

## Article

# Blended Learning as a Tool for Promoting Academic Literacy in Arabic: Effects on Students with Different Ability Levels

Bahaa' Makhoul <sup>1,2,\*</sup>  and Elit Olshtein <sup>3</sup><sup>1</sup> M.Ed. Program in Languages, Graduate Studies Faculty, Oranim College of Education, Tivon 3600600, Israel<sup>2</sup> Arabic Section, Literacy and Language Arts Department, The Center for Educational Technology (CET), Tel Aviv 61394, Israel<sup>3</sup> School of Education, The Hebrew University of Jerusalem, Jerusalem 9190501, Israel

\* Correspondence: baham@cet.ac.il

## Abstract

This study investigated the effects of “The Golden Triangle”, a blended-learning intervention program designed to promote academic language literacy among Arabic-speaking seventh-grade students in Israel. Receptive and productive academic vocabulary and academic reading comprehension were assessed before and after the intervention (Grade 7), and again five months later at the beginning of the eighth grade. Participants were 70 Arabic-speaking students from four middle schools, divided into a high/average reading comprehension (HRC) group and a low reading comprehension (LRC) group. Results showed that the LRC group improved its reading comprehension immediately after the intervention but declined at follow-up, whereas the HRC group improved during the program and maintained gains. Moreover, in the HRC group, baseline receptive academic vocabulary predicted reading comprehension at follow-up, whereas no similar predictive effect was found in the LRC group. These findings highlight the need to address persistent heterogeneity in reading abilities and to design differentiated interventions that support the long-term consolidation of academic language among struggling readers.

**Keywords:** blended learning; middle school; learning differences; reading comprehension; intervention; Arabic

## 1. Introduction

Academic language refers to the decontextualized language through which teachers and students learn, construct, and demonstrate knowledge in school settings (Schleppegrell, 2001; Snow & Uccelli, 2009). Academic language proficiency—beyond everyday conversational fluency—includes the ability to understand and express school-relevant concepts and ideas in both oral and written modes (Cummins & Yee-Fun, 2007). A prominent component of academic language is academic vocabulary, which we conceptualize here as general academic words/roots that recur across multiple content areas (analogous to Tier 2 vocabulary), rather than subject-specific technical terms (Tier 3) (Beck et al., 2002). Academic literacy is critical to success in high school and beyond (Heller & Greenleaf, 2007; Guthrie et al., 2012), and interventions that deepen students' word knowledge can improve academic vocabulary and comprehension (Lesaux et al., 2010).

Arab middle schoolers in Israel are exposed to factors putting them at literacy risk, including structural resource gaps and broader socio-economic inequalities in schooling (Resh, 2008; Blass, 2020), as well as the added complexity of multilingualism (Arabic, Hebrew, and English). Moreover, Arabic itself has unique characteristics that can complicate



Academic Editor: Lawrence  
Jun Zhang

Received: 28 August 2025

Revised: 5 February 2026

Accepted: 11 March 2026

Published: 27 March 2026

**Copyright:** © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

literacy development, most notably the diglossic gap between spoken and written varieties, reflected in vocabulary, phonology, grammar, syntax, linguistic forms, and forms of expression (Saiegh-Haddad, 2022). Together, these features can make acquiring proficient academic literacy more complex and demanding.

To address these challenges, we developed and implemented “The Golden Triangle”, a station-rotation blended-learning program that integrates three complementary components: (1) explicit instruction and practice in academic vocabulary through authentic content-area texts in Arabic; (2) blended learning through teacher-led, small-group, and individualized digital stations; and (3) guided digital-literacy activities via the “Ofek” educational platform. This blended design draws on evidence that combining face-to-face instruction with structured technology-mediated practice can support engagement and provide opportunities for differentiation (Saavedra & Opfer, 2012; Staker & Horn, 2012).

We examined the program’s efficacy both immediately after the intervention and at a five-month follow-up, focusing on receptive and productive academic vocabulary and reading comprehension. In addition, because blended learning can support differentiated instruction, we investigated whether outcomes and long-term maintenance differed between students with high/average reading comprehension (HRC) and those with low reading comprehension (LRC), and the extent to which the intervention helped lower-achieving students narrow gaps with their peers.

## 2. Theoretical Background

### 2.1. Academic Literacy

There are distinct acute linguistic transitions in children’s literacy development. Children’s linguistic development initially occurs in their natural environment and in their native language, spoken language particularly. With the beginning of formal education, children are exposed to written language more frequently, later encountering language in a higher register, which is referred to as academic language. Academic language proficiency is defined as the mastery of vocabulary, specialized or complex grammatical structures, and multifarious linguistic functions (Bailey, 2007). Hence, the development of academic language involves flexibility in the use of language, requiring children to make use of language in new ways to learn and demonstrate knowledge (Schleppegrell, 2012).

Textbooks are replete with academic texts characterized by complex syntactic structures, conciseness, abstract vocabulary, and high lexical compression; the cognitive load increases in the middle school and so does the compression of information (Snow & Uccelli, 2009). In order to succeed, students must have good verbal skills, high reading comprehension skills and academic language skills, which will be able to support them when dealing with educational material in all content areas independently (Snow & Uccelli, 2009; Nagy & Townsend, 2012; Scarcella, 2003). Due to the complexity of academic language, studies show that a variety of students can benefit from focused intervention programs aimed at improving their acquisition of academic vocabulary specifically and academic literacy in general (Carlo et al., 2004; Lawrence et al., 2012; Lesaux et al., 2010).

Although academic words are not prominent in academic texts (between 8% and 10% of the text), they have a great impact on understanding texts but usually have abstract meanings, posing difficulties for the learner (Baumann & Graves, 2010; Nation, 2001). Academic vocabulary often overlaps with Tier 2 words—high-utility, cross-disciplinary words used to express relationships and concepts—whereas Tier 3 words are discipline-specific technical terms; both can contribute to comprehension but may require different instructional approaches (Beck et al., 2002). Their meaning may change between contexts and have abstract or concrete meanings, thus “inflating” their lexical value. Research shows that acquiring academic vocabulary demands frequent exposure to new words in many

different texts and contexts (Nagy & Townsend, 2012), while simply defining the academic words and memorizing them is not sufficient (Coxhead, 2000; McKeown et al., 2013; Nation, 2001; Nagy & Townsend, 2012).

Intervention programs that deconstruct the complexity of academic vocabulary often target the depth of word knowledge and morphological structure, thereby increasing vocabulary and comprehension skills. For example, Lesaux et al. (2010) developed a text-based intervention program for the acquisition of academic vocabulary intended for sixth-grade minority students with low literacy skills in English. The cyclical teaching model encompasses exposure to the word through a text, activating students' prior knowledge, working on meaning in context, introducing additional meanings, conducting morphological analysis of different word forms, and writing sentences using the target word. Such cycles are consistent with broader traditions of concept analysis and semantic mapping in vocabulary instruction (e.g., the Frayer model; Frayer et al., 1969). The program had significant effects on students' knowledge of words, knowledge of words' meanings in context, and morphological skills. These findings suggest that an academic vocabulary program based on promoting multiple language components may facilitate vocabulary expansion and improve reading comprehension.

## 2.2. Teaching and Learning Differently: Approaches to Blended Learning

Blended learning is an umbrella concept for a variety of teaching and learning approaches (and associated models) that integrate digital technologies with face-to-face classroom instruction. A central goal is to optimize learning by enabling differentiation in pace and content (Powell et al., 2015; Staker & Horn, 2012).

Blended learning emphasizes technology as a valuable tool that can offer students ways to solve problems and develop critical thinking and communication skills (Saavedra & Opfer, 2012). In addition, technology can help students practice transferring these skills to different contexts, as well as reflecting and sharing their thinking and that of their peers. The dynamics of technological systems interwoven with other learning modes can provide teachers, in real time, tools to adapt instruction according to students' needs and progress (Staker & Horn, 2012).

The components of blended learning are learning environment, media, and teaching. The learning environment can be synchronous or asynchronous, and each environment has its advantages or disadvantages. The goal of blended learning is to leverage the positive features of each environment to ensure maximum use of resources in accordance with the teaching and learning goals. The media is the tool through which the learning content is conveyed. The third component of teaching brings forth the dynamic and multidimensional nature of the teaching-learning process, thus enabling the mixing of several teaching-learning tools, for example, by combining face-to-face learning with learning through digital media sources.

Many blended-learning approaches combine traditional classroom instruction with computer-assisted platforms. Kazakoff et al. (2018) demonstrated that a blended-learning program was successful in advancing students' reading skills among both English and non-English learners. Hurwitz and Macaruso (2021) similarly found that struggling middle school students who participated in a blended-learning literacy program made greater progress than peers in traditional settings. Overall, the literature supports the potential of blended learning and technology-enhanced instruction to promote literacy, particularly following the widespread shift to online and hybrid learning during COVID-19 (Topping et al., 2022; Ferdig & Kennedy, 2018).

### 2.3. *The Arab Student in Middle School: Numerous Hardships*

Middle school is a setting designed for students in early adolescence, an age when individuals experience simultaneous changes in various personal areas. Students move from elementary school, usually a smaller and more intimate environment, to middle school, which is typically a more complex system with many subjects and teachers. In Israel, middle schools are also exposed to structural constraints such as limited resources and overcrowding, alongside sometimes inconsistent educational reforms and programs (Resh, 2008; Blass, 2020).

In the 2021–2022 school year, 25.9% of students studying in middle schools in Israel were Arabic speakers (Israeli Central Bureau of Statistics, 2022b). Large achievement gaps between Hebrew- and Arabic-speaking students have been documented across national and international assessments. For instance, in PISA 2018 (15-year-olds), Hebrew-speaking students in Israel scored above the OECD average in reading (506 vs. 487), whereas Arabic-speaking students scored substantially lower (362) (Blass, 2020). Similar gaps were also reported in PISA 2015 (Israeli Ministry of Education, 2016). National assessments (MEITZAV; administered in Grades 5 and 8) similarly indicate persistent gaps between Arab and Jewish students in literacy, mathematics, English, and science and technology (Israeli Ministry of Education, 2013a, 2013b; Knesset of Israel, 2013a, 2013b).

While in Hebrew-language education about 13% of the study hours are devoted to the field of language and literature, the scope of these hours in Arabic-language education is 27% (Addi-Raccah et al., 2011). However, in Arab middle schools, five hours a week are dedicated to the teaching of Arabic (language and literature) and four hours are dedicated to learning Hebrew. Thus, while in the Arab sector more hours are dedicated to language learning, they are divided almost equally between Arabic and Hebrew.

In and of itself, the Arabic language presents many challenges to its speakers due to its unique features (Saiegh-Haddad & Henkin-Roitfarb, 2014; Saiegh-Haddad, 2018). The most striking is the “diglossic phenomenon”, which describes the gap between the spoken language and the written language. The gap is reflected in vocabulary, phonology, grammar, syntax, linguistic forms, and forms of expression (Saiegh-Haddad, 2022; Eviatar & Ibrahim, 2014). Many studies investigating the effect of Arabic diglossia on the acquisition of reading and writing in Arabic claimed that Arabic-speaking students have great difficulty constructing accurate phonological and lexical representations for words in standard Arabic (Saiegh-Haddad & Lina, 2018). This slows down the decoding of the letters that represent phonemes that are heard but that do not exist in the spoken language (Ibrahim, 2010; Eviatar & Ibrahim, 2014; Saiegh-Haddad, 2003).

Another challenging feature of Arabic is its orthography, mainly the visual complexity of the letters themselves and the orthographic depth, in which the written representation of words does not always directly or accurately correspond with their pronunciation (Eviatar & Ibrahim, 2014). These features make reading difficult and therefore impair reading comprehension (Eviatar et al., 2004; Ibrahim et al., 2007).

The many hardships Arab middle schoolers in Israel face can put them at literacy risk. Therefore, students who already struggle with existing learning deficits and hardships, either in one area of learning or throughout all subjects, can accumulate further learning gaps that can make assimilating and progressing an almost impossible task. The Diagnostic and Statistical Manual of Mental Disorders (DSM) estimates that 5–15% of the population is characterized as learning disabled (American Psychiatric Association, 2013), highlighting the importance of relating and adapting learning to the intricate needs of these students both inside and outside the classroom.

## 2.4. The Current Study

The present study aimed to examine the effectiveness of “The Golden Triangle” blended-learning intervention program in developing Arabic-speaking middle school students’ academic vocabulary and reading comprehension of academic texts, while also testing whether the program’s effects differed between students with high/average reading comprehension (HRC) and those with low reading comprehension (LRC). This study expands on a previous study that used a control group to test the main effects of the intervention.

Taken together, the theoretical and empirical literature reviewed above highlights three main gaps that the present study seeks to address. First, despite the centrality of academic vocabulary to reading comprehension, relatively few intervention studies have targeted academic vocabulary in Arabic as a diglossic language, particularly in authentic middle-school content-area texts. Second, while blended-learning models have shown promise for promoting literacy, their differential effects on higher- and lower-achieving readers in Arabic remain underexplored. Third, most existing studies focus on short-term outcomes, with limited evidence regarding the sustainability of gains once interventions end. The current study addresses these gaps by implementing a blended-learning program (“The Golden Triangle”) that explicitly targets academic vocabulary and reading comprehension in Arabic, and by tracking both short-term and longer-term outcomes among high/average (HRC) and low-achieving (LRC) readers.

Research questions were as follows:

1. How do students’ reading comprehension skills change following participation in the intervention program, both immediately after the program (short term) and at follow-up (long term)?
2. To what extent do gains in receptive and productive academic vocabulary and reading comprehension differ between students with high/average reading comprehension (HRC) and those with low reading comprehension (LRC)?
3. To what extent do baseline academic vocabulary and reading comprehension predict long-term reading comprehension outcomes following the intervention, and does this relation differ by group?

## 3. Method

### 3.1. Participants

Seventy students who speak Arabic as their mother tongue participated in the study: 30 boys (42.9%) and 40 girls (57.1%). All major Arab groups in Israel were represented (Christian Arabs, Druzes and Muslims/Bedouins). Four Arab schools participated in the study, one class per school. Schools were located in three Arab cities and villages (Socioeconomic index: 2–5, [Israeli Central Bureau of Statistics, 2022a](#)) and one mixed Jewish–Arab city (Socioeconomic index: 7, [Israeli Central Bureau of Statistics, 2022a](#), although this can be attributed to the city’s mixed population). Because each of the four participating schools contributed a single class, the number of level 2 units (schools) was too small to support reliable multilevel modeling of school-level variance; we therefore focused our analyses on student-level variables.

Students were divided into two groups according to their teachers’ assessments of their reading comprehension performance at the beginning of the year. Teachers scored students’ performance on a scale ranging from 32 to 100 ( $M = 81.86$ ,  $SD = 15.99$ ,  $Md = 87.50$ ). Given the relatively small sample ( $N = 70$ ), we opted to classify students into higher/average and lower reading comprehension groups based on these global teacher ratings rather than on cut-offs derived from a single standardized test score. This approach was intended to capture teachers’ holistic views of students’ day-to-day reading comprehension across the curriculum, while also allowing for the formation of two roughly equal-sized groups for

analysis. The division cut-off was the median. Pearson correlation showed a significant positive correlation ( $r = 0.25$ ,  $p < 0.05$ ) between the teachers' assessments and students' scores on the reading comprehension test before the intervention, corroborating the validity of using teacher evaluation for group division.

The first group included 34 students (20 boys and 14 girls) characterized as students with low reading comprehension ability (LRC). The second group included 36 students (10 boys and 26 girls) characterized as students with high or average reading comprehension ability (HRC). A significant difference was found in the gender distribution between the two groups of students ( $\chi^2(1) = 6.88$ ,  $p < 0.01$ ), indicating more girls in the HRC group compared to the LRC group.

### 3.2. Research Tools

#### 3.2.1. The "Golden Triangle" Intervention Program

The "Golden Triangle" intervention program is based on three main components, serving as the cornerstones of the triangle: the acquisition and strengthening of academic vocabulary, blended learning, and digital literacy.

**Academic Vocabulary.** The program included ten units aimed at promoting the academic literacy of middle school students by addressing three main areas (reading, writing and linguistic knowledge), while using academic vocabulary and informational texts from different subjects. The materials of the program were developed in the Arabic Section, Literacy and Language Arts Department at the Center for Educational Technology (CET). The goals of the program and the developed content were adapted to the Arabic language curriculum for middle school (Israeli Ministry of Education, 2013c). Texts learned in the program were taken from textbooks of text-heavy subjects studied in the seventh grade and processed according to the goals of the program and the abilities of the students. Each unit included five sub-units: introductory unit, reading unit, linguistic knowledge unit, writing unit, and summary unit (see Table 1).

**Table 1.** Description of the five sub-units in the intervention program.

| Sub-Unit | Goal                 | Description                                                                                                                                                                                                                                                    |
|----------|----------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1        | Introduction         | Activities designed to deepen morphological, semantic, and syntactic knowledge about the roots were given to students. These exercises included root introduction and then identification, application, and production.                                        |
| 2        | Reading              | Students read an informational text containing academic words belonging to the five roots that they learned at the beginning of the unit and in previous units. Then, they were asked to identify the words within the text and asked comprehension questions. |
| 3        | Linguistic knowledge | Students practiced morphological and syntactic issues in accordance with the curriculum and in relation to the text.                                                                                                                                           |
| 4        | Writing              | Students were given examples of correctly written texts, from which the learned academic words have been omitted. Students were asked to fill in the correct words. Then, they were asked to write a text individually.                                        |
| 5        | Summary              | Students were administered tasks aimed at testing and evaluating the depth and breadth of the academic vocabulary acquired so far in the program. Teachers checked these tasks according to a guidance document built by the intervention developers.          |

Through the course of ten units, students were presented with the 55 most common academic roots in the Arabic middle school textbooks in Israel, as mapped by Makhoul and Olshtain (2015). The learning of roots and words in the program was cyclical and cumulative; the roots learned in the previous unit were also included in the new unit.

The units and sub-units were fixed and structured. At the beginning of each sub-unit, students encountered the new roots they would learn and deepened their familiarity with them and with the words built around them, including different morphological patterns for each basic root. At the end of the unit there was a dictionary containing all of the

academic words in the unit (including interpretations in several contexts and examples). Each sub-unit ended with reflective questions about the learning process. Delivery of each unit in Arabic language classes took about a week.

The study units were accompanied by a teacher's guide detailing the theoretical and structural principles behind them. It provided a summary of the unit's themes, an explanation of the unit's structure, the literacy skills addressed in the unit, a dictionary for new words, suggestions for lesson structure; and a guiding document for the teacher, assisting them in evaluating questions that follow the texts and activities.

**Digital literacy.** The digital and technological aspect of learning was achieved using "Ofek", an educational platform developed by the Center for Educational Technology (CET) for the Israeli Ministry of Education. "Ofek" contains interactive activities, study units, videos and games, divided by subjects and grade levels. It uses different modes of learning: learning through seeing, learning through listening, learning through reading and writing, and learning through doing and experiencing. It allows the teacher to independently build new units and assign individual and group tasks to their students. Ofek can be accessed by both teachers and students in the school and at home.

**Blended learning.** The "Golden Triangle" program was based on the Station-Rotation Model (Staker & Horn, 2012), in which students rotate between several learning stations during the same lesson: a frontal learning station with the teacher, a small group learning station, and an individual learning station using a computer. The individualized-computerized aspect supplied each student with the ability to learn in their own pace while encouraging motivation and engagement via playfulness and content accessibility. The small groups assisted the teacher in acknowledging and adapting their teaching to the different needs of each student in the classroom. The use of the "Ofek" website assisted in developing a more personalized teaching approach and enabled the teacher to adapt their teaching to students in need of more support, while also enabling the easier monitoring of students' progress.

### 3.2.2. Assessment Tests

**Receptive Academic Vocabulary Test (RAVT).** This test was adapted from the Vocabulary Levels Test (VLT) (Schmitt et al., 2001). To construct the test, the 55 most common academic roots in Arabic textbooks in Israeli middle schools were used (Makhoul & Olshtain, 2015). The test consisted of 46 matrices; each matrix was divided into two columns. In the right column were three academic words that served as the target words next to three words that were distractors. In the left column three definitions were presented; students were asked to match the three definitions with the three target words on the right side. All words were matched in their syntactic function (nouns, verbs, and adjectives) and in their inflected morphology to avoid morphological bias toward the target words. Each correct answer earned the student one point, with a possible score of three points for each matrix; the highest possible score for this test was 138. The test was administered twice: pre and post intervention. The internal consistency of the test pre-intervention ( $\alpha = 0.95$ ) and post-intervention ( $\alpha = 0.86$ ) was high.

**Productive Academic Vocabulary Test (PAVT).** This test examined the application of academic vocabulary in context and was adapted to Arabic based on the VLT (Schmitt et al., 2001). In this test, the students had to complete 137 sentences according to a given root and according to context. Due to the rich morphological structure of the Arabic language, students were shown the first suffix of the word in order to direct them to the desired response, and to prevent confusion with other words. All the words that were omitted from the sentences were constructed according to the common academic roots in the Arabic textbooks in Israeli middle schools (Makhoul & Olshtain, 2015). Correct

answers awarded students one point, while wrong answers received zero points. The highest possible score for this test was 137. The test was administered twice: pre and post intervention. The internal consistency of the test was found to be high (pre-intervention:  $\alpha = 0.95$ ; post-intervention:  $\alpha = 0.93$ ).

Academic Reading Comprehension Tests—ARCT7, ARCT8. Reading comprehension tests were administered three times: before the beginning of the intervention in the seventh-grade school year, after the intervention (Grade 7), and at the beginning of the eighth grade. Students received an informational text, after which they were required to answer comprehension questions (Grade 7 assessments: 19 questions—2 open-ended and 17 closed-ended; Grade 8 assessment: 21 questions—3 open-ended and 18 closed-ended). The texts consisted of 815 words for the seventh-grade assessments and 1166 words for the eighth-grade assessment and were aligned with the goals of the Arabic language curriculum for middle school (Israeli Ministry of Education, 2013c): understanding the visible layer of the text; interpretation and inference; evaluation and criticism; and meta-textual understanding. Closed-ended questions were scored as 0 (incorrect) or 1 (correct), whereas open-ended questions were scored as 0 (incorrect), 1 (partial answer), or 2 (correct answer). Cronbach's alphas were high ( $\alpha = 0.92$ – $0.93$ ).

### 3.3. Research Procedure

This study was approved by the Chief Scientist's Office, Ministry of Education, Israel (research permit no. 9675, approved on 4 September 2017). Prior to the beginning of the program, teachers participated in four group training sessions aimed at familiarizing them with the program, its principles and its contents and teaching materials.

The intervention program took place in the first school semester. The PAVT, RAVT, and ARCT7 tests were administered before (T1) and after (T2) the start of the intervention program in the seventh grade. The ARCT8 reading comprehension test was administered at the beginning of the eighth grade (T3). All tests were administered to the entire class via a computer platform. The research team performed three observations to assure the teachers followed the program's protocol. The first observation was conducted in the program's first week, the second observation was conducted in the middle of the program (12–13 weeks), and the last one was conducted two weeks prior to the end of the program. It should also be noted that during the program each teacher had personal meetings with the researchers to discuss and assist with hardships arising from program implementation.

Implementation monitoring and dosage of the intervention are summarized in Table 2.

**Table 2.** Implementation monitoring and dosage of the intervention program.

| Monitoring Element  | Description/Frequency                                                                                                                                                                             |
|---------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Intervention dose   | Ten instructional units delivered during the first semester (approximately one unit per week) as part of the regular Arabic language instruction.                                                 |
| Teacher preparation | Four group training sessions prior to implementation, focusing on program principles, instructional routines, and materials.                                                                      |
| Fidelity monitoring | The research team conducted three classroom observations during the intervention to monitor adherence to the station-rotation structure and instructional routines.                               |
| Technology use      | Students used the “Ofek” platform during the individualized digital-learning station; platform usage logs were available but were not systematically extracted for analysis in the present study. |

## 4. Results

Preliminary analyses indicated that most of the dependent variables were not normally distributed. Therefore, for each main comparison we conducted both parametric analyses and non-parametric robustness checks. Specifically, we used Mann–Whitney U tests for between-group comparisons at single time points and Wilcoxon signed-rank/Friedman tests for within-group comparisons across time (depending on the number of time points). In all cases, the non-parametric tests yielded the same pattern of statistically significant and non-significant effects as the parametric analyses. For the sake of parsimony, we therefore report only the parametric results below.

### 4.1. Differences in Test Scores at the Beginning of the Seventh Grade According to Group (T1)

Due to significant correlations between the scores on the three tests ( $r = 0.91$ ,  $p < 0.001$  between RAVT and PAVT;  $r = 0.48$ ,  $p < 0.001$  between RAVT and ARCT7;  $r = 0.54$ ,  $p < 0.001$  between PAVT and ARCT7), a one-way MANOVA was conducted to examine the group differences in the scores on the three tests before the intervention. The results of the MANOVA indicated a significant difference in the performance on the tests between the HRC and LRC groups ( $F(3, 66) = 16.59$ ,  $p < 0.001$ ,  $\eta^2 = 0.43$ ). In order to examine in which tests the two groups differed, an additional one-way ANOVA for each test was conducted. Cohen's  $d$  was also calculated in order to examine the effect size of the differences between the two groups on the three dependent variables (see Table 3).

**Table 3.** Mean and SD of scores in the RAVT, PAVT and ARCT7 at T1 (N = 70).

|       | LRC (n = 34) |       | HRC (n = 36) |       | F         | p      | $\eta^2$ | d [95% CI]        |
|-------|--------------|-------|--------------|-------|-----------|--------|----------|-------------------|
|       | M            | SD    | M            | SD    |           |        |          |                   |
| RAVT  | 42.84        | 25.94 | 75.48        | 12.54 | 45.48 *** | <0.001 | 0.40     | 1.60 [1.07, 2.15] |
| PAVT  | 45.17        | 23.20 | 73.86        | 10.36 | 45.74 *** | <0.001 | 0.40     | 1.60 [1.07, 2.15] |
| ARCT7 | 44.29        | 27.23 | 59.75        | 30.70 | 4.95 *    | 0.03   | 0.07     | 0.53 [0.05, 1.01] |

\*  $p < 0.05$ , \*\*\*  $p < 0.001$ . RAVT = Receptive academic vocabulary test; PAVT = Productive academic vocabulary test; ARCT7 = Academic reading comprehension test (7th grade).

As Table 3 shows, the scores on the three tests were significantly higher in the HRC group compared to the LRC group, indicating substantial differences between students whose teachers assessed them as having low reading comprehension and students whom their teachers assessed as having high or average reading comprehension. The very large effect sizes for the vocabulary measures (RAVT:  $d = 1.60$ , 95% CI [1.07, 2.15]; PAVT:  $d = 1.60$ , 95% CI [1.07, 2.15]) indicate that LRC students entered the intervention with markedly weaker academic vocabulary knowledge. This pattern further supports the validity of using teachers' assessments as a basis for group division in the present sample.

### 4.2. Differences in Reading Comprehension Scores by Study Group Across Three Time Points

In order to examine the differences between the two study groups according to the three time points, two-way mixed ANOVA was conducted. The independent variables were Group (as a between-subjects factor) and Time (as a within-subjects factor). The dependent variable was the reading comprehension scores.

The main effects of Group and Time were significant [ $F(2, 67) = 5.45$ ,  $p = 0.001$ ,  $\eta^2 = 0.07$  and  $F(2, 67) = 27.56$ ,  $p < 0.001$ ,  $\eta^2 = 0.47$ , respectively] as well as the interaction between Group and Time [ $F(2, 67) = 4.67$ ,  $p = 0.013$ ,  $\eta^2 = 0.12$ ]. In order to examine the source of the interaction, a one-way repeated measures ANOVA was conducted to explore the differences between the three time points in each study group. Significant differences between the three time points were found in both study groups [ $F(2, 32) = 26.64$ ,

$p < 0.001$ ,  $\eta^2 = 0.63$  for the LRC group and  $F(2, 34) = 6.80$ ,  $p < 0.01$ ,  $\eta^2 = 0.29$  for the HRC group]. Bonferroni analyses indicated that both study groups increased their performance on the ARCT7 after the intervention (T2) compared to their performance on the ARCT7 before the intervention (T1). However, while the performance on the ARCT8 was stable at the beginning of the eighth grade (T3) compared to T2 in the HRC group, a significant decrease was found in ARCT8 performance at T3 compared to T2 in the LRC group ( $p < 0.001$ ). The results also indicated that both study groups increased their performance on the ARCT8 at T3 compared to their performance on the ARCT7 at T1 ( $p < 0.05$ ) (see Figure 1). Taken together, these results indicate that the Group  $\times$  Time interaction ( $\eta^2 = 0.12$ ) represents a moderate effect. In practical terms, both groups benefited from the intervention in the short term, as reflected in higher ARCT7 scores at T2 relative to T1. However, only the HRC group maintained these gains into eighth grade, whereas the LRC group showed a pattern of initial improvement followed by a marked decline once the program ended. Thus, the intervention produced more durable benefits for higher/average readers, while only partially closing the gap for struggling readers.

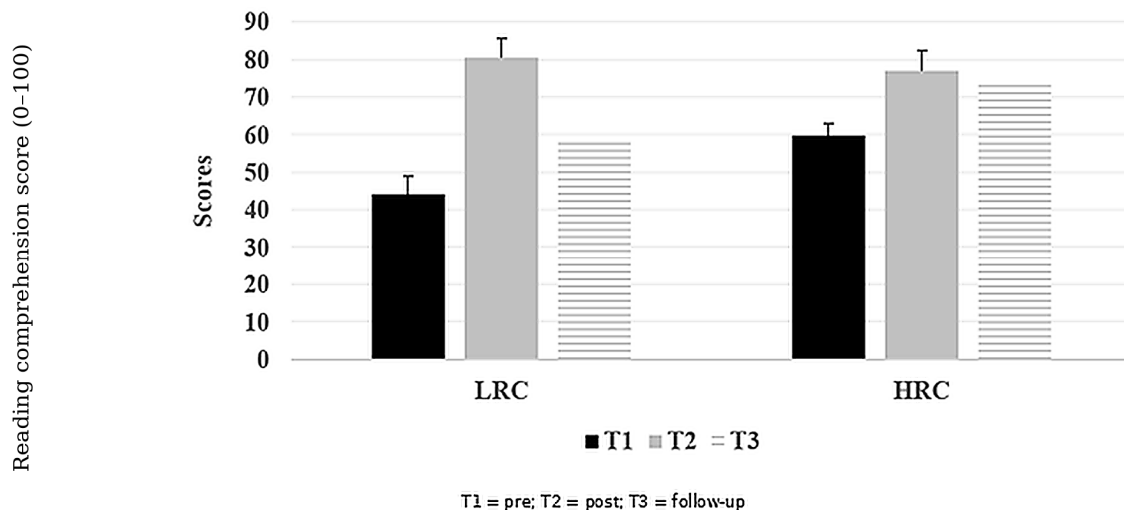


**Figure 1.** Mean ( $\pm$ SE) reading comprehension scores (ARCT7/ARCT8) for low reading comprehension (LRC) and high/average reading comprehension (HRC) groups across the three time points (T1–T3).

#### 4.3. The Contribution of RAVT and PAVT Scores to the Prediction of ARCT8 Scores (T3)

To examine the contributing factors for the prediction of the prolonged literacy effects of the program, gender and baseline performance on the RAVT and PAVT was examined as predictors of performance on the ARCT8 at the beginning of the eighth grade. Three hierarchical regression analyses were conducted, one analysis for the prediction of ARCT8 test scores at T3 among the overall sample and two analyses for the prediction of ARCT8 test scores at T3 in each study group. In the first step, gender was entered. In the second step, the RAVT and PAVT performance at T1 was entered (see Table 4).

Table 4 indicates that gender contributed a significant 6.4% to the explained variance (EPV) of the ARCT8 scores, with higher ARCT8 performance among girls compared to boys. The performances at T1 on the RAVT added a significant 13.2% to the EPV beyond the participants' gender to the prediction of ARCT8 scores in the overall sample. However, the PAVT scores did not contribute significantly to the EPV of the ARCT8 performance.

Regression analyses conducted in each study group revealed that performance at T1 on the RAVT added a significant 33% to the EPV beyond the participants' gender to the prediction of the performance on the ARCT8 among the HRC group. However, in the

LRC group neither gender nor the RAVT scores contributed significantly to the EPV of the performance on the ARCT8.

**Table 4.** Results of the hierarchical regressions for the performance on the ARCT8 by participants' gender and the performance on T1 RAVT and PAVT.

|                               | Steps | Independent Variables | B    | SE.B | $\beta$ | T        | R <sup>2</sup> | $\Delta R^2$ |
|-------------------------------|-------|-----------------------|------|------|---------|----------|----------------|--------------|
| Overall sample<br>(df = 2.67) | 1     | Gender                | 7.38 | 3.41 | 0.25    | 2.16 *   | 0.06 *         | 0.06 *       |
|                               | 2     | Gender                | 2.99 | 3.45 | 0.10    | 0.87     |                |              |
|                               |       | RAVT (T1)             | 0.22 | 0.07 | 0.39    | 3.31 *** | 0.19 ***       | 0.13 ***     |
| LRC (df = 1.32)               | 1     | Gender                | 0.01 | 5.03 | 0.00    | 0.00     | 0.000          | 0.000        |
| HRC (df = 2.33)               | 1     | Gender                | 6.42 | 3.89 | 0.27    | 1.65     | 0.07           | 0.07         |
|                               | 2     | Gender                | 4.93 | 3.29 | 0.21    | 1.50     |                |              |
|                               |       | RAVT (T1)             | 0.47 | 0.12 | 0.55    | 3.92 *** | 0.36 ***       | 0.33 ***     |

\*  $p < 0.05$ , \*\*\*  $p < 0.001$ . LRC = Low reading comprehension group; HRC = High reading comprehension group; RAVT = Receptive academic vocabulary test; PAVT = Productive academic vocabulary test; ARCT8 = Academic reading comprehension test (8th grade).

#### 4.4. Moderation Analysis

Following the regression results, we conducted a moderation analysis using the PROCESS macro (Hayes, 2013) to test whether the group (LRC/HRC) moderated the association between baseline receptive academic vocabulary (RAVT; T1) and follow-up reading comprehension (ARCT8; T3). The analysis revealed a significant Group  $\times$  RAVT interaction,  $F(1, 66) = 7.21$ ,  $p < 0.001$ ,  $R^2 = 0.07$ , indicating that the association between RAVT at T1 and ARCT8 at T3 differed between groups. For the HRC group, higher RAVT scores were associated with higher ARCT8 scores ( $b = 0.48$ ,  $t(61) = 3.03$ ,  $p < 0.01$ ), whereas in the LRC group, the association was not significant ( $b = 0.00$ ,  $t(61) = 0.06$ ,  $p = 0.94$ ).

## 5. Discussion and Conclusions

The present study examined the effects of an intervention program aimed at fostering academic language literacy among seventh-grade students who speak Arabic as their mother tongue, while also examining the differences in intervention efficacy between students with low and high/average reading comprehension abilities. Across the three time points, both groups showed short-term gains in reading comprehension following the intervention; however, their trajectories diverged in the longer term. Among students with low reading comprehension abilities, an improvement was found from the first time point to the second, but a significant decrease was found at the third time point, although their scores at T3 still exceeded those at T1. Among the students with high reading comprehension skills, an improvement was observed from the first time point to the second and was also maintained at the third time point. From this point of view, it seems that the intervention program was successful for both groups, but its longer-term effects were more beneficial to students with high or average reading abilities.

These findings underline the different effects that intervention programs have on students with different abilities. In addition, the findings indicate the need for long-term, specialized intervention for students with low reading skills, as the gaps between skilled students and struggling students widen with each year of school (Fryer & Levitt, 2004, 2006).

The strength of the intervention program described in the current study stems from several sources. The first is related to the characteristics of Arabic-speaking learners in Israel. As traditional teaching approaches do not necessarily facilitate the building of

knowledge for all students in heterogeneous classes and may widen the gap between students (Valiande, 2010), the adaptation made in the “Golden Triangle” intervention to the Arab middle schoolers’ needs was imperative. Specifically, the “Golden Triangle” intervention program was built based on findings originating from that population highlighting the unique Arab academic lexicon prevalent in the Arab education system in Israel (see Makhoul, 2017; Makhoul et al., 2018).

The second strength of the current intervention, which allowed teachers to address the heterogeneity of students’ needs, is the different teaching setup made possible thanks to the blended-learning model, and specifically to the Station-Rotation Model. Teachers who participated in the program were able to create small groups of students according to each student’s personal characteristics. In addition, the intervention allowed the teachers to use innovative teaching methods and provide the students with personalized learning tools that allow the teacher and students in tandem to create a new learning experience. Tomlinson (2003) explains that teachers who changed their teaching system according to the needs of the students and their characteristics (e.g., socio-economic status, learning profile, reading levels, interest) have a positive impact on the students’ ability to acquire new information.

The current study highlighted differential trajectories of intervention effects as a function of students’ initial reading ability. While higher/average readers (HRC) showed improvements that remained stable through the follow-up measurement in eighth grade, this was not the case for the less proficient readers (LRC), whose gains faded over time. Moreover, only in the HRC group did baseline receptive academic vocabulary predict reading comprehension at follow-up, whereas no similar predictors emerged for the LRC group. For lower-achieving students, the program activities designed to cultivate academic vocabulary and reading comprehension did not yield sustained effects, underscoring the need for future research to identify factors and supports that facilitate long-term consolidation among struggling readers.

The question of recommended intervention length for effective reading outcomes has been widely discussed previously, with mixed conclusions. Moreover, findings focusing solely on middle schoolers are scarce. Vaughn et al. (2012) suggested that increasing the duration of intervention may lead to more useful results, while a large-scale meta-analysis conducted by Suggate (2010) found no added effect for program duration for at-risk readers. A meta-analysis conducted by Wanzek et al. (2018) suggested that it is not necessarily the number of overall hours dedicated to early intervention that is important, but rather giving an extended number of sessions for instruction. A similar meta-analysis conducted on upper elementary school struggling readers (Donegan & Wanzek, 2021) found that longer interventions did succeed in predicting significant, albeit small, effects for reading comprehension outcomes. These mixed findings suggest that overall duration alone may not be sufficient, and that, especially for struggling readers, the distribution and intensity of support and explicit reinforcement of skills after the formal program ends may be critical. Therefore, it may be the case that in the current study, a more intense or longer implementation of the “Golden Triangle” program, possibly with structured follow-up support, could have assisted in advancing the struggling students in the longer term and in achieving patterns more similar to their more skilled peers.

The development of the intervention examined in the current study took into account two major aspects: the first is the added literacy risk factor presented by the unique features of the Arabic language, and the second is the added advantage that a blended, interactive learning model can achieve. The Arabic language incorporates many unique features that can impose challenges for its speakers and readers (Saiegh-Haddad & Henkin-Roitfarb, 2014; Saiegh-Haddad, 2018), and this may warrant a different approach to literacy

interventions in Arabic. A cross-sectional study examined the contribution of different language elements (i.e., phonology, orthography, morphology, syntax, etc.) to the two main components of reading: decoding and fluency (Asadi et al., 2017). The study found that while some of the language elements successfully affect both decoding and fluency, other elements affect only one or the other, or neither. Thus, it can be concluded that in building intervention for struggling Arabic readers, who are at a higher risk than average- or high-achieving students, more thought should be given to the way each language element is treated in the intervention, in order to alleviate some of the hardships already presented by Arabic itself.

The use of blended learning was a key cornerstone of the “Golden Triangle” intervention, and especially its focus on media and technology as tools for promoting student engagement and involvement. However, it can be suggested that encouraging these socio-emotional constructs is not equally effective in readers with different ability levels. Klauda and Guthrie (2015) performed a longitudinal study comparing struggling and advanced readers in regard to the ability of reading motivation and engagement to promote reading achievement in early adolescence. They found that both motivation and engagement predicted achievement more strongly for advanced readers than struggling readers, despite the fact that both groups did not necessarily differ in their willingness to engage and put effort into learning. The authors suggest that due to their stronger cognitive functioning, skilled readers have a greater capacity for change in reading achievement, whereas struggling readers’ capacity for such change may be limited by weaker cognitive abilities. Therefore, the ability of their engagement and motivation to generate robust reading progress is more limited. Our findings are consistent with this account: although LRC students participated in the same blended, engaging activities and showed short-term progress, their weaker baseline skills and the linguistic challenges of Arabic may have limited their capacity to consolidate and generalize the newly acquired academic vocabulary to independent reading over time.

In conclusion, the current research suggests the need for intervention programs that are thought out and adapted to the unique needs of the Arab middle-schooler. Similarly to other studies, the research findings show the ability of vocabulary to distinguish between skilled readers and struggling readers. However, despite the improvement in the students’ knowledge of vocabulary, the students with low reading abilities still do not use their newly gained vocabulary knowledge in order to increase their achievement in reading comprehension in the longer term. This suggests the need to take each student’s personal difficulties into account when trying to teach and acquire literacy in the Arab classroom, in order to promote literacy in a more sustainable and consistent manner.

There are several limitations to the current study. First, the classification of students into the HRC and LRC groups was based on the median of teachers’ personal assessments, due to the small sample size. Although we were able to validate these assessments via their correlation with test scores and the robust differentiation between the two groups on the T1 measures, using a median split on teacher ratings is a relatively coarse procedure. Future studies with larger samples should rely on more sensitive, test-based cut-offs (e.g., PAVT, RAVT, ARCT7) and/or model baseline performance as a continuous moderator rather than dichotomizing it, in order to identify struggling readers more precisely and to test the robustness of the present findings. Second, the overall sample size and the number of participating schools were modest, limiting statistical power and generalizability; because each school contributed a single class, the number of level 2 units was too small to support reliable multilevel modeling of school-level variance. As a result, potential contextual influences of school characteristics on students’ outcomes could not be examined. Third, the study did not include a no-intervention control group, which limits

causal inferences about short-term gains and longer-term maintenance (e.g., maturation or concurrent instruction may have contributed to outcomes). Fourth, the gender distribution differed between groups; although gender was included as a covariate in the predictive analyses, future research with larger and more balanced samples is needed to disentangle intervention effects from gender-related variability in literacy development. Finally, while the intervention relied on the “Ofek” platform for individualized digital practice, usage logs were not systematically analyzed in the present study; future work should incorporate dosage measures to better link individual exposure to outcomes.

**Author Contributions:** Conceptualization, B.M. and E.O.; Methodology, B.M. and E.O.; Investigation, B.M. and E.O.; Formal analysis, B.M. and E.O.; Data curation, B.M.; Writing—original draft preparation, B.M.; Writing—review and editing, B.M. and E.O.; Project administration, B.M.; Supervision, B.M. and E.O. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Chief Scientist’s Office, Ministry of Education, Israel.

**Institutional Review Board Statement:** The study was approved by the Chief Scientist’s Office, Ministry of Education, Israel (research permit no. 9675, approved on 4 September 2017).

**Informed Consent Statement:** Written informed consent was obtained from participating teachers, and parental/guardian consent for student participation was obtained in accordance with the approved research protocol.

**Data Availability Statement:** The datasets generated and/or analyzed during the current study are not publicly available due to privacy, confidentiality, and ethical restrictions related to research involving minors.

**Conflicts of Interest:** B.M. is employed by the Center for Educational Technology (CET) and serves as Director of the Arabic Section in the Literacy and Language Arts Department. The instructional units used in this study were implemented on CET’s Ofek platform. E.O. declares no conflict of interest. The authors declare no other conflicts of interest.

## References

- Addi-Raccah, A., Biran, H., & Freedman-Goldberg, S. (2011). *Middle schools: Characteristics and challenges*. Review Commissioned by the Language and Literacy Committee, The Initiative for Applied Education Research, School of Education, Tel Aviv University. Available online: [https://education.tau.ac.il/sites/education.tau.ac.il/files/media\\_server/education/dl/mediniut\\_veminhal/dochot\\_mechkar/Language-middle-school-review-Audrey-Heb.pdf](https://education.tau.ac.il/sites/education.tau.ac.il/files/media_server/education/dl/mediniut_veminhal/dochot_mechkar/Language-middle-school-review-Audrey-Heb.pdf) (accessed on 10 March 2026).
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Association.
- Asadi, I. A., Khateb, A., Ibrahim, R., & Taha, H. (2017). How do different cognitive and linguistic variables contribute to reading in Arabic? A cross-sectional study from first to sixth grade. *Reading and Writing*, 30, 1835–1867. [CrossRef]
- Bailey, A. L. (Ed.). (2007). Introduction. In *The language demands of school: Putting academic English to the test* (pp. 1–26). Yale University Press.
- Baumann, J. F., & Graves, M. F. (2010). What is academic vocabulary? *Journal of Adolescent and Adult Literacy*, 54(1), 4–12. [CrossRef]
- Beck, I. L., McKeown, M. G., & Kucan, L. (2002). *Bringing words to life: Robust vocabulary instruction*. Guilford Press.
- Blass, N. (2020). *Achievements and gaps: The status of the Israeli education system*. Policy paper No. 01.2020. Taub Center for Social Policy Studies in Israel. Available online: <https://www.taubcenter.org.il/wp-content/uploads/2020/12/achievementsangapseng-1.pdf> (accessed on 10 March 2026).
- Carlo, M. S., August, D., McLaughlin, B., Snow, C. E., Dressler, C., Lippman, D. N., Lively, T. G., & White, C. E. (2004). Closing the gap: Addressing the vocabulary needs of English-language learners in bilingual and mainstream classrooms. *Reading Research Quarterly*, 39(2), 188–215. [CrossRef]
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213–238. [CrossRef]
- Cummins, J., & Yee-Fun, E. M. (2007). Academic language: What is it and how do we acquire it? In J. Cummins, & C. Davison (Eds.), *International handbook of English language teaching* (pp. 797–810). Springer. [CrossRef]
- Donegan, R. E., & Wanzek, J. (2021). Effects of reading interventions implemented for upper elementary struggling readers: A look at recent research. *Reading and Writing*, 34(8), 1943–1977. [CrossRef] [PubMed]

- Eviatar, Z., & Ibrahim, R. (2014). Why is it hard to read Arabic? In E. Saiegh-Haddad, & M. Joshi (Eds.), *Handbook of arabic literacy: Insights and perspectives* (pp. 77–96). Springer. [CrossRef]
- Eviatar, Z., Ibrahim, R., & Ganayim, D. (2004). Orthography and the hemispheres: Visual and linguistic aspects of letter processing. *Neuropsychology*, 18(1), 174–184. [CrossRef]
- Ferdig, R. E., & Kennedy, K. (2018). *Handbook of research on K-12 online and blended learning* (2nd ed., pp. 1–516). ETC Press.
- Freyer, D. A., Frederick, W. C., & Klausmeier, H. J. (1969). *A schema for testing the level of concept mastery (technical report No. 16)*. Wisconsin Research and Development Center for Cognitive Learning, University of Wisconsin. Available online: <https://books.google.com/books?id=pSRFGwAACAAJ> (accessed on 10 March 2026).
- Fryer, R. G., Jr., & Levitt, S. D. (2004). Understanding the black-white test score gap in the first two years of school. *Review of Economics and Statistics*, 86(2), 447–464. [CrossRef]
- Fryer, R. G., Jr., & Levitt, S. D. (2006). The black-white test score gap through third grade. *American Law and Economics Review*, 8(2), 249–281. [CrossRef]
- Guthrie, J. T., Wigfield, A., & Klauda, S. L. (2012). *Adolescents' engagement in academic literacy*. University of Maryland. Available online: <https://region7comprehensivecenter.org/wp-content/uploads/2020/07/4.0-CORI-Adolescents%27-Engagement-in-Academic-Literacy.pdf> (accessed on 10 March 2026).
- Hayes, A. F. (2013). Mediation, moderation, and conditional process analysis. In A. F. Hayes (Ed.), *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Publications.
- Heller, R., & Greenleaf, C. (2007). *Literacy instruction in the content areas: Getting to the core of middle and high school improvement*. Alliance for Excellent Education. Available online: <https://pdfs.semanticscholar.org/96e1/19757b1d6b26e296b13f9ba4181e822bf79f.pdf> (accessed on 10 March 2026).
- Hurwitz, L. B., & Macaruso, P. (2021). Supporting struggling middle school readers: Impact of the Lexia® PowerUp Literacy® program. *Journal of Applied Developmental Psychology*, 77, 101329. [CrossRef]
- Ibrahim, R. (2010). Psycholinguistic challenges in processing the Arabic language. In W. F. Caldwell (Ed.), *Bilinguals: Cognition, education and language processing* (pp. 81–102). Nova Science Publishing.
- Ibrahim, R., Eviatar, Z., & Aharon-Peretz, J. (2007). Metalinguistic awareness and reading performance: A cross language comparison. *The Journal of Psycholinguistic Research*, 36(4), 297–317. [CrossRef]
- Israeli Central Bureau of Statistics. (2022a). *Characterization and classification of geographical units by the socio-economic level of the population 2019*. Available online: [https://www.cbs.gov.il/he/mediarelease/DocLib/2022/375/24\\_22\\_375b.pdf](https://www.cbs.gov.il/he/mediarelease/DocLib/2022/375/24_22_375b.pdf) (accessed on 10 March 2026).
- Israeli Central Bureau of Statistics. (2022b). *Schools, classes and students in secondary education*. Available online: [https://education.tau.ac.il/sites/education.tau.ac.il/files/media\\_server/education/dl/mediniut\\_vemi...](https://education.tau.ac.il/sites/education.tau.ac.il/files/media_server/education/dl/mediniut_vemi...) (accessed on 10 March 2026).
- Israeli Ministry of Education. (2013a). *Madadei ya'ilut v'tzmikha beit sifrit meitzav hatashag: Mimitzai'im ikari'im* [School efficiency and growth indicators of the 2013 (Jewish year 5773) Meitzav–Main findings]. The National Authority for Measurement and Evaluation in Education (RAMA). Available online: [https://meyda.education.gov.il/files/rama/2013/Report\\_Hesegim\\_2013.pdf](https://meyda.education.gov.il/files/rama/2013/Report_Hesegim_2013.pdf) (accessed on 10 March 2026).
- Israeli Ministry of Education. (2013b). *PISA 2012: Orianut b'kerev talmidim b'nei 15 b'madaim, b'kriah u'b'matematika: Mabat yisraeli* [PISA 2012: Literacy among 15-year-old students in science, reading and mathematics: An Israeli perspective]. The National Authority for Measurement and Evaluation in Education (RAMA). Available online: [https://meyda.education.gov.il/files/Rama/PISA\\_2012\\_Report\\_ed\\_f\\_nh.pdf](https://meyda.education.gov.il/files/Rama/PISA_2012_Report_ed_f_nh.pdf) (accessed on 10 March 2026).
- Israeli Ministry of Education. (2013c). *Tochnit Halimudim BeAravit: Safa, havanat hanikra vebitui bachinuch aAravii vebachinuch ha Druzi-Kitot 7–12* [Arabic curriculum: Language, comprehension and expression in Arab education and Druze education 7th grade–12th grade]. Ministry of Education, Pedagogical Office, Department of Languages and Supervision of Arabic Instruction in Arab and Druze Schools.
- Israeli Ministry of Education. (2016). *PISA 2015: Orianut b'kerev talmidim b'nei 15 b'madaim, b'kriah u'b'matematika: Mabat yisraeli* [PISA 2015: Literacy among 15-year-old students in science, reading and mathematics: An Israeli perspective]. The National Authority for Measurement and Evaluation in Education (RAMA). Available online: [https://meyda.education.gov.il/files/Rama/PISA\\_2015\\_Full\\_Report.pdf](https://meyda.education.gov.il/files/Rama/PISA_2015_Full_Report.pdf) (accessed on 10 March 2026).
- Kazakoff, E. R., Macaruso, P., & Hook, P. (2018). Efficacy of a blended learning approach to elementary school reading instruction for students who are English learners. *Education Technology Research Development*, 66, 429–449. [CrossRef]
- Klauda, S. L., & Guthrie, J. T. (2015). Comparing relations of motivation, engagement, and achievement among struggling and advanced adolescent readers. *Reading and Writing*, 28, 239–269. [CrossRef] [PubMed]
- Knesset of Israel. (2013a). *Nitunim al ma'arekhet hakhinukh bamigzar ha'aravi* [Data on the education system in the Arab sector]. Research and Information Center. Available online: <https://main.knesset.gov.il/activity/info/research/pages/incident.aspx?businessstype=1&docid=29556b58-e9f7-e411-80c8-00155d010977&ver=2> (accessed on 10 March 2026).

- Knesset of Israel. (2013b). *Nitunim al ma'arekhet hakhinukh bamizgar habedu'i* [Data on the education system in the Bedouin sector]. Research and Information Center. Available online: <https://main.knesset.gov.il/activity/info/research/pages/incident.aspx?businessstype=1&docid=cd536b58-e9f7-e411-80c8-00155d010977&ver=2> (accessed on 10 March 2026).
- Lawrence, J., Capotosto, L., Branum-Martin, L., White, C., & Snow, C. (2012). Language proficiency, home-language status, and English vocabulary development: A longitudinal follow-up of the Word Generation program. *Bilingualism: Language and Cognition*, 15(3), 437–451. [CrossRef]
- Lesaux, N. K., Kieffer, M. J., Faller, S. E., & Kelley, J. G. (2010). The effectiveness and ease of implementation of an academic vocabulary intervention for linguistically diverse students in urban middle schools. *Reading Research Quarterly*, 45(2), 196–228. [CrossRef]
- Makhoul, B. (2017). *Tochnit hitarvut lekidum lexicon academi vehashpaata ak kishurim shonim-oryaniyim shel talmidim Aravim bechativat habeynayim* [An intervention program to promote the academic lexicon and its influence on the language skills of Arab students in junior high school]. Mofet Institute, The Center for Educational Technology (Metach) and the Hebrew University of Jerusalem.
- Makhoul, B., & Olshtain, E. (2015). *Eem mah ani bah v'eem mah ani yotzeh mi khativat habeynaim: Mekhkar al odot otzar hamilim ha'akademi b'kerev talmidim aravi'im—Kamut v'eikhut* [What I come with and what I leave middle school with: A study regarding academic vocabulary among Arab students—Quality and quantity]. Center for Educational Technology (METACH) and the Hebrew University of Jerusalem.
- Makhoul, B., Olshtain, E., Sabah, K., & Copti-Mshael, T. (2018). The development of academic vocabulary among Arabic native speaking middle school pupils: How much do they really know? *Psychology*, 9(3), 323–339. [CrossRef]
- McKeown, M. G., Crosson, A. C., Artz, N. J., Sandora, C., & Beck, I. L. (2013). In the media: Expanding students' experience with academic vocabulary. *Reading Teacher*, 67(1), 45–53. [CrossRef]
- Nagy, W., & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly*, 47(1), 91–108. [CrossRef]
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.
- Powell, A., Watson, J., Staley, P., Patrick, S., Horn, M., Fetzer, L., Hibbard, L., Oglesby, J., & Verma, S. (2015). *Blending learning: The evolution of online and face-to-face education from 2008–2015*. International Association for K-12 Online Learning. Available online: <https://files.eric.ed.gov/fulltext/ED560788.pdf> (accessed on 10 March 2026).
- Resh, N. (2008). *Khativat beinayim b'yisrael. Ba'ayot, k'sha'im v'pitronot efshari'im: Misgeret l'rituakh kivunei hakhra'a* [Junior high schools in Israel. Problems, difficulties and possible solutions: A framework for analyzing the means for reaching decisions]. Van Leer Institute. Available online: [www.vanleer.org.il/wp-content/uploads/2020/04/118\\_PDF.pdf](http://www.vanleer.org.il/wp-content/uploads/2020/04/118_PDF.pdf) (accessed on 10 March 2026).
- Saavedra, A. R., & Opfer, V. D. (2012). Learning 21st-century skills requires 21st-century teaching. *Phi Delta Kappan*, 94(2), 8–13. [CrossRef]
- Saiegh-Haddad, E. (2003). Linguistic distance and initial reading acquisition: The case of Arabic diglossia. *Applied Psycholinguistics*, 24(3), 431–451. [CrossRef]
- Saiegh-Haddad, E. (2018). MAWRID: A model of Arabic word reading in development. *Journal of Learning Disabilities*, 51(5), 454–462. [CrossRef] [PubMed]
- Saiegh-Haddad, E. (2022). A psycholinguistic-developmental approach to the study of reading in Arabic diglossia: Assumptions, methods, findings and educational implications. In E. Saiegh-Haddad, L. Laks, & C. McBride (Eds.), *Handbook of literacy in diglossia and in dialectal contexts. literacy studies* (Vol. 22). Springer. [CrossRef]
- Saiegh-Haddad, E., & Henkin-Roitfarb, R. (2014). The structure of Arabic language and orthography. In E. Saiegh-Haddad, & M. Joshi (Eds.), *Handbook of Arabic literacy: Insights and perspectives* (pp. 3–28). Springer. [CrossRef]
- Saiegh-Haddad, E., & Lina, H. A. J. (2018). Does phonological distance impact quality of phonological representations? Evidence from Arabic diglossia. *Journal of Child Language*, 45(6), 1377–1399. [CrossRef]
- Scarcella, R. (2003). *Academic English: A conceptual framework*. Linguistic Minority Research Institute, University of California. Available online: <https://escholarship.org/uc/item/6pd082d4> (accessed on 10 March 2026).
- Schleppegrell, M. J. (2001). Linguistic features of the language of schooling. *Linguistics and Education*, 12(4), 431–459. [CrossRef]
- Schleppegrell, M. J. (2012). Academic language in teaching and learning: An introduction to the special issue. *The Elementary School Journal*, 112(3), 409–418. [CrossRef]
- Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two versions of the vocabulary levels test. *Language Testing*, 18(1), 55–88. [CrossRef]
- Snow, C. E., & Uccelli, P. (2009). The challenge of academic language. In D. R. Olson, & N. Torrance (Eds.), *The Cambridge handbook of literacy* (pp. 112–133). Cambridge University Press.
- Staker, H., & Horn, M. (2012). *Classifying K–12 blended learning*. Innosight Institute. Available online: <https://www.christenseninstitute.org/wp-content/uploads/2013/04/Classifying-K-12-blended-learning.pdf> (accessed on 10 March 2026).
- Suggate, S. P. (2010). Why what we teach depends on when: Grade and reading intervention modality moderate effect size. *Developmental Psychology*, 46(6), 1556. [CrossRef]
- Tomlinson, C. (2003). *Fulfilling the promise of the differentiated classroom: Strategies and tools for responsive teaching*. ASCD Publishing.

- Topping, K. J., Douglas, W., Robertson, D., & Ferguson, N. (2022). Effectiveness of online and blended learning from schools: A systematic review. *Review of Education*, 10(2), e3353. [[CrossRef](#)]
- Valiande, S. (2010). *Differentiated teaching in mixed ability classrooms/Learning processes and effectiveness* [Unpublished doctoral dissertation]. University of Cyprus. (In Greek)
- Vaughn, S., Wanzek, J., Murray, C. S., & Roberts, G. (2012). *Intensive interventions for students struggling in reading and mathematics: A practice guide*. RMC Research Corporation, Center on Instruction.
- Wanzek, J., Stevens, E. A., Williams, K. J., Scamacca, N., Vaughn, S., & Sargent, K. (2018). Current evidence on the effects of intensive early reading interventions. *Journal of Learning Disabilities*, 51(6), 612–624. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.