



Characteristics of preschoolers' early spelling in Hebrew

Dorit Aram¹ · Hadar Hazan² · Michal Zohar¹

Accepted: 8 April 2022 / Published online: 30 May 2022
© The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract

Before formal instruction, preschoolers represent words in print in various degrees of conventionality. Unlicensed letters are letters that have no connection to the word that the child is aiming to write; they are neither licensed by phoneme-grapheme rules nor by orthographical representations in the mental lexicon. In the current paper, we explore the characteristics of unlicensed letters in the written products of Hebrew-speaking children. Specifically, we examined the role of statistical learning in predicting specific categories of unlicensed letters in preschoolers' spelling, focusing on letters that are present/absent in the child's first name, letters that are more/less frequent in the Hebrew scripts, letters that can spell vowels/consonant, letters that are visually similar/dissimilar, and letters that are easy/difficult to produce graphically. We also evaluated the role of the children's writing level and individual indices (age, gender, socioeconomic status, length of the first name) in predicting the use of these categories. The writing outputs ($N=733$ words), written by 152 preschoolers ($M=63.9$ months, $SD=6.90$), were analyzed and yielded 2109 unlicensed letters. Results indicated that the unlicensed letters in children's early spellings contained significantly more letters with high frequency in Hebrew texts, consonant letters, letters that are visually similar to other letters, and letters that are easy to produce graphically. The child's writing level, age, gender, and length of the first name, uniquely explained the use of each of the categories of unlicensed letters. Parents and teachers should learn about children's writing and spelling development to support their writing and spelling appropriately.

Keywords Early writing · Invented spelling · Prephonological spelling · Random letters · Statistical learning

✉ Dorit Aram
dorita@tauex.tau.ac.il

¹ Constantiner School of Education, Tel Aviv University, Tel Aviv, Israel

² Department of Psychology, Hong Kong University, Pok Fu Lam, Hong Kong

Introduction

Children practice writing and have some ideas about the writing system before they systematically learn to read or write (Gerde et al., 2012; Tolchinsky, 2003). When writing words without formal instruction, children often use *unlicensed letters* that have no connection to the word that they are writing—(see Fig. 2). These letters are referred to as *unlicensed* because they are neither licensed by phoneme-grapheme rules nor by orthographical representations in the mental lexicon. The current study investigated the use of these letters along children’s early writing development in Hebrew—an abjad writing system. We aimed to learn about the role of statistical learning in predicting specific categories of unlicensed letters during the early writing development. Also, we studied how children’s writing level, age, gender, length of the first name, Socio-Economic Status (SES, represented by the mothers’ education predict children’s use of unlicensed letters in each of these categories.

Early writing development

Before the formal study of writing, children participate in writing activities such as writing their name or writing a birthday card (e.g., Neumann et al., 2009; Puranik & Al Otaiba, 2012; Puranik & Lonigan, 2011). When children practice early writing, they first learn the basics of their writing system (e.g., words’ length and letter shapes), and gradually convert the sounds of the word (phoneme/sub-syllable) into the appropriate letters and eventually build the letters into words (Jones, 2015; Levin & Aram, 2013; Martins et al., 2013; Tolchinsky, 2005). This process is positively associated with the development of other literacy skills (e.g., Kessler et al., 2013; Martins et al., 2013; Ouellette & Sénéchal, 2008, 2017).

When studying children’s writing development, Levin and Bus (2003) and Levin et al. (2005) described a developmental scale that begins with nonrepresentational scribbles, progresses to letterlike scribbles, and then shifts to using letters with no phonological connection (“random” letters which we refer to in our study as *unlicensed letters*). Later, children transition into a basic phonetic representation of the word when they use at least a single letter that is phonologically related to the target word, accompanied by unlicensed letters. Gradually, children use more phonologically appropriate letters in their writing, and words become readable. They eventually write all the letters with a phonological connection to the word that they are spelling, even if the spelling is somewhat incorrect (e.g., spelling mistakes). Finally, children use conventional spelling.

Between the time of using letters with no phonological connection and using letters that are phonologically related to the word, there is a bridge in which children add unlicensed letters to their writing. These letters are not related to the word that the child aims to write but are used to varying degrees alongside other phonologically correct letters. Sometimes the child uses only unlicensed letters to write a word. For example, a child can write the word גזר GZR (carrot), pronounced *gezer*, as a string of letters that are entirely unrelated phonologically like ליבלד, or as letters that are partly phonologically related like גדק where the letters ג is phonologically related, but

the other two letters are not, or גלמ where the letters λ and σ (λ for G and σ for Z) are phonologically related (we accepted phonologically close letters as phonologically related), and the other two letters are not. In the first example, all the letters are considered unlicensed; in the second and the third examples, only the two letters that are not related to the word are considered unlicensed letters. In the current study, we examined the characteristics of these unlicensed letters in Hebrew. We investigated the role of statistical learning in predicting characteristics of preschoolers' use of unlicensed letters throughout their writing development.

Statistical learning and early writing development

Statistical learning is a form of implicit learning that occurs due to the internalization of patterns in the environment without conscious effort. The more frequent a pattern is, the better it is internalized (Perruchet & Pacton, 2006; Saffran et al., 2004). Statistical learning requires pattern recognition and relies on the human ability to identify rules and regularity in everyday life. Sensitivity to the regularities and structures is the building block for our expectations in various areas such as cognition, social behavior, visual domains, and language (Aslin & Newport, 2012; Monroy et al., 2017).

In the context of language, salient regular visual characteristics of the local orthography are those that children internalize (Turk-Browne et al., 2009). Statistical learning does not require learners to understand why they make certain writing choices. Rather, learners unknowingly pick up patterns in their written environment and use them in their spelling (Kessler et al., 2013). For example, children are exposed to letters in their names from a young age. Due to this exposure, they tend to overuse these letters in their spellings when writing different words (Treiman et al., 2001). In languages that capitalize the first letter of first names, children show a special preference for the first (capitalized) letter of their name and use it frequently, even when it is not correct phonologically (Both-de Vries & Bus, 2008). Similarly, letters with high frequency in the local orthography tend to be over-represented in children's early spellings (Pollo et al., 2009). In languages with a high proportion of consonants to vowels such as English, children include more consonant letters in their early writing than in languages with a low consonant to vowel ratio, such as Portuguese (Pollo et al., 2005). Despite the growing evidence that statistical learning forms some of the characteristics of early spellings, research has mainly stemmed from Romance and Germanic languages like Portuguese, French, and English (see Read & Treiman, 2013). In the current study, we explored a Semitic script, that of Hebrew, which is an abjad writing system.

The Hebrew writing system

The Hebrew writing system is an *abjad* consonantal script. It consists of 22 consonant letters that are written from right to left. Five letters (*final letters*) have an allograph when placed as the final letter of a word (Ravid, 2012). Four letters can take the role of both consonant and vowel. For example, in the word שיר ('song'), the

letter <׳> takes the role of a vowel, while in the word **דַּי** ('day') it represents a consonant. However, these letters usually indicate a vowel. The twinned role of these letters makes them highly frequent in the Hebrew script, albeit difficult for young spellers to use correctly (Levin et al., 2013). Vowels are only partly represented by letters, and so a word's length is relatively short (Frost et al., 1987) and generally ranges between 2–6 letters, as in the phrases "hello mother" «שָׁלוֹם אִמָּא» "good morning" «בֹּקֶר טוֹב» or "give me" «תֵּן לִי» (for an example of number of letters per word in a corpus of children's books, see the methods section).

The Hebrew alphabet is known as a *square script*—a block alphabet with many 90° angle letters. Letters tend to have high visual similarity to one another. They have more horizontal and vertical strokes and fewer curves and diagonals (י, ד, ו, ת, ג, ח) than letters in the Latin script (Shatil et al., 2000; Treiman et al., 2012). As in other languages, Hebrew-speaking children use unlicensed letters in their early writing, gradually moving towards phonologically correct spelling with unlicensed letters slowly decreasing (Aram & Levin, 2001; Aram, 2005; Aram et al., 2014; Aram & Chorowicz Bar-Am, 2016; Aram et al., 2016).

In the current paper, we studied five categories of unlicensed letters. We focused on letters that are present/absent in the child's first name, letters that are more/less frequent in Hebrew texts, letters that spell consonants/vowels, letters that are visually similar/dissimilar, and letters that are more/less difficult to graphically produce. We explored the role of children's writing measures (writing level and mean sum of written letters per word) and individual indices (age, gender, mother's education as a proxy for family SES and length of the child's first name) in predicting the amount of children's use of each of the abovementioned five categories of unlicensed letters during early writing development.

Children's writing measures

Writing level Children's writing level is related to their use of unlicensed letters. During the process of early writing development, a higher level of writing is associated with the lower use of letters from the child's first name and higher use of phonologically representing letters (Both-de Vries & Bus, 2008; Share & Levin, 1999). Since preschoolers keep using some unlicensed letters in their spelling as their general writing level progresses (Aram & Levin, 2001; Tolchinsky et al., 2012), we were interested in the amount of use of each of the five categories of unlicensed letters throughout writing development.

Number of letters per word Another aspect that reflects children's understanding of their orthography is the number of letters they write to represent a word. The more children know about their orthography, the more the number of letters in their written products resembles the mean length of words commonly used in their language (Pollo et al., 2009). Aram and Levin (2001) found that children who were more advanced in their literacy skills (letter knowledge, phonological awareness, and orthographic awareness) were more attuned to the phonological length of Hebrew words in their early writing. They very rarely used only one letter to represent a word and frequently

wrote about 3–5 letters per word, even when all the letters were unlicensed and not related to the target word. Children with lower literacy skills used up to 10 letters to write one word, far from the mean letters per word in Hebrew.

Individual indices

Age Children's early literacy skills, including spelling, are related to age (e.g., Puranik & Lonigan, 2011). With the increase in age, children are more exposed to their orthography, have more experience, and learn more from their parents and teachers. For example, the nonphonological spellings of older children include a wider variety of letters compared with younger children (Treiman et al., 2007).

Gender Some studies suggest that girls outperform boys in early literacy skills (Lee & Al Otaiba, 2015; McTigue et al., 2020). Additionally, a few studies found that girls are more interested in literacy than boys (Alexander et al., 2008; Baroody & Diamond, 2013; Meece et al., 2006; Peterson & Parr, 2012) and progress more in their reading during kindergarten (Chatterji, 2006).

Length of child's first name The length of the child's first name may be related to spelling patterns due to the size of the pool of familiar letters from which children can choose. Preschoolers with a long first name tend to be familiar with more letter names and practice writing more letters than children whose names are short (Diamond et al., 2008; Puranik & Lonigan, 2011; Welsch et al., 2003).

Socioeconomic status (SES) Children's literacy skills are related to their socioeconomic background across cultures and orthographies (e.g., Aram et al., 2014; Ergül et al., 2017). Families from lower SES practice fewer literacy activities with their children, and parent–child interactions are characterized by lower levels of literacy support (Aram & Levin, 2001; Puranik & Al Otaiba, 2012; Robins et al., 2014). The spelling skills of children from a higher SES are significantly higher than those of children from a lower SES (Lee & Al Otaiba, 2015). Maternal education is a widely used indicator of SES (Dingemann et al., 2019), as it is associated with other indicators such as income (Rendall et al., 2021), and it is positively related to children's literacy skills (e.g., Vernon-Feagans et al., 2020; Younger et al., 2019). Thus, we used maternal education as an indicator of family SES.

We opted to learn how a child's writing level, age, gender, SES (measured via mother's education), and length of the child's first name predict the amount of use of each of five categories of unlicensed letters: Letters from the child's first name, letters that are more/less frequent in Hebrew texts, consonant/vowel letters, letters that are visually similar/dissimilar, and letters that are more/less difficult to graphically produce.

Based on the literature, we hypothesized that children's use of unlicensed letters when spelling words in Hebrew would include greater use of letters from their first name, letters with high frequency in Hebrew texts, as well as greater use of

consonants and letters that are easy to produce graphically. Unlike some other scripts, a capital letter at the beginning of one’s first name is not used in Hebrew; therefore, we asked if Hebrew-speaking children would show greater use of the first letter of their first name in comparison to other letters. In addition, acknowledging the high degree of visual similarity among letters in Hebrew, we asked if when using unlicensed letters in their spelling, children would use more letters that are visually similar to other letters than letters that are visually dissimilar. We aimed to study these five categories of unlicensed letters that children use at different writing levels (writing level and mean sum of letters per word) and across individual indices (age, gender, SES, and length of the first name).

Method

Participants

In this study, we used the writing outputs of Hebrew-speaking preschool children who participated in three studies in which they were asked to write between six to eight words (nouns) that are part of young children’s vocabulary (Table S1). Participants were 152 children (85 girls and 67 boys) recruited through convenience sampling. Their parents had volunteered to participate and reported that their children had no developmental disabilities. Children’s age ranged from 4 to 7 years ($M=63.9$ months; $SD=6.90$). They lived mainly in the center of Israel and studied in different preschools. In Israel, preschools are separated from elementary schools. Formal reading and writing instruction begin in first grade. Preschool teachers focus on language and communication skills, read books to children and introduce the alphabet, with little time devoted to writing activities (Aram et al., 2014). Mothers’ education ranged from high school diploma (27.1%), through BA (45.2%), to MA and Ph.D. (27.7%).

The analysis of the children’s 1205 words revealed that 115 of the word writing products (9.54%) were only scribbles (i.e., signs that are not identified as a Hebrew letter; see Fig. 1) and 357 writing products (29.63%) were words fully spelled phonologically or conventionally. This study did not analyze these 472 writing products (scribbles, phonologically or conventionally spelled).





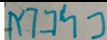
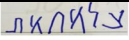

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|--|---|--|
| Scribble: a shape | Scribble: pseudo letters | Nonphonological spelling: All the letters are unlicensed letters (אויאקיי) | Basic consonantal writing - the letter ך (for L) and two unlicensed letters (ה׳) | Partial consonantal spelling - the letter ך (L) and ץ (for H with a spelling mistake) and three unlicensed letters (אוי) | Advanced consonantal spelling (with additions) - the two ך are unlicensed letters | Correct spelling |
|  |  |  |  |  |  |  |

Fig. 1 Examples of Children’s writing on the 7-point writing development scale: the children wrote the word צלחת, (CLHT, tsalaxat, ‘plate’)

The study focused on the use of unlicensed letters. Hence, we analyzed only written words that contained unlicensed letters ($N=733$ written words). These 733 words included 2109 unlicensed letters—recognizable letters that have no connection to the target word. These letters appeared either as the only letters in the word or alongside phonologically related letters (see Fig. 2). In the current study, we only analyzed the characteristics of these unlicensed letters. We studied if they are present/absent in the child's first name, more/less frequent in the Hebrew scripts, spell vowel/consonant, visual similar/dissimilar, and if they are easy/difficult to graphically produce.

Procedure

Education M.A. students collected data. Following the parents' signing a consent form, the student came to the children's home and asked the children to independently write the words. The student gave the child an A4 paper and a pencil, showed a drawing of an object, and asked the child to write the word. The oral instructions were straightforward, for example: "Please write the word *tsalaxat* (plate)." To decipher the children's spellings and decide which letter was written and if it is an unlicensed letter, we had to reach an agreement regarding the identification of the letters. Inter-rater reliability was performed by the leading researcher and three M.A. students in Education. All raters coded 884 letters that appeared in 250 words written by 38 children (20% of the data). Each rater was given pictures of all the scanned writing outputs of a particular participant (including a scan of the written form of the child's first name). Raters were told that the letters did not necessarily match the target words and that preschoolers might write the letters in a "mirror writing," tilted or not perfectly graphically formed. The raters were only required to




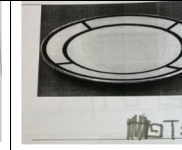
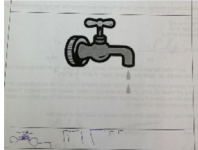
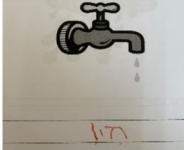
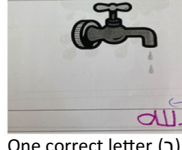
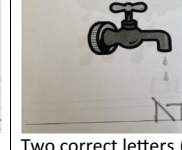
| Examples of different spelling of the word 'plate' - צלחת | | | |
|---|---|---|---|
|  |  |  |  |
| Eleven unlicensed letters (איאשרישיאירמ) | Five unlicensed letters (ריודע) | One correct letter (ח) and two unlicensed letters (שס) | Two correct letters and two unlicensed letters (פד) |
| Examples of different spelling of the word 'faucet' - ברז | | | |
|  |  |  |  |
| Five unlicensed letters (ההוריי) | Three unlicensed letters (חיי) | One correct letter (ב) and two unlicensed letters (שס) | Two correct letters (ר ב) and two unlicensed letters (שא) |

Fig. 2 Spelling development in Hebrew: examples of use of unlicensed letters

identify the letters by name in each writing output. A good percentage of agreement was achieved (inter-rater reliability of 94%). Any coding discrepancies between the raters were resolved through discussion.

Measures

Writing level and number of written letters per word

We evaluated the child's general writing level on a 7-point scale adapted from Levin et al. (1996) and Levin and Bus (2003). This scale was successfully used in previous studies in Hebrew (Aram, 2007; Aram et al., 2021). The scale ranged from scribbles and pseudo letters, through the use of only unlicensed letters and basic consonantal spellings (one consonant letter that is phonologically represented), partial and full consonantal spellings, to conventional spelling, including vowel letters (see Fig. 1). Higher scores indicated a more advanced writing level. The mean score across the words served as the word writing score. Inter-judge reliability computed on the scores of the writings of 20% of the sample by two independent judges (M.A. students in Special Education), resulted in a highly significant Cohen's κ of 0.90. Reliability across items was excellent (Cronbach's $\alpha=0.99$). The words that the children were asked to write included a mean of 3.35 letters per word. The letters that the child wrote for each word were counted, and the mean score across the words served as the number of written letters per word score (Table S1).

Characteristics of unlicensed letters

We were interested in children's use of different letters as unlicensed letters in Hebrew, focusing on five characteristics: Child's name, frequency in texts, consonants and vowels, visual similarity, and ease of production. For each characteristic, we grouped letters into categories (e.g., frequent vs. infrequent letters), studied the chance of using unlicensed letters from the category and how age, gender, SES, and length of the first name predict the use of each category.

Letters from the child's first name The number of times the children used any letter from their first name or the first letter of their first name as an unlicensed letter was counted.

Letters' frequency in Hebrew texts Each letter that was used in children's spellings as an unlicensed letter was categorized as belonging to one of four frequency groups (low, medium–low, medium–high, high) based on its frequency of appearance in children's books. The Hebrew letters were categorized into these groups using a letter counter. We used this counter and calculated the frequency of Hebrew letters in eight popular narrative children's books (e.g., *Itamar Meets a Rabbit* by David Grossman or *Frog is Frog* by Max Velthuijs). The number of words ranged from 330 to 728 ($M=528$, $SD=143$) and the number of letters ranged from 1387 to 3676 letters ($M=2342$, $SD=750$). The number of letters per word ranged between 2

and 7 ($M = 3.68$, $SD = 1.20$). The correlation between our categorization of letters' frequency and the letters' frequency count in Israeli adults' newspapers (Shoken & Shor, 2010) was high ($r = 0.90$, $p < 0.001$). Letters' frequency was categorized as: low ($n = 6$ letters appeared between 0.1 and 1%); medium–low ($n = 7$ letters appeared between 1.1 and 3%); medium–high ($n = 8$ letters appeared between 1 and 6%); and high ($n = 6$ letters appeared between 6.1 and 10.65%) (See Table S2).

Consonants and vowel letters Each unlicensed letter used in the children's spellings was categorized as being a consonant or vowel letter. Notice that all the vowel letters in the words in our study served only as vowel letters [except for the letter /Y/ in the word *yad* (a hand), See Table S1].

Letter's visual similarity versus dissimilarity Each unlicensed letter used in the children's spellings was categorized as either visually similar to other letters or visually dissimilar. For example, the letters ך, ך, ך, ך have some visual similarities, while the letters ך or ך do not look like other letters. To determine this, a visual similarity test was used (based on Treiman, 2006). The test was administered via email to 90 adults (60 Hebrew and 30 English-speakers). Participants were asked to rank pairs of letters for their level of similarity on a 7-point scale (0 = *no similarity* to 7 = *very similar*). Pairs of letters with an average score greater than 3.5 were defined as visually similar. Pairs of letters with an average score lower than 3.5 were defined as visually dissimilar. The test showed good reliability (Cronbach's $\alpha = 0.82$). Studying the possible differences between the Hebrew and the English-speaking groups showed no significant differences for the similar ($t(88) = 0.712$, $p > 0.05$) and the dissimilar letters ($t(88) = 0.534$, $p > 0.05$). The letters that show visual similarity (ת, ה, ן, ן, ק, כ, ך, י, ח, ז, ו, ה, ד, ב) have at least one letter that is visually similar to them. The dissimilar letters were ץ, ף, ש, צ, פ, ע, ן, ן, ל, ט, ג, ס, ך. We ran ANOVAs with repeated measures to verify the difference between the two groups of letters. We found a significant difference ($F(89, 1) = 229.33$, $p = 0.000$) between the letters rated as having high visual similarity ($M = 4.20$, $SD = 1.04$) and the letters rated as visually dissimilar ($M = 2.10$, $SD = 0.086$).

Letter's complexity of graphic production Each unlicensed letter used in the children's spellings was categorized into one of three levels of graphic production (easy, medium, and difficult). Graphically writing letters is a demanding, acquired fine motor skill. According to the literature, children's handwriting starts with vertical strokes, followed by horizontal strokes and circles. Last to appear are the diagonal lines (Feder & Majnemer, 2007). After controlling the direction of single lines children learn how to use a combination of lines. When studying children's copying skills, the hardest shapes for preschoolers are the square and the triangle (Beery & Buktenica, 1989). Following these ideas, we based our scoring on Shatil's (1993) mapping of the Hebrew letters according to psychographic development to three groups of letters according to the number of strokes, their directions, and combinations: Easy—one stroke, mainly vertical lines (ס, ך, כ, י, ח, ן); Medium—two

strokes, vertical and horizontal lines (פ, ת, ק, נ, ה, ד, ב); and Difficult—two or three strokes that include diagonal lines (ג, א, ש, ז, ע, מ, ל, ו, ר).

Data analysis

For baseline comparisons, we used the generalized estimating equations (GEE) procedure to test differences in letter counts of different characteristics (e.g., the number of times similar or dissimilar letters appeared in the child's word writing). Specifically, we applied the Negative Binomial Distribution (NBD; SPSS V.25.0) to these counted letters. The NBD is a corrected form of the Poisson distribution, where the mean is equal to the variance assumption is not assumed (Hilbe, 2017). A count outcome represents discrete numbers and is usually controlled by a total count, in case this total varies from one observation to another. The advantage of this statistical procedure is twofold: It applies various effects to predicted letter counts, i.e., differences between two or more letter characteristics that appear in the child's written words (Wald's Chi-Square test); and predicted marginal means, which are the predicted frequencies of each tested letter's characteristics. In our study, each child wrote a few words. The GEE procedure integrates two words per child, that is, the child's name and the non-name word (Horton & Lipsitz, 1999), namely, the count of letters in the name, and the count of letters in other words.

The more advanced analysis aimed to explain children's use of various categories of unlicensed letters (e.g., vowel letters, highly frequency letters, graphically similar letters) by the child's writing level, personal variables (e.g., age, gender), and the target words (Baayen et al., 2008). To study this question, we used the HLM 7.01 program (Raudenbush et al., 2013). To account for the single word and the child's features of the study, we estimated a two-level random effects model, in which level one was assigned to words within children, and level two was assigned to children. Specifically, in level 1, we entered the child's writing measures (writing level and mean sum of written letters) as assessed per word. In Level 2, we included the child's variables (age, gender, mother's education, and length of the first name). The HLM program improved the sensitivity of the model by allowing a random slope, whereas GEE was limited to random intercept only. The random slope means that regression coefficients varied randomly across children, namely, the association between the sum of letters per word, for example, and the outcome could vary across children. In the original analyses, we included a third level, which was comprised of the target word. We excluded it from the final modeling framework to remain with a two-level analysis due to the small variance of the third level. In this analysis, we assumed a normal distribution, yet the original letter count distributions were asymmetric—higher frequencies of smaller counts, and lower frequencies of higher counts, thus, a natural log transformation was applied to these letter counts with the minimum addition to avoid the undefined $\ln(0)$. Our interpretation of the estimates was more focused on the direction, that is, a negative sign indicated a negative association between the independent and the outcome variables and vice versa for a positive sign. In

log-transformed models, exponentiated estimates should be interpreted as the percent change in the letter counts in response to a change in one unit of the independent variable $((\exp(b) - 1) * 100)$, and this interpretation refers to additional effects beyond the intercept.

Results

Descriptive statistics

Writing level and number of written letters per word

Evaluated on a 7-point scale, children's average writing level was 3.60 ($SD=0.94$), reflecting frequent use of phonological unlicensed letters and the beginning of phonological spelling (writing one or more letters phonologically correct in a word). It should be noted that even children who wrote at a higher writing level (basic or partial consonantal spelling) still used some unlicensed letters (for example, see Fig. 1). Studying children's spellings, we found that all the outputs included at least two letters. On average, children wrote 3.63 letters per word ($SD=1.53$), with 0.75 letters spelled phonologically correct per word ($SD=1.06$), and 2.88 unlicensed letters per word ($SD=1.76$).

The spelling outputs revealed a range of use of unlicensed letters from one or two per word (47.30%), through three or four (38.10%), five or six (10.50%), to more than seven letters in one written word (4.10%), even though the longest word the children had to write consisted of only six letters. The ratio was one phonologically correct letter for every four unlicensed letters. Children who used unlicensed letters in their spelling tended to use them more often than phonologically correct letters (See Fig. 2 for examples).

We counted the appearance of each of the Hebrew letters (including final letters) in the children's unlicensed letters (2109 letters) to learn which letters appeared more frequently. We found that the first letter in the Hebrew script—א, was most frequently used ($n=226$), followed by ה ($n=223$), while the least used letters were י ($n=1$) and ך ($n=2$).

Predicting characteristics of unlicensed letter use

Table 1 presents the GEE modeling results; these are the predicted frequencies of use of each of the five categories of unlicensed letters based on children's outputs.

Table 1 shows that Hebrew-speaking children were generally less likely to use letters from their names ($M=6.24$; $SD=0.60$) as unlicensed letters compared with other letters ($M=7.66$; $SD=0.53$). They were also less likely to use the first letter of their name ($M=1.85$; $SD=0.22$) compared with other letters in their name ($M=4.39$; $SD=0.46$). Children tended to use more consonants ($M=8.12$; $SD=0.60$) than vowels ($M=5.77$; $SD=0.50$), and more visually similar letters ($M=8.40$; $SD=0.63$) than visually dissimilar letters ($M=5.50$; $SD=0.45$). Post-hoc

Table 1 Generalized estimating equations modeling results: differences in frequencies of use of unlicensed letters by their characteristics in children's spellings (N=2109 unlicensed letters)

| | <i>M (SE)</i> | 95% CI | Wald χ^2 | <i>p</i> |
|---|---------------|--------------|---------------|----------|
| <i>Letters from first name</i> | | | | |
| From the child's first name | 6.24 (0.60) | 5.17 7.54 | 4.32 | 0.04 |
| Not from the child's first name | 7.66 (0.53) | 6.70 8.80 | | |
| <i>The first letter in the child's first name</i> | | | | |
| First letter | 1.85 (0.22) | 1.50 2.30 | 57.50 | <0.0001 |
| Remaining letters in the name | 4.39 (0.46) | 3.57 5.40 | | |
| <i>Consonant vs. vowel letters</i> | | | | |
| Consonants | 8.12 (0.60) | 7.00 9.40 | 15.54 | <0.0001 |
| Vowels | 5.77 (0.50) | 4.86 6.84 | | |
| <i>Letters that are visually similar to other letters</i> | | | | |
| Yes | 8.40 (0.63) | 7.25 9.70 | 29.90 | <0.0001 |
| No | 5.50 (0.45) | 4.70 6.50 | | |
| <i>Complexity of graphical production</i> | | | | |
| Easy | 4.87 (0.42) | 4.10 5.77 | 7.40 | 0.025 |
| Moderate | 4.00 (0.31) | 3.50 4.70 | | |
| Difficult | 5.00 (0.43) | 4.20 5.90 | | |

Values in cells represent predicted marginal frequencies of use resulting from the GEE outcome; standard errors are in parentheses

tests revealed that children were more likely to use letters that were easy to produce graphically ($M=4.87$; $SD=0.42$) than letters that were moderately complex to produce ($M=4.00$; $SD=0.31$), however, the difference was not significant ($p=0.059$).

To learn about the correspondence between the frequency that the children used specific letters as unlicensed letters in their spellings, and the letters' frequency in Hebrew texts, we summed up their appearance in the children's outputs. We found that the children's mean use of letters with high, moderate-high, moderate-low, and low frequency in Hebrew texts as unlicensed letters was 7.60 ($SD=0.60$), 3.73 ($SD=0.33$), 1.86 ($SD=0.21$), and 0.69 ($SD=0.11$), respectively. There was a strong Spearman's correlation ($r_p=0.91$; $p<0.000$) between the letters' frequency in Hebrew texts and their frequency of appearance as unlicensed letters in the children's spellings. That is, the more frequently the letters appeared in Hebrew texts, the more frequently they were used as unlicensed letters in the children's spellings and vice versa.

Table 2 Two-level modeling results predicting characteristics of unlicensed letters in children's spelling

| Characteristics of usage | First name letters | | High frequency letters | | Low frequency letters | | Consonant letters | | Vowel letters | |
|-----------------------------------|--------------------|--------------------|------------------------|--------------------|-----------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
| | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 |
| Model ^a | | | | | | | | | | |
| Fixed effects | <i>b</i> (SE) | <i>b</i> (SE) | <i>b</i> (SE) | <i>b</i> (SE) | <i>b</i> (SE) | <i>b</i> (SE) | <i>b</i> (SE) | <i>b</i> (SE) | <i>b</i> (SE) | <i>b</i> (SE) |
| Intercept | -2.74*** (0.42) | -1.99*** (0.58) | -2.30*** (0.28) | -2.06*** (0.34) | -8.03*** (0.17) | -8.08*** (0.24) | -1.02*** (0.19) | -0.82*** (0.21) | -3.18*** (0.29) | -2.94*** (0.34) |
| <i>First level: Child X Word</i> | | | | | | | | | | |
| Writing level | | -1.41*** (0.34) | | -1.04*** (0.22) | | -0.52*** (0.15) | | -1.63*** (0.21) | | -0.94*** (0.23) |
| Sum of letters in the word | | .52*** (0.13) | | .93*** (0.11) | | 0.17* (0.08) | | 0.31*** (0.08) | | 1.05*** (0.11) |
| <i>Second level: Child</i> | | | | | | | | | | |
| Age | | -0.13* (0.06) | | -0.09** (0.11) | | 0.05* (0.02) | | 0.06** (0.02) | | -0.12*** (0.03) |
| Gender | | -0.73 (0.72) | | 0.03 (0.44) | | 0.20 (0.32) | | 0.14 (0.29) | | 0.05 (0.44) |
| Name length | | 1.61*** (0.46) | | 0.25 (0.29) | | 0.25 (0.20) | | -0.27 (0.19) | | 0.27 (0.29) |
| SES | | -0.45 (0.51) | | -0.42 (0.29) | | -0.11 (0.22) | | -1.63 (0.20) | | -0.05 (0.30) |
| <i>Variance Components b (SD)</i> | | | | | | | | | | |
| First level | 21.68 (4.66) | 21.80 (4.67) | 10.31 (3.12) | 9.50 (3.08) | 7.30 (2.70) | 7.16 (2.68) | 7.33 (2.71) | 6.65 (2.58) | 12.12 (3.48) | 11.40 (0.38) |
| Second level | 21.76 (4.66) | 13.78*** (3.71) | 8.81*** (2.96) | 4.83*** (2.20) | 2.54*** (1.59) | 2.21*** (1.49) | 3.97*** (1.99) | 1.27*** (1.61) | 9.41*** (3.06) | 4.54*** (2.13) |
| ICC | 0.50 | | 0.46 | | 0.26 | | 0.35 | | 0.40 | |
| Deviance | 4587.35 | 4538.61 | 4023.55 | 3907.77 | 3677.73 | 3653.46 | 3723.49 | 3580.02 | 4131.19 | 4016.99 |
| Deviance Δ | | 47.74*** | | 115.78*** | | 24.27*** | | 143.47*** | | 114.20*** |

Table 2 (continued)

| Characteristics of usage | First name letters | High frequency letters | Low frequency letters | Consonant letters | Vowel letters |
|----------------------------------|--------------------|------------------------|---------------------------|-------------------------------|--------------------------------|
| Model ^a | 1 2 | 1 2 | 1 2 | 1 2 | 1 2 |
| Fixed effects | <i>b</i> (SE) | <i>b</i> (SE) | <i>b</i> (SE) | <i>b</i> (SE) | <i>b</i> (SE) |
| Pseudo R ² | 0.18 | 0.25 | 0.05 | 0.27 | 0.26 |
| | Visually similar | Visually dissimilar | Easy graphical production | Moderate graphical production | Difficult graphical production |
| Model # ^a | 1 2 | 1 2 | 1 2 | 1 2 | 1 2 |
| Fixed effects | <i>b</i> (SE) | <i>b</i> (SE) | <i>b</i> (SE) | <i>b</i> (SE) | <i>b</i> (SE) |
| Intercept | -1.11** (0.19) | -2.99*** (0.27) | -3.53*** (0.27) | -3.69*** (0.24) | -3.91*** (0.30) |
| <i>First level: Child X Word</i> | | | | | |
| Writing level | -1.20*** (0.18) | -1.47*** (0.21) | -1.52*** (0.20) | -1.11*** (0.20) | -1.46*** (0.21) |
| Sum of letters | 0.61** (0.09) | 0.53*** (0.12) | 0.96*** (0.10) | 0.78*** (0.13) | 0.58*** (0.13) |
| <i>Second level: Child</i> | | | | | |
| Age | -0.05** (0.2) | 0.09* (0.03) | -0.05 (0.03) | -0.05 (0.03) | 0.09** (0.03) |
| Gender | 0.83** (0.30) | -0.55 (0.45) | 0.72 (0.43) | 0.69 (0.42) | -0.65 (0.47) |
| Name length | 0.03 (0.22) | 0.05 (0.29) | 0.36 (0.27) | -0.11 (0.29) | -0.12 (0.28) |
| SES | 0.003 (0.19) | -0.40 (0.36) | -0.32 (0.29) | -0.8 (0.30) | -0.15 (0.34) |

Table 2 (continued)

| | Visually similar | | Visually dissimilar | | Easy graphical production | | Moderate graphical production | | Difficult graphical production | |
|--------------------------------------|-------------------|-------------------|---------------------|-------------------|---------------------------|-------------------|-------------------------------|-------------------|--------------------------------|-------------------|
| | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 |
| Model # ^a | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 |
| Fixed effects | <i>b</i> | | <i>b</i> | | <i>b</i> | | <i>b</i> | | <i>b</i> | |
| | (SE) | | (SE) | | (SE) | | (SE) | | (SE) | |
| <i>Components of variance b (SD)</i> | | | | | | | | | | |
| First level | 9.40 (3.07) | 8.75 (2.96) | 12.54 (3.54) | 11.76 (3.43) | 14.03 (3.74) | 12.64 (3.55) | 16.83 (4.10) | 15.54 (3.94) | 13.29 (3.65) | 12.60 (3.55) |
| Second level | 3.30*** (1.80) | 1.08*** (1.16) | 7.50*** (2.74) | 4.73*** (2.18) | 7.83*** (2.80) | 3.88*** (1.97) | 4.71*** (2.17) | 3.30*** (1.82) | 7.96*** (2.82) | 4.92*** (2.22) |
| ICC | 0.26 | | 0.37 | | 0.36 | | 0.22 | | 0.37 | |
| Deviance | 3863.82 | 3741.15 | 4127.84 | 4040.73 | 4202 | 4069.31 | 4271.73 | 4191.89 | 4170.33 | 4088.74 |
| Deviance Δ | | 122.67*** | | 87.11*** | | 133.11*** | | 79.84*** | | 81.59*** |
| Pseudo R ² | | 0.22 | | 0.18 | | 0.25 | | 0.13 | | 0.18 |

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

^a1 = baseline model; 2 = conditional model

Letter counts were natural log transformed

Table 2 presents the results regarding the role of the child's word writing characteristics (writing level and mean sum of written letters—first level) and personal variables (age, gender, mother's education, and length of the first name—second level) in predicting the percent change of using each of the five categories of unlicensed letters that we studied via HLMs. The estimated coefficients give the direction in which the independent factor affects the outcome number.

Counts of letters from the child's first name

The conditional model provided a significantly better fit to the data than the baseline model (Δ Deviance = 47.74, $p < 0.001$). In the conditional model, a lower writing level ($b = -1.49$, $p < 0.001$), more letters per word ($b = 0.52$, $p < 0.001$), younger age ($b = -0.13$, $p < 0.05$), and a longer name ($b = 1.60$, $p < 0.001$), were significantly related to a higher usage of unlicensed letters from the child's first name. In other words, a unit change of lower writing was associated with a 75% decrease in applying the first name letters, and an additional letter in the word was associated with a 68% increase in the use of the first name letter.

Letters' frequency in Hebrew texts

Overall, the conditional model provided a significantly better fit to the data than the baseline models for high (Δ Deviance = 115.8, $p < 0.001$) and low (Δ Deviance = 24.30, $p < 0.001$) frequency letters.

In the conditional model, a higher writing level was associated with lower use of both high ($b = -1.00$, $p < 0.001$) and low frequency letters ($b = 0.52$, $p < 0.001$). A higher total mean sum of written letters per word was associated with greater use of both high ($b = 0.90$, $p < 0.001$) and low frequency ($b = 0.17$, $p < 0.05$) letters as unlicensed letters. Older children used fewer high frequency letters ($b = -0.05$, $p < 0.05$) and more low frequency letters as unlicensed letters.

Counts of consonants and vowels

The conditional models provided a significantly better fit to the data than the baseline models for consonants (Δ Deviance = 143.5, $p < 0.001$) and vowels (Δ Deviance = 114.20, $p < 0.001$). This means that adding the level 1 and level 2 explanatory variables improved the model fit for both the counts of consonants and vowels.

In the conditional models, a higher writing level was associated with lower use of both consonant ($b = -1.60$, $p < 0.001$) and vowel ($b = -0.94$, $p < 0.001$) letters as unlicensed letters. A higher mean sum of written letters per word was associated with greater use of both consonants ($b = 0.30$, $p < 0.001$) and vowels ($b = 1.00$, $p < 0.001$) as unlicensed letters. Older children used more consonants ($b = -0.06$, $p < 0.01$) and fewer vowels ($b = -0.10$, $p < 0.001$).

Letters with visual similarity versus visual dissimilarity

The conditional models provided a significantly better fit to the data than the baseline models for visually similar (Δ Deviance = 122.67, $p < 0.001$) and visually dissimilar (Δ Deviance = 87.11, $p < 0.001$) letters. A higher writing level was associated with lower use of both visually similar ($b = -1.20$, $p < 0.001$) and dissimilar ($b = -1.50$, $p < 0.001$) letters as unlicensed letters. A higher mean sum of written letters per word was associated with greater use of both visually similar ($b = 0.60$, $p < 0.001$) and visually dissimilar ($b = 0.50$, $p < 0.001$) letters as unlicensed letters. Older children used fewer visually similar letters ($b = 0.05$, $p < 0.01$), and more visually dissimilar letters ($b = 0.09$, $p < 0.05$) as unlicensed letters. Girls were more likely to use visually similar letters than boys ($b = 0.83$, $p < 0.01$).

Letters' complexity of graphical production

The conditional models provided a significantly better fit to the data than the baseline models for the use of letters with easy (Δ Deviance = 133.10, $p < 0.001$), moderate (Δ Deviance = 79.80, $p < 0.001$), and difficult (Δ Deviance = 81.60, $p < 0.001$) levels of graphical production. A higher writing level was associated with lower use of unlicensed letters with easy ($b = -1.50$, $p < 0.001$), moderate ($b = -1.10$, $p < 0.001$), and difficult ($b = -1.46$, $p < 0.001$) levels of graphical production. A higher mean sum of written letters per word was associated with greater use of unlicensed letters with easy ($b = 0.96$, $p < 0.001$), moderate ($b = 0.78$, $p < 0.001$), and difficult ($b = 0.58$, $p < 0.001$) graphical production. Older children used more unlicensed letters with difficult graphical production ($b = 0.09$, $p < 0.05$).

Discussion

Studying literacy development within a wide range of writing systems promotes our understanding of their uniqueness and provides implications for children's learning (Share, 2020). In the current study, we examined the nature of unlicensed letters—written letters that have no exact or close phonological connection to the word's letters that the child is writing in Hebrew. Before formally learning to write, children start to use unlicensed letters with varying frequency (from solely using these letters to using one or two letters alongside phonologically correct letters). In this study, we (a) explored characteristics of unlicensed letters in children's early spellings in Hebrew, and (b) evaluated the role of children's level of word writing and individual indices (age, gender, family SES, and length of the first name) in predicting the use of these categories of letters. Results indicated that children's early spelling products contained significantly more unlicensed letters that have a high frequency in Hebrew texts, consonant letters, letters that are visually similar to other letters, and letters that are easy to produce graphically. However, children showed no general preference for letters from their first name, including the first letter of the name. The child's writing level as well as their age, gender, and length of the first name (but not

SES), uniquely predicted the use of the studied characteristics of unlicensed letters in early spellings.

Children's writing measures significantly predicted each of the five categories of unlicensed letters. Specifically, a higher writing level and a lower mean sum of written letters per word in the children's spelling products corresponded with the general lower use of unlicensed letters. This negative link is in accordance with the idea that higher writing skills go hand-in-hand with an increase in children's ability to encode phonemes into graphemes and use fewer unlicensed letters (Both-de Vries & Bus, 2008). Since words in Hebrew are generally short and include 2–6 letters (Velan & Frost, 2011), and the mean length of letters per word in the present study was 3.35 letters, the use of more letters per word indicates a lower level of understanding of the writing system. This result supports a previous study that showed that Hebrew-speaking children who had higher early literacy skills used fewer letters per word (3–5) systematically in their spellings, even if all of them were unlicensed letters than children with lower literacy skills who used more letters (6–7) (Aram & Levin, 2001).

Children's age also predicted characteristics of unlicensed letters' use. Older age was linked to lower usage of letters from the child's first name, letters with high frequency, and vowel letters, as well as greater use of low-frequency letters, consonant letters, letters with high visual dissimilarity, and letters with difficult graphical production as unlicensed letters. These findings align with the idea that age predicts writing level; thus, older children tended to use more sophisticated characteristics of unlicensed letters. Children who are older (and thus, have more exposure to the Hebrew orthography) used unlicensed letters that are different from the letters in their first name, fewer vowel letters that are easy to graphically produce in Hebrew, more consonant letters as the orthography is consonant based, and more letters with low frequency and high visual dissimilarity. All of these characteristics require a better understanding of the writing system.

Characteristics of unlicensed letters

Children's first name

Based on the result, Hebrew-speaking children did not overuse letters from their first name compared with the rest of the alphabet. Yet, the length of the child's name predicted the use of letters from the first name as unlicensed letters, with a longer first name associated with greater use of letters from first names as unlicensed letters. Since children use their names as their first pool of letters (Read & Treiman, 2013), the more letters one's first name contains, the greater the pool of letters and possibilities, which is reflected in the greater use of letters from the first name as an unlicensed letter.

Based on the statistical learning principle of exposure, as well as previous findings (Bloodgood, 1999; Kessler et al., 2013; Pollo et al., 2009), we expected children to use letters from their first names more frequently than any other letters. Nonetheless, no such pattern was found. A possible explanation for this finding may

be the relatively short length of first names in Hebrew compared with English names (Treiman et al., 2007). In our study, most of the children's first names ranged between two or three (e.g., גל or תמר) to four letters (e.g., אילן). Six children (out of 152) had the longest names at five letters (e.g., אריאל). In their early writing, before understanding the phoneme to grapheme relations, children tend to use a variety of letters. They assume that the same letter cannot appear multiple times in a word, especially not one after the other (Ferreiro & Teberosky, 1982). We think that young Hebrew-speaking children try to diversify, and so if their name is short, they use other letters as well.

Also, unlike the Latin alphabet (Treiman & Kessler, 2011), in Hebrew, the first letter of the name is not capitalized and thus does not draw more visual attention. We suggest that when the first letter is the same size as the rest of the letters in the child's name, children tend to use all the letters from their name with similar frequency. This pattern in Hebrew is different from other languages in which the capitalized first letter is used more frequently as an unlicensed *letter*, probably due to its salience (Both-De Vries & Bus, 2008; Levin & Aram, 2005; Treiman & Kessler, 2004, 2013).

Frequent/infrequent letters and consonants/vowels

As native speakers of languages such as English and Portuguese (Kessler et al., 2013; Pollo et al., 2009), Hebrew-speaking children used significantly more letters with high frequency in texts as unlicensed letters. This finding is consistent with the statistical learning approach, which states that young readers capture the orthographic regularities based on the relative frequency of their occurrence in written texts (Chetail, 2015).

Children in our sample used more letters that represent consonants than letters that represent vowels. Indeed, while using more consonants, they maintained a similar ratio of consonants to vowels as observed in Hebrew texts (Levin et al., 2013). Our results support those of Pollo et al. (2005), who found that Portuguese-speaking children tended to use more vowel letters than English-speaking children did, which is consistent with the higher percentage of vowels in Portuguese compared with English. Hebrew orthography is primarily consonantal, with consonantal phonemes fully represented by letters and vowels represented in a deficient manner (Shatil et al., 2000). It is plausible that even when spelling with unlicensed letters, children are affected by this consonantal nature of the Hebrew language.

Visual similarity versus visual dissimilarity

Children's unlicensed letters contained more letters that are visually similar to other letters than letters that are visually dissimilar. This finding exemplifies the role of statistical learning in detecting visual regularities (Turk-Browne et al., 2009). The exposure to print in literate societies enables preschoolers to learn implicitly about the visual characteristics of written words. These findings are consistent with studies in which children also focused on the visual similarity of letters (Bourke et al., 2014; Treiman & Kessler, 2011; Treiman et al., 2012, 2014). While naming visually

similar letters is a difficult task for preschoolers (Bourke et al., 2014; Treiman, 2006), using letters that are visually similar to other letters as unlicensed letters in their writing is easier. Since children learn from an early age to identify letters as members of a set, visual similarity seems to be one way that they link letters together (Lavine, 1977). By forming sets of patterns, children can make slight changes and write different letters.

Girls were more likely than boys to use visually similar letters. Beyond this single result, we found no gender differences. The results regarding gender differences in early literacy are divided, with some studies showing differences favoring girls' literacy achievements while others showing no differences (McTigue et al., 2020). Our findings support the studies that did not find differences between genders in terms of literacy skills like writing letters and words (Puranik et al., 2013; Ritchey, 2008), as well as in the quality of graphically forming the letters (Graham et al., 2001, 2006; Weintraub & Graham, 2000).

Complexity of letters' graphical production

Children's use of unlicensed letters contained significantly more letters that were graphically easy to form than letters that were moderately complex to form. However, contrary to our assumption, no difference was found between letters that are easy and difficult to form. It is possible that the division of letters based on the psychographic theory (Shatil, 1993) is insufficient to explain the choices children make when using unlicensed letters. Some letters such as א (A) or ב (B) are more difficult to produce graphically, yet these letters are very familiar to children due to their location as the first letters of the alphabet and they use them frequently. This result is similar to studies that showed children favoring the first letters of the English ABC (Justice et al., 2006; Treiman et al., 2012). In addition, most Hebrew letters are comprised of a vertical line and one or more attached horizontal lines (e.g., ט, י, ך, ם, ן). This means that the difference in graphical production is not great and maybe this is why it is not reflected in the use of unlicensed letters.

Finally, SES did not contribute to any of the unlicensed letters' characteristics. Although previous studies have demonstrated the role of children's SES in writing development (Elimelech et al., 2020; Aram et al., 2014), the current findings suggest that statistical learning is less related to SES measures and more to the specifics of the orthography and other personal variables. Another possible explanation is that the sample was not varied enough, as about 73% of the mothers held a university degree compared to the national average of 55% (Taub Center, 2019).

Limitations and future direction

This study is the first to provide a window into children's statistical learning and patterns of use of unlicensed letters during spelling development in Hebrew. However, some limitations should be noted. Participants' SES background was relatively homogeneous, the children wrote different words, and linguistic measures such as vocabulary and other early literacy measures (e.g., letter knowledge, phonological awareness) that might have also accounted for the characteristics of their writing

were not taken into account. Also, we used categorical rather than continuous scales. Some information was possibly lost in treating variables in such a categorical way.

Future studies should include a more heterogeneous sample and a smaller pool of target words while maintaining a clear distinction between consonant-only words versus consonant and vowel words. The measures should be continuous, and beyond writing measures, it may be useful to include linguistic measures to learn further about the contributions of these characteristics to the nature of children's use of unlicensed letters.

In conclusion, this study supports the role of statistical learning in writing development in Hebrew—an abjad writing system. It exemplifies the non-random way in which unlicensed letters are used and corresponds with universal (e.g., frequent letters in the orthography) as well as the language-specific features of Hebrew orthography (e.g., consonant vs. vowel letters, letters with visual similarity vs. visual dissimilarity).

Finally, our results could alert policymakers in the field of early literacy to the role of statistical learning and orthography in children's literacy development in Hebrew. Sometimes, literacy teaching methods are based on literacy studies in English and are less adapted to the specifics of different orthographies (e.g., Lipka et al., 2016). Children need adults' direction and support to learn to write and read (Ehri et al., 2001). Adults (parents and teachers) can use the fact that children use characteristics of unlicensed letters in their spelling to initiate conversations to support children's literacy development. Adults who are engaged in writing activities with young children should be aware of children's statistical learning and the nature of orthography-specific characteristics. Adults asking questions regarding letters when writing with children may promote children's understanding of the writing system.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11145-022-10305-2>.

References

- Alexander, J. M., Johnson, K. E., Leibham, M. E., & Kelley, K. (2008). The development of conceptual interests in young children. *Cognitive Development, 23*(2), 324–334.
- Aram, D. (2005). Continuity in children's literacy achievements: A longitudinal perspective from kindergarten to school. *First Language, 25*(3), 259–289.
- Aram, D. (2007). Sensitivity and consistency of maternal writing mediation to twin kindergartners. *Early Education and Development, 18*(1), 71–92. <https://doi.org/10.1080/10409280701274733>.
- Aram, D., Abiri, S., & Elad, L. (2014). Predicting early spelling: The contribution of children's early literacy, self-regulation, private speech during spelling, and parental spelling support. *Reading and Writing: An Interdisciplinary Journal, 27*(4), 685–707. <https://doi.org/10.1007/s11145-013-9466-z>.
- Aram, D., & Chorowicz Bar-Am, O. (2016). Parent-kindergartner joint writing: A comparison between computer use and pencil and paper. *International Journal of Child-Computer Interaction, 7*, 15–21. <https://doi.org/10.1016/j.ijcci.2016.03.001>.
- Aram, D., Cohen Meidan, I., & Deitcher, D. A. (2016). Comparison between homeschooled and formally schooled kindergartners: Children's early literacy, mothers' beliefs, and writing mediation. *Reading Psychology, 37*(7), 995–1024. <https://doi.org/10.1080/02702711.2016.1157537>.

- Aram, D., Hazan, H., & Levin, I. (2021). Preschoolers' private speech during spelling in Hebrew: the role of word structure and letter position. *Reading and Writing, 34*(5), 1171–1190. <https://doi.org/10.1007/s11145-020-10102-9>.
- Aram, D., & Levin, I. (2001). Mother-child joint writing in low SES: Socio-cultural factors, maternal mediation and emergent literacy. *Cognitive Development, 16*(3), 831–852.
- Aslin, R. N., & Newport, E. L. (2012). Statistical learning: From acquiring specific items to forming general rules. *Current Directions in Psychological Science, 21*(3), 170–176. <https://doi.org/10.1177/0963721412436806>
- Baroody, A. E., & Diamond, K. E. (2013). Measures of preschool children's interest and engagement in literacy activities: Examining gender differences and construct dimensions. *Early Childhood Research Quarterly, 28*(2), 291–301. <https://doi.org/10.1016/j.ecresq.2012.07.002>
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412. <https://doi.org/10.1016/j.jml.2007.12.005>.
- Beery, K. E., & Buktenica, N. A. (1989). *Developmental test of visual-motor integration-revised*. Follett.
- Bloodgood, J. W. (1999). What's in a name? Children's name writing and literacy acquisition. *Reading Research Quarterly, 34*, 342–367. <https://doi.org/10.1598/RRQ.34.3.5>
- Both-De Vries, A. C., & Bus, A. G. (2008). Name writing: A first step to phonetic writing? Does the name have a special role in understanding the symbolic function of writing? *Literacy Teaching and Learning, 12*, 37–55. <https://doi.org/10.1598/RRQ.34.3.5>
- Bourke, L., Davies, S. J., Sumner, E., & Green, C. (2014). Individual differences in the development of early writing skills: Testing the unique contribution of visuo-spatial working memory. *Reading and Writing: An Interdisciplinary Journal, 27*, 315–335. <https://doi.org/10.1007/s11145-013-9446-3>
- Chatterji, M. (2006). Reading achievement gaps, correlates, and moderators of early reading achievement: Evidence from the Early Childhood Longitudinal Study (ECLS) kindergarten to first grade sample. *Journal of Educational Psychology, 98*(3), 489–507. <https://doi.org/10.1037/0022-0663.98.3.489>
- Chetail, F. (2015). Reconsidering the role of orthographic redundancy in visual word recognition. *Frontiers in Psychological Science, 18*, 645. <https://doi.org/10.3389/fpsyg.2015.00645>
- Diamond, K. E., Gerde, H. K., & Powell, D. R. (2008). Development in early literacy skills during the pre-kindergarten year in Head Start: Relations between growth in children's writing and understanding of letters. *Early Childhood Research Quarterly, 23*(4), 467–478. <https://doi.org/10.1016/j.ecresq.2008.05.002>
- Dingemann, C., Sonne, M., Ure, B., Bohnhorst, B., von Kaisenberg, C., & Pirr, S. (2019). Impact of maternal education on the outcome of newborns requiring surgery for congenital malformations. *PLoS ONE, 14*(4), e0214967. <https://doi.org/10.1371/journal.pone.0214967>
- Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Review of Educational Research, 71*(3), 393–447.
- Elimelech, A., Shaanan, M., & Aram, D. (2020). Evaluating preschoolers' references to characteristics of the Hebrew orthography via a computerized early spelling game. *Reading Research Quarterly, 55*(2), 235–250. <https://doi.org/10.1002/rrq.264>
- Ergül, C., Sarica, A. D., Akoglu, G., & Karaman, G. (2017). The home literacy environments of Turkish kindergarteners: Does SES make a difference? *International Journal of Instruction, 10*, 187–202. <https://doi.org/10.12973/iji.2017.10112a>
- Feder, K. P., & Majnemer, A. (2007). Handwriting development, competency, and intervention. *Developmental Medicine & Child Neurology, 49*, 312–317. <https://doi.org/10.1111/j.1469-8749.2007.00312.x>
- Ferreiro, E., & Teberosky, A. (1982). *Literacy before schooling*. Heinemann Educational Books Inc., 70 Court St., Portsmouth, NH 03801.
- Frost, R., Katz, L., & Bentin, S. (1987). Strategies for visual word recognition and orthographical depth: A multilingual comparison. *Journal of Experimental Psychology: Human Perception and Performance, 13*(1), 104–115. <https://doi.org/10.1037/0096-1523.13.1.104>
- Gerde, H. K., Bingham, G. E., & Wasik, B. A. (2012). Writing in early childhood classrooms: Guidance for best practices. *Early Childhood Education Journal, 40*, 351–359. <https://doi.org/10.1007/s10643-012-0531-z>
- Graham, S., Struck, M., Santoro, J., & Berninger, V. W. (2006). Dimensions of good and poor handwriting in first and second graders: Motor programs, visual-spatial arrangement, and letter formation

- parameter setting. *Developmental Neuropsychology*, 29, 43–60. https://doi.org/10.1207/s15326942dn2901_4
- Graham, S., Weintraub, N., & Berninger, V. (2001). Which manuscript letters do primary grade children write legibly? *Journal of Educational Psychology*, 93, 488–497. <https://doi.org/10.1037/0022-0663.93.3.488>
- Hilbe, J. M. (2017). The statistical analysis of count data. *Culture and Education*, 29(3), 409–460.
- Horton, N. J., & Lipsitz, S. R. (1999). Review of software to fit generalized estimating equations regression models. *The American Statistician*, 53(2), 160–169.
- Jones, C. D. O. (2015). Effects of writing instruction on kindergarten students' writing achievement: An experimental study. *The Journal of Educational Research*, 108(1), 35–44. <https://doi.org/10.1080/00220671.2013.836466>
- Justice, L. M., Pence, K., Bowles, R. B., & Wiggins, A. (2006). An investigation of four hypotheses concerning the order by which 4-year-old children learn the alphabet letters. *Early Childhood Research Quarterly*, 21, 374–389. <https://doi.org/10.1016/j.ecresq.2006.07.010>
- Kessler, B., Pollo, T. C., Treiman, R., & Cardoso-Martins, C. (2013). Frequency analyses of nonphonological spellings as predictors of success in conventional spelling. *Journal of Learning Disabilities*, 46, 252–259. <https://doi.org/10.1177/0022219412449440>
- Lavine, L. O. (1977). Differentiation of letterlike forms in prereading children. *Developmental Psychology*, 13(2), 89–94. <https://doi.org/10.1037/0012-1649.13.2.89>
- Lee, J. A. C., & Al Otaiba, S. (2015). Socioeconomic and gender group differences in early literacy skills: A multiple-group confirmatory factor analysis approach. *Educational Research and Evaluation*, 21(1), 40–59. <https://doi.org/10.1080/13803611.2015.1010545>
- Levin, I., & Aram, D. (2005). Children's names contribute to early literacy: A linguistic and social perspective. In D. D. Ravid & H. B. Z. Shyldkrot (Eds.), *Perspectives on language and language development* (pp. 219–239). Springer. https://doi.org/10.1007/1-4020-7911-7_17
- Levin, I., & Aram, D. (2013). Promoting early literacy via practicing invented spelling: A comparison of different mediation routines. *Reading Research Quarterly*, 48(3), 221–236. <https://doi.org/10.1002/rrq.48>
- Levin, I., Aram, D., Tolchinsky, L., & McBride, C. (2013). Maternal mediation of writing and children's early spelling and reading: The Semitic abjad versus the European alphabet. *Writing System Research*, 5, 134–155. <https://doi.org/10.1080/17586801.2013.797335>
- Levin, I., Both-de Vries, A. C., Aram, D., & Bus, A. (2005). Writing starts with own name writing: From scribbling to conventional spelling in Israeli and Dutch children. *Applied Psycholinguistics*, 26, 463–477.
- Levin, I., & Bus, A. G. (2003). How is emergent writing based on drawing? Analyses of children's products and their sorting by children and mothers. *Developmental Psychology*, 39(5), 891–905. <https://doi.org/10.1037/0012-1649.39.5.891>
- Levin, I., Share, D. L., & Shatil, E. (1996). A qualitative-quantitative study of preschool writing: Its development and contribution to school literacy. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 271–293). Lawrence Erlbaum Associates Inc.
- Lipka, O., Katzir, T., & Shaul, S. (2016). The basis of reading fluency in first grade of Hebrew speaking children. In A. Khateb & I. Bar (Eds.), *Reading fluency* (pp. 91–94). Springer. https://doi.org/10.1007/978-3-319-30478-6_6
- Martins, M. A., Albuquerque, A., Salvador, L., & Silva, C. (2013). The impact of invented spelling on early spelling and reading. *Journal of Writing Research*, 5, 215–237. <https://doi.org/10.17239/jowr-2013.05.02.3>
- McTigue, E. M., Schwippert, K., Upstad, P. H., Lundetræ, K., & Solheim, O. J. (2020). Gender differences in early literacy: Boys' response to formal instruction. *Journal of Educational Psychology*. <https://doi.org/10.1037/edu0000626>
- Meece, J. L., Glienke, B. B., & Burg, S. (2006). Gender and motivation. *Journal of School Psychology*, 44(5), 351–373. <https://doi.org/10.1016/j.jsp.2006.04.004>
- Monroy, C., Meyer, M., Gerson, S., & Hunnius, S. (2017). Statistical learning in social action contexts. *PLoS ONE*, 12(5), e0177261.
- Neumann, M. M., Hood, M., & Neumann, D. L. (2009). The scaffolding of emergent literacy skills in the home environment: A case study. *Early Childhood Education Journal*, 36(4), 313–319. <https://doi.org/10.1007/s10643-008-0291-y>

- Ouellette, G., & Sénéchal, M. (2008). Pathways to literacy: A study of invented spelling and its role in learning to read. *Child Development*, 79(4), 899–913. <https://doi.org/10.1111/j.1467-8624.2008.01166.x>
- Ouellette, G., & Sénéchal, M. (2017). Invented spelling in kindergarten as a predictor of reading and spelling in grade 1: A new pathway to literacy, or just the same road, less known? *Developmental Psychology*, 43(1), 77–88. <https://doi.org/10.1037/dev0000179>
- Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences*, 10, 233–238. <https://doi.org/10.1016/j.tics.2006.03.006>
- Peterson, S. S., & Parr, J. (2012). Gender and literacy issues and research: Placing the spotlight on writing. *Journal of Writing Research*, 3, 151–161. <https://doi.org/10.17239/jowr-2012.03.03.1>
- Pollo, T. C., Kessler, B., & Treiman, R. (2005). Vowels, syllables, and letter names: Differences between young children's spelling in English and Portuguese. *Journal of Experimental Child Psychology*, 92, 161–181. <https://doi.org/10.1016/j.jecp.2005.01.006>
- Pollo, T. C., Kessler, B., & Treiman, R. (2009). Statistical patterns in children's early writing. *Journal of Experimental Child Psychology*, 104, 410–426. <https://doi.org/10.1016/j.jecp.2009.07.003>
- Puranik, C. S., & Al Otaiba, S. (2012). Examining the contribution of handwriting and spelling to written expression in kindergarten children. *Reading and Writing: An Interdisciplinary Journal*, 25(7), 1523–1546. <https://doi.org/10.1007/s11145-011-9331-x>
- Puranik, C. S., & Lonigan, C. J. (2011). From scribbles to scrabble: Preschool children's developing knowledge of written language. *Reading and Writing: An Interdisciplinary Journal*, 24, 567–589. <https://doi.org/10.1007/s11145-009-9220-8>
- Puranik, C. S., Petscher, Y., & Lonigan, C. J. (2013). Dimensionality and reliability of letter writing in 3-to 5-year-old preschool children. *Learning and Individual Differences*, 28, 133–141. <https://doi.org/10.1016/j.lindif.2012.06.011>
- Raudenbush, S. W., Bryk, A. S., & Congdon, R. (2013). *HLM 7.01 for Windows. Computer software*. Skokie, IL: Scientific Software International Inc.
- Ravid, D. D. (2012). Spelling morphology: The psycholinguistics of Hebrew spelling. *Springer*. <https://doi.org/10.1007/978-1-4419-0588-8>
- Read, C., & Treiman, R. (2013). Children's invented spelling: What we have learned in forty years. In M. Piattelli-Palmarini & R. C. Berwick (Eds.), *Rich languages from poor inputs* (pp. 197–211). Oxford University Press.
- Rendall, M. S., Weden, M. M., & Brown, J. (2021). Family and household sources of poverty for Black, Hispanic, and White newborns. *Journal of Marriage and Family*. <https://doi.org/10.1111/jomf.12781>
- Raudenbush, S.W., Bryk, A.S., & Congdon, R. (2013). *HLM7.01 for Windows [Computersoftware]*. Skokie, IL: Scientific Software International Inc
- Ritchey, K. D. (2008). The building blocks of writing: Learning to write letters and spell words. *Reading and Writing: An Interdisciplinary Journal*, 21, 27–47. <https://doi.org/10.1007/s11145-007-9063-0>
- Robins, S., Ghosh, D., Rosales, N., & Treiman, R. (2014). Letter knowledge in parent–child conversations: Differences between families differing in socio-economic status. *Frontiers in Psychology*, 5, 632. <https://doi.org/10.3389/fpsyg.2014.00632>
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (2004). Statistical learning by 8-month-old infants. In D. A. Balota & E. J. Mrash (Eds.), *Cognitive psychology: Key readings* (pp. 538–542). Psychology Press.
- Share, D. L. (2020). Extricating reading science from entrenched Anglocentricism, Eurocentricism, and Alphabetism and embracing global diversity: A personal journey. *International Journal for Research in Learning Disabilities*, 4(2), 3–14.
- Share, D. L., & Levin, I. (1999). Learning to read and write in Hebrew. In M. Harris & G. Hatano (Eds.), *Learning to read and write: A cross-linguistic perspective* (pp. 89–111). Cambridge University Press.
- Shatil, E., Share, D. L., & Levin, I. (2000). On the contribution of kindergarten writing to grade 1 literacy: A longitudinal study in Hebrew. *Applied Psycholinguistics*, 21, 1–21. <https://doi.org/10.1017/S0142716400001016>
- Shatil, J. (1993). *The psychography of the child*. Modan Press. in Hebrew.
- Shoken, S., & Shor, B. (2010). How to send secret information. (in Hebrew). <http://www.cs.tau.ac.il/~bchor/CSU/Crypto.pdf>
- Taub Center. (2019). A picture of the nation. <http://taubcenter.org.il/a-picture-of-the-nation-2020-eng/>
- Tolchinsky, L. (2003). *The cradle of culture and what children know about writing and numbers before being taught*. Erlbaum.

- Tolchinsky, L. (2005). The emergence of writing. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 83–96). Guilford.
- Tolchinsky, L., Levin, I., Aram, D., & McBride-Chang, C. (2012). Building literacy in alphabetic, abjad and morphosyllabic systems. *Reading and Writing: An Interdisciplinary Journal*, 25(7), 1573–1598. <https://doi.org/10.1007/s11145-011-9334-7>
- Treiman, R. (2006). Knowledge about letters as a foundation for reading and spelling. In R. M. Joshi & P. G. Aaron (Eds.), *Handbook of orthography and literacy* (pp. 581–599). Erlbaum.
- Treiman, R., Gordon, J., Boada, R., Peterson, R. L., & Pennington, B. F. (2014). Statistical learning, letter reversals, and reading. *Scientific Studies of Reading*, 18, 383–394. <https://doi.org/10.1080/10888438.2013.873937>
- Treiman, R., & Kessler, B. (2004). The case of case: Children's knowledge and use of upper-and lower-case letters. *Applied Psycholinguistics*, 25(3), 413–428. <https://doi.org/10.1017/S0142716404001195>
- Treiman, R., & Kessler, B. (2011). Similarities among the shapes of writing and their effects on learning. *Written Language and Literacy*, 14, 39–57. <https://doi.org/10.1075/wll.14.1.03tre>
- Treiman, R., & Kessler, B. (2013). Learning to use an alphabetic writing system. *Language Learning and Development*, 9(4), 317–330.
- Treiman, R., Kessler, B., & Bourassa, D. (2001). Children's own name influence their spelling. *Applied Psycholinguistics*, 22, 555–570. <https://doi.org/10.1017/S0142716401004040>
- Treiman, R., Levin, I., & Kessler, B. (2007). Learning of letter names follows similar principles across languages: Evidence from Hebrew. *Journal of Experimental Child Psychology*, 96, 87–106. <https://doi.org/10.1016/j.jecp.2006.08.002>
- Treiman, R., Levin, I., & Kessler, B. (2012). Linking the shapes of alphabet letters to their sounds: The case of Hebrew. *Reading and Writing: An Interdisciplinary Journal*, 25, 569–585. <https://doi.org/10.1007/s11145-010-9286-3>
- Turk-Browne, N. B., Scholl, B. J., Chun, M. M., & Johnson, M. K. (2009). Neural evidence of statistical learning: Efficient detection of visual regularities without awareness. *Journal of Cognitive Neuroscience*, 21, 1934–1945. <https://doi.org/10.1162/jocn.2009.21131>
- Velan, H., & Frost, R. (2011). Words with and without internal structure: What determines the nature of orthographic and morphological processing? *Cognition*, 118, 141–156. <https://doi.org/10.1016/j.cognition.2010.11.013>
- Vernon-Feagans, L., Bratsch-Hines, M., Reynolds, E., & Willoughby, M. (2020). How early maternal language input varies by race and education and predicts later child language. *Child Development*, 91(4), 1098–1115. <https://doi.org/10.1111/cdev.13281>
- Weintraub, N., & Graham, S. (2000). The contribution of gender, orthographic, finger function, and visual-motor processes to the prediction of handwriting status. *Occupation, Participation and Health*, 20, 121–140. <https://doi.org/10.1177/153944920002000203>
- Welsch, J. G., Sullivan, A., & Justice, L. M. (2003). That's my letter! What preschoolers' name writing representations tell us about emergent literacy knowledge. *Journal of Literacy Research*, 35(2), 757–776. https://doi.org/10.1207/s15548430jlr3502_4
- Younger, J. W., Lee, K. W., Demir-Lira, O. E., & Booth, J. R. (2019). Brain lateralization of phonological awareness varies by maternal education. *Developmental Science*, 22(6), e12807. <https://doi.org/10.1111/desc.12807>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.