and posttest, a time gap variable was created by the PI by subtracting the date of pretest assessment from the data of posttest. This resulted in a number of days between pretest and posttest for each child. This variable was continuous.

Years in Jumpstart. There were a small number of children in the treatment group who experienced the Jumpstart program the year prior to the year of this study ($n = 21$, reported by teachers). These children typically had one extra year of experience with Jumpstart, in addition to the year under study. Generally, these children were in a classroom the previous year that had the Jumpstart program (no children in the comparison group had ever experienced the Jumpstart program). Because it seems that children who experienced the Jumpstart program for two years in a row would outperform children who had only experienced the Jumpstart program for one year, this data is used to look at the influence of two years in the Jumpstart program compared to one year (though the very small number of children who have more than the one year of Jumpstart does limit statistical power to detect differences). Children were coded as 1 if they had previous experience with Jumpstart before the current year and as 0 if they had not.

Training

Child Assessments. Research assistants were trained in administering the Peabody Picture Vocabulary Test-4 (PPVT-4) (Dunn & Dunn, 2007) and the Expressive Vocabulary Test, Second Edition (EVT-2)(Williams, 2007) by an outside paid consultant. Research assistants were trained to administer the Jumpstart Target Word Assessment by the PI and a research assistant who worked on developing the measure. Beyond the initial trainings, research assistants practiced the assessment with four adults of their choosing, engaged in a staged administration of the assessments with the PI, were given feedback on their performance, and were observed in the field by the PI to ensure that they were following standard procedures for tests administration.

Classroom Observations. To conduct these observations, observers were trained to use a coding protocol adapted from one developed by researchers at the University of Michigan (Wright & Neuman, 2010; in addition, the PI contributed to the initial design and piloting of this measure while working with the authors at University of Michigan). Training was conducted by the PI and a research assistant and consisted of a three hour session where research assistants were taught the procedures and coding protocols to follow in the observation. During this training session, research assistants watched a video that was previously coded by the trainers. As a form of modeling, they were shown the video and were simultaneously shown the coding. Next, they were asked to actually engage in coding a short video clip and compare their coding to previous coding done on that clip. Finally, they were given an assignment to complete after the training—they were to engage independently in coding a longer video clip and send that coding to the PI for review. Feedback was then sent to each research assistant on their performance. The final part of the training was to conduct their initial classroom observations with a partner. Each observer independently coded during the entire observation, went home and cleaned up their coding using the audio from the observation, and then met with their partner to come to a consensus on the coding. Finally, they sent

along any questions they had to the PI. These partner observations acted as a continuation of training, while at the same time the consensus coding document was accurate enough that it can be used for actual data.

Procedure

Child Assessments. Children were individually administered all three assessments in one testing sessions. Testing generally took place either in a quiet room near the child's classroom or in the hallway outside of the child's classroom. The assessment battery was always administered to each child in the following order: Jumpstart Target Word Assessment, EVT-2, and PPVT-4. The entire battery took approximately 20-25 minutes per child, with variability depending on a child's ability level and attention span. When a child would become very fidgety, efforts were made to refocus the child. If a child became upset or unwilling to participate, gentle efforts were made to regain the child's attention and motivation. If these efforts did not work, testing was discontinued.

Classroom Observations. Observations were two hours long. After helping the teacher/Corps member place the wireless microphone on their person, observers were to place themselves somewhere in the classroom to be as unobtrusive as possible, but still have a clear view of the teacher/Corps member as he/she interacted with children. This sometimes required the observer to move during the observation period. Observers used personal laptops and an electronic protocol.

Observations conducted in Jumpstart classrooms during non-Jumpstart time and in comparison classrooms only required one observer. Observations conducted in Jumpstart classrooms during Jumpstart time required two observers. One observer focused on the Jumpstart Team Leader for the duration of the observation and a second observer focused on an "average" Jumpstart Corps Member that was identified for observation for the duration of each observation.

Analytic Method

Preparing Child Outcomes for Analysis. To prepare data for analysis, all pre- and posttest vocabulary outcomes were standardized (z-scored) such that they have a mean of 0 and a standard deviation of 1. This allows for comparison of the size of the effect across outcomes.

Investigating Mean Differences. At the beginning of analysis for each child outcome, mean differences between the treatment and comparison group were explored. T-tests and ANOVAs were used to determine if these differences were statistically significant. These analyses provided evidence of pre-existing differences between the two groups (in any case where there were significant differences between the treatment and comparison group on pretest measures, these pretest measures were used as covariates in every analysis in order to statistically control for important incoming differences) and initial suggestive evidence of the efficacy of the Jumpstart program in fostering general receptive vocabulary, general expressive vocabulary, and Jumpstart target vocabulary knowledge.

Multivariate Analyses. Because of the limited sample size (both at the classroom and the child level) and concomitant lack of statistical power, it was not advisable to use a multi-level analytic tool like Hierarchical Linear Modeling (HLM) (Raudenbush & Bryk, 2002). Although imperfect for use with nested data, as the data from this study are, ordinary least squared (OLS) regression was used for all multivariate analyses in this report.

Child's age in months at pretest, gender, language background, time gap, and experience with Jumpstart were included as covariates in every initial multivariate analysis. However, if they were not significant, they were removed to keep models as parsimonious as possible.

Compensation for Participants

To compensate teachers for their time and efforts in participating in this study, preschool teachers in both Jumpstart treatment classrooms and comparison classrooms were given a \$25

bookstore gift card at the conclusion of the study. Additionally, all preschool directors were offered professional development sessions during the summer after the study concluded in order to share the results of the study and to discuss how to provide research-based support for vocabulary development in their preschool classrooms. Thus far, only one preschool has taken advantage of this offer.

Results

Research Question 1: Do preschool children in Jumpstart classrooms acquire more receptive, expressive, and Jumpstart target vocabulary during the year than children in comparison classrooms? Means analyses and OLS regressions were run for each of the three vocabulary outcomes: general receptive vocabulary, general expressive vocabulary, and Jumpstart target word vocabulary. Results from each analysis are listed separately.

General Receptive Vocabulary-PPVT. Overall, there was a statistically significant difference on both the general receptive vocabulary pretest and posttest between the Jumpstart treatment group and the comparison group (pretest differences: $t = 2.5$, $p \leq .05$; posttest differences: $t = 2.0$, $p \leq .05$) (Table 7 presents Jumpstart treatment and comparison group means and standard deviations on the pre- and posttest PPVT). On average at pretest, children in the Jumpstart treatment group were less than 1/3 of a standard deviation from the mean for the norming sample on this measure (for PPVT, this is a score of 100). In contrast, children in the comparison group were almost 2/3 of a standard deviation from the mean. In an ideal quasi-experimental design, treatment and comparison groups would be equal before the treatment began on all child assessment measures. Unfortunately, this is not the case here. To statistically control for these initial differences, pretest scores on receptive vocabulary are used as a predictor in all analyses.

Table 7

Mean Difference at Pre- and Posttest Between Jumpstart Treatment and Comparison Groups on General Receptive Vocabulary

	Jumpstart Treatment Group Mean (SD)	Comparison Group Mean (SD)
Pretest Receptive Vocabulary	96 ^{**a} (18)	89 (13)
Posttest Receptive Vocabulary	97 [*] (16)	91 (13)

* $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$

a Indicates that there was a significant difference in mean PPVT scores at pretest between the Jumpstart treatment group and the comparison group, favoring the Jumpstart treatment group.

Without controlling for any other factors, it appears that neither the treatment group nor the comparison group significantly increased their general receptive vocabulary from pre- to posttest. That said, there were larger standard deviations in the treatment group at both pretest and posttest than in the comparison group, indicating that there was more variability in children’s scores in the treatment group than in the comparison group.

To investigate changes in receptive vocabulary more rigorously, multivariate analyses were conducted. These analyses confirmed what means analyses suggested—there was no significant difference between the treatment group and the comparison group in growth⁶ over the course of the year in general receptive vocabulary (see Table 8 for all results). This is not surprising, as the majority of vocabulary intervention studies that have been conducted have also failed to detect change on this measure. Girls seemed to gain more receptive vocabulary during the year than boys ($\beta = .39, p \leq .01$). In addition, children who entered the year with higher general receptive and expressive vocabulary ended the year with higher general receptive vocabulary ($\beta = .35, p \leq .01$ and

⁶ In this and subsequent analyses the word growth is used to refer to children’s learning over the course of the year. Growth is referred to rather than posttest achievement because when statistically controlling for children’s pretest achievement, the analysis is essentially looking at growth from pretest to posttest.

$\beta = .44, p \leq .001$, respectively). This replicates previous research that suggests that the higher a child's vocabulary, the more vocabulary they are able to acquire (Penno, Wilkinson, & Moore, 2002; Stanovich, 1986).

Table 8

OLS Regression Investigating the Effect of the Jumpstart Treatment on General Receptive Vocabulary at Posttest

Variable	Posttest General Receptive Vocabulary
	β (S.E.)
Gender	.39** (.15)
Pretest General Receptive Vocabulary (PPVT)	.35** (.11)
Pretest General Expressive Vocabulary (EVT)	.44*** (.12)
Treatment condition	-.09 (.15)
F	35***
R ²	.60

* $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$

General Expressive Vocabulary-EVT. There was a statistically significant difference both at the pretest and posttest between the Jumpstart treatment group and the comparison group on general expressive vocabulary (pretest differences: $t = 2.6, p \leq .01$; posttest differences: $t = 3.9, p \leq .001$) (see Table 9). Both groups seemed to increase in their expressive vocabulary knowledge over the duration of the study, but it is unclear from this analysis if there are statistically significant differences between groups in learning from pre- to posttest. It is important to note that, as with expressive vocabulary, it is less than ideal that there are significant differences between groups in this expressive measure. For this reason, children's initial expressive vocabulary score was included in all analyses to control for these initial differences.

Table 9

Mean Difference at Pre- and Posttest Between Jumpstart Treatment and Comparison Groups on General Expressive Vocabulary

	Jumpstart Treatment Group Mean (SD)	Comparison Group Mean (SD)
Pretest General Expressive Vocabulary	101 ^{*a} (18)	93 (15)
Posttest General Expressive Vocabulary	105 ^{***} (13)	96 (12)

* $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$

a Indicates that there was a significant difference in mean EVT scores at pretest between the Jumpstart treatment group and the comparison group, favoring the Jumpstart treatment group.

To investigate this in a more rigorous fashion, OLS regression was used. These analyses suggested that children in the treatment group did outperform children in the comparison group on expressive vocabulary knowledge at posttest ($\beta = .30$, $p \leq .01$), even after controlling for important pretest differences in expressive vocabulary knowledge (and pretest receptive vocabulary knowledge) (see Table 10). This indicates that children in the Jumpstart treatment group gained more expressive vocabulary than children in the comparison group during the year.

Table 10

OLS Regression Investigating the Effect of the Jumpstart Treatment on General Expressive Vocabulary at Posttest

Variable	Posttest General Expressive Vocabulary
	β (S.E.)
Pretest General Receptive Vocabulary (PPVT)	.30*** (.09)
Pretest General Expressive Vocabulary (EVT)	.55*** (.09)
Treatment condition	.30** (.11)
F	79***
R ²	.68

* $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$

Jumpstart Target Word Assessment (JTWA). On average, there was a small but statistically significant difference between the Jumpstart treatment group and the comparison group on the JTWA pretest ($t = 2.0, p \leq .05$) (see Table 11). There was a larger and statistically significant difference between the two groups at posttest ($t = 2.9, p \leq .01$). This is suggestive evidence that children in the treatment group learned more of the Jumpstart target words than children in the comparison group. As was true with the other vocabulary outcomes, it is not ideal that children in the treatment group appeared to know somewhat more Jumpstart vocabulary words than children in the comparison group did before the intervention began. However, pretest scores on this measure will be used to statistically control for this difference.

Table 11

Mean Difference at Pre- and Posttest Between Jumpstart Treatment and Comparison Groups on Jumpstart Target Word Assessment

	Jumpstart Treatment Group Mean	Comparison Group Mean
	(SD)	(SD)
Pretest Jumpstart Target Word Assessment	15 (5)	14 (4)
Posttest Jumpstart Target Word Assessment	20 (5)	17 (6)

* $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$

a Indicates that there was a significant difference in mean JTWA scores at pretest between the Jumpstart treatment group and the comparison group, favoring the Jumpstart treatment group.

A more rigorous multivariate analysis bears out suggestive evidence from mean differences. OLS regression analyses suggest that children in the Jumpstart treatment group acquired significantly more of the Jumpstart target vocabulary words than children in the comparison group ($\beta = .29, p \leq .05$) (see Table 12). This was true even after controlling for worrisome initial pretest differences on all vocabulary outcomes that favored the Jumpstart group, as well as controlling for gender and age. In addition, females and children who were older at pretest scored significantly higher than children who were younger and males on the Jumpstart Target Vocabulary Word Assessment at posttest ($\beta = .48, p \leq .001$ and $\beta = .03, p \leq .01$, respectively). Interestingly, a child's score on the general expressive language assessment did not predict children's learning on the Jumpstart target words.

Table 12

OLS Regression Investigating the Effect of the Jumpstart Treatment on Jumpstart Target Word

Vocabulary at Posttest

Variable	Posttest Jumpstart Target Word Vocabulary
	β (S.E.)
Age in Months at Pretest	.03** (.01)
Gender	.48*** (.13)
Pretest Target Word Vocabulary (TWA)	.34** (.10)
Pretest Receptive Vocabulary (PPVT)	.34** (.12)
Pretest Expressive Vocabulary (EVT)	.07 (.11)
Treatment condition	.29* (.14)
F	28***
R ²	.65

*p ≤ .05; **p ≤ .01; *** p ≤ .001

Overall, analyses demonstrate that after controlling for important variables, children in Jumpstart treatment classrooms acquired more general expressive vocabulary and Jumpstart target word vocabulary than their peers in the comparison group during the course of the school year. There were not differences, however, in children's growth on general receptive vocabulary.

Research Question 2: Is the extent of support for vocabulary development different in Jumpstart sessions, Jumpstart preschool classrooms during non-Jumpstart time, and comparison preschool classrooms during the usual course of business? In the process addressing this research question, it was important to first determine if the Jumpstart classrooms and the comparison classrooms were similar to one another in the amount of support for vocabulary development

provided for children. Doing so establishes that there is relative parity between the two groups on the total extent of support for vocabulary development that is happening outside of Jumpstart sessions. Similar amounts of support for vocabulary development between the treatment and comparison group allows for greater inferences in subsequent analyses that it is the Jumpstart program that may be making the difference in children’s vocabulary growth.

Analysis suggest that there were no statistically significant differences between the two groups on the proportion of vocabulary support that is occurring outside of Jumpstart sessions (see Table 13). It is possible that this null finding was due to insufficient statistical power, since there are only 10 classrooms, each with an individual average score. It is also possible that the null finding was due to large standard deviations indicating that there was great variability across classrooms in the amount of support provided outside of Jumpstart sessions (ranging from 0-26 instances per hour in the treatment group to 4.25-22 instances per hour in the comparison group). However, it should be noted that each group had one teacher that had a substantially higher mean support for vocabulary than the rest of the teachers in that group—which could have been the cause of the very large standard deviation.

Table 13

Mean Differences in Support for Vocabulary Between Treatment and Comparison Groups

	Jumpstart Classrooms During Non-Jumpstart Time (SD) (N = 6)	Comparison Classrooms During Business-as-Usual (SD) (N = 4)
Average Number of Instances of Support for Vocabulary Per Hour	6.89 (9.65)	10.37 (7.95)

*p ≤ .05; **p ≤ .01; *** p ≤ .001

Given that relative parity (at least statistically) seems to exist between treatment and comparison classrooms on the amount of vocabulary support provided *outside* of Jumpstart sessions, it is important to carefully consider what was occurring *during* Jumpstart sessions. Each observation was conducted by two trained observers. One observer focused exclusively on the

Jumpstart Team Leader and one observer focused only on a designated “average” Jumpstart Corps Member on the team. That team leader and average corps member were observed twice during the year.

Analyses indicate that there are large differences between Jumpstart Team Leaders and Average Corps Members in the amount of support they were providing to children during the course of Jumpstart sessions (with a mean difference of 13.41, $p \leq .05$) (see Table 14).

Table 14

Mean Differences in Support for Vocabulary Between Jumpstart Team Leaders and Jumpstart Average Corps Members

	Jumpstart Team Leaders During Jumpstart Sessions (SD) (N = 6)	Jumpstart Average Corps Members During Jumpstart Sessions (SD) (N = 6)
Average Number of Instances of Support for Vocabulary Per Hour	20.84* (6.31)	7.44 (5.03)

* $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$

Finally, differences between all classroom settings were investigated—that is, comparing the amount of support provided by the following individuals: Jumpstart Team Leaders during Jumpstart sessions, Jumpstart Average Corps members during Jumpstart sessions, Lead Teachers in Jumpstart classrooms during business-as-usual, and Lead Teachers in comparison classroom during business-as-usual. It is important to note that the means illustrated in Table 15 are just that-means. Only the standard deviations can illustrate the sometimes substantial variability that was found within each type of individual. For example, there was significant variability among the Jumpstart Team Leaders in the amount of support for vocabulary that they provided. Team Leaders ranged from an average of 12 instances of support per hour to an average of 31 instances of support per hour. There was also significant variability between average Corps Members on the amount of support they provided, ranging from 3 instances per hour to 16 instances per hour.

With that in mind, Table 15 illustrates mean differences in support for vocabulary development between these individuals and settings.

Table 15

Mean Differences in Support for Vocabulary Between Jumpstart Team Leaders and Lead Teachers in Both Groups

	Jumpstart Team Leaders During Jumpstart Sessions (SD) (N = 6)	Jumpstart Average Corps Members During Jumpstart Sessions (SD) (N = 6)	Jumpstart Classrooms During Business-as-Usual (SD) (N=6)	Comparison Classrooms During Business-as-Usual (SD) (N = 4)
Average Number of Instances of Support for Vocabulary Per Hour	20.84 ^{**a} (6.31)	7.44, ns ^b (5.03)	6.89 (9.65)	10.37, ns ^c (7.95)

*p ≤ .05; **p ≤ .01; *** p ≤ .001

- a. To be interpreted that Jumpstart team leaders provided significantly more instances of support per hour than lead teachers in the treatment classrooms.
- b. To be interpreted that there were no significant difference between Jumpstart Average Corps Members and Jumpstart Lead Teachers.
- c. To be interpreted that there were no significant differences between the amount of support provided by lead teachers in comparison classrooms and either Jumpstart average Corps members or Jumpstart lead teachers.

There were large differences in the support for vocabulary development that was provided for children depending on the person with whom they were interacting. Jumpstart Team Leaders were providing by far the most support for children’s vocabulary development—well above average Corps Members and both treatment and comparison lead teachers. Average Corps Members were providing about the same amount of support for vocabulary development as lead teachers in both treatment and comparison classrooms. Finally, lead teachers in each group appear to be providing about the same amount of support.

Research Question 3: Is there a relationship between the extent of support for vocabulary development found in Jumpstart and comparison classrooms and growth in children’s receptive, expressive, and Jumpstart target vocabulary size? This research question will be considered in two

parts, as follows: Research Question 3a) What is the relationship between the support for vocabulary development provided by regular classroom lead teachers and children's growth in receptive, expressive, and Jumpstart target vocabulary size? and Research Question 3b) What is the relationship between the support for vocabulary development by Jumpstart Team Leaders and Corps Members and children's growth in receptive, expressive, and Jumpstart target vocabulary size?

These two questions are addressed separately due to issues of statistical power and interest. Question 3a can be addressed using the entire sample because all children in the sample had a lead classroom teacher who provided a certain amount of support for vocabulary development during observations. This analysis allows for the maximum amount of statistical power, while addressing the issue of the influence of the regular classroom teacher on children's vocabulary growth *above and beyond any effect of the Jumpstart treatment*. When addressing question 3b, it is necessary to limit the sample to only children in the Jumpstart treatment group, as children in the comparison group did not experience support from Jumpstart Team Leaders and Corps Members. This limits the sample size, decreasing the ability to detect effects that may be there. However, it addresses the issue of the influence of Team Leaders and Corps Members on Jumpstart children's vocabulary growth during the year. These research questions will be addressed separately.

Research Question 3a) What is the relationship between the support for vocabulary development provided by regular classroom lead teachers and children's growth in receptive, expressive, and Jumpstart target vocabulary size?

General Receptive Vocabulary. Initial investigation by correlation into the relationship between children's general posttest receptive vocabulary at pretest, posttest, or growth over the course of the year and the support for vocabulary development to which they are exposed by their lead teacher suggests that there may not be a relationship between general receptive vocabulary and

support for vocabulary development (using these measures). There were no statistically significant correlations between pretest, posttest, or growth in receptive vocabulary and support for vocabulary provided by the lead teacher.

More rigorous investigation using OLS regression and the entire sample confirms exploratory analyses. After controlling for important covariates, there was no influence of support for vocabulary development provided by the lead teacher on children’s growth on receptive vocabulary over the course of the year (see Table 16).

Table 16

OLS Regression Investigating the Effect of the Jumpstart Treatment and Support for Vocabulary on General Receptive Vocabulary at Posttest

Variable	Posttest General Receptive Vocabulary
	β (S.E.)
Gender	.37** (.13)
Pretest General Receptive Vocabulary (PPVT)	.32** (.10)
Pretest General Expressive Vocabulary (EVT)	.46*** (.11)
Treatment condition	.02 (.13)
Average Instances of Vocabulary Support by Lead Teacher During Business as Usual	.01 (.01)
F	30***
R ²	.59

*p ≤ .05; **p ≤ .01; *** p ≤ .001

General Expressive Vocabulary. Initial investigation into the relationship between general posttest receptive vocabulary at pretest, posttest, or growth over the course of the year and the support for vocabulary development to which they are exposed by their lead teacher suggests that, as with receptive vocabulary, there may not be a relationship between general expressive

vocabulary and support for vocabulary development (using these measures). There were no statistically significant correlations between pretest, posttest, or growth in expressive vocabulary and support for vocabulary provided by their lead teacher.

Regression analyses suggested the same thing as correlation analyses. It seems that there was no relationship between the average instances of support for vocabulary development to which children are exposed during the usual course of business and their acquisition of general receptive vocabulary over the course of the year (above and beyond the effect of the Jumpstart program) (see Table 17).

Table 17

OLS Regression Investigating the Effect of the Jumpstart Treatment and Support for Vocabulary on General Expressive Vocabulary at Posttest

Variable	Posttest General Expressive Vocabulary β (S.E.)
Pretest General Receptive Vocabulary (PPVT)	.30*** (.09)
Pretest General Expressive Vocabulary (EVT)	.53*** (.09)
Treatment condition	.33** (.11)
Average Instances of Vocabulary Support by Lead Teacher During Business as Usual	.01 (.01)
F	60***
R ²	.69

*p ≤ .05; **p ≤ .01; *** p ≤ .001

Jumpstart Target Word Vocabulary. Initial investigation into the relationship between support for vocabulary development and Jumpstart target word knowledge at pretest, posttest, and growth over the year suggested that there may be some significant relationships. There was a positive and significant correlation between the average instances of support for vocabulary

provided by classroom teachers and growth in target vocabulary over the course of the year ($r = .2$, $p \leq .05$).

Regression analyses indicate that the amount of support provided by the lead teacher had a significant influence on children’s acquisition of Jumpstart vocabulary during the year ($\beta = .02$, $p \leq .05$), above and beyond the influence of the Jumpstart program (see Table 18). This is somewhat perplexing, as anecdotal evidence suggests that lead teachers were doing little to support children’s acquisition of specific Jumpstart target vocabulary.

Table 18

OLS Regression Investigating the Effect of the Jumpstart Treatment and Support for Vocabulary on Jumpstart Target Word Vocabulary at Posttest

Variable	Posttest Jumpstart Target Word Vocabulary
	β (S.E.)
Age in Months at Pretest	.03** (.01)
Gender	.42*** (.12)
Pretest Target Word Vocabulary (TWA)	.38*** (.10)
Pretest Receptive Vocabulary (PPVT)	.26* (.11)
Pretest Expressive Vocabulary (EVT)	.15 (.10)
Treatment condition	.37*** (.10)
Average Instances of Vocabulary Support by Lead Teacher During Business as Usual	.02* (.13)
F	30***
R ²	.66

* $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$

Taken together, it seems that the support provided by lead teachers in these preschool classrooms (as measured in this study) had no effect above and beyond the effect of the Jumpstart

program on children's growth on norm-referenced measures of general receptive and expressive vocabulary. However, perhaps surprisingly, it appears that children's growth on Jumpstart target word vocabulary is positively influenced by the amount of support provided by the lead teacher outside of Jumpstart sessions-and this is true after accounting for the effect of the Jumpstart program.

Research Question 3b) What is the relationship between the support for vocabulary development by Jumpstart Team Leaders and Corps Members and children's growth in receptive, expressive, and Jumpstart target vocabulary size? Previous analyses looked only at the influence of the support for vocabulary development of the lead teacher during the usual course of business (due to issues with statistical power when considering support for vocabulary provided only to treatment children by Jumpstart Team Leaders and Corps Members, which cuts the sample in half). For this reason, additional regression analyses were run for each outcome to investigate the influence of the support for vocabulary development provided by Team Leaders and average Corps Members.

General Receptive Vocabulary. There was no detectable influence of either the amount of support for vocabulary provided by the Team Leader ($\beta = .01$, ns) or the average Corps member ($\beta = .01$, ns) on treatment children's growth in general receptive vocabulary during the year (see Table 19). However, it is important to note that statistical power to detect effects was quite limited when only using children in the treatment group.

Table 19

OLS Regression Using Only The Treatment Group Investigating the Effect of Support for Vocabulary Provided by Jumpstart Team Leaders and Corps Members on General Receptive Vocabulary at Posttest

Variable	Posttest General Receptive Vocabulary
	β (S.E.)
Gender	.26 (.18)
Pretest General Receptive Vocabulary (PPVT)	.28* (.14)
Pretest General Expressive Vocabulary (EVT)	.54*** (.16)
Average Instances of Vocabulary Support By Jumpstart Team Leader	.01 (.02)
Average Instances of Vocabulary Support By Jumpstart Average Corps Member	.01 (.02)
F	19***
R ²	.62

* $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$

General Expressive Vocabulary. Results with respect to general expressive vocabulary appear to be somewhat similar to those around receptive vocabulary. Results indicate that, within the treatment group, there seemed to be no influence of the amount of support provided by the Team Leader ($\beta = .02$, ns) (see Table 20). However, there was a marginally significant influence of the support provided by the average Corps Member ($\beta = .03$, $p \leq .1$) on children's general expressive vocabulary.

Table 20

OLS Regression Using Only The Treatment Group Investigating the Effect of Support for Vocabulary Provided by Jumpstart Team Leaders and Corps Members on General Expressive Vocabulary at Posttest

Variable	Posttest General Expressive Vocabulary
	β (S.E.)
Pretest General Receptive Vocabulary (PPVT)	.36*** (.10)
Pretest General Expressive Vocabulary (EVT)	.50*** (.11)
Average Instances of Vocabulary Support of Jumpstart Team Leaders	.02 (.01)
Average Instances of Vocabulary Support of Jumpstart Average Corps Members	.03~ (.01)
F	48***
R ²	.76

* $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$

Jumpstart Target Word Vocabulary. There were marginally significant correlations between the support for vocabulary provided by the Team Leader and by the average Corps Member and children's target word knowledge at posttest (both $r = .1$, $p \leq .1$). Regression analyses of the influences of Jumpstart Team Leaders and average Corps Members support for vocabulary on children's acquisition of Jumpstart target vocabulary, using only the Jumpstart treatment group, indicated that there were no significant relationships (see Table 21). This is surprising, given that anecdotal evidence suggests that both Team Leaders and Corps Members were consistently and frequently providing children with definitions of Jumpstart target words. However, the very small sample size makes it very difficult to determine if there is an effect that cannot be detected due to lack of statistical power, or if there is truly no effect.

Table 21

OLS Regression Using Only The Treatment Group Investigating the Effect of Support for Vocabulary Provided by Jumpstart Team Leaders and Corps Members on Jumpstart Target Vocabulary at Posttest

Variable	Posttest Jumpstart Target Word Vocabulary
	β (S.E.)
Age in Months at Pretest	.03 (.02)
Gender	.54** (.12)
Pretest Target Word Vocabulary (TWA)	.41*** (.11)
Pretest Receptive Vocabulary (PPVT)	.26~ (.14)
Pretest Expressive Vocabulary (EVT)	.07 (.14)
Average Instances of Vocabulary Support by Jumpstart Team Leaders	.02 (.01)
Average Instances of Vocabulary Support by Jumpstart Average Corps Members	.00 (.02)
F	16***
R ²	.67

*p ≤ .05; **p ≤ .01; *** p ≤ .001

In sum, there was very little discernible effect of the support for vocabulary development provided by Jumpstart Team Leaders and average Corps Members on any of children's vocabulary outcomes. The only exception was a marginally significant effect of average Corps Members' support for vocabulary on children's general expressive vocabulary.

Discussion

This study was designed to investigate the influence and extent of the support for vocabulary development found in the Jumpstart program on children's growth in general receptive, general expressive, and Jumpstart target word vocabulary. Data and analyses revealed several important findings, both for the Jumpstart program and for the field of early childhood vocabulary learning.

Overall Effect of Jumpstart on Children's Vocabulary. Overall, children who experienced the Jumpstart program increased their general expressive vocabulary and Jumpstart target vocabulary knowledge more over the course of the year than children in the comparison group. This suggests that the Jumpstart program is effective in meeting one of their goals—increasing children's vocabulary development. Given that the Jumpstart program trains young adults with very little experience in classrooms and that the program only works with children for four hours per week, these gains are quite impressive.

Despite growth on general expressive vocabulary and target vocabulary, there was no discernible effect of the Jumpstart program on children's growth over the year in general receptive vocabulary. This is perhaps unsurprising, given that this measure is notoriously difficult to change as a result of intervention. That said, it would still be prudent to consider why the Jumpstart program does seem to foster increases in general expressive vocabulary and target word knowledge, but not in general receptive vocabulary. Some vocabulary researchers have suggested that acquisition of receptive and expressive vocabulary requires different instructional strategies (Ard & Beverly, 2004; Hargrave & Senechal, 2000). Perhaps Jumpstart session plans and implementation of these session plans currently include more instructional strategies designed to foster expressive vocabulary than receptive vocabulary. Future research should include efforts to determine how Jumpstart can provide instruction that effectively fosters each area of vocabulary development.

Extent of Support for Vocabulary Development. Results suggested that there were no differences between lead teachers in Jumpstart classrooms and lead teachers in comparison classrooms on the amount of support for vocabulary development that they are providing for children during the usual course of business. This finding suggests that simply partnering with Jumpstart and having a Jumpstart team implement the program in a classroom does not seem to have an obvious influence on the amount of support that classroom teachers are providing for children. Classroom teachers don't seem to be changing their practice as a result of interacting with the Jumpstart program, at least with respect to support for vocabulary development.

This was not necessarily an unexpected finding, given that the Jumpstart program model does not obviously include any professional development for classroom teachers. However, it seems that the power and reach of the Jumpstart program could only be increased by working closely with classroom teachers by including professional development designed to create synergy between what is happening in Jumpstart sessions and what is happening outside of Jumpstart sessions.

Another major finding with respect to support for vocabulary development was the high level of support for vocabulary development provided by Jumpstart Team Leaders as compared to average Corps Members. These large differences could be due to several reasons. First, Team Leaders were all at least in their second year working in the Jumpstart program, whereas average Corps Members were all in their first year working with Jumpstart. These differences in experience could make a difference in how much support an individual would be likely to provide. Second, average Corps Members were specifically chosen *because* they were determined to be "average". In contrast, Team Leaders were appointed to their leadership position because of their performance in the Jumpstart program, which likely included a propensity for working effectively with children. Therefore, it could be that the differences in support observed between "average" Corps Members and Team Leaders was a result of a difference in general ability in working with children. A third

possible contributing explanation is that Team Leaders and average Corps Members have different designated roles and responsibilities during the course of each Jumpstart session. It is possible that Team Leaders assume roles that are more likely to provide opportunities for them to support children's vocabulary development, such as "Let's Find Out About It" and whole group meetings. It is possible that each of these explanations is partially responsible for differences seen between Jumpstart Team Leaders and average Corps Members. Though it is a positive thing (and perhaps unsurprising) that Team Leaders are providing such extensive support for vocabulary, it would be interesting to investigate ways that the amount of support provided by the average Corps Member could be increased to levels similar to that of Team Leaders. Increased training for first year Corps members specific to vocabulary development and/or observation coupled with supportive feedback might be avenues worth investigating.

Though Team Leaders, on average, did provide a much greater amount of support for vocabulary than any other type of adult in this study, it is important to note that there was significant variability among the Jumpstart Team Leaders in the amount of support for vocabulary that they provided. In fact, Team Leaders ranged from an average of 12 instances of support per hour to an average of 31 instances of support per hour. It is beyond the scope of this study to investigate the reasons behind these large differences—but it is important to consider that not all Jumpstart Team Leaders seem to be created equal when it comes to provision of support for vocabulary development. Future research should delve into the characteristics and experiences of particularly successful and supportive Team Leaders—knowing this might help Jumpstart to create training and experiences that contributes to this type of team leadership.

Team Leaders clearly provided the greatest amount of support for vocabulary development. However, it was also true that Jumpstart Average Corps members provided a statistically equivalent amount of support for vocabulary development as classroom lead teachers in both the treatment

classrooms and the comparison classrooms during the usual course of business. For Jumpstart's purposes, it may be welcome news to learn that even first year, average Corps members are providing similar amounts of support for vocabulary as typical preschool teachers (at least in this sample). Average Corps Members as a group have much less experience with children and no certifications in early childhood as compared to lead teachers—and yet they are providing the same amount of support for children.

This could be taken to indicate that Jumpstart is providing experiences and training around vocabulary learning and teaching that are equivalent to the experience and certification of classroom teachers—and are therefore successful. However, it is critical to recognize that the overall level of support for vocabulary development provided by both average Corps Members and lead teachers is relatively low (particularly compared to that provided by Team Leaders). If we wish to increase the vocabulary knowledge of disadvantaged children, both Jumpstart average Corps members and classroom lead teachers could stand to increase the support for vocabulary that they are providing.

Finally, it is concerning that lead teachers in preschool classrooms (both in classrooms where Jumpstart is happening and in comparison classrooms) are providing the lowest amount of support for children of all of the individuals observed in this study. This is particularly concerning for several reasons. First, classroom teachers in this sample have many more years of experience and more education in early childhood than any of the Jumpstart Team Leaders or Corps Members. One would think that this edge in experience and education would translate into greater instructional support for children. That does not appear to be the case in this study, at least with respect to support for vocabulary development. Secondly, the level of support (or lack thereof) being provided by lead classroom teachers is concerning because this sample is comprised of low income children, many of whom entered preschool and this study with below average vocabulary knowledge and are the very children that could benefit the most from effective instruction.

Although it is a positive thing that Jumpstart is present in some of these classrooms to provide additional support, it is concerning that these classroom sites seem to be providing such a small amount of support for vocabulary development outside of the Jumpstart program. Future research with a larger sample of preschool teachers is necessary to replicate this finding. In addition, research should investigate how teacher's level of support could be increased.

Relationship Between Support for Vocabulary Development and Children's Growth on Vocabulary Knowledge. An important part of these analyses was determining if there was an influence, above and beyond the effect of the Jumpstart program, of the support for vocabulary development provided by lead classroom teachers on children's vocabulary growth. Findings indicated that a) there was no relationship between lead teachers' support for vocabulary and children's growth on general expressive and receptive vocabulary, and b) there was a significant and positive relationship between lead teachers' support for vocabulary development and children's growth on Jumpstart target vocabulary. This is a curious finding, given that at no time in the duration of the study was there any indication (anecdotally) that classroom teachers were aware of the Jumpstart target vocabulary, were carefully examining Jumpstart session plans, or were making efforts to integrate Jumpstart vocabulary into their classroom instruction outside of Jumpstart time. In fact, teachers seemed to be relatively uninvolved in the activities occurring during Jumpstart sessions. Given this, it is interesting that lead teacher's support for vocabulary development is related to children's acquisition of Jumpstart target vocabulary. More research is necessary to tease out the explanation for this finding.

When investigating the influence of the support for vocabulary development provided by Jumpstart Team Leaders and average Corps Members on children's vocabulary growth, it was important to only look at children in the treatment group, as they are the only children who had access to the support provided by Team Leaders and Corps Members. Findings indicated that there

was no direct influence of support for vocabulary provided by Jumpstart Team Leaders on general receptive, general expressive, or target word vocabulary knowledge. This is unexpected, considering the extensive amount of support provided by Jumpstart Team Leaders that was documented during observations. In addition, there was very little influence of support provided by average Corps Members on children's vocabulary growth. The exception was a marginally significant influence of average Corps Members' support for vocabulary on children's growth in general expressive vocabulary. Given the fact that Team Leaders provided much more support for vocabulary and did not have an influence on this measure, this seems that it may have been a spurious finding.

It is not immediately clear why there was overall little detectable influence of support for vocabulary development (whether it be provided by lead classroom teachers, Jumpstart Team Leaders, or average Jumpstart Corps Members) on children's vocabulary growth. It seems logical that the more children are exposed to support for vocabulary development, in the form of an adult providing them with the meaning of a word, the more they should learn. The fact that this was not the case is perplexing. Perhaps there was something about the measure of support for vocabulary development that was faulty and not powerful enough to demonstrate a significant relationship. It could be that it is too limiting to measure support for vocabulary development by simply counting adult provision of definitions. An alternative explanation is that the limited statistical power available in these analyses (particularly those that only investigated children in the treatment group) was responsible for the null effects of adult support for vocabulary development. More research is necessary to determine why this measure of support for vocabulary development was statistically unrelated to children's outcomes.

Target Population. One of the main stated goals of the Jumpstart program is to work with children who come from disadvantaged backgrounds. Several findings from this study suggest that

Jumpstart may not always be working with their intended target population. First, children in the treatment group had near average scores on well-respected norm-referenced vocabulary assessments, compared to below average mean scores on these assessments in the comparison group (all of whom attend Head Start). Because vocabulary size has been determined to be highly related to socioeconomic status (Hart & Risley, 1995), it follows that the children with whom Jumpstart is working (at least in this site) are likely more advantaged than the Head Start children in the comparison group. Secondly, data from the family background questionnaires that were returned provide suggestive support for this contention—children in the treatment group appear to come from families that have a *much* higher yearly income and parental level of education than children in the comparison group. These findings suggest that if Jumpstart hopes to work with children from the most disadvantaged backgrounds, procedures for recruiting and partnering with preschools may need to be carefully considered.

Limitations and Areas for Future Research

Though there are many interesting findings that came from this study, there are some limitations that are important to mention. These include issues with sample size, initial differences between groups, and missing data. Each of these issues will be considered in turn.

Sample Size and Statistical Power. The initial sample size proposed for this study was 120 children in the treatment group and 120 children in the comparison group. For various reasons described in earlier progress reports and due to a limited budget, the sample size in this study was considerably smaller than intended. A small sample size has the potential to limit statistical power, which in turn limits the ability to detect differences that may truly exist or may underestimate effects that are detected. For example, significant and positive effects of Jumpstart on children's vocabulary growth over the course of the study may have been underestimated due to the small sample size and concomitant limited statistical power. However, in situations where there are no

effects, the results can be somewhat difficult to interpret. With limited statistical power, it is difficult to determine if null findings are a result of limited statistical power to detect an effect that really does exist (a Type II error) or a result of there actually being no effect of the Jumpstart program or support for vocabulary on children's vocabulary growth. Future studies on the efficacy of Jumpstart should be larger, with more classrooms and children. This will ensure adequate statistical power to detect effects. Increasing the amount awarded by Jumpstart for research grants would make this possible.

Initial Vocabulary Differences Between Treatment Groups. As mentioned previously, there were initial differences between children in the treatment group and children in the comparison group on all vocabulary outcomes. These initial differences were accounted for in all analyses by statistically controlling for initial vocabulary knowledge scores. However, it is well documented that initial vocabulary can have an influence on how much children gain from vocabulary instruction (Reese & Cox, 1999; Silverman & Crandell, 2010). This research indicates that even after controlling for initial vocabulary knowledge, it is possible that there are some other unmeasured ways that the children in the treatment group with higher incoming vocabulary knowledge are different from the children in the comparison group. Future research should look carefully at interactions between incoming vocabulary knowledge and learning to determine if children who enter the Jumpstart program with higher vocabulary do acquire more vocabulary during the course of the year.

Missing Demographic Data. Background questionnaires were designed and distributed to families to collect critical information about the socioeconomic status of the child's family, language background, and information about literacy activities in the home. These data were to be used in analyses to look at differences in the efficacy of Jumpstart for children from different backgrounds and to control for demographic differences between groups. After three rounds of

sending out background questionnaires to families, we had a disappointingly low response rate (about 49%). This was problematic in two ways. First, if these data were used in multivariate analyses, sample size and statistical power would have been greatly reduced. Second, it is possible that the families that did not send these questionnaires back were different from those that did in important ways. These differences could bias the information garnered from these demographic questionnaires. For these reasons, demographic data collected in these questionnaires was not used in multivariate analyses. Because data that was collected suggest that there were large differences between the treatment and comparison groups on important demographic characteristics, it is quite unfortunate that these variables could not be used in analyses to control for these differences. Future research on the efficacy of Jumpstart should engage in even more efforts to collect these demographic data.

Conclusion

Taken together, findings in this study indicate that the Jumpstart program is effective in increasing children's growth on general expressive vocabulary and Jumpstart target word vocabulary. When looking more closely at the levels of support for vocabulary development provided by different individuals in the study, it was clear that Jumpstart is training their Team Leaders to provide a high level of support for vocabulary development to children. However, the program may be less successful at doing so for average Corps Members and classroom teachers. Though it was not possible in this study to detect direct effects of support for vocabulary development on children's vocabulary growth, the fact that children in the Jumpstart program did outperform children in the comparison group on two out of three vocabulary outcomes suggests that there is something about the program that fosters vocabulary learning. Additional research can go further toward determining exactly what Team Leaders and Corps Members are doing to foster

these increases by increasing sample size and honing measures of support for vocabulary development.

What seems clear is that Jumpstart is a program that has the potential to foster vocabulary development in young children who come from somewhat disadvantaged backgrounds. Considering the minimal amount of contact that children have with Jumpstart Corps Members each week, these findings are impressive. Jumpstart appears to be the type of program that can begin to narrow the vocabulary gaps that exist between children prior to school entry and to work toward providing all children with an equal chance for school success.

References

- Ard, L.M., & Beverly, B.L. (2004). Preschool word learning during joint book reading: Effect of adult questions and comments. *Communication Disorders Quarterly*, 26, 1-29.
- Biemiller, A. (2010). *Words worth teaching*. Columbus, OH: McGraw-Hill SRA.
- Beck, I.L., McCaslin, E.S., & McKeown, M.G. (1980). *The rationale and design of a program to teach vocabulary to fourth-grade students*. Pittsburgh: University of Pittsburgh, Learning Research and Development Center.
- Beck, I.L., & McKeown, M.G. (2007). Increasing young low-income children's oral vocabulary repertoires through rich and focused instruction. *The Elementary School Journal*, 107, 251-271.
- Biemiller, A., & Boote, C. (2006). An effective method for building meaning vocabulary in primary grades. *Journal of Educational Psychology*, 98, 44-62.
- Blachowicz, C. & Fisher, P. (2000). Vocabulary instruction. In M. Kamil, P. Mosenthal, D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3)(pp. 503-523). Mahwah, NJ: Lawrence Erlbaum Associates.
- Chall, J.S., Jacobs, V.A., & Baldwin, L.E. (1990). *The reading crisis: Why poor children fall behind*. Cambridge, MA: Brookline Books.
- Cunningham, A.E., & Stanovich, K.E. (1997). Early reading acquisition and its relation to reading experience and ability ten years later. *Developmental Psychology*, 33, 934-45.
- Denton, K., West, J., & Walston, J. (2003). *Young children's achievement and classroom experiences: Findings from the condition of education*. Washington, D.C.: National Center for Education Statistics.
- Dunn, L.M., & Dunn, D.M. (2007). *Peabody picture vocabulary test* (4th edition). Minneapolis, MN: Pearson Assessment.

- Hargrave, A.C., & Senechal, M. (2000). Book reading intervention with language-delayed preschool children: The benefits of regular reading and dialogic reading. *Journal of Child Language, 15*, 765-790.
- Hart, B. & Risley, T.R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Brookes Publishing.
- Jencks, C., & Phillips, M. (Eds.)(1998). *The black-white test score gap*. Washington, D.C.: Brookings Institution Press.
- Jumpstart for Young Children (2010). *Jumpstarting school success: A toolkit for corps members*. New York, NY: Learning Solutions.
- Lonigan, C.J., & Whitehurst, G. (1998). Relative efficacy of a parent and teacher involvement in a shared-reading intervention for preschool children from low-income backgrounds. *Early Childhood Research Quarterly, 13*, 263-290.
- Mol, S.E., Bus, A.G., & deJong, M.T. (2009). Interactive book reading in early education: A tool to stimulate print knowledge as well as oral language. *Review of Educational Research, 79*, 979-1007.
- National Early Literacy Panel (2009). *Developing early literacy: A scientific synthesis of early literacy development and implications for intervention*. Washington, DC: National Institute for Literacy.
- National Reading Panel (2000). *Teaching children to read: An evidence-based assessment of the scientific literature on reading and its implications for reading instruction*. Washington, DC: US Department of Education.
- Neuman, S.B., & Dwyer, J. (2009). Missing in action: Vocabulary instruction in pre-k. *Reading Teacher, 62*, 384-392.

- Raudenbush, S., & Bryk, A. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Newbury Park, CA: Sage.
- Reese, E., & Cox, A. (1999). Quality of adult book reading affects children's emergent literacy. *Developmental Psychology*, 35, 20-28.
- Scarborough, H. S. (1998). Early identification of children at risk of reading disabilities: Phonological awareness and some other promising predictors. In Shapiro, Accerdo, & Capute (Eds.). *Specific reading disability: A view of the spectrum* (pp. 75-199). Timonium, MD: York Press.
- Senechal, M., Ouelette, G. & Rodney, D. (2006). The misunderstood giant: On the predictive role of early vocabulary to future reading. In D. K. Dickinson & S. B. Neuman (Eds.), *Handbook of early literacy research* (Vol. 2) (pp. 173-182).
- Silverman, R., & Crandell, J.D. (2010). Vocabulary practices in prekindergarten and kindergarten classrooms. *Reading Research Quarterly*, 45, 318-340.
- Snow, C.E., Burns, S., & Griffin, P. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.
- Stahl & Fairbanks (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research*, 56, 72-110.
- Wasik, B.A., Bond, M.A., & Hindman, A. (2006). The effects of a language and literacy intervention on Head Start children and teachers. *Journal of Educational Psychology*, 98, 63-74.
- Williams, K.T. (2007). *Expressive vocabulary test* (2nd edition). Minneapolis, MN: Pearson Assessment.

Wright, T. S., & Neuman, S. B. (2010). *Opportunities to learn vocabulary at school: The contribution of teachers*. Paper presented at the Annual Meeting of the National Reading Conference, Fort Worth, TX.

Appendix A

Analysis of the Relationship Between Family Income, Parental Education, and Child Race/Ethnicity and Children's Pre- and Posttest Vocabulary Outcomes

	Pretest	Posttest
General Receptive Vocabulary		
Family Income		
\$25,000 or under	86	89
Over \$70,000 ^a	107*** ^b	104**
Parental Education		
High School/GED	86	86
Bachelor's Degree	99	107**
Master's Degree	116***	106
Child Race/Ethnicity		
Black/African	93	101
White	99, ns	106~
Hispanic/Latino/a	84	86
General Expressive Vocabulary		
Family Income		
\$25,000 or under	89	97
Over \$70,000	113***	117***
Parental Education		
High School/GED	88	92
Bachelor's Degree	105	114
Master's Degree	112***	114***
Child Race/Ethnicity		
Black/African	99	107
White	104*	110~
Hispanic/Latino/a	83	95
Jumpstart Target Word Assessment		
Family Income		
\$25,000 or under	13	17
Over \$70,000	17*	21*
Parental Education		
High School/GED	14	16
Bachelor's Degree	17	22
Master's Degree	19*	25***
Child Race/Ethnicity		
Black/African	15, ns	21, ns
White	15	20
Hispanic/Latino/a	13	15

~ p ≤ .1; *p ≤ .05; **p ≤ .01; *** p ≤ .001

^a To save space and to demonstrate maximum variation in scores, several categories for family income, parental income, and child race ethnicity are not included on this table.

^b This should be interpreted that there were statistically significant differences between different categories of income in mean expressive vocabulary pretest (though means for all categories are not displayed on this table).

Appendix B

Instructional Moves Identified in Jumpstart Session Plans to Enhance Vocabulary Development

Instructional Move	Description of Instructional Move	Example of the Instructional Move from Jumpstart Session Plans
Definition	Plans suggest that the word be defined for children using a child-friendly definition.	“The author means the person who wrote the story”
Gesture	Plans suggest using gestures to help children understand the meaning of a word. This generally would include suggestions about acting out, pantomiming, or mimicking a word. Generally speaking, this does not include pointing to an object (that would be labeling).	“Use gestures and tone of voice, pointing to illustrations, and/or providing brief explanations soon after reading words in the story. For example, mimic Peter’s stretching while reading the word ‘stretched’”.
Tone of Voice	Plans suggest using tone of voice to emphasize vocabulary words.	“Use gestures and tone of voice , pointing to illustrations, and/or providing brief explanations soon after reading the words in the story.
Labeling	Plans suggest labeling an object/action to help children learn the word associated with the object/action. This can include suggesting that they point to the object/action or a picture of the object/action and name that object/action.	“Point to the cradle in the illustration while reading the text and then briefly explain that a cradle is a small bed used for babies-like Peter’s sister Susie” (NOTE: In addition to being coded as Labeling, this would also be coded as Definition).
Use in Sentence: High Contextual Support	Plans suggest using the word in a sentence that includes a lot of support for understanding the meaning of the word.	“Wow, she was crying so hard that her <i>wails</i> were shaking the pictures off the walls. She must have been really sad.”
Use in Sentence: Low Contextual Support	Plans suggest using the word in a sentence that does not include a lot of support for understanding the meaning of the word.	“Her <i>wails</i> were shaking the pictures off the walls.”
Demonstrate use/provide real example of item	Plans suggest that the use of an item be demonstrated for children or that the actual object be shown to children.	“Show a real harmonica and play it. Explain what you are doing (blowing into the holes in it, moving from one end to the other to make different sounds)”.
Making connections to previous experiences/background knowledge	Plans suggest that connections be made to children’s previous experiences, prior knowledge, or background knowledge	Refer back to <i>Oonga Boonga</i> and use some of the vocabulary to talk about the harmonica. For example, ‘Remember how Grandpa played a ‘happy tune’ on the harmonica to try to get Baby Louise to stop crying and that he did a little jig, a

		lively dance. I can play a happy tune and can change the sound of it by blowing gently or with more force”.
Suggests that Corps Members use words from specific categories	Plans suggest that certain categories of words should be used in interactions with children, including the following: target words from the story, general words, sophisticated words, descriptive words, color words, target words from “Let’s Find Out About It”, or target words from a center activity.	“Introduce children to each center activity by showing selected objects from each and briefly demonstrate one to two activities. Use rich vocabulary and full explanations but do not engage children in a discussion.
Encourage children to label generally	Plans suggest asking the child/children to label a general object/action. Note that this is not encouraging use of target words (either from the book or from Let’s Find Out About It).	“Invite children to draw and label a picture of their families. Ask children to name the people and/or animals in the picture, both in English and in his or her first language (where appropriate).”
Encourage children to label generally	Plans suggest asking the child/children to use target words from either the story or from Let’s Find Out About It.	“Encourage children to use vocabulary from the core storybook and make comments about what they say”

Appendix C

Proportion of Total Vocabulary Instructional Techniques Devoted to Each Type of Instruction

Instructional Technique Suggested by Jumpstart^a	Percent of Total Instances of Vocabulary Instruction That the Technique Was Suggested
Use in a sentence	32% ^b
Labeling	23%
Defining	21%
Making connections	17%
Reminders to use target words from center activities	13%
Reminders to use target words	13%
Demonstrate use/provide artifact	13%
Gestures	8%
Encourage children to use target words	8%
Reminders to use sophisticated words	8%
Encourage children to label generally	7%
Tone of voice	6%
Categorized as vocabulary development but not clearly support for vocabulary	11%

a. Eleven percent of the total number of instances of vocabulary instruction found in the session plans were those that were categorized in the lesson plan as supportive of vocabulary instruction but were not clearly supportive of vocabulary development.

b. Percentages do not add up to 100 because several instances of vocabulary instruction included several different instructional techniques.