

Online Supplemental Materials, Interactive Book Reading

Method: Participants

Supplemental Tables S1a and S1b provide detailed information about the participant characteristics in each of the four intensity conditions. Table S1a provides percentile scores on clinical tests as well as the percentage of children in each group who scored at or below the 10th percentile. A Kruskal–Wallis test is shown for each variable. A significant effect of intensity was not obtained for any of the variables, suggesting that the groups did not differ significantly. However, the probability of the intensity effect was low (.10 or less) for two measures: RIAS Nonverbal IQ and CREVT Expressive Vocabulary. For the RIAS, children in intensities 36 and 48 tended to have lower percentile ranks than children in intensities 12 and 24. For the CREVT, children in intensity 36 tended to have lower percentile ranks than children in the other intensities. Notably, correlation analyses show that these two measures were not significantly correlated with treatment outcomes (see Supplemental Table S3a). Moreover, the condition with the lowest scoring individuals on these measures (i.e., intensity 36) was actually the intensity with the best treatment response.

Supplemental Table S1a. Percentile scores for participants in each intensity condition on standardized clinical tests. Results of Kruskal–Wallis test are indicated for each variable.

Standardized clinical test		Intensity 12	Intensity 24	Intensity 36	Intensity 48
Age (years;months) $\chi^2 (3, N = 27) = 0.48, p = .92$	<i>M</i>	5;10	5;9	5;8	5;8
	<i>SD</i>	0;6	0;6	0;5	0;6
	Range	5;0–6;5	5;1–6;4	5;2–6;5	5;0–6;5
RIAS Nonverbal IQ $\chi^2 (3, N = 27) = 6.35, p = .10$	<i>M</i>	75	60	43	46
	<i>SD</i>	27	26	21	20
	Range	45–99	23–99	23–79	30–82
% at or below 10 th percentile		0%	0%	0%	0%
CELF Core Language $\chi^2 (3, N = 27) = 3.69, p = .30$	<i>M</i>	2	3	4	5
	<i>SD</i>	2	3	3	3
	Range	0–5	0–7	0–10	0–8
% at or below 10 th percentile		100%	100%	100%	100%
Vocabulary: DELV Semantic $\chi^2 (3, N = 27) = 0.58, p = .90$	<i>M</i>	9	7	8	10
	<i>SD</i>	11	6	8	8
	Range	0–25	1–16	2–25	1–25
% at or below 10 th percentile		60%	75%	86%	71%
Vocabulary: CELF Word Classes $\chi^2 (3, N = 27) = 5.50, p = .14$	<i>M</i>	22	34	21	18
	<i>SD</i>	17	17	16	25
	Range	9–50	9–63	1–50	5–75
% at or below 10 th percentile		40%	13%	29%	71%
Vocabulary: CREVT Expressive $\chi^2 (3, N = 27) = 6.24, p = .10$	<i>M</i>	32	24	15	35
	<i>SD</i>	23	15	10	14
	Range	10–63	1–50	5–32	18–50
% at or below 10 th percentile		20%	13%	43%	0%

Standardized clinical test		Intensity 12	Intensity 24	Intensity 36	Intensity 48
CELF Concepts & Following Directions $\chi^2(3, N = 27) = 1.88, p = .60$	<i>M</i>	4	8	7	8
	<i>SD</i>	4	8	8	6
	Range	0–9	0–25	0–25	0–16
	% at or below 10 th percentile	100%	88%	86%	71%
CELF Word Structure $\chi^2(3, N = 27) = 1.06, p = .79$	<i>M</i>	8	10	12	12
	<i>SD</i>	7	9	8	9
	Range	1–16	1–25	0–25	2–25
	% at or below 10 th percentile	60%	63%	43%	43%
CELF Recalling Sentences $\chi^2(3, N = 27) = 3.90, p = .27$	<i>M</i>	3	4	4	9
	<i>SD</i>	4	3	4	9
	Range	0–9	0–9	0–9	0–25
	% at or below 10 th percentile	100%	100%	100%	71%
CELF Formulating Sentences $\chi^2(3, N = 27) = 2.45, p = .48$	<i>M</i>	8	5	10	9
	<i>SD</i>	10	5	7	8
	Range	1–25	0–16	0–16	2–25
	% at or below 10 th percentile	80%	88%	57%	86%
CELF Understanding Spoken Paragraphs $\chi^2(3, N = 27) = 0.20, p = .98$	<i>M</i>	7	7	8	7
	<i>SD</i>	6	8	9	8
	Range	0–16	0–25	0–25	0–25
	% at or below 10 th percentile	80%	88%	71%	86%
CTOPP Nonword Repetition $\chi^2(3, N = 27) = 0.32, p = .96$	<i>M</i>	25	24	17	23
	<i>SD</i>	29	24	15	21
	Range	1–75	1–63	1–37	1–63
	% at or below 10 th percentile	40%	50%	43%	43%
CTOPP Phonological Memory $\chi^2(3, N = 27) = 3.43, p = .33$	<i>M</i>	7	16	9	28
	<i>SD</i>	6	15	6	26
	Range	1–16	1–37	1–16	1–75
	% at or below 10 th percentile	80%	50%	57%	29%
CTOPP Phonological Awareness $\chi^2(3, N = 27) = 4.60, p = .20$	<i>M</i>	7	13	5	6
	<i>SD</i>	10	10	4	7
	Range	1–25	3–30	1–12	2–21
	% at or below 10 th percentile	80%	50%	86%	86%
GFTA $\chi^2(3, N = 27) = 2.06, p = .56$	<i>M</i>	36	22	30	36
	<i>SD</i>	22	18	24	21
	Range	10–65	1–53	1–67	5–67
	% at or below 10 th percentile	20%	25%	14%	14%

Note. RIAS = Reynolds Intellectual Assessment Scale (Reynolds & Kamphaus, 2003); CELF = Clinical Evaluation of Language Fundamentals–Fourth Edition (Semel et al., 2003); DELV = Diagnostic Evaluation of Language Variation (Seymour, Roeper, de Villers, & de Villers, 2005); CREVT = Comprehensive Receptive and Expressive Vocabulary Test–Third Edition (Wallace & Hammill, 2013); CTOPP = Comprehensive Test of Phonological Processing–Second Edition (Wagner, Torgesen, Rashotte, & Pearson, 2013); GFTA = Goldman-Fristoe Test of Articulation–Second Edition (Goldman & Fristoe, 2000).

Supplemental Table S1b shows how the demographic characteristics of the children were distributed across the four intensities: 12, 24, 36, and 48 exposures. A likelihood ratio test is shown for each variable to determine whether the effect of intensity was significant. A significant effect of intensity was not obtained for any of the variables, suggesting that the groups did not differ significantly. However, the probability of an intensity effect of was low (.10 or less) for one characteristic: father’s education. Specifically, children of fathers who were college graduates tended to be clustered in the intensity 36 condition. In addition, children of fathers whose education was unknown were more prevalent in intensity 24 and 48. The impact of this potential asymmetry on the results is unknown given that minimal attention has been paid to the contribution of father’s education on vocabulary learning and/or treatment outcomes. Most of the literature has focused more on mother’s education, which was better matched across conditions. Moreover, analyses of the data from the current study suggest that father’s education did not significantly affect treatment response (see Supplemental Table S3b).

Supplemental Table S1b. Demographic characteristics of participants in each intensity condition on standardized clinical tests. Results of Likelihood Ratio Test are indicated for each variable.

Demographic characteristics	Intensity 12	Intensity 24	Intensity 36	Intensity 48
Gender: % male $\chi^2 (3, N = 27) = 3.87, p = .28$	20%	38%	71%	57%
Race: % White $\chi^2 (6, N = 27) = 7.74, p = .26$	80%	88%	100%	71%
Ethnicity: % Non-Hispanic $\chi^2 (6, N = 27) = 7.54, p = .27$	80%	50%	71%	100%
Parent Marital Status: % married $\chi^2 (6, N = 27) = 4.87, p = .56$	80%	63%	86%	57%
Mother’s education $\chi^2 (15, N = 27) = 12.48, p = .64$				
% College Graduates	20%	13%	43%	43%
% Partial College	20%	50%	43%	29%
% High School Graduates	20%	25%	14%	29%
Father’s education $\chi^2 (18, N = 27) = 26.12, p = .10$				
% College Graduates	0%	0%	29%	0%
% Partial College	20%	25%	29%	14%
% High School Graduates	40%	0%	29%	29%
% Unknown	0%	63%	14%	57%

Method: Treatment Materials

Because items were randomly assigned to treatment or a no-treatment control condition for each child, each word was used in treatment for 13 children. One set of words was used for a fourteenth child but this child did not complete the definition post-test. To examine item effects, the percentage of children (out of 13) who defined the target word correctly at post-test when the word received treatment was computed. These values are shown in the second column of Supplemental Table S2. These values were significantly correlated with those reported by Justice and colleagues (2005, Table 4, summed post-test scores) for the subset of 30 words in Justice's elaborated treatment condition, which is similar to the treatment form in the current study, $r(30) = 0.72$, $p < .001$, $r^2 = 0.52$. For the current study, the percentage of children (out of 13) who defined untreated control words correctly also was computed to use as a reference point in interpreting learning of the treated items. On average, 0% (range = 0–23%) of children defined untreated control words correctly. Treated words were then classified at post-test by those that were learned beyond the *range* of the control words (i.e., more than 0–23% of children learned the word). These words are considered to be the frequently learned words (17% of treated words) and are shaded in green in Supplemental Table S2. Treated words that were learned beyond the *average* of the control words (i.e., by more than 0% of children) also are noted and are shaded in blue in Supplemental Table S2 (53% of treated words). Finally, treated words that were never learned (30% of treated words) are shaded in orange in Supplemental Table S2.

The remaining columns of Supplemental Table S2 show the characteristics of the materials used to treat each target word. This includes the number of words in each treatment exposure as well as the average frequency of those words from the Storkel and Hoover (2010) online calculator based on word frequency in child corpora of spoken American English. Possible associations between the percentage of children who learned a word and exposure variables (shown in Supplemental Table S2) and word characteristics (phonotactic probability, neighborhood density) were explored through correlation analyses. However, no significant effects were obtained. The only potentially promising observation was that verbs were rarely present in the frequently learned words category: 8% frequently learned (green shading), 52% learned (blue shading), 40% never learned (orange shading). Likewise, the majority of nouns fell in the middle/learned category: 13% frequently learned (green shading), 69% learned (blue shading), 19% never learned (orange shading). Adjectives were relatively equally distributed across learning categories: 32% frequently learned (green shading), 42% learned (blue shading), 26% never learned (orange shading). Overall, these results should be viewed with caution because it is likely that the nature of this study may not lend itself to a thorough understanding of word or exposure effects. That is, some children received treatment in less effective conditions (e.g., intensity 12), limiting the number of words learned.

Supplemental Table S2. *Characteristics of words learned during treatment based on correct definition scores at post-treatment testing.*

Target Word	%Children Defined Word Correctly at Post	Classification of Word at Post	Book Title	Part of Speech	Frequency of Word	# of Words in Definition	Average Frequency of Words in Definition	# of Words in Synonym	Frequency of Synonym	# of Words in Context Sentence	Average Frequency of Words in Context Sentence	# of Words in Book Text Sentence	Average Frequency of Words in Book Text Sentence
pouted	46%	Above Control Range (0-23%)	Book! Book! Book!	verb	0	8	1952	1	0	9	313	3	60
haddock	38%	Above Control Range (0-23%)	The Bear Under the Stairs	noun	0	8	134	1	217	11	1501	7	20
horrified	38%	Above Control Range (0-23%)	The Caterpillar That Roared	adjective	0	5	107	1	7	11	1563	3	0
tight	38%	Above Control Range (0-23%)	The Bear Under the Stairs	adjective	58	5	275	1	94	13	232	20	731
crept	31%	Above Control Range (0-23%)	The Bear Under the Stairs	verb	1	6	143	1	4	11	1725	5	1311
flashing	31%	Above Control Range (0-23%)	What Do You Do With a Kangaroo	adjective	12	8	446	1	1	9	990	27	1075
gulp	31%	Above Control Range (0-23%)	Swimmy	noun	0	9	779	1	8	13	1779	10	2954
invisible	31%	Above Control Range (0-23%)	Swimmy	adjective	18	5	295	1	125	15	1307	7	123
overjoyed	31%	Above Control Range (0-23%)	Imogene's Antlers	adjective	0	5	321	1	10	10	1805	16	2088
worn	31%	Above Control Range (0-23%)	What Do You Do With a Kangaroo	adjective	8	7	173	1	9	9	240	22	543

Target Word	%Children Defined Word Correctly at Post	Classification of Word at Post	Book Title	Part of Speech	Frequency of Word	# of Words in Definition	Average Frequency of Words in Definition	# of Words in Synonym	Frequency of Synonym	# of Words in Context Sentence	Average Frequency of Words in Context Sentence	# of Words in Book Text Sentence	Average Frequency of Words in Book Text Sentence
damp	23%	Above Mean Control (0%)	Harry and the Terrible Whatzit	adjective	8	4	41	1	1	9	494	12	486
glared	23%	Above Mean Control (0%)	Imogene's Antlers	verb	0	8	2049	1	1	12	669	6	195
hooves	23%	Above Mean Control (0%)	Otis	noun	0	11	202	1	189	12	1646	20	807
swaying	23%	Above Mean Control (0%)	Swimmy	verb	0	8	547	1	7	10	175	13	178
clamor	15%	Above Mean Control (0%)	Possum and the Peeper	noun	0	9	1666	1	107	11	579	10	985
furnace	15%	Above Mean Control (0%)	Harry and the Terrible Whatzit	noun	13	8	1304	1	42	11	270	6	72
gathered	15%	Above Mean Control (0%)	Book! Book! Book!	verb	2	7	1983	1	3	12	3105	21	376
midday	15%	Above Mean Control (0%)	Swimmy	noun	0	6	375	1	204	11	1295	19	593
racket	15%	Above Mean Control (0%)	Possum and the Peeper	noun	0	6	52	1	8	12	274	16	1063
ripe	15%	Above Mean Control (0%)	Otis	adjective	9	8	537	1	529	9	1638	24	257
sidelines	15%	Above Mean Control (0%)	Otis	noun	0	10	49	2	107	14	196	14	113
spotless	15%	Above Mean Control (0%)	Otis	adjective	0	5	471	1	624	11	507	9	315
squawked	15%	Above Mean Control (0%)	Book! Book! Book!	verb	0	5	370	1	10	13	1624	13	358
swat	15%	Above Mean Control (0%)	Harry and the Terrible Whatzit	verb	6	10	396	1	91	12	1845	9	205
swift	15%	Above Mean Control (0%)	Swimmy	adjective	0	5	192	1	316	12	358	16	827
tailor	15%	Above Mean Control (0%)	What Do You Do With a Kangaroo	noun	0	7	244	1	2	10	795	16	1030
awful	8%	Above Mean Control (0%)	The Bear Under the Stairs	adjective	26	5	772	1	354	15	438	6	855
gloomy	8%	Above Mean Control (0%)	Harry and the Terrible Whatzit	adjective	0	4	2047	1	254	10	125	6	553
grumbling	8%	Above Mean Control (0%)	Possum and the Peeper	verb	0	5	20	1	0	11	2859	14	452
heaved	8%	Above Mean Control (0%)	Book! Book! Book!	verb	0	6	1529	1	12	10	621	5	1
marsh	8%	Above Mean Control (0%)	Possum and the Peeper	noun	1	9	90	1	1	12	1530	15	1863
marvel	8%	Above Mean Control (0%)	Swimmy	noun	0	9	259	1	86	14	385	19	236
nervous	8%	Above Mean Control (0%)	Shy Charles	adjective	8	5	80	1	1	13	716	7	2085
noticed	8%	Above Mean Control (0%)	The Bear Under the Stairs	verb	2	7	1848	1	5241	9	241	11	143
ripples	8%	Above Mean Control (0%)	The Caterpillar That Roared	noun	0	5	1167	1	8	9	877	10	996
ruffle	8%	Above Mean Control (0%)	Book! Book! Book!	verb	0	10	1594	1	8	10	1349	11	68
scarlet	8%	Above Mean Control (0%)	Shy Charles	noun	9	5	343	1	365	11	289	7	25
snuggled	8%	Above Mean Control (0%)	The Caterpillar That Roared	verb	0	9	543	1	1	11	1226	15	729
stale	8%	Above Mean Control (0%)	What Do You Do With a Kangaroo	adjective	9	6	4207	1	513	13	525	16	952
swung	8%	Above Mean Control (0%)	Harry and the Terrible Whatzit	verb	2	8	551	1	2	10	223	5	390
wandered	8%	Above Mean Control (0%)	Imogene's Antlers	verb	1	7	848	2	58	12	353	4	79
whinnied	8%	Above Mean Control (0%)	Book! Book! Book!	verb	0	9	1613	1	0	10	774	5	33

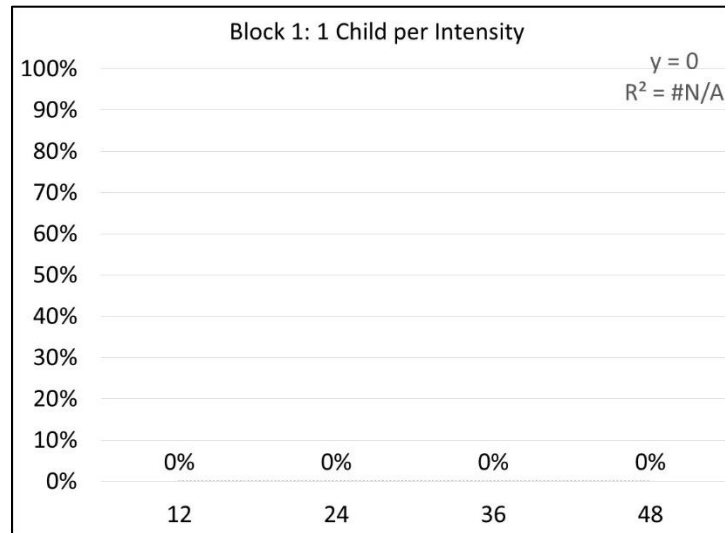
Target Word	% Children Defined Word Correctly at Post	Classification of Word at Post	Book Title	Part of Speech	Frequency of Word	# of Words in Definition	Average Frequency of Words in Definition	# of Words in Synonym	Frequency of Synonym	# of Words in Context Sentence	Average Frequency of Words in Context Sentence	# of Words in Book Text Sentence	Average Frequency of Words in Book Text Sentence
advice	0%	Never Learned (0%)	Imogene's Antlers	noun	0	9	1296	1	624	12	1575	6	1621
decided	0%	Never Learned (0%)	The Bear Under the Stairs	verb	13	5	4799	1	4	9	626	14	262
discovered	0%	Never Learned (0%)	Harry and the Terrible Whatzit	verb	0	7	771	1	2	11	266	15	1155
embarrassed	0%	Never Learned (0%)	Shy Charles	adjective	1	6	29	1	6	13	778	6	4100
frayed	0%	Never Learned (0%)	What Do You Do With a Kangaroo	adjective	0	7	980	1	22	11	547	21	33
gaze	0%	Never Learned (0%)	The Caterpillar That Roared	verb	0	8	2240	1	1	14	598	22	1841
hauled	0%	Never Learned (0%)	Otis	verb	58	5	101	1	8	12	2077	24	167
murmured	0%	Never Learned (0%)	Shy Charles	verb	0	8	89	1	1	11	363	10	1287
peering	0%	Never Learned (0%)	Possum and the Peeper	verb	0	8	1313	1	16	11	1550	15	194
prodded	0%	Never Learned (0%)	Imogene's Antlers	verb	0	6	334	1	82	11	216	9	28
rare	0%	Never Learned (0%)	Imogene's Antlers	adjective	1	5	2758	2	4987	11	1744	20	736
silky	0%	Never Learned (0%)	Otis	adjective	9	8	1340	1	23	11	1272	9	77
smooth	0%	Never Learned (0%)	What Do You Do With a Kangaroo	adjective	23	4	159	1	411	10	530	22	1427
squinting	0%	Never Learned (0%)	Possum and the Peeper	verb	0	7	427	1	9	11	464	12	86
success	0%	Never Learned (0%)	Shy Charles	noun	0	9	1000	1	0	12	186	5	6324
surface	0%	Never Learned (0%)	The Caterpillar That Roared	noun	0	5	659	1	440	12	548	10	70
trembled	0%	Never Learned (0%)	Shy Charles	verb	0	9	143	1	1	10	509	6	4
twitch	0%	Never Learned (0%)	The Caterpillar That Roared	verb	0	5	91	1	16	11	1217	23	598

Results: Treatment Intensity—Definition task: Percentage of children responding to treatment

Supplemental Figure S1 shows the percentage of children who responded to treatment based on the definition task (i.e., post-treatment score of 5 or higher for treated words), for each intensity as blocks accumulated. Block-by-block data were used to determine when the pattern of the effect of intensity had stabilized. Once a stable pattern has been reached, data collection can be discontinued and the adequate intensity can be identified. Supplemental Figure S1 notes the pattern for each block and provides a detailed description of each pattern.

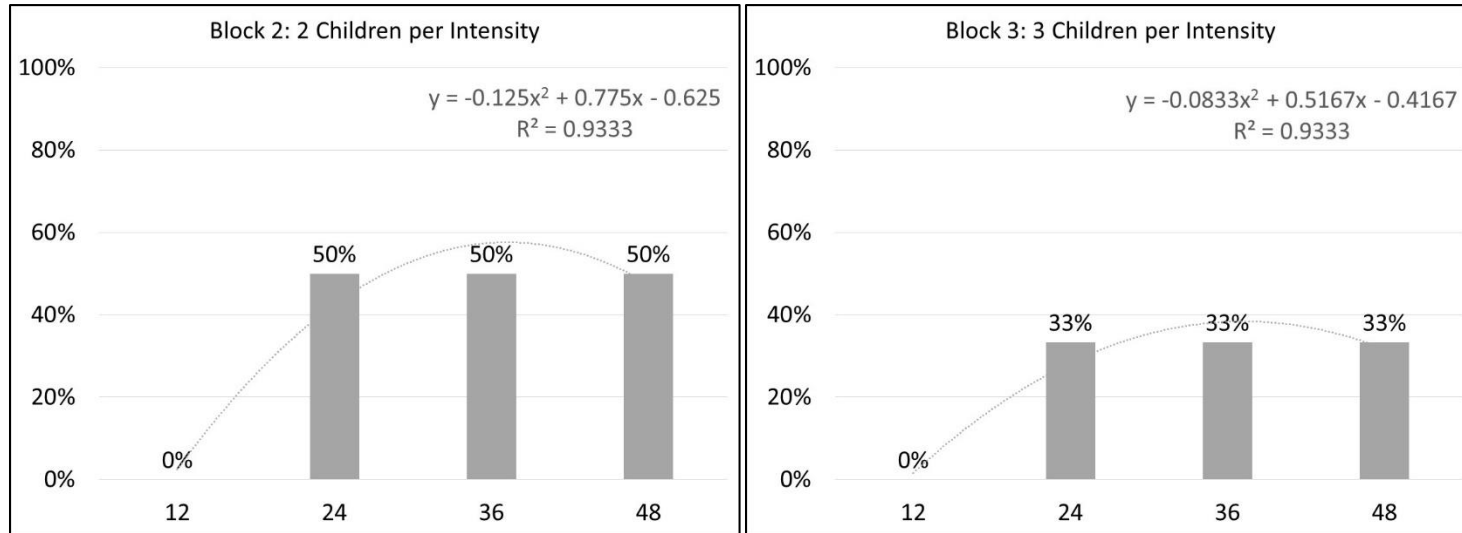
Supplemental Figure S1. Percentage of children responding to the treatment (i.e., post-treatment definition score of 5 or higher for treated words) for each treatment intensity condition (12, 24, 36, 48) by enrollment block. The trendline illustrates the polynomial trend also depicted by the regression equation noted in each panel.

Pattern 1: No Treatment Response



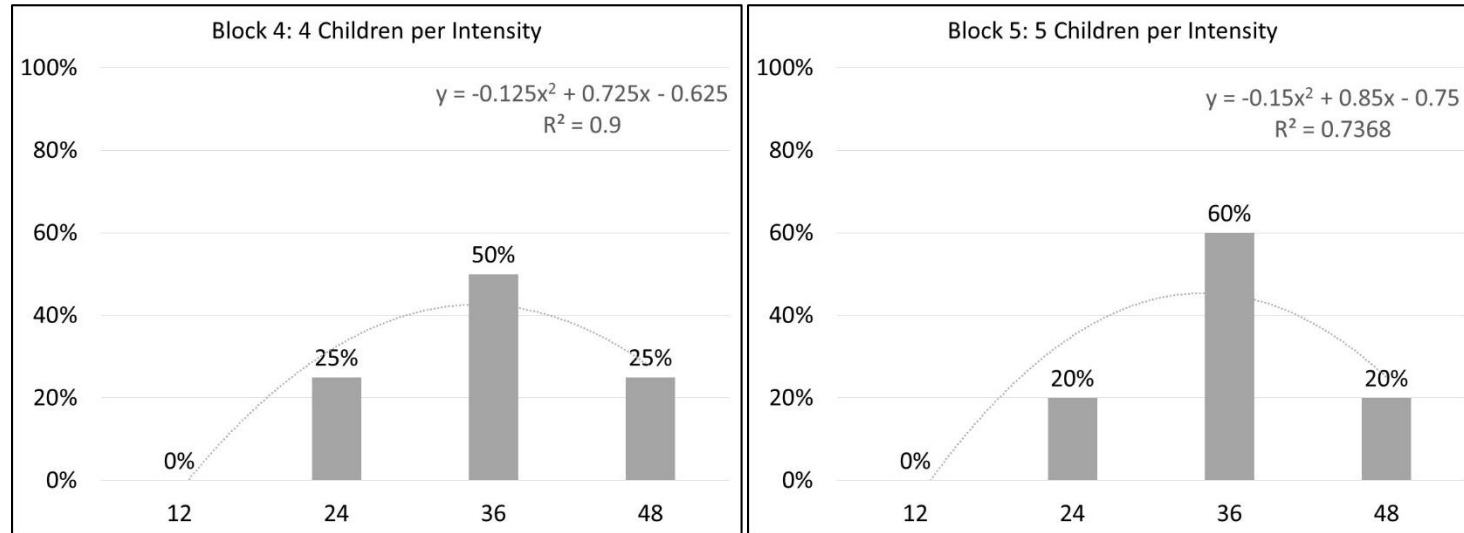
As shown above, no child in any treatment intensity condition showed a treatment response in the first block.

Pattern 2: Better Treatment Response in Intensities 24, 36, 48 Than in Intensity 12

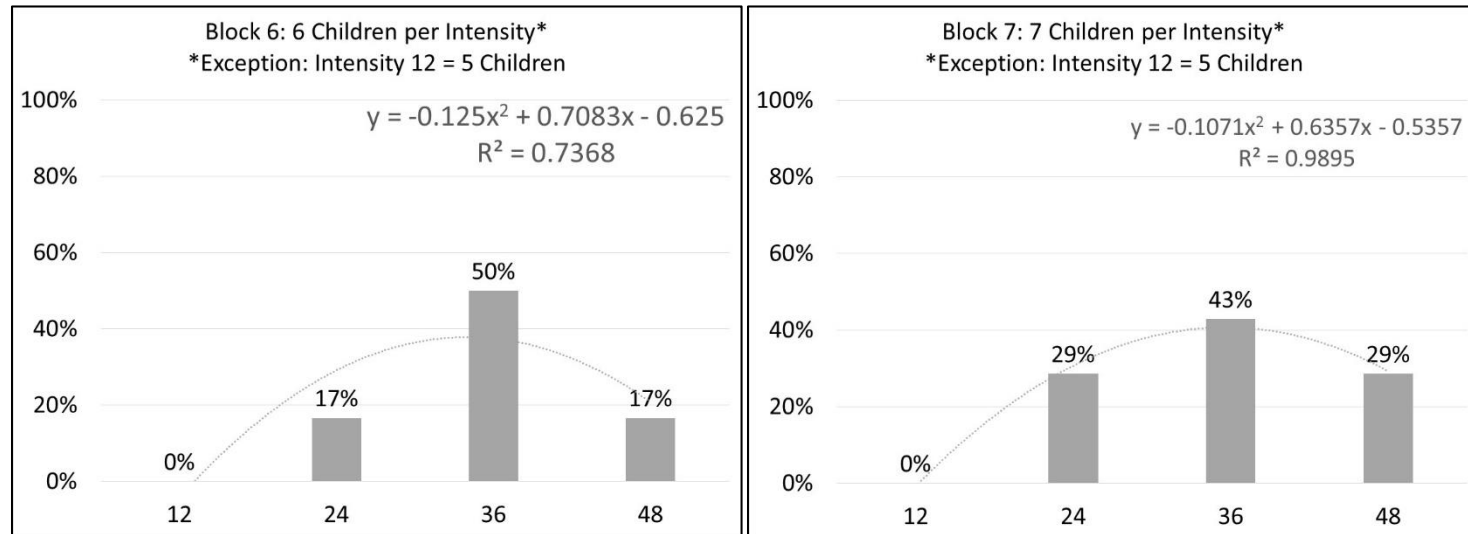


As shown above, a different pattern emerged at block 2 and was maintained in block 3. Specifically, within each of these blocks, no children showed a treatment response in intensity 12, but children in intensities 24, 36, and 48 did show a treatment response. Thus, intensity 12 appeared less effective than intensities 24, 36, and 48, which were not differentiated from one another within block 2 or block 3.

Pattern 3/Final Pattern: Treatment Response Increases From Intensity 12 to 24 to 36 and Then Decreases for Intensity 48.



A different pattern emerged within block 4 whereby the number of children responding to treatment increased as intensity increased but then plateaued. Specifically, within block 4, the percent of children responding to treatment was 0% at intensity 12 but increased to 25% of children as intensity increased to 24 exposures and further increased to 50% of children as intensity increased to 36 exposures. Then, the percent of children responding to treatment decreased to 25% as intensity increased to the maximum of 48 exposures. This is a desirable pattern in an escalation design because it indicates that the benefit of merely increasing the intensity of the treatment has potentially plateaued. That is, further increases in intensity do not lead to an increase in the percent of children responding to the treatment. Thus, a plateau is indicative that the promising or adequate intensity has been identified. This pattern is replicated in block 5 with intensity 36 showing greatest response to treatment.



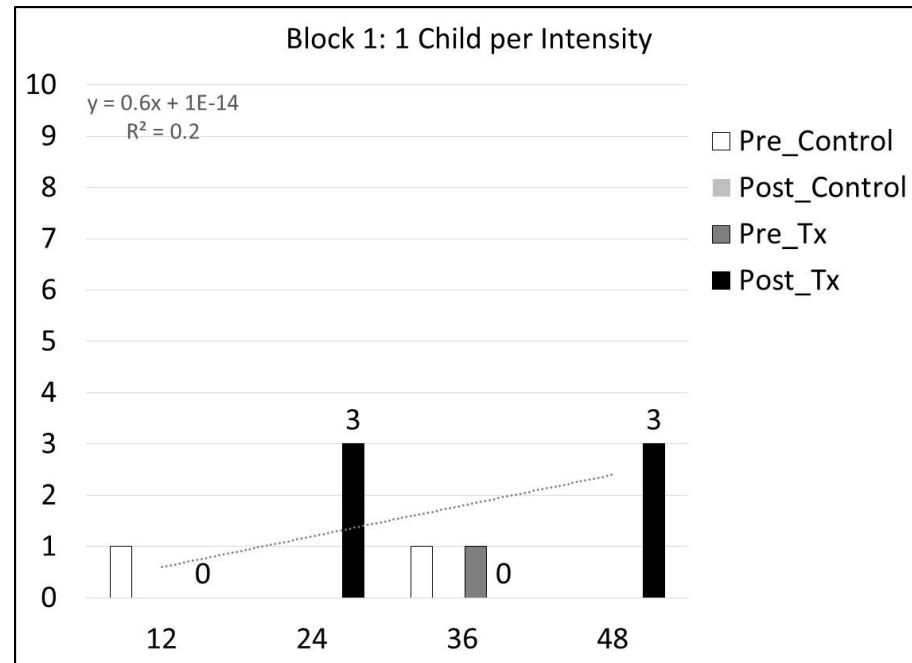
Following block 5, no additional children were randomized to intensity 12 because 0% of children responded to the treatment, indicating that intensity 12 was not sufficient for word learning by children with specific language impairment. Block 6 replicated the pattern observed in blocks 4 and 5. Data collection likely could have been discontinued after this block because the pattern observed was replicated across three blocks; however, block 7 was already in progress so it was completed. Block 7 showed the same pattern as the immediately prior blocks. Therefore, data collection for all conditions was discontinued. The final pattern was that the percentage of children responding to treatment increased as intensity increased from 12 to 24 to 36 exposures, but then the percentage of children responding to treatment decreased as intensity increased from 36 to 48 exposures, indicating that additional exposures beyond 36 did not improve treatment responding. Thus, the primary analysis identified 36 exposures as the adequate intensity out of the four intensities tested.

Results: Treatment Intensity—Definition task: Number of words defined correctly

Similar to Supplemental Figure S1, Supplemental Figure S2 provides block-by-block data for each treatment intensity for the definition task. However, rather than showing the percentage of children who responded to the treatment, Supplemental Figure S2 shows the number of words with correct definitions for control words at pre-treatment (open bars) and post-treatment (lightly shaded bars) as well as treated words at pre-treatment (medium shaded bars) and post-treatment (black bars). The average number of treated words correct at post-treatment also is noted at the top of the black bars. The goal of this analysis was to provide converging evidence with the primary analysis to ensure that a more detailed look at the actual number of treated words learned would identify the same intensity as adequate or promising.

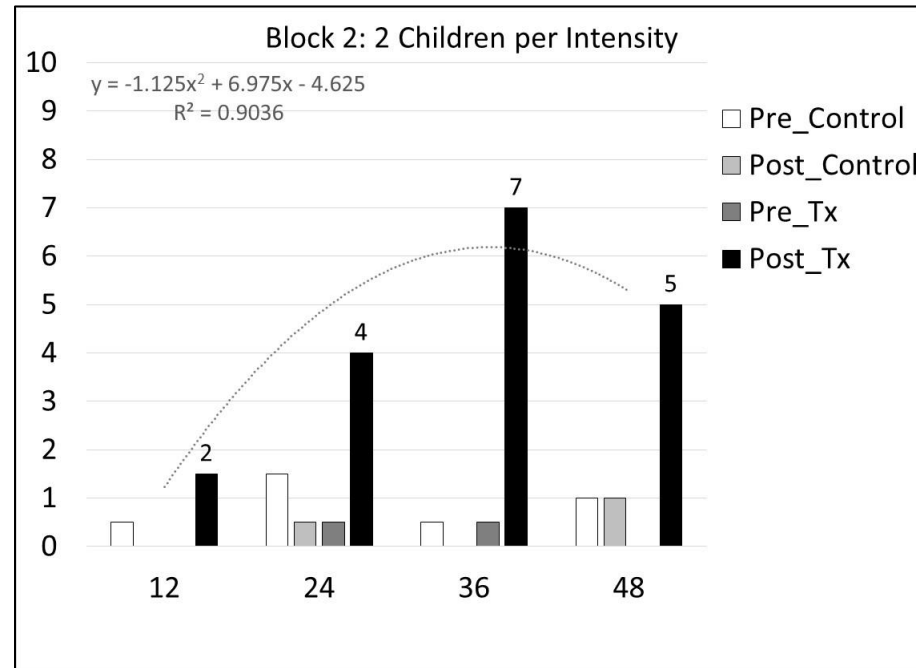
Supplemental Figure S2. The average number of words with correct definitions for each treatment intensity (12, 24, 36, 48) by enrollment block. Pre-treatment control words are shown with 0% shaded bars. Post-treatment control words are shown with 25% shaded bars. Pre-treatment treated words are shown with 50% shaded bars. Post-treatment treated words are shown with 100% shaded bars. The trendline illustrates the polynomial trend for post-treatment treated words, also depicted by the regression equation noted in each panel.

Pattern 1: No clear pattern



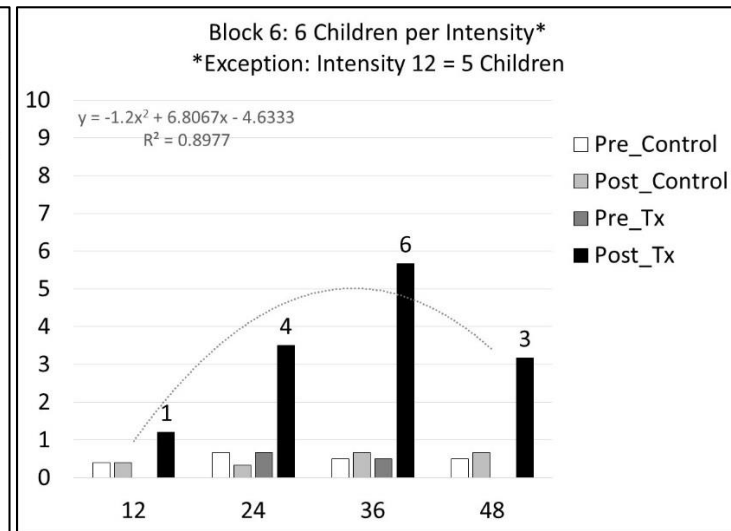
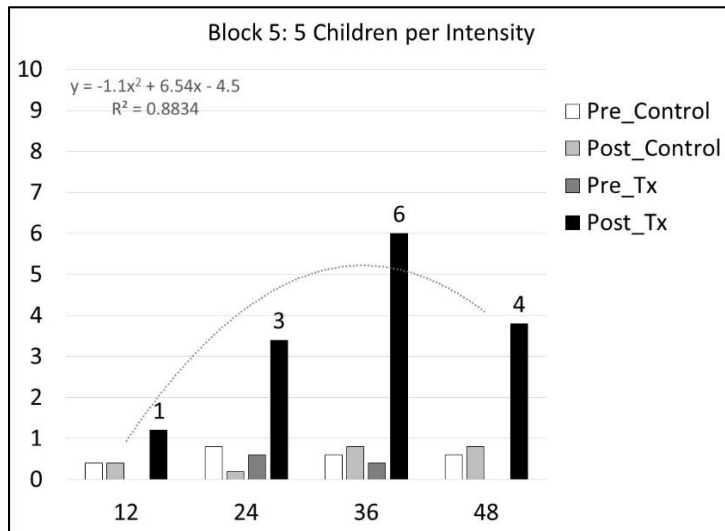
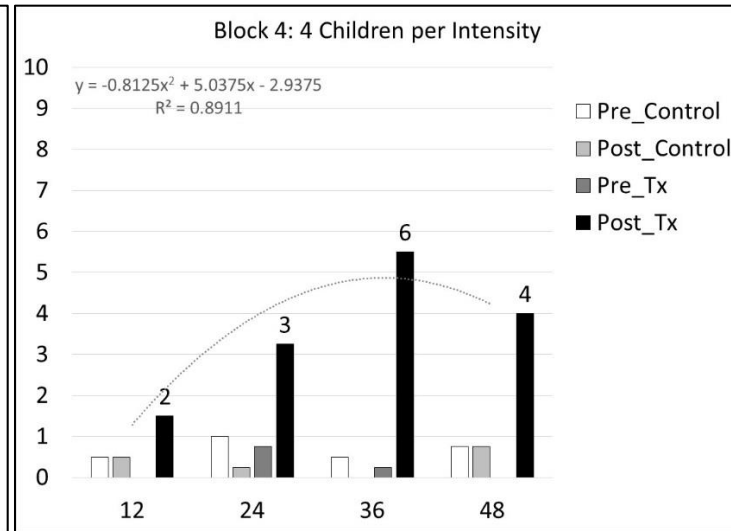
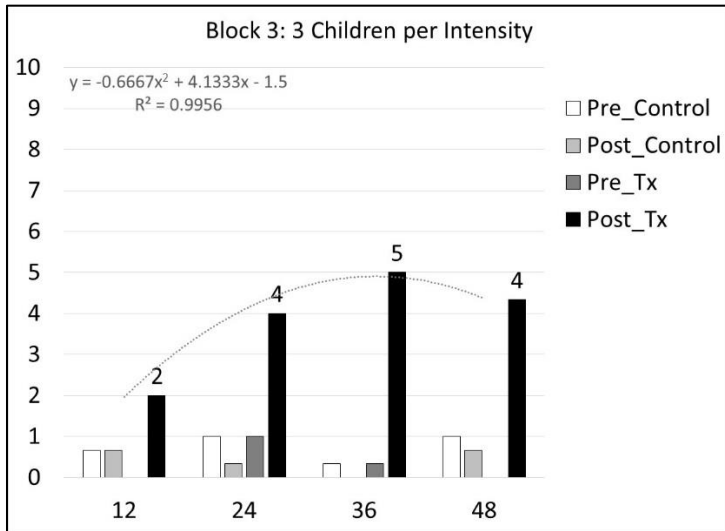
No clear pattern emerged in block 1. Children in intensities 12 and 36 did not define any words correctly at the post-treatment definition test. In contrast, children in intensities 24 and 48 defined 3 treated words correctly at the post-treatment definition test.

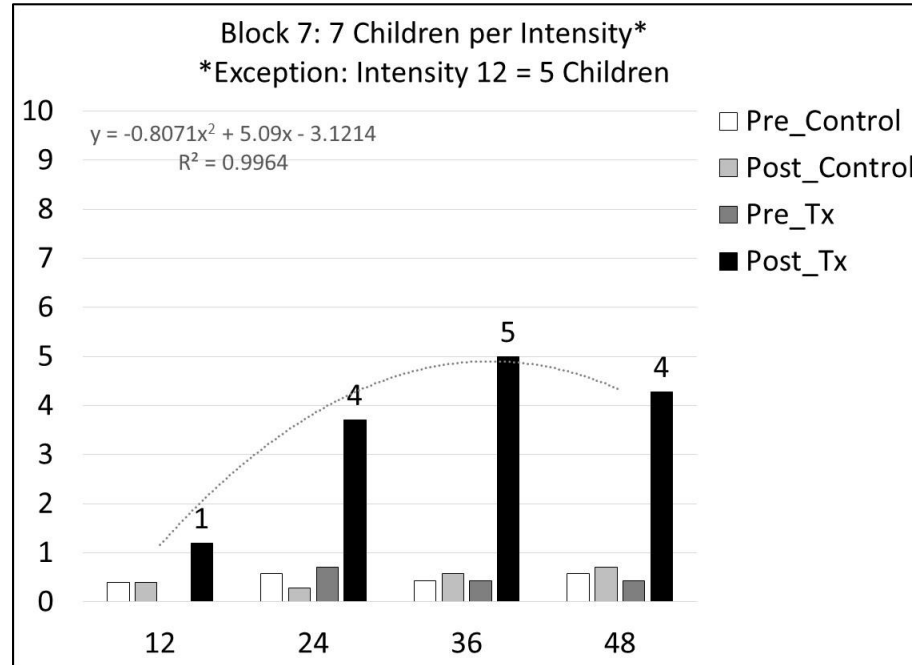
Pattern 2/Final Pattern: Number of treated words learned by post-treatment increases from intensity 12 to 24 to 36 and then decreases for intensity 48.



In the second block, a clear pattern emerged for the treated words. Children in intensity 12 defined 2 treated words correctly post-treatment. The number of words defined correctly increased from 12 to 24 exposures. Children in intensity 24 defined 4 treated words correctly post-treatment. Further increases in the number of treated words defined correctly are observed as intensity increased from 24 to 36. Children in intensity 36 defined 7 treated words correctly post-treatment. The number of treated words defined correctly then reduces as intensity increases from 36 to 48 exposures. Children in intensity 48 defined 5 treated words correctly post-treatment. Note that the number of treated words defined correctly post-treatment in intensities 24, 36, and 48 is appreciably higher than the number of control words defined correctly pre-treatment and post-treatment and higher than the number of treated words defined correctly pre-treatment.

This pattern of the number of treated words defined correctly post-treatment increasing from 12 to 24 to 36 exposures and then decreasing as exposures increased to 48 was observed in all remaining blocks, as shown below. This pattern also mirrors the pattern observed in Figure S1. Thus, the percentage of children showing a response to treatment and the number of treated words defined correctly post-treatment converge on the conclusion that intensity 36 is the most promising of the four intensities tested.



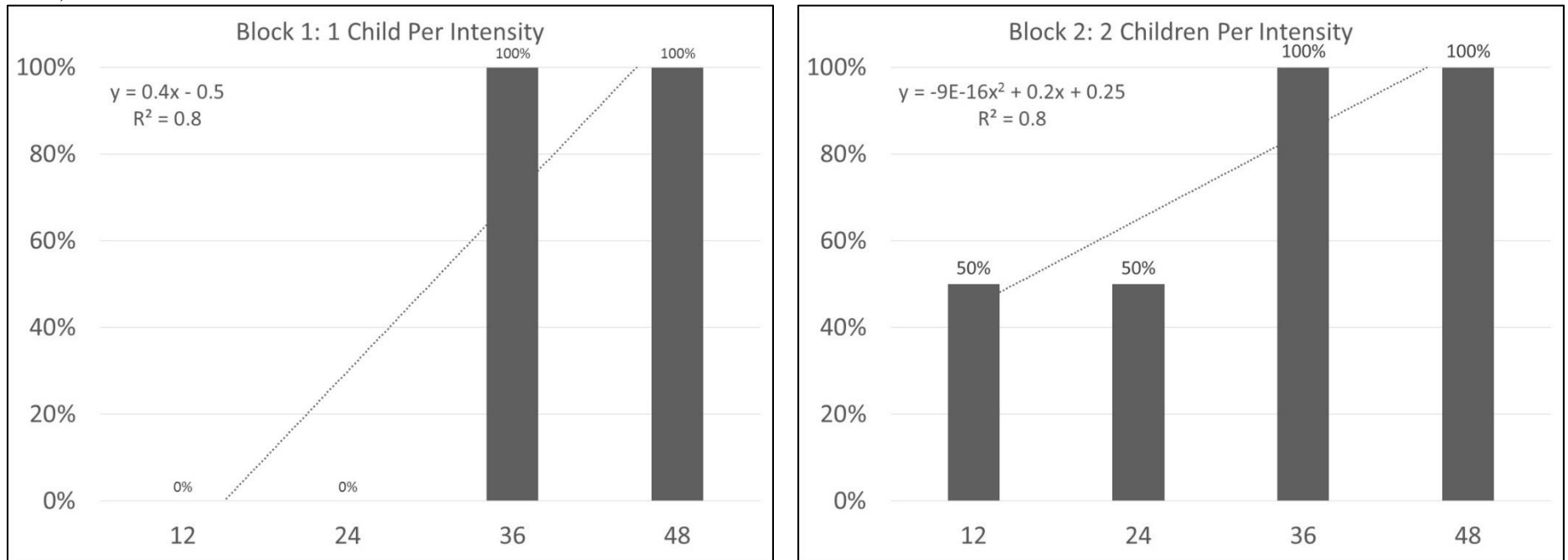


Results: Treatment Intensity—Naming task (final): Percentage of children responding to treatment

Similar to Supplemental Figures S1 and S2, Supplemental Figure S3 provides block-by-block data for each treatment intensity. However, rather than showing data from the definition task, Supplemental Figure S3 shows data from the naming task administered at the last treatment session. Data from the naming task are examined to once again provide converging evidence of the adequate treatment intensity. Supplemental Figure S3 shows the percent of children responding to the treatment based on a post-treatment naming score of 4 or higher for treated words.

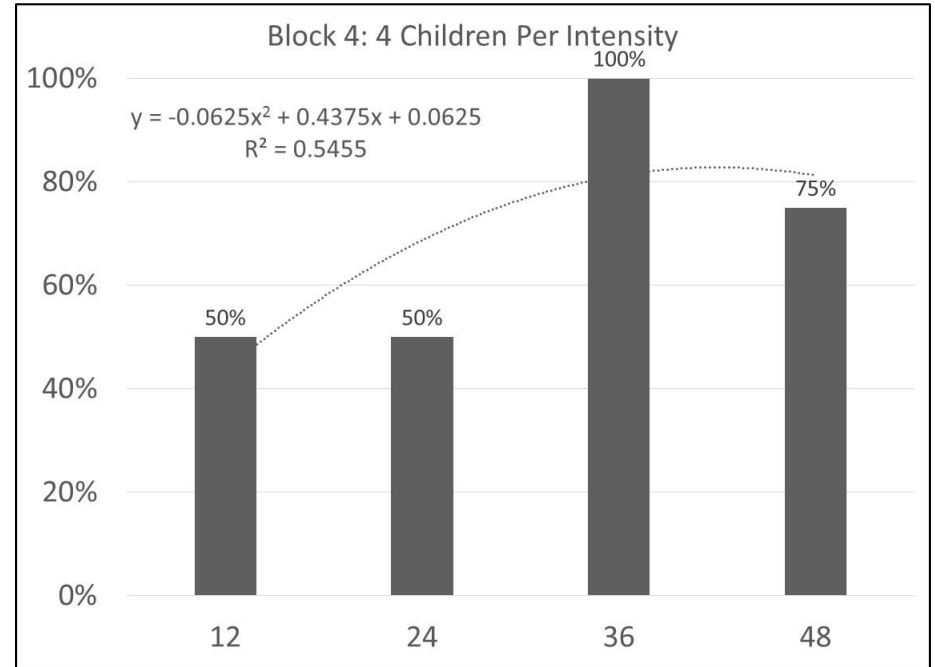
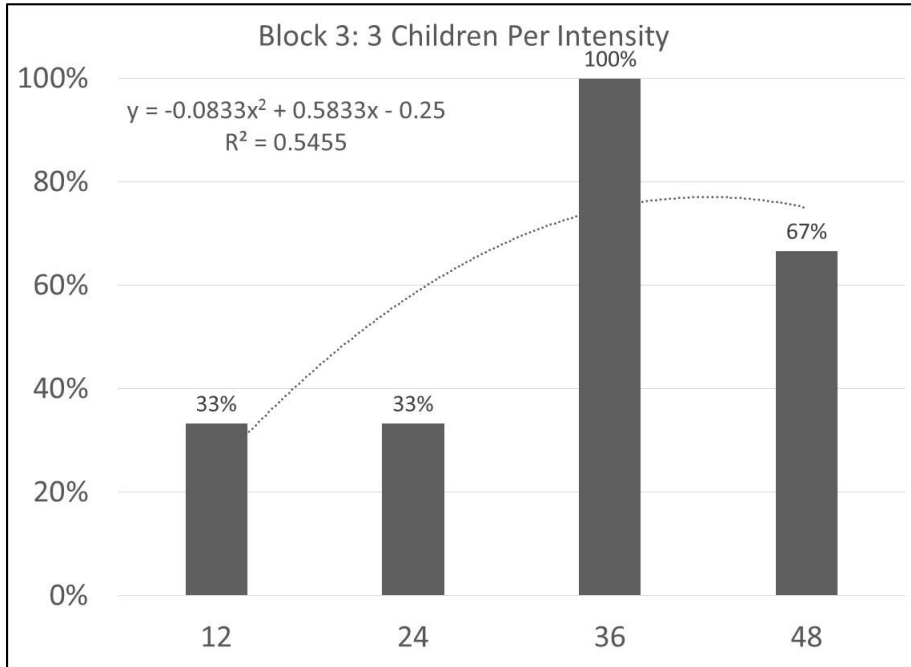
Supplemental Figure S3. Percent of children responding to the treatment based on naming data (i.e., post-treatment naming score of 4 or higher for treated words) for each treatment intensity (12, 24, 36, 48) by enrollment block. The trendline illustrates the polynomial trend also depicted by the regression equation noted in each panel.

Pattern 1: Percent of children responding to intensity 12 & 24 is similar but increases as intensity increases to 36 & 48 (which are similar to each other)

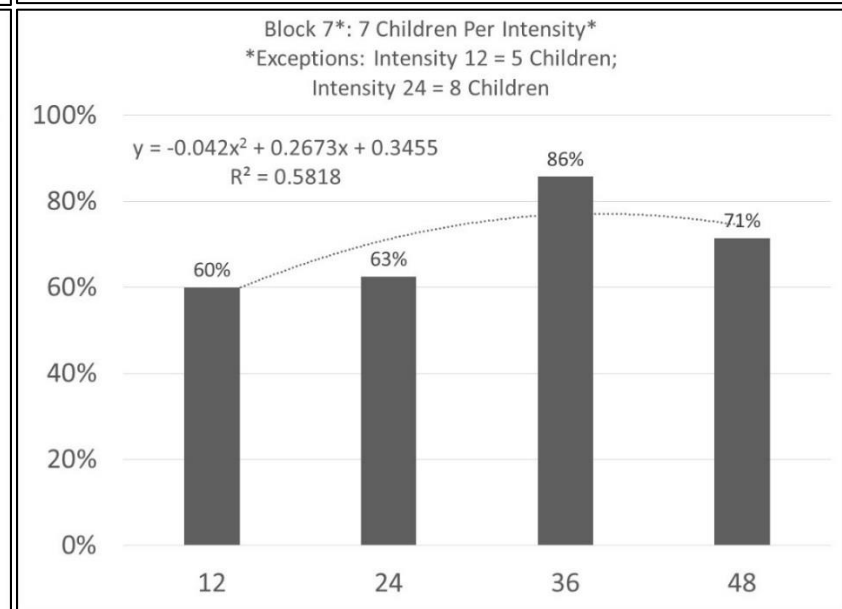
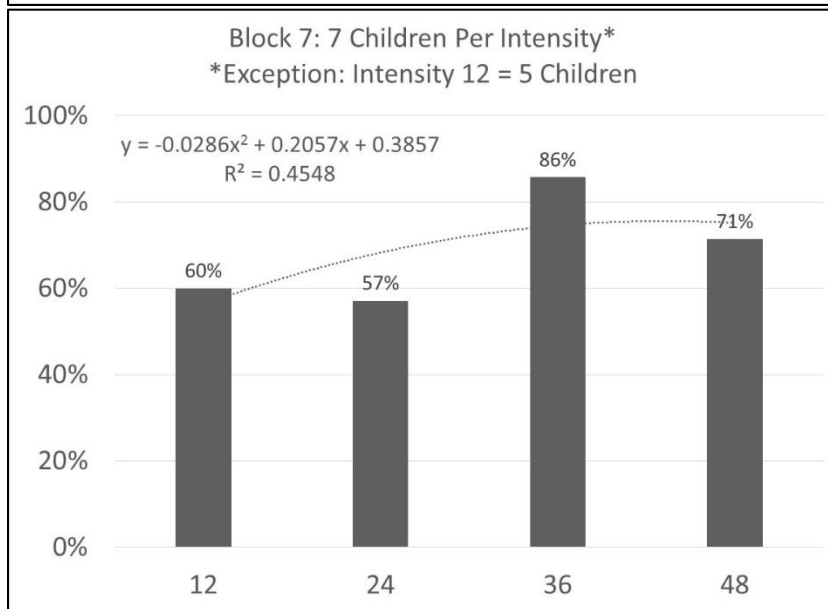
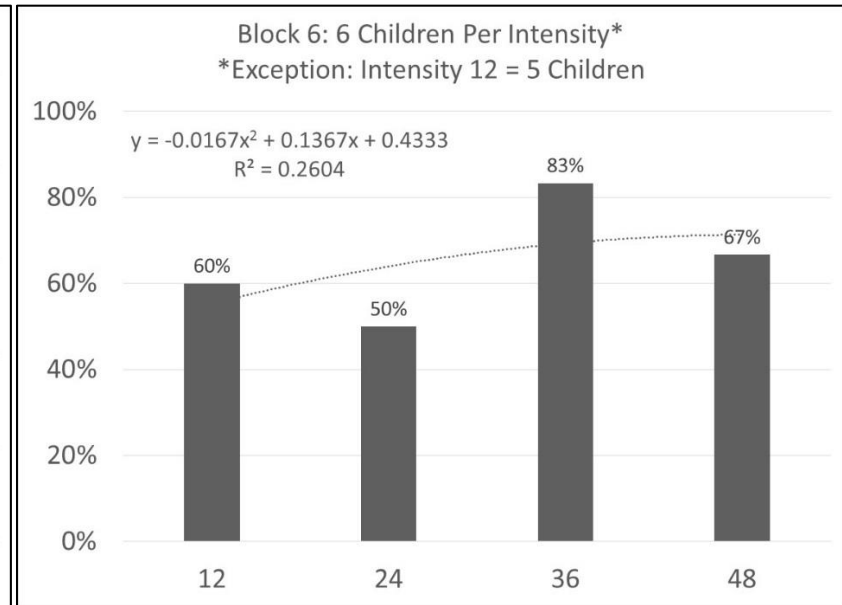
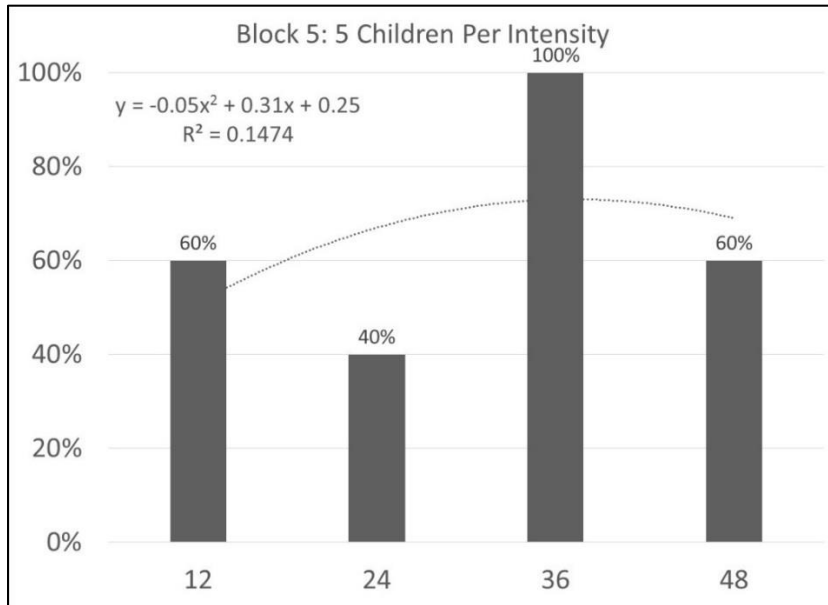


In block 1, none of the children in intensity 12 or 24 responded to treatment. In contrast, the children in intensity 36 and 48 responded to treatment based on naming data. This pattern was replicated in block 2 with fewer children responding to treatment in intensity 12 and 24 and more children responding to treatment in intensity 36 and 48. Thus, within the first two blocks there was evidence that 36 and 48 exposures was more beneficial than 12 and 24 exposures.

Pattern 2/Final Pattern: Percent of children responding to intensity 12 & 24 is similar but increases as intensity increases to 36 and then decreases at intensity 48.



In block 3, 33% of children responded to treatment in intensity 12. As intensity increased to 24 exposures, response to treatment remained the same with 33% of children responding to treatment. In contrast, as intensity increased from 24 to 36 exposures, the percent of children responding to treatment increased from 33% to 100%. As intensity increased from 36 to 48 exposures, there was no further increase in the percent of children responding to treatment. In fact, the percent of children responding to treatment fell from 100% to 67%. In general, this pattern was replicated across all remaining blocks. Generally, the percent of children responding to treatment was relatively similar across 12 and 24 exposures and then increased from 24 to 36 exposures. The percent of children responding to treatment then decreased as exposures increased further from 36 to 48 exposures. This final pattern for the naming data converges with the patterns observed for the definition data.



Results: Variability in Response to Treatment—Pretreatment characteristics associated with post-treatment definition scores

To examine potential pre-treatment predictors of treatment response, correlations between pre-treatment measures (i.e., age and all standardized test scores), early treatment performance (i.e., number of words named correctly at the first naming test) and treatment outcome (i.e., number of treated words defined correctly post-treatment) were examined for the 21 children in the intensities that produced a treatment response (i.e., intensities 24, 36, and 48). The full set of correlations is shown in Table S3a with the first column of correlations addressing the research question. The remaining columns show correlations among measures. Significant and marginally significant correlations are marked. Significant effects are described in the article. Two measures showed promising but non-significant correlations: the overall CTOPP phonological memory score [$r(21) = 0.42, p = .06, r^2 = 0.18$] and the treated word score at the first naming test during treatment [$r(21) = 0.37, p = .10, r^2 = 0.14$]. Here, children with higher phonological memory scores or greater treated words named correctly early in treatment tended to achieve higher scores on the post-treatment definition test than children with lower phonological memory scores or fewer treated words named correctly early in treatment.

Supplemental Table S3a. Correlations among measures.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. Treated Word Score on Post-Treatment Definition Test	1.00															
2. Age (in months)	.14	1.00														
3. RIAS Nonverbal IQ	.11	-.29	1.00													
4. CELF Core Language	.25	-.06	-.16	1.00												
5. Vocabulary: DELV Semantic	.52*	.11	.03	.50*	1.00											
6. Vocabulary: CELF Word Classes	-.14	-.34	.40~	.17	.05	1.00										
7. Vocabulary: CREVT Expressive	.17	.23	.12	.23	.01	-.16	1.00									

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
8. CELF Concepts & Following Directions	.12	-.02	.09	.76**	.42~	.31	.22	1.00								
9. CELF Word Structure	.27	.09	-.17	.72**	.03	-.12	.28	.33	1.00							
10. CELF Recalling Sentences	.33	0.00	-.03	.69**	.76**	.35	.15	.69**	.10	1.00						
11. CELF Formulating Sentences	.09	-.15	-.36	.73**	.28	-.15	.07	.19	.73**	.21	1.00					
12. CELF Understanding Spoken Paragraphs	.24	-.33	.26	.40~	.50*	.03	.04	.30	.17	.45*	.23	1.00				
13. CTOPP Nonword Repetition	.44*	.43~	.20	.18	.15	-.09	.47*	.42~	.15	.39~	-.30	.19	1.00			
14. CTOPP Phonological Memory	.42~	.52*	.12	.42~	.48*	.02	.41~	.59**	.14	.68**	-.10	.16	.84*	1.00		
15. CTOPP Phonological Awareness	.48*	.38~	-.09	.26	.16	.10	.19	.30	.25	.22	-.02	.05	.38~	.36	1.00	
16. GFTA	.34	.37~	-.08	.31	.02	-.02	.57**	.25	.44*	.19	.10	-.16	.59*	.54*	.21	1.00
17. Treated Word Score at First Naming Test	.37~	.25	-.10	.47*	.47*	-.36	.44*	.26	.44*	.30	.40~	.13	.22	.38~	.09	.41~

Note. RIAS = Reynolds Intellectual Assessment Scale (Reynolds & Kamphaus, 2003); CELF = Clinical Evaluation of Language Fundamentals–Fourth Edition (Semel et al., 2003); DELV = Diagnostic Evaluation of Language Variation (Seymour, Roeper, de Villers, & de Villers, 2005); CREVT = Comprehensive Receptive and Expressive Vocabulary Test–Third Edition (Wallace & Hammill, 2013); CTOPP = Comprehensive Test of Phonological Processing–Second Edition (Wagner, Torgesen, Rashotte, & Pearson, 2013); GFTA = Goldman-Fristoe Test of Articulation–Second Edition (Goldman & Fristoe, 2000).

** $p < .01$, * $p < .05$, ~ $.05 \leq p \leq .10$. Standard scores were used for items 3–16.

The relationship between treatment response versus non-response also was examined for demographic characteristics for the 21 children in conditions that yielded some treatment responding (i.e., intensities 24, 36, and 48). The first column of Table 3b shows the overall characteristics of this sample of 21 children. The second column shows the percentage of children who responded to treatment that had the majority characteristic of the subsample. For example, 71% of the children who responded to treatment (i.e., 5 of 7) were boys. The third column shows the percentage of children who did not respond to treatment that had the majority characteristic. For example, 50% of children who did not respond to treatment (i.e., 7 of 14) were boys. If the demographic characteristic is not associated with treatment response, then we would expect each of the subgroups (i.e., children who responded to treatment vs. children who did not respond to treatment) to show a similar percentage of the characteristic as the overall sample (e.g., approximately 57% of children in each group should be boys). This was tested statistically using a chi square test, which is shown in the first column of Supplemental Table S3b. For example, for gender, the chi square test indicates that the association between gender and presence versus absence of a treatment response is not significant.

Supplemental Table S3b. Demographic characteristics of children who responded to treatment compared to children who did not respond to treatment.

Demographic characteristics	Children who responded to treatment (<i>n</i> = 7)	Children who did not respond to treatment (<i>n</i> = 14)
Gender: $\chi^2(1, N = 21) = 0.90, p = .34$		
Boys (<i>n</i> = 12, 57%)	71% (<i>n</i> = 5)	50% (<i>n</i> = 7)
Race: $\chi^2(2, N = 21) = 5.46, p = .07$		
White (<i>n</i> = 18, 86%)	71% (<i>n</i> = 5)	93% (<i>n</i> = 13)
Ethnicity: $\chi^2(2, N = 21) = 4.84, p = .09$		
Non-Hispanic (<i>n</i> = 15, 71%)	57% (<i>n</i> = 4)	79% (<i>n</i> = 11)
Parent marital status: $\chi^2(2, N = 21) = 4.84, p = .09$		
Married (<i>n</i> = 15, 71%)	57% (<i>n</i> = 4)	79% (<i>n</i> = 11)
Mother’s education: $\chi^2(3, N = 21) = 6.08, p = .11$		
% College graduates (<i>n</i> = 7, 33%)	43% (<i>n</i> = 3)	29% (<i>n</i> = 4)
% Partial college (<i>n</i> = 8, 38%)	57% (<i>n</i> = 4)	29% (<i>n</i> = 4)
% High school graduates (<i>n</i> = 5, 24%)	0% (<i>n</i> = 0)	36% (<i>n</i> = 5)

Demographic characteristics	Children who responded to treatment (<i>n</i> = 7)	Children who did not respond to treatment (<i>n</i> = 14)
Father’s education: $\chi^2(4, N = 21) = 7.64, p = .11$		
% College graduates (<i>n</i> = 2, 10%)	0% (<i>n</i> = 0)	14% (<i>n</i> = 2)
% Partial college (<i>n</i> = 5, 24%)	43% (<i>n</i> = 3)	14% (<i>n</i> = 2)
% High school graduates (<i>n</i> = 4, 19%)	0% (<i>n</i> = 0)	29% (<i>n</i> = 4)
% Unknown (<i>n</i> = 9, 43%)	57% (<i>n</i> = 4)	36% (<i>n</i> = 5)

None of the demographic characteristics were significantly associated with treatment responding. However, several characteristics approached significance: race, ethnicity, parent marital status. In terms of race, fewer children who were white responded to treatment than expected. An alternative way of looking at race is to note that the two children in the sample who were African American (10% of sample) both responded to the treatment. In terms of ethnicity, fewer children who were non-Hispanic responded to treatment than expected. Alternatively, of the four children who were Hispanic (19% of sample), three responded to treatment, indicating a greater than expected response to treatment in this group. In terms of parent marital status, fewer children whose parents were married responded to treatment than expected. Alternatively, of the four children whose parents were single (19% of sample), three responded to treatment. It is notable that the variables that approached significance indicated that treatment responding was observed in children from non-majority groups. This suggests that the treatment may be effective for a diverse population of children, although this study did not address efficacy or effectiveness of the treatment. However, it should be noted that these more diverse children who responded to treatment also may have had a protective factor in terms of mother’s education. Although mother’s education was not significantly related to treatment outcomes, children who responded to treatment had mothers who either graduated from college or completed partial college. Also, recall that children whose fathers were college graduates were more prevalent in intensity 36. However, the data in Supplemental Table S3b shows that these children were actually the ones that did not respond to treatment. These patterns highlight the difficulty of evaluating the relationship between nominal demographic characteristics and treatment response in a small sample of children. More research clearly is needed to better understand the relationship between child demographics and treatment responding but Supplemental Table S3b provides a foundation for future inquiry.

References

- Goldman, R., & Fristoe, M. (2000). *Goldman-Fristoe Test of Articulation—Second Edition*. Circles Pines, MN: American Guidance Service.
- Reynolds, C. R., & Kamphaus, R. W. (2003). *Reynolds Intellectual Assessment Scales and the Reynolds Intellectual Screening Test*. Lutz, FL: Psychological Assessment Resources.
- Semel, E., Wiig, E. H., & Secord, W. A. (2003). *Clinical Evaluation of Language Fundamentals—Fourth Edition*. Bloomington, MN: Pearson.
- Seymour, H. N., Roeper, T. W., de Villers, J., & de Villers, P. A. (2005). *Diagnostic Evaluation of Language Variation—Norm-Referenced*. San Antonio, TX: The Psychological Corporation.
- Wagner, R. K., Torgesen, J., Rashotte, C., & Pearson, N. A. (2013). *Comprehensive Test of Phonological Processing—Second Edition*. Austin, TX: Pro-Ed.
- Wallace, G., & Hammill, D. D. (2013). *Comprehensive Receptive and Expressive Vocabulary Test—Third Edition*. Austin, TX: Pro-Ed.