LSHSS

Research Article

Moved by Reading in a Spanish-Speaking, Dual Language Learner Population

Ashley M. Adams,[a] Arthur M. Glenberg,[b,c,d] and Ma Adelaida Restrepo[a]

Purpose: The aim of this study was to evaluate the effectiveness of an English-only version and a Spanish-support version of an embodied reading comprehension intervention (Moved by Reading) consisting of 3 stages (physical manipulation, imagined manipulation, and transfer) for Spanish–English dual language learners.
Method: Sixty-one dual language learners in Arizona were randomly assigned to 4 groups (Spanish-support control, Spanish-support intervention, English-only control, and English-only intervention). Analyses of variance were used to compare control and intervention groups and to compare groups according to the language of the intervention.
Results: Children in the Spanish-support intervention group significantly outperformed both control groups during the physical manipulation stage, whereas children in the English-only intervention group outperformed both control groups in the imagined manipulation stage, but there was little transfer to a new, unrelated text.
Conclusions: The Moved by Reading intervention, in both its English-only and Spanish-support versions, improved performance on comprehension questions, but in different stages of the intervention. The Spanish-support version of the intervention was most effective in the physical manipulation stage, whereas the English-only version was more effective in the imagined manipulation stage. Neither version was effective in producing significant transfer effects.

Learning to read is undoubtedly one of the most important skills a child will learn during the school years. Learning the code-based component skills of reading (i.e., decoding, phonological awareness, and alphabetic knowledge) is a necessary part of achieving literacy; however, the ultimate goal of reading is to understand the meaning of what is written. Reading comprehension involves a constellation of higher-level skills that allow readers to make connections between elements of the text and their background knowledge (Kendeou, Van den Broek, White, & Lynch, 2009). According to simulation theory, these connections are accomplished by the use of one's own motor, perceptual, and emotional systems to simulate the situations described in the text (Fischer & Zwaan, 2008; Glenberg, 2007; Kiefer & Pulvermüller, 2012). For example, when one hears or reads a sentence such as "He was doing somersaults down the hallway," one would activate the visual perceptual areas necessary to visualize a hallway and motion, the motor areas necessary to accomplish the action of somersaulting, and the emotional systems that would respond to the unusual sight of a person attempting to somersault down a hallway.

Simulation Theory, Language, and Reading

Simulation theory of language comprehension postulates that comprehending language requires using words and syntax to drive changes in sensorimotor and emotional systems that simulate the content of the language (Glenberg & Gallese, 2012; Glenberg & Robertson, 2000; Kaschak & Glenberg, 2000; Pulvermüller, 2012; Zwaan, Stanfield, & Yaxley, 2002). That is, these systems take on states homologous to those produced by observing, acting in, and emotionally responding to the actual events, not only their descriptions. Listeners create a dynamic mental image (although not necessarily conscious) of the content of the sentence.

Glenberg and Gallese's (2012) model of action-based language builds on simulation theory. According to their model, children's language develops as they learn how a word sounds, how to execute the speech motor commands necessary to say the word, and how to interact with that

[a]Department of Speech and Hearing Sciences, Arizona State University, Tempe
[b]Department of Psychology, Arizona State University, Tempe
[c]Department of Psychology, University of Wisconsin-Madison
[d]University of Salamanca, Spain

Correspondence to Ashley M. Adams: amadam13@asu.edu

Editor: Krista Wilkinson
Associate Editor: Marleen Westerveld

Received April 21, 2016
Revision received October 25, 2016
Accepted February 26, 2018
https://doi.org/10.1044/2018_LSHSS-16-0032

Disclosure: *Arthur M. Glenberg is a partial owner of Moved by Reading, LLC, which may in the future commercialize the intervention described in this article. The co-authors have declared that no competing interests existed at the time of publication.*

word's referent. For example, parents consistently and repeatedly label objects in a child's environment. A parent might say, "here is your teddy bear," while handing the teddy bear to the child or "wave bye-bye" while moving the child's hand in a wave gesture. Thus, the process of language comprehension is supported by the actual presence of the items or actions being discussed. When a child is learning to understand oral language, the objects and actions referred to in speech are frequently and immediately indexed to their referent, thus producing the connection between language and sensorimotor systems. According to this theory, Hebbian learning ("what fires together, wires together") occurs such that speech motor areas, auditory perceptual areas, and controllers for goal-directed interaction with the environment form the basis of language acquisition (Glenberg & Gallese, 2012).

Behavioral data (e.g., Glenberg & Kaschak, 2002) and imaging data (e.g., Hauk, Johnsrude, & Pulvermüller, 2004) demonstrate how the motor system is involved in language comprehension. For example, Hauk et al. (2004) used functional magnetic resonance imaging to test simulation theory by measuring brain activity while participants listened to action verbs. The results of the study showed somatotopic activation of areas of the motor strip while listening to action verbs. Thus, while listening to "lick," there was enhanced activity in the motor cortex that controls the mouth, and while listening to "pick," activity in hand areas increased. This type of grounded language is not specific to only action-related words. In fact, semantic information is proposed to be stored in networks that include traditional language areas and category-specific sensory areas. So hearing a word such as "jasmine" or "garlic" activates not only traditional language centers but also the olfactory cortex (González et al., 2006). The same has been shown for processing of sound-related words in the superior temporal cortex (Kiefer & Pulvermüller, 2012), spatial words in the parietal cortex (Pulvermüller & Fadiga, 2010), and sentences describing motion in parts of the brain responsible for actual visual motion processing (Rueschemeyer, Glenberg, Kaschak, Mueller, & Friederici, 2010). As such, simulation theory postulates that language processing includes speech motor areas, auditory perceptual areas, and category-specific sensorimotor areas.

Written language, in contrast to oral language, relies less on the perceptual environment: When a child is reading, it is unlikely that the objects referred to in the text are present. Even when pictures are available, as in a child's storybook, the child may not accurately index the words to the pictures (Glenberg, 2011). As children progress through school, the language they read becomes progressively more decontextualized (i.e., less likely to be supported by illustrations and using language that is increasingly dissimilar to spoken language; Snow, 1991). Moreover, even for proficient decoders, it is unlikely that the language they produce when reading out loud contains the prosodic and tonal cues to the same extent as fluent speech. Thus, creating a sensorimotor simulation from the written text alone may be a challenge for young readers. However, when children are unable to make the link between their own experiences and the text, that is, when they are unable to create a simulation, they are less likely to understand the meaning of the text (Glenberg, 2011; Herrera, Perez, & Escamilla, 2015). Therefore, the idea of Moved by Reading (MbR) is to bring the focus back to creating these simulations through movement of pictures on a computer screen during the intervention.

Although simulation appears to be a natural outcome of learning an oral language, as suggested above, some children may need additional instruction in how to produce simulations while reading. The challenge may be even greater for at-risk dual language learners (DLLs) who have limited experience in the second language and thus limited vocabulary and background knowledge in that language (Alderson & Urquhart, 1988; Barry & Lazarte, 1995; Carrell, 1987, 1992). Background knowledge and vocabulary are important predictors of reading comprehension in general and especially in DLLs (e.g., Lesaux & Kieffer, 2010; Perfetti, Landi, & Oakhill, 2005; Rydland, Aukrust, & Fulland, 2012), and lack of background knowledge can be at the root of poor reading comprehension (Elbro & Buch-Iversen, 2013). Thus, using simulation to activate native language vocabulary and background knowledge should specifically help children who are still learning English to build stronger connections across motor, sensory, and linguistic systems necessary for reading comprehension. Recent research indicates that doing so facilitates acquisition of second language vocabulary and translation equivalents (Bilson, Yoshida, Tran, Woods, & Hills, 2015; Kaushanskaya, Gross, & Buac, 2014).

MbR

MbR is a computerized intervention designed to teach children to simulate while reading. The first phase of the intervention uses physical manipulation (PM). Children hear or read a sentence and move images on a computer screen to match the action described in sentences they have just heard or read. A small green stoplight at the end of the action sentences signals that children should manipulate the pictures to demonstrate understanding. For example, if a child reads the sentence "The farmer drove the tractor to the barn," she would click on the farmer, move him to the tractor, and then move the conjoined image of the farmer and the tractor to the barn. PM improves comprehension for two reasons. First, the child relates nouns to images on the computer screen, thereby relating the noun to the sensorimotor activity generated by the image. Second, the child uses the images on the screen to act out the sentence, thereby relating the syntax of the sentence, the who does what to whom, to the child's own motor activity. In effect, the child externalizes the simulation that produces comprehension (see Figure 1 for a sample story).

The second phase of MbR, imagined manipulation (IM), scaffolds the child toward independent reading. In this stage, children are asked to imagine the manipulations of the objects instead of physically manipulating them as

Figure 1. Sample story.

they did in the previous stage. Thus, the practice of manipulation in the PM stage gives the child a specific strategy for creating mental imagery. In addition, because the IM instructions specifically refer to imagining movement of the objects on the screen, the instructions encourage the use of the motor system in text comprehension.

MbR has been shown to be effective in improving reading comprehension for both free and cued recall of text content (for a review, see Glenberg, 2011), although most of the work has been conducted with monolingual English-speaking children in the early elementary years. Glenberg, Gutierrez, Levin, Japuntich, and Kaschak (2004) examined an early version of MbR that involved manipulation of physical toys with first and second graders. Compared with a control group that read stories with the toys visible but did not manipulate them, the group that engaged in manipulation recalled more action sentences in a free recall task ($d = 1.39$) and correctly answered more inference comprehension questions ($d = 0.81$) during the PM stage, the IM stage ($d = 1.87$ and $d = 1.50$, respectively), and a transfer stage when no specific instructions were given on how to read ($d = 1.23$ and $d = 0.95$, respectively). Glenberg, Brown, and Levin (2007) implemented a PM strategy in small groups of children 6–8 years old in which children took turns manipulating real objects to demonstrate understanding of the story. Children in a control group simply reread target sentences with no manipulation. The authors found large advantages ($d = 1.72$) for children in the intervention groups on answering forced choice questions related to story content. Finally, Glenberg, Goldberg, and Zhu (2011) showed that first and second graders using a computerized version of the MbR intervention, very similar to that used in this study, outperformed children in a control condition ($p = .01$, $d = 1.16$) when answering open-ended comprehension questions.

The current study advances our knowledge of how MbR works because it targets a new population (Spanish–English DLLs). This study also allows for the comparison of the effectiveness of two versions of the MbR intervention, one that provides first language support and one that is entirely in English, in improving English reading outcomes for children with Spanish as their first/native language.

Potential Advantages of Using a Spanish-Support Version of MbR

The Spanish-support intervention may assist children in activating relevant background knowledge before reading the stories. According to simulation theory, when we read a word, our knowledge about that word is not a list of arbitrary features but instead a neural representation of past experience with the object or action the word describes. Because experience with the world often takes place in a specific language environment, hearing or reading a word in one's native language may more strongly activate this neural pattern. In fact, Jared, Pei Yun Poh, and Paivio (2013) found that bilingual (Mandarin and English) participants were faster to identify culturally biased items (e.g., a traditional Chinese mask) when they read the word in Mandarin rather than in English. From the perspective of simulation theory, this advantage occurred because hearing the word "mask" in Mandarin activated a perceptual memory that was specific to the participant's experience with a mask in their Chinese language environment. If two languages are learned in different environments at different times, there may be differences in the background knowledge that is activated depending on the language of input.

Another component of the intervention that may be particularly useful among DLLs is the use of actual movement of objects on the screen during comprehension.

The idea of using the body as a tool for second language learning through movement and gesture is not a new one. As early as the 1960s, Asher (1966) developed the Total Physical Response approach to second language teaching, which involved having learners act out commands given to them in their second language. Several studies have shown that moving the body to demonstrate understanding of a new word or concept improves later recall of that same word or concept over a control group taught in a traditional second-language classroom format (Asher, 1966; Elliott & Yountchi, 2009; Hwang et al., 2014; Kariuki & Bush, 2008). Relatedly, recent studies have shown that accompanying new word learning in a second language with corresponding gestures improves acquisition of these new words (Bergmann & Macedonia, 2013; Macedonia & Knösche, 2011; Porter, 2012). Requiring a child to “act out” the content of short stories by manipulating images on a screen to demonstrate understanding (as MbR does) is supported by these previous findings.

When considering the best evidence-based practices for literacy intervention in a group of DLLs, the research community has called for more experimental research that examines the efficacy of specific interventions with this group (Gutierrez, Zepeda, & Castro, 2010), specifically in terms of ways to incorporate technology (Proctor, Dalton, & Grisham, 2007). The current research does exactly that. Two versions of the MbR intervention were administered to a group of Spanish-speaking DLLs in second and third grades. The first version of the intervention was entirely in English, whereas the second version provided scenario familiarization, instructions, and practice in Spanish before transitioning to English-only reading. Two questions were addressed: (a) For Spanish-speaking DLLs, does either version (English-only or Spanish-support) of the MbR intervention result in significantly more comprehension questions answered correctly compared with respective control groups? On the basis of previous evidence of success of the intervention discussed above, we hypothesize that both versions of the intervention will be effective at improving accuracy on comprehension questions; and (b) does a Spanish-support version of MbR provide any additional increase in comprehension scores over and above the English-only version? We hypothesize that children who receive the Spanish-support MbR intervention will perform even better than children who receive the English-only version based on the literature discussed above suggesting that providing native language support can improve outcomes for the DLL population. For both questions, we will measure reading comprehension through responses to cued recall questions.

Method

Participants

Children were recruited from local schools and after-school programs in the Phoenix metropolitan area. This area has a particularly high concentration of Spanish–English DLLs. Because these children were attending public school in Arizona, they were receiving English-only education and no formal literacy instruction in Spanish. Sixty-seven children and their parents from three different sites (one school and two after-school programs) consented to participate in the study. As this was our first study using the Spanish-support intervention and the focus is reading comprehension, we chose to evaluate its efficacy using typically developing bilingual children who could decode at a high enough level to avoid frustration when reading the intervention stories. Therefore, to participate in the study, the children had to (a) attend second or third grade; (b) come from a primarily Spanish-speaking home based on parent report; (c) present with no significant history of speech, language, or cognitive delay per parent report; and (d) have a score > 50% correct on grade level words using the decoding portion of the Qualitative Reading Inventory–Fifth Edition (QRI-5; Leslie & Caldwell, 2010). Children were given a $10 gift card in exchange for their participation in the project. Six children (two first graders and four second graders) were eliminated because of difficulty reading the stories and poor decoding skills. The final sample consisted of 21 third graders and 40 second graders. We randomly assigned 15 children to the English-only control group, 13 children to the Spanish-support control group, 18 children to the English-only intervention group, and 15 children to the Spanish-support intervention group (see Table 1 for group demographic descriptions). There were no significant differences in age among the four control and experimental groups. Uneven group numbers were due to the elimination of children with decoding difficulties and inclusion of the Spanish-support control group after subject recruitment and random assignment had already started at Site 1. Nevertheless, the integrity of random assignment was maintained, as there were participants in all four groups at each site. This project involving research with human subjects was approved by the Arizona State University Institutional Review Board on January 28, 2013.

Materials

MbR Stories

MbR consists of 14 interactive, online stories. Seven stories take place in a farm scenario, and seven stories take place in a house scenario. For all stories, the interactive images were based on Fisher-Price toys (see Figure 1 for a sample story). Each individual child was exposed to seven stories: six intervention stories and one transfer story. Of the six intervention stories, two were read aloud using a recorded voice, and the child read the other four stories. Before the stories were presented, the research assistant familiarized each child with all relevant objects and characters in the stories using an introduction screen. A standard computer mouse was used to manipulate the images.

Open-Ended Cued Recall Outcome Measure

After completion of every story, the computer screen was turned off and the experimenter asked four to five

Table 1. Participant characteristics.

Condition	*n*	Gender	Mean age in years;months (*SD* in months)	% Mothers with high school education or greater	Number of times children read at home per week	Decoding (out of 40)	Baseline reading (out of 6)
English-only control	15	3 M, 12 F	8;3 (12.3) Range = 6;11–10;5	77%	3.99 Range = 0–7	35.93 (5.19) Range = 26–40	3.06 (0.93) Range = 2–5
Spanish-support control	13	5 M, 8 F	8;7 (9.3) Range = 7;9–9;9	58%	4.02 Range = 2–7	35.31 (5.02) Range = 24–40	3.85 (1.57) Range = 0–6
English-only intervention	18	9 M, 9 F	8;2 (12.2) Range = 7;0–10;5	89%	4.00 Range = 0–7	35.94 (5.07) Range = 24–40	3.44 (0.96) Range = 2–5
Spanish-support intervention	15	7 M, 8 F	8;1 (7.9) Range = 7;2–9;4	70%	3.93 Range = 1–7	36.53 (3.76) Range = 29–40	2.88 (1.17) Range = 1–5

Note. Maternal education and number of times children read at home were obtained via parent questionnaire. Decoding and baseline reading scores were obtained using subtests of the Qualitative Reading Inventory–Fifth Edition (Leslie & Caldwell, 2010). There were no significant between-group differences on the participant demographics reported above ($p > .05$). F = female; M = male.

open-ended cued recall questions assessing the children's comprehension of the story (see Figure 2 for sample questions). If the child answered incorrectly, the experimenter asked the child a follow-up two-alternative forced choice question. Only questions from the stories that children read themselves were considered (the first story in each stage was read aloud by a recorded voice to familiarize children with the intervention strategies, and therefore questions related to this story were not included in the analysis). This resulted in a maximum possible score of 10 for the PM and IM stages (two stories read independently in each stage, with four to five corresponding questions per story) and a maximum possible score of 5 for the transfer stage (one story read independently, with four to five corresponding questions). Each correct answer on the open-ended cued recall questions counted as 1 point. All scores were converted to percentages to account for the small differences in number of questions answered. The first author scored all open-ended cued recall questions. A second Spanish–English bilingual individual scored responses for 10 participants (16% of the data), and interrater reliability was 88%.

We used the number of correct answers to the open-ended questions and did not consider the forced choice answers as the dependent variable, for two reasons. First, performance on the two-alternative questions can be strongly influenced by guessing. Second, Keenan, Betjemann, and Olson (2008) found that, when compared with three other standardized tests of reading comprehension, the test containing open-ended questions contained the most variance explained by listening comprehension and the least variance explained by decoding.

Figure 2. Sample questions.

Q1: When Martin gets home, who does he see in the kitchen?
When Martin gets home, does he see Rosa or Sara in the kitchen?

Q2: Where does Martin go to see Sara?
Does Martin go to the playroom or the kitchen to see Sara?

Q3: When Martin goes downstairs with Sara, where does he put her?
When Martin goes downstairs with Sara, does he put her in the kitchen or in her stroller?

Q4: After Martin and Sara go outside, who do they see?
After Martin and Sara go outside, do they see Rosa or the rabbit?

Q5: When Martin and Sara get back, who will get a treat?
When Martin and Sara get back, will the rabbit get a treat or will Sara get a treat?

Parent Consent and Questionnaire

A parental consent form and a brief questionnaire were collected for each child participant. The questionnaire gathered information about the child's reading and language skills in both Spanish and English as well as the child's home literacy environment. The questionnaire consisted of 12 questions and was filled out with an experimenter's support when needed.

QRI-5

This criterion-referenced measure (Leslie & Caldwell, 2010) was used to assess children's word decoding abilities. The measure assesses word reading ability, oral reading accuracy, and comprehension of passages read orally and silently for children with a wide range of reading abilities. Although the QRI-5 is not a norm-referenced test, the authors suggest that it is a useful tool for researchers as a measure of comprehension (Leslie & Caldwell, 2010). The authors report that the correlation between comprehension performance on grade-level QRI texts and standardized tests of reading achievement as a function of grade are $r = .65$, $p < .0.01$, for second grade, and $r = .55$, $p < .05$, for third grade.

Spanish–English Language Proficiency Scale

The Spanish–English Language Proficiency Scale (SELPS; Smyk, Restrepo, Gorin, & Gray, 2013) is designed to measure the level of Spanish and English oral language proficiency in bilingual children 5–8 years old. The test uses modifications of the wordless storybooks *Frog on his Own* (Mayer, 1967a) and *A Boy, a Dog, a Frog, and a Friend* (Mayer, 1967b). The child looks at the pictures while listening to an oral telling of the story in either Spanish or English. Immediately after listening, the child retells the story in the same language. The retellings are scored on four domains: syntactic complexity, grammatical accuracy, verbal fluency, and lexical diversity. Each domain is scored between 1 and 5 according to how well the child can speak the target language, and the SELPS total score is the sum of scores obtained in the domains. The mean score is used to classify the participant as with high proficiency (4–5), with medium proficiency (2.5–3.5), or with low proficiency (1–2). The sample for standardization included 500 Latino children in the Phoenix metropolitan area.

Scores for the English proficiency portion of the SELPS have been validated for identifying different English developmental language levels in sequential Spanish–English bilingual children (Smyk et al., 2013). Construct validity was demonstrated by moderate and significant correlations between the English SELPS scores, teacher report of language proficiency, and language sample analyses using SALT software (Miller & Iglesias, 2008). These correlations ranged in magnitude from moderate ($r = .59$, $p < .005$, for the verbal fluency domain) to strong ($r = .82$, $p < .005$, for the lexical diversity domain). Test–retest reliability was demonstrated based on correlations between the overall scaled scores on two different story retells. These correlations were statistically significant, $r = .78$, $p < .005$ (Smyk et al., 2013).

For the purposes of this study, each child retold two frog stories, one in Spanish and another in English, resulting in a language proficiency score for each language. The first author (English–Spanish bilingual) scored all language samples in both Spanish and English. A second native speaker of each language also scored 10 samples in each language (15% of the data). For English, the interrater reliability was 88.4%, and for Spanish, the interrater reliability was 89.0%. These levels of reliability are generally considered good to excellent (McHugh, 2012).

Procedure

Children were randomly assigned to the four conditions (English-only control, English-only intervention, Spanish-support control, and Spanish-support intervention) such that, for every four children, one would be assigned to each condition. Separate assignment sheets were made for second and third graders to ensure that an approximately equal number of children from each grade participated in each experimental condition. In addition, stories were counterbalanced so that each story appeared in every order an equal number of times. For example, Child A might have seen, in sequence, Farm Story 2, Farm Story 4, Farm Story 3, Farm Story 1, Farm Story 5, Farm Story 6, and House Story 4, and Child B might have seen House Story 1, House Story 4, House Story 3, House Story 2, House Story 6, House Story 5, and Farm Story 2.

Preintervention Procedure (Day 1)

On Day 1 of testing, each child came from his or her class or after-school activity to a quiet testing room. All children participated in the word reading measure from the QRI-5. Children read aloud 40 words from the primer level of the QRI word reading test. This list was chosen because the words were of approximately the same difficulty as those found in the texts. There were three possible scores for each word read aloud. If the child was able to identify the word immediately, it was scored as "automatically identified." If the child was able to identify the word after sounding it out, it was scored as "identified," and if the child was unable to identify the word, it was scored as "incorrect." Next, the children read the primer short story from the QRI-5 out loud. This level was chosen as it was judged to be of a similar difficulty level as the intervention stories. Then, the experimenter asked six comprehension questions after the story (a format very similar to the type of questions asked during the intervention). Accuracy on these questions was used as a baseline measure of reading comprehension. Finally, each child listened to a frog story in each language and retold the stories for language proficiency analysis with SELPS. The procedures for the first day were identical for every child and took place in a single session lasting between 20 and 40 min.

Intervention Procedure (Days 2 and 3)

Spanish-support intervention. On Day 2, the first day of the intervention, children randomly assigned to the Spanish-support intervention group were first familiarized with the objects and characters from the scenario in which the stories were going to take place. For example, from the house scenario, the crib, high chair, living room, and people living in the house were identified. The experimenter read the narrative that introduced the characters to the child and asked the child to point with the mouse to each item on the screen as the experimenter said the words and pointed with her finger to each item. This ensured that the child was able to use the mouse and was attending to each item as it was introduced. This introduction procedure was followed once in Spanish and then repeated in English.

After the scenario was introduced, the experimenter instructed the child (in Spanish) on how to use the PM strategy, that is, how to use the mouse to move the toys on the screen to show what was happening in the story. Then, the experimenter repeated the instructions in English. Next, the experimenter played the recorded version of one story in Spanish and paused at each green light (as depicted in Figure 1) while the child manipulated the pictures. Thus, the child could practice the intervention using the native language. Finally, all children read two stories in English using the PM procedure.

On Day 3, children began to use the IM procedure. The experimenter instructed the children in Spanish that, at each green light, they would need to imagine that they were moving the toys on the screen. To practice IM, the children listened to a three-sentence Spanish practice story. Children were instructed to move the mouse to the green light when they were done imagining. Then, children heard another longer Spanish story during which they used IM. Next, the child read aloud two stories in English while using

IM. The stories read using IM were always from the same scenario (e.g., house or farm) as the stories read using PM on the previous day.

The last story the child read was the transfer story with a new scenario. For example, if the child had practiced PM and IM with the farm scenario, then the transfer story came from the house scenario. The experimenter gave the child the same IM instructions as described above (in English), but no additional vocabulary support was provided for the new story scenario. After completion of every story, the experimenter turned off the computer screen and asked four to five open-ended recall questions assessing the children's comprehension of the story (see Figure 2 for sample questions).

English-only intervention. All intervention procedures for the English-only intervention group were identical to the Spanish-support intervention group with three exceptions. On Day 2, the initial familiarization with the scenario took place twice in English. The experimenter also administered the instructions for the PM stage and the initial practice story only in English. On Day 3, the experimenter administered the IM instructions and oral practice story in English. As in the other conditions, the child read all stories in the PM and IM stages in English, and comprehension questions followed each story.

Control groups. A Spanish-support control group and an English-only control group received the same pattern of language input (i.e., introduction to vocabulary, instructions on how to read the stories, and orally presented stories) as the respective intervention groups described above. All procedures were identical to the intervention groups, except that the experimenter instructed the children that, when they saw the green lights, they should stop and think about the sentence they had just read and, once they had finished thinking carefully about the sentence, move on to read the next sentence. Importantly, the children in the control groups could see the same images as the children in the intervention group, but they did not manipulate any objects on the screen, and they did not receive any explicit instruction to imagine the objects moving on the screen.

Results

A 2 × 2 analysis of covariance of Language (English-only or Spanish-support) × Treatment Condition (control or intervention) was conducted for each stage of the intervention while using age, decoding ability, baseline reading comprehension performance, maternal education, number of times children read in the home per week, English language proficiency, and Spanish language proficiency as covariates. The use of multiple covariates controlled for several of the complex demographic characteristics inherent in the school-age DLL population. Although the selection criteria excluded very poor decoders, we decided to still use decoding as a covariate because of its significant contribution to reading comprehension in these early grades and the variability in decoding ability within the groups (Lesaux & Kieffer, 2010; Lindsey, Manis, & Bailey, 2003). The dependent variable was the percentage correct score on the cued recall questions asked after the completion of each story read in English.

The effect sizes reported (Cohen's *d*) are standard deviations of between-group differences. Generally, an effect size of 0.2 is considered a small effect, 0.5 is considered a medium effect, and 0.8 or greater is considered a large effect (Cohen, 1988). Where partial η^2 is reported for the effect size, it refers to the percentage of variance in the dependent variable accounted for by an independent variable, holding constant the effect of other independent variables. Generally, for partial η^2, .01 is considered a small effect size, .06 is considered a medium effect size, and .14 or greater is considered a large effect size (Richardson, 2011). Missing data were missing completely at random and were handled using multiple imputations in SPSS. Pooled F values across five imputations with corresponding p values are reported.

An initial set of univariate analyses of variance were conducted to explore the balance of demographic characteristics of the participants between groups. The groups did not differ significantly in English language proficiency, $F(3,57) = 0.76$, $p = .52$, or Spanish language proficiency, $F(3, 57) = 0.98$, $p = .41$. However, within-group paired t tests revealed that the children in the English-only control and intervention groups were stronger in English than in Spanish ($p = .04$ and $p = .06$, respectively). In contrast, the Spanish-support groups were statistically equivalent in terms of English and Spanish proficiency. See Table 2 for descriptive statistics and paired t tests.

There were no significant between-group differences on participant demographics such as age, maternal education, decoding ability, and baseline reading comprehension ability. Nonsignificant between-group differences were found for age, $F(3, 57) = 0.76$, $p = .52$; number of times children read at home per week, $F(3, 57) = 0.10$, $p = .96$; decoding, $F(3, 57) = 0.15$, $p = .93$; and baseline reading comprehension, $F(3, 57) = 1.72$, $p = .17$. Mothers with at least a high school graduation were also equally distributed across groups, $\chi^2(3) = 2.7$, $p = .44$. Surprisingly, none of the covariates was significant, and age was not a significant predictor of performance during any stage of the intervention, which could possibly be because of the simplicity of the texts or unique characteristics of the subjects from this sample (all ps > .4). The assumption of homogeneity of variance across groups was met according to Levene's test for the PM stage, $F(3, 57) = 1.381$, $p = .26$; the IM stage, $F(3, 57) = 0.384$, $p = .77$; and the transfer stage, $F(3, 57) = 0.322$, $p = .81$. An examination of skewness and kurtosis statistics as well as Q-Q plots revealed approximately normal distributions within each group.

The raw recall data from the texts read in English as well as the estimated marginal means with all covariates included in the model and the between-group effect sizes are presented in Table 3. Below, in each stage of the intervention, the main effect of treatment is examined to provide the answer to the first research question (whether either version of the intervention is effective in improving accuracy

Table 2. SELPS language proficiency by group.

Condition	SELPS English	SELPS Spanish	*t* Test	*d*
English-only control	4.33 (0.83)	3.95 (0.89)	$t(14) = 2.25, p = .04$	0.57
Spanish-support control	4.52 (0.51)	4.28 (0.83)	$t(12) = 1.06, p = .31$	0.29
English-only intervention	4.18 (0.52)	3.93 (1.04)	$t(17) = 2.02, p = .06$	0.47
Spanish-support intervention	4.15 (0.69)	4.43 (0.75)	$t(14) = -0.36, p = .72$	–0.09

Note. Mean Spanish–English Language Proficiency Scale (SELPS) score was provided out of a maximum of 5. Paired samples *t* test examines within-group differences in language proficiency. The effect size calculated for the paired samples *t* test is the mean of the within-group difference divided by the standard deviation of the within-group difference (Cohen, 1988).

in responses to reading comprehension questions among DLLs). The planned pairwise comparisons performed when appropriate provide the response to the second research question (whether the Spanish-support version provided an additional improvement in responses to reading comprehension questions over and above that of the English-only version).

PM Stage

In the PM stage, there was a statistically significant main effect of treatment condition, $F(3, 50) = 9.48$, $p < .001$, partial $\eta^2 = .13$, with the groups that received the MbR intervention performing significantly better overall than the control groups. Although the means were in the expected direction, neither the main effect of language, $F(3, 50) = 1.25$, $p = .30$, nor the interaction between language and treatment condition, $F(3, 50) = 0.91$, $p = .44$, was significant. Planned pairwise comparisons using least significant difference indicated that the overall difference in treatment condition emerged because of the advantage of the Spanish-support intervention group over both the English-only control group, $t(27) = 3.0$, $p = .004$, $d = 1.21$, and the Spanish-support control group, $t(29) = 2.81$, $p = .007$, $d = 0.84$. The comparison between the English-only intervention and control groups was not significant. Although the difference

Table 3. Summary of accuracy statistics and effect sizes for performance in all intervention stages.

Stage	Condition	Percentage correct mean accuracy (PM: out of 10, IM: out of 10, transfer: out of 5)	EMM	*SD*	Effect size compared with English-only control	Effect size compared with English-only intervention	Effect size compared with Spanish-support control	Effect size compared with Spanish-support intervention
PM	English-only control	.59, range = .2–.9	.59	.18	0			
	English-only intervention	.69, range = .4–1.0	.68	.20	0.53	0		
	Spanish-support control	.60, range = .2–1.0	.61	.28	0.04	–0.37	0	
	Spanish-support intervention	.79, range = .6–1.0	.79	.15	1.21**	0.56	0.84**	0
IM	English-only control	.60, range = .2–1.0	.59	.24	0			
	English-only intervention	.73, range = .5–1.0	.73	.17	0.63*	0		
	Spanish-support control	.58, range = .3–0.9	.59	.18	–0.09	–0.86*	0	
	Spanish-support intervention	.72, range = .3–1.0	.71	.21	0.53	–0.05	0.72	0
Transfer	English-only control	.49, range = 0.0–1.0	.50	.27	0			
	English-only intervention	.52, range = 0.0–1.0	.51	.29	0.11	0		
	Spanish-support control	.54, range = 0.0–0.8	.56	.28	0.18	0.07	0	
	Spanish-support intervention	.65, range = .2–1.0	.65	.32	0.54	0.43	0.37	0

Note. Effect sizes reported are Cohen's *d*s. PM = physical manipulation; IM = imagined manipulation; EMM = estimated marginal mean with all covariates in the model.

$^*p < .05$. $^{**}p < .01$.

was not significant between the Spanish-support and English intervention groups, the Spanish-support intervention group performed over half a standard deviation above the English-only intervention group, $t(32) = 1.50$, $p = .14$, $d = 0.56$. See Table 3 for a summary of the effect sizes.

IM Stage

There was a statistically significant main effect of treatment condition, $F(3, 50) = 5.97$, $p = .01$, partial $\eta^2 = .12$, with the groups that received the intervention performing significantly better overall than the control groups. However, in this case, the planned comparisons using least significant difference revealed that the overall difference in treatment condition emerged because of the advantage of the English-only intervention group over the English-only control group, $t(32) = 2.23$, $p = .31$, $d = 0.63$, and the Spanish-support control group, $t(30) = 2.31$, $p = .024$, $d = 0.86$. The Spanish-support intervention group also outperformed the Spanish-support control group, although this difference did not reach statistical significance, $t(27) = 1.67$, $p = .10$, $d = 0.72$. Neither the effect of language, $F(3, 50) = 0.13$, $p = .94$, nor the interaction between language and treatment condition, $F(3, 50) = 0.12$, $p = .95$, was significant.

Transfer Stage

In the transfer stage, the main effect of treatment condition was not significant, $F(3, 50) = 0.50$, $p = .68$. In addition, as in the other stages of the intervention, the effect of language was also not significant, $F(3, 50) = 1.78$, $p = .16$. As neither of the main effects was significant, no further pairwise comparisons or interaction effects were examined statistically. Nonetheless, numerically, children in the Spanish-support intervention group outperformed children in the other conditions by a sizeable margin (see Table 3).

In general, those in the English-only MbR condition outperformed those in the English-only control condition in the IM (but not the PM) stage, and those in the Spanish-support MbR condition outperformed those in the Spanish-support control condition in the PM (but not the IM) stage. Importantly, as indicated by the effect sizes in Table 3, children in the Spanish-support MbR condition generally outperformed the children in the English-only MbR condition, although this was not a statistically significant difference.

Discussion

The first research question examined whether either version of the MbR intervention would be effective in improving performance on reading comprehension questions for typically developing Spanish-speaking DLLs. Results indicate that children in the MbR intervention groups performed better at answering the reading comprehension questions than children in the control groups in both the PM and IM stages of the intervention, but not in the transfer stage. However, some interesting patterns emerged. During the PM stage, the intervention group that received support in Spanish outperformed both control groups. This suggests that the combination of physical movement and first language support conveyed a specific advantage in reading comprehension. In contrast, in the IM stage, children in the treatment group still outperformed those in the control group, but this difference was driven by the statistical advantage of the English-only intervention group over the control groups. However, the raw average percentage of questions answered correctly was essentially identical in both intervention groups (73% correct for the English-only intervention group and 72% correct for the Spanish-support intervention group) and control groups (60% for the English-only control group and 58% for the Spanish-support control group), suggesting that both versions of the intervention conveyed an advantage in comprehension in the IM stage, at least as it related to the intervention stories.

The second question tested whether the Spanish-support intervention would be more effective in improving reading comprehension performance than the English-only version. The main effect of language of the intervention was not significant in any stage of the intervention; therefore, we cannot draw the conclusion that the Spanish-support intervention was more effective than the English-only intervention at improving reading comprehension. However, this may be an issue of power due to small sample sizes; therefore, we turn to an examination of effect sizes (see Table 3). The planned comparisons showed that, for the PM stage, the Spanish-support intervention group outperformed the English-only intervention group by only half a standard deviation. Although this translates to a small practical effect (the difference of answering one additional question correctly), it suggests the potential for benefit from a comprehension strategy that combines native language use and physical movement. However, in the IM stage, whereas there was still an advantage of being in the intervention groups over the control groups, there was no additional advantage for the Spanish-support intervention.

It should be noted that the total instructional time included in the intervention was approximately 2 hr, and the number of questions per stage was relatively small. As a result, the statistically significant results in this study, in some cases, represent a small practical difference between the control and intervention groups (one to two questions answered correctly). Such a short intervention is unlikely to alter the long-term course of a construct as complex as reading comprehension, and more questions in each stage would provide a more reliable measure of any between-group differences. Nonetheless, we interpret the improvement in responses to comprehension questions as a result of the MbR intervention as a proof of concept that the embodied approach to reading can be effective for DLLs and merits further exploration into the use of embodied strategies with this population. Our results are consistent with previous research pertaining to using the body as a tool for second language learning (Asher, 1966; Bergmann & Macedonia, 2013). These results also add to a growing body of knowledge suggesting that embodied approaches to learning in the classroom, especially as they relate to interaction with technology, are a promising, new area of research (Chang, Lee, Wang, & Chen, 2010; Hughes, Stapleton, Hughes, & Smith, 2005; Johnson-Glenberg, Birchfield, & Uysal, 2009; Pan, Cheok, Yang, Zhu, & Shi, 2006).

The finding that the English-only intervention was effective in the IM stage is consistent with previous findings with monolingual English and Native American populations (Glenberg et al., 2004; Marley, Levin, & Glenberg, 2010). In addition, we found that the Spanish-support version of the intervention was at least as effective as the English-only version and possibly more effective specifically when combined with the PM strategy, although future research with larger groups will be necessary to substantiate this claim. However, neither intervention group showed a statistically significant advantage on the final (transfer) story, suggesting that children were unable to generalize the strategy to a novel scenario. Given the short term of this research, this is not surprising. In accordance with these findings, a recent meta-analysis of transfer effects of reading comprehension interventions found small or no effects (Suggate, 2010, 2016).

The advantage in the Spanish-support intervention group in the PM stage suggests that presenting information in Spanish and English may have activated more relevant background knowledge (Lesaux & Kieffer, 2010; Rydland et al., 2012), which in turn facilitated English comprehension (Bilson et al., 2015; Kaushanskaya et al., 2014). By including both languages in the current intervention, we may have facilitated activation of the representations corresponding to each language, thereby allowing the child to access more relevant background knowledge and vocabulary needed to understand the text, which in turn facilitates word learning in the second language. For example, R. T. Jiménez, Garcia, and Pearson (1996) found that successful bilingual readers are able to implement knowledge they have learned in their first language while reading in their second language. In fact, many researchers in the area of bilingual literacy suggest that using the child's first language to support acquisition of reading skills and strategies is key to second language reading outcomes (August & Shanahan, 2006; Castro, Páez, Dickinson, & Frede, 2011; Goldenberg, 2008). Being able to express this knowledge through physical movement appears to have been a key component to the effectiveness of the Spanish-support intervention, as there was not a similar advantage seen in the IM stage where children simply imagined they were moving objects.

We expected to see an advantage of the Spanish-support version of the intervention across the PM, IM, and transfer stages of the intervention, but this advantage only emerged in the PM stage. It is possible that the IM and transfer stages were more cognitively demanding than the PM stage for several reasons: (a) The child did not create or observe any movement in these stages, removing visual and motor supports that were present in the PM stage; (b) the transfer stage came from an unfamiliar scenario in which the child had not had any previous practice; (c) the use of only a single text with only five related questions in the transfer stage greatly reduced power for the between-group comparison; and/or (d) the intervention was not long enough to have a more generalized effect. The embodied theory of comprehension posits that activation of the motor system while reading is critical for comprehension (Glenberg, 2011). This activity must be present in the PM stage because children are using their bodies to interact with the story content. However, in the IM and transfer stages, the child must generalize that strategy to imagined movement, which was likely a more difficult task. When combined with the already cognitively demanding process of decoding an orthographically opaque language (the child's second language), the added cognitive load of the IM and transfer may have prevented the potentially positive effects of the intervention from being observed.

Finally, it is important to consider the sociocultural situation of the children who participated in Arizona. These children were provided legally mandated English-only education. In other scenarios, such as transitional bilingual programs or two-way dual language programs, where children are developing literacy in both languages, this intervention may also be particularly effective. Regardless of the setting, the results reported here suggest that the MbR strategy is a promising intervention for improving reading comprehension in the Spanish–English DLL population.

Limitations

The primary limitation of the intervention used in this study was that there were several issues related to power. There was limited practice with each stage of the intervention, as the entire intervention was administered over only 3 days, resulting in a limited number of stories and questions to use as the outcome variable. Because of the small number of stories and questions in each stage, even when statistically significant differences were found between groups, it was difficult to draw conclusions about the practical nature of these effects. Statistically, the use of a large number of covariates with a relatively small number of participants certainly diminishes the overall power of the analysis. Future research can build on these findings by including a larger number of participants to make stronger statements about between-group differences. However, considering this short period and limited power, the results are promising for future larger studies using this intervention.

The lack of significant transfer of the strategy in DLLs suggests that further exploration into longer and more diverse training is necessary to determine the ability to internalize and generalize the MbR strategy to improve comprehension. However, this suggestion must be moderated by (a) the sizeable numerical advantage of the Spanish-support intervention over the other conditions in the transfer stage and (b) the low power in these comparisons. Nonetheless, evidence-based instructional programs targeted at reading comprehension generally include sustained opportunities for skill practice in multiple contexts (Denton, Fletcher, Taylor, Barth, & Vaughn, 2014; Shanahan et al., 2010). If children had more practice with the PM stage of the intervention or more explicit instruction related to the PM strategy, we may have seen more transfer effects. It is also possible that this strategy-based intervention may be more effective with older children with more well-developed metacognitive skills (Baker, 2006) or children who have better decoding skills, as this would free up cognitive capacity for implementing higher-level strategies (Brunswick, 2015; Hoover & Gough, 1990).

In addition, the children in this study were approximately balanced in Spanish and English proficiency, with the exception of the English-only control group where English was statistically stronger than Spanish ($p < .05$). This phenomenon was accounted for by using language proficiency in both languages as a covariate in all analyses. Nonetheless, future studies that involve children who are more dominant in Spanish would allow for a more complete examination of the effectiveness of the Spanish-support intervention for this population. Finally, the texts used in the current research were all narrative texts. As recent research suggests (e.g., Catts & Kamhi, 2015), reading comprehension is not a unidimensional construct. A more diverse pretest and posttest measure of comprehension that includes free recall, cued recall, and cloze tasks for both narrative and expository texts may help to build a portfolio of a child's comprehension that allows for a more detailed analysis of the kinds of comprehension that the intervention may improve.

Clinical Implications

This study focused on typically developing DLLs. Thus, the finding that combining native language support and the PM strategy was effective for this population has implications for teachers and other school-based practitioners. The primary implication is that embodied learning, or involving children in physical activity that builds conceptual understanding, is a promising method of improving reading comprehension. The premise of the intervention was to emphasize that children should ground their understanding in their own experiences in their bodies and in the physical world, and this principle can be widely applied in the school setting (e.g., see Kontra, Lyons, Fischer, & Beilock, 2015, for application to learning abstract principles in physics). In addition, teaching and providing practice with comprehension strategies in the native language with structured support for carryover in English were shown to be at least as effective as teaching these same strategies in English alone. Conceptually, this has important implications for educational and intervention models for DLLs.

Acknowledgments

The research was partially supported by NSF Grant 1324807, awarded to Arthur Glenberg (principal investigator), Erin Walker (co-principal investigator), and M. Adelaida Restrepo (co-principal investigator). In addition, Arthur Glenberg was partially supported by NSF Grant 1020367. Many thanks to the elementary schools in the Phoenix metropolitan area that participated and to Sarah Fialko, Albertina Albouchar, and Giener Rojas who assisted with data collection. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding agencies.

References

Alderson, J. C., & Urquhart, A. H. (1988). This test is unfair: I'm not an economist. In P. L. Carrell, J. Devine, & D. Eskey (Eds.), *Interactive approaches to second language reading* (pp. 168–182). Cambridge, United Kingdom: Cambridge University Press.

Asher, J. J. (1966). The learning strategy of the total physical response: A review. *The Modern Language Journal, 50*(2), 79–84.

August, D. L., & Shanahan, T. (2006). *Developing literacy in a second language: Report of the National Literacy Panel*. Mahwah, NJ: Lawrence Erlbaum.

Baker, L. (2006). Developmental differences in metacognition: Implications for metacognitively oriented reading instruction. In S. E. Israel, C. C. Block, K. L. Bauserman & K. Kinnucan-Welsch (Eds.), *Metacognition in literacy learning: Theory, assessment, instruction, and professional development* (pp. 61–80). New York, NY: Routledge.

Barry, S., & Lazarte, A. (1995). Embedded clause effects on recall: Does high prior knowledge of context domain overcome syntactic complexity in students of Spanish? *The Modern Language Journal, 79,* 491–504.

Bergmann, K., & Macedonia, M. (2013, August). A virtual agent as vocabulary trainer: Iconic gestures help to improve learners' memory performance. In R. Aylett, B. Krenn, C. Pelachaud, & H. Shimodaira (Eds.), *International Workshop on Intelligent Virtual Agents* (pp. 139–148). Berlin, Germany: Springer-Verlag Berlin Heidelberg.

Bilson, S., Yoshida, H., Tran, C. D., Woods, E. A., & Hills, T. T. (2015). Semantic facilitation in bilingual first language acquisition. *Cognition, 140,* 122–134.

Brunswick, N. (2015). Metacognition and reading: Comparing three forms of metacognition in normally developing readers and readers with dyslexia. *Dyslexia, 21*(3), 273–284.

Carrell, P. L. (1987). Content and formal schemata in ESL reading. *TESOL Quarterly, 21*(3), 461–481.

Carrell, P. L. (1992). Awareness of text structure: Effects on recall. *Language Learning, 42*(1), 1–20.

Castro, D. C., Páez, M. M., Dickinson, D. K., & Frede, E. (2011). Promoting language and literacy in young dual language learners: Research, practice, and policy. *Child Development Perspectives, 5*(1), 15–21.

Catts, H., & Kamhi, A. (2015, November). *Reading comprehension is knowledge acquisition*. Paper presented at the annual convention of the American Speech-Language-Hearing Association, Denver, CO.

Chang, C. W., Lee, J. H., Wang, C. Y., & Chen, G. D. (2010). Improving the authentic learning experience by integrating robots into the mixed-reality environment. *Computers & Education, 55*(4), 1572–1578.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Denton, C. A., Fletcher, J. M., Taylor, W. P., Barth, A. E., & Vaughn, S. (2014). An experimental evaluation of guided reading and explicit interventions for primary-grade students at-risk for reading difficulties. *Journal of Research on Educational Effectiveness, 7*(3), 268–293.

Elbro, C., & Buch-Iversen, I. (2013). Activation of background knowledge for inference making: Effects on reading comprehension. *Scientific Studies of Reading, 17*(6), 435–452.

Elliott, E., & Yountchi, L. (2009). Total physical response and Russian multi- and unidirectional verbs of motion: A case study in acquisition. *The Slavic and East European Journal, 53*(3), 428–450.

Fischer, M. H., & Zwaan, R. A. (2008). Embodied language: A review of the role of the motor system in language comprehension. *The Quarterly Journal of Experimental Psychology, 61*(6), 825–850.

Glenberg, A. M. (2007). Language and action: Creating sensible combinations of ideas. In G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 361–370). Oxford, England: Oxford University Press.

Glenberg, A. M. (2011). How reading comprehension is embodied and why that matters. *International Electronic Journal of Elementary Education, 4*(1), 5–18.

Glenberg, A. M., Brown, M. & Levin, J. R. (2007). Enhancing comprehension in small reading groups using a manipulation strategy. *Contemporary Educational Psychology, 32*(3), 389–399.

Glenberg, A. M., & Gallese, V. (2012). Action-based language: A theory of language acquisition, comprehension, and production. *Cortex, 48*(7), 905–922.

Glenberg, A. M., Goldberg, A. B., & Zhu, X. (2011). Improving early reading comprehension using embodied CAI. *Instructional Science, 39*(1), 27–39.

Glenberg, A. M., Gutierrez, T., Levin, J. R., Japuntich, S., & Kaschak, M. (2004). Activity and imagined activity can enhance young readers' reading comprehension. *Journal of Educational Psychology, 96,* 424–436.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin and Review, 9*(3), 558–565.

Glenberg, A. M., & Robertson, D. A. (2000). Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning. *Journal of Memory and Language, 43*(3), 379–401.

Goldenberg, C. (2008, Summer). Teaching English language learners: What the research does—and does not—say. *American Educator,* 8–44.

González, J., Barros-Loscertales, A., Pulvermüller, F., Meseguer, V., Sanjuán, A., Belloch, V., & Ávila, C. (2006). Reading cinnamon activates olfactory brain regions. *NeuroImage, 32*(2), 906–912.

Gutierrez, K., Zepeda, M., & Castro, D. (2010). Advancing early literacy learning for all children: Implications of the NELP report for dual-language learners. *Educational Researcher, 39,* 334–339.

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron, 41*(2), 301–307.

Herrera, S., Perez, D., & Escamilla, K. (2015). *Teaching reading to English language learners* (2nd ed.). Hoboken, NJ: Pearson Education.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing, 2*(2), 127–160.

Hughes, C. E., Stapleton, C. B., Hughes, D. E., & Smith, E. M. (2005). Mixed reality in education, entertainment, and training. *IEEE Computer Graphics and Applications, 25*(6), 24–30.

Hwang, W. Y., Shih, T. K., Yeh, S. C., Chou, K. C., Ma, Z. H., & Sommool, W. (2014). Recognition-based physical response to facilitate EFL learning. *Educational Technology & Society, 17*(4), 432–445.

Jared, D., Pei Yun Poh, R., & Paivio, A. (2013). L1 and L2 picture naming in Mandarin–English bilinguals: A test of Bilingual Dual Coding Theory. *Bilingualism: Language and Cognition, 16*(2), 383–396.

Jiménez, R. T., Garcia, G. E., & Pearson, P. D. (1996). The reading strategies of Latina/o students who are successful English readers: Opportunities and obstacles. *Reading Research Quarterly, 31,* 90–112.

Johnson-Glenberg, M. C., Birchfield, D., & Uysal, S. (2009). *SMALLab*: Virtual geology studies using embodied learning with motion, sound, and graphics. *Educational Media International, 46*(4), 267–280.

Kariuki, P. N., & Bush, E. D. (2008). *The effects of total physical response by storytelling and the traditional teaching styles of a foreign language in a selected high school*. Paper presented at the Annual Conference of the Mid-South Educational Research Association, Knoxville, TN.

Kaschak, M. P., & Glenberg, A. M. (2000). Constructing meaning: The role of affordances and grammatical constructions in sentence comprehension. *Journal of Memory and Language, 43*(3), 508–529.

Kaushanskaya, M., Gross, M., & Buac, M. (2014). Effects of classroom bilingualism on task-shifting, verbal memory, and word learning in children. *Developmental Science, 17*(4), 564–583.

Keenan, J. M., Betjemann, R. S., & Olson, R. K. (2008). Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension. *Scientific Studies of Reading, 12*(3), 281–300.

Kendeou, P., Van den Broek, P., White, M. J., & Lynch, J. S. (2009). Predicting reading comprehension in early elementary school: The independent contributions of oral language and decoding skills. *Journal of Educational Psychology, 101*(4), 765–778.

Kiefer, M., & Pulvermüller, F. (2012). Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions. *Cortex, 48*(7), 805–825.

Kontra, C., Lyons, D. J., Fischer, S. M., & Beilock, S. L. (2015). Physical experience enhances science learning. *Psychological Science, 26,* 737–749.

Lesaux, N. K., & Kieffer, M. J. (2010). Exploring sources of reading comprehension difficulties among language minority learners and their classmates in early adolescence. *American Educational Research Journal, 47*(3), 596–632.

Leslie, L., & Caldwell, J. S. (2010). *Qualitative Reading Inventory–Fifth Edition*. Boston, MA: Allyn & Bacon.

Lindsey, K., Manis, F., & Bailey, C. (2003). Prediction of first-grade reading instruction for second-grade English language learners with reading difficulties. *Journal of Educational Psychology, 95,* 482–494.

Macedonia, M., & Knösche, T. R. (2011). Body in mind: How gestures empower foreign language learning. *Mind, Brain, and Education, 5*(4), 196–211.

Marley, S., Levin, J., & Glenberg, A. (2010). What cognitive benefits does an activity-based reading strategy afford young native American readers? *The Journal of Experimental Education, 78*(3), 395–417.

Mayer, M. (1967a). *Frog on his own*. New York, NY: Dial Press.

Mayer, M. (1967b). *A boy, a dog, a frog, and a friend*. New York, NY: Dial Press.

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica, 22*(3), 276–282.

Miller, J. F., & Iglesias, A. (2008). Systematic Analysis of Language Transcripts, Bilingual SE (Version 2008) [Computer software]. Madison, WI: SALT Software, LLC.

Pan, Z., Cheok, A. D., Yang, H., Zhu, J., & Shi, J. (2006). Virtual reality and mixed reality for virtual learning environments. *Computers & Graphics, 30*(1), 20–28.

Perfetti, C. A., Landi, N., & Oakhill, J. (2005). The acquisition of reading comprehension skills. In M. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 227–247). Oxford, United Kingdom: Blackwell.

Porter, A. (2012). A helping hand with language learning: Teaching French vocabulary with gesture. *Language Learning Journal, 44,* 236–256. https://doi.org/10.1080/09571736.2012.750681

Proctor, C. P., Dalton, B., & Grisham, D. L. (2007). Scaffolding English language learners and struggling readers in a universal literacy environment with embedded strategy instruction and vocabulary support. *Journal of Literacy Research, 39*(1), 71–93.

Pulvermüller, F. (2012). Meaning and the brain: The neurosemantics of referential, interactive, and combinatorial knowledge. *Journal of Neurolinguistics, 25*(5), 423–459. https://doi.org/10.1016/j.jneuroling.2011.03.004

Pulvermüller, F., & Fadiga, L. (2010). Active perception: Sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience, 11*(5), 351–360.

Richardson, J. T. E. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review, 6*(2), 135–147. https://doi.org/10.1016/j.edurev.2010.12.001

Rueschemeyer, S., Glenberg, A., Kaschak, M., Mueller, K., & Friederici, A. (2010). Top–down and bottom–up contributions to understanding sentences describing objects in motion. *Frontiers in Psychology, 1,* 183.

Rydland, V., Aukrust, V. G., & Fulland, H. (2012). How word decoding, vocabulary and prior topic knowledge predict reading comprehension. A study of language-minority students in Norwegian fifth grade classrooms. *Reading and Writing, 25*(2), 465–482.

Shanahan, T., Callison, K., Carriere, C., Duke, N. K., Pearson, P. D., Schatschneider, C., & Torgesen, J. (2010). *Improving reading comprehension in kindergarten through 3rd grade: IES practice guide* (NCEE 2010-4038). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Smyk, E., Restrepo, M. A., Gorin, J., & Gray, S. (2013). Development and validation of the Spanish–English Language Proficiency Scale (SELPS). *Language, Speech, and Hearing Services in Schools, 44,* 252–265.

Snow, C. E. (1991). The theoretical basis for relationships between language and literacy in development. *Journal of Research in Childhood Education, 6*(1), 5–10.

Suggate, S. P. (2010). Why "what" we teach depends on "when": Grade and reading intervention modality moderate effect size. *Developmental Psychology, 46,* 1556–1579. https://doi.org/10.1037/a0020612

Suggate, S. P. (2016). A meta-analysis of the long-term effects of phonemic awareness, phonics, fluency, and reading comprehension interventions. *Journal of Learning Disabilities, 49*(1), 77–96.

Zwaan, R. A, Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science, 13*(2), 168–171.