Journal of Experimental Child Psychology 108 (2011) 786–801

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

# Early differentiation between drawing and writing in Chinese children

Rebecca Treiman [a], Li Yin [b,*]

[a] Department of Psychology, Washington University in St. Louis, St. Louis, MO 63130, USA
[b] Department of Foreign Languages and Berkeley Center for Advanced Psychological Studies, Tsinghua University, Haidian District, 100084 Beijing, China

ARTICLE INFO

*Article history:*
Received 21 April 2010
Revised 28 August 2010
Available online 23 November 2010

*Keywords:*
Writing
Drawing
Emergent literacy
Chinese
Iconicity
Cultural transmission
Name writing

ABSTRACT

Children under 3½ years of age or so are often thought to produce the same types of scribbles for writing and drawing. We tested this idea by asking Chinese 2- to 6-year-olds to write and draw four targets. In Study 1, Chinese adults judged the status of the productions as writings or drawings. The adults performed significantly above the level expected by chance even with the productions of 2- to 2½-year-olds. In Study 2, we examined specific characteristics of the children's writings and drawings. Although the younger children's scribbles bore little resemblance to the correct characters, they tended to be smaller, sparser, and more angular than their artwork, with less filling in. Differences were also found in paper use and implement use. Children did not appear to distinguish writing from drawing for their own names before they did so for other targets.



## Introduction

Written words and pictures are similar in some ways and different in others. They are similar in that they both involve marks that appear on the surfaces of objects. Both are artificial rather than natural. However, some aspects of the surface forms are different. For example, writing is often composed of small black marks on a white background. Pictures are often larger and more colorful, and they are more likely to consist of outlines that are filled in. Writing and pictures also differ in their symbolic function. The marks of writing designate specific units of a language. For example, the written form

* Corresponding author.
*E-mail address:* yinl@mail.tsinghua.edu.cn (L. Yin).

0022-0965/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jecp.2010.08.013

*cat* derives its meaning from the fact that its units represent a linguistic form, which in turn represents a meaning. One must know the conventions of the English writing system, as well as the English language, to produce and interpret the marks. A realistic picture of a cat, in contrast, derives its meaning from the similarities between its form and that of its referent. The picture is iconic in a way that the written form is not.

Early knowledge about the characteristics of writing, which some researchers have called print awareness or concepts of print, appears to be related to the later development of conventional literacy (e.g., Tunmer, Herriman, & Nesdale, 1988). Some researchers have examined young children's knowledge about print by assessing their ability to distinguish between writing and drawing in perceptual tasks (e.g., Lavine, 1977). In the current study, we took another approach: examining children's productions when asked to write and to draw. For reasons to be explained here, we examined this ability among young Chinese children.

Adults often use the same term to refer to young children's writings and drawings, calling them *scribbles* in English or *tú yā* (涂鸦) in Mandarin Chinese. They often assume that scribbling is purely a motor activity and that children do not concern themselves with the appearance of the product. Indeed, a number of researchers have suggested that most children younger than 3½ or 4 years of age produce similar marks for writing and drawing, failing to differentiate the two. Gombert and Fayol (1992) studied the writings of French children between ages 2;10 (years;months) and 3;10, observing that the majority of the writings "exhibit similar workmanship" (p. 32) to the children's pictures. Only a few of the children in their study, most on the older end of the age range, were judged to make a distinction. Bader and Hildebrand (1991) stated that nearly 60% of U.S. children between ages 3;6 and 4;6 failed to distinguish writing from drawing in their production, and Noyer and Baldy (2005) characterized approximately half of the written productions of a group of French children with a mean age of 3;4 as showing no such differentiation.

Other investigators have reported that children sometimes include pictorial elements in their writing, potentially another way in which they confuse it with drawing. This phenomenon was first reported by Luria (1929/1978), who asked Russian children to write down sentences so as to remember them. He mentioned that 4- and 5-year-olds sometimes represented the colors, sizes, or shapes of objects mentioned in the sentences in their writing, for example, scribbling with a black pencil to write *smoke* but using other colors for other words. Levin and Tolchinsky Landsmann (1989) reported that Israeli 5- and 6-year-olds, when asked to write pairs of words such as *tomato* and *cucumber,* sometimes wrote tomato in red and cucumber in green. The use of referential color was fairly common in this study, more common than writing a word for a large object with more marks or larger marks than writing a word for a small object. Tolchinsky-Landsmann and Levin (1985) made similar observations among Israeli 3- to 5-year-olds, and Ferreiro and Teberosky (1982) observed that Argentinean children sometimes produced longer lines of scribble when writing words that stood for large objects than when writing words that stood for small objects.

Although some researchers have claimed that children younger than 3½ or 4 years of age do not usually distinguish between writing and drawing in production, others (e.g., Karmiloff-Smith, 1992) have suggested that they do. The U.S. children in several case studies appeared to make some distinctions between writing and drawing before 3 years of age (Baghban, 1984; Grinnell & Burris, 1983; Martens, 1996; Schickedanz, 1990). In another study, a teacher researcher at a U.S. preschool concluded that by 3 years of age, 15 of 18 children produced different types of marks for writing and for art (Rowe, 2008). However, the characteristics of the marks were not examined quantitatively in these studies, and statistical analyses were not reported.

Levin and Bus (2003) went beyond the previous work by asking adults who were not present during young children's mark making to classify the products as writing or drawing and by subjecting the results to statistical analysis. Levin and Bus examined drawings and writings produced by Israeli and Dutch children ages 2;4 to 3;0, 3;1 to 3;9, and 3;10 to 4;5. The children were given a choice of felt-tipped markers of different colors with which to write and draw specified targets. In a task that we refer to as sorting by target, mothers with children in the same age range as those in the study were shown all of the children's productions of a given target and were asked to classify each as writing or drawing. According to the statistical tests that Levin and Bus presented, significantly above-chance performance on both writings and drawings was not consistently found for children in either of the

two younger groups. In another task, sorting by child, mothers saw the productions of one child at a time and were asked to put them into two piles: one for drawing and one for writing. Levin and Bus reported that performance was not significantly above chance for the youngest group when examining the results for the Dutch and Israeli children separately but that it was significantly above chance for the combined data. Given the statistical tests that the researchers reported, above-chance results could have been found for the combined group of children at the youngest age level in the by-child sorting task because just a few children differentiated writing from drawing.

The results just reviewed suggest that at least some children distinguish writing from drawing before 3½ years of age. However, we do not know how prevalent this is, and we have little information about the specific distinctions that children of this age may make. Based on the results of Levin and Bus (2003), for example, young children may distinguish between writing and drawing primarily on the basis of the writing implements they choose, such as colored markers for drawing and black ones for writing. It is not clear from this study whether young children make different types of marks in the two tasks. The current study was designed to determine whether naive adults could tell which of children's products were produced in response to requests to write and which ones were produced in response to requests to draw. In addition, we went beyond previous studies by examining the characteristics of the productions in a quantitative manner.

The current study also went beyond previous work by examining a group of children whose writing and drawing have been little studied: children from China. The characteristics of the Chinese writing system make this a particularly interesting population to study. Chinese characters evolved from pictures, and although many have lost their original iconicity, some retain visual features that provide a clue to their meaning. For instance, 人 'human' looks rather like a person who is walking on two legs, and university students with no knowledge of Chinese perform above the level of chance when shown two photographs and asked to select the one that depicts the character's meaning (Luk & Bialystok, 2005). Such characters have been reported to constitute approximately a quarter of those in Chinese first-grade reading materials (Shu, Chen, Anderson, Wu, & Xuan, 2003). We know of no tabulations of the characters that Chinese children see before the first-grade level, but we suspect that iconic characters of the sort just mentioned are fairly common for them as well. If some of the writing that Chinese children see has a degree of iconicity, as pictures do, then Chinese children might be late to learn about the distinction between writing and drawing and might often incorporate pictorial elements in their writing.

The only evidence to date on writing–drawing differentiation in children exposed to Chinese comes from a study by Chan and Louie (1992) in which Hong Kong 3- to 5-year-olds were asked to draw pictures of themselves and then to write their names. The majority of the children, including 75% of the 3-year-olds, were judged to make different types of marks in the two tasks. The researchers stated that the written names tended to be smaller and less circular than the pictures, but they did not quantify these impressions.

In the study of Chan and Louie (1992), as in a number of other studies, the writing that children did was often their own names. Studies carried out in cultures that use alphabetic writing systems suggest that the ability to distinguish written words from pictures may emerge first for children's own names. In such cultures, children produce conventional or near-conventional versions of their names at a time when they produce primitive renditions of other words. For example, the Israeli 4- and 5-year-olds studied by Tolchinsky-Landsmann and Levin (1985) were more likely to use real Hebrew letters when writing their names than when writing other targets and were more likely to use correct letters. Levin, Both-de Vries, Aram, and Bus (2005) also reported that Israeli and Dutch children were more advanced at writing their own names than at writing other words. Moreover, young children who are exposed to alphabetic writing systems sometimes write other words using letters from their own names (French children: Gombert & Fayol, 1992; Brazilian children: Pollo, Kessler, & Treiman, 2009; U.S. children: Pollo et al., 2009, and Treiman, Kessler, & Bourassa, 2001). However, a superiority for names has not always been observed in learners of alphabetic writing systems. Tolchinsky-Landsmann and Levin (1985) reported that products that were made up of units that were small, separated by blanks, and arranged along a line were not more common when Israeli 3- to 5-year-olds wrote their names than when they wrote other targets.

In the current study, we compared Chinese children's writing of their names and other targets. Even if differentiation between writing and drawing takes place earlier for their names than for other words in learners of alphabetic writing systems, this might not be true for Chinese. The set of Chinese characters is much larger than the set of alphabet letters. Producing their names might be difficult for Chinese children, especially as compared with producing the simple one-character words that were also included in the current study.

Before introducing our study, we present some further information about the Chinese writing system and about how children are exposed to it. Chinese characters are composed of lines and dots, rarely of circular forms as in the Latin letters C and O. Each character generally occupies a square-shaped space. In the government-funded kindergartens that Chinese children attend beginning at 2 or 3 years of age, children are read to often, see characters on posters on walls and in other locations, and have the opportunity to look at books on their own. In the kindergarten attended by the children in the current study, children had a small amount of exposure to English starting at 4 years of age but were not exposed to pinyin. In this and other kindergartens, children's full names, which usually consist of one character for their family names and two characters for their given names, appear on their cups, cots, kerchiefs, and so on. Teachers identify characters for children if they ask but do not provide systematic formal instruction in character reading and writing. Such instruction begins during the first year of primary school at 6 or 7 years of age. However, some parents send children 1 or 2 years younger than this to private preschools that offer such instruction. Formal instruction in character writing during primary school involves writing in black pencil on grid paper, the squares of which are used to help children form the characters. As children get older, they gradually switch to writing in pen on lined paper. Although children are not formally taught to write characters in kindergarten, they are encouraged to write their full names on greeting cards, artwork, on so on, with teachers helping as needed. Kindergartners also draw a good deal, often with colored crayons. They are also exposed to traditional black-on-white Chinese line drawings, and children age 4 years or over often receive some instruction in how to produce them.

The Chinese 2- to 6-year-olds in the current study were asked to write four targets, one of which was their names, and also to draw pictures of the targets. We gave the children a choice of writing implements and types of paper. This allowed us to ask such questions as whether children use object-appropriate colors to draw objects and the same object-appropriate colors to write them and also whether children select different types of paper for writing and drawing. In Study 1, Chinese adults judged the status of the productions as writings or drawings. In Study 2, we examined the properties of the productions quantitatively, looking both at the nature of the marks and the implements and papers that the children chose. In both studies, we were primarily interested in whether 2- and 3-year-olds distinguished between drawing and writing and, if so, how they did so. The older children were included mainly to verify whether they made distinctions in the ways we expected.

The results of Levin and Bus (2003) suggest that any differences that young children may produce between writing and drawing are subtle and that adults need experience with an individual child's productions so as to notice them. In one of the subexperiments of Study 1, Study 1a, we gave the adults even more such experience than Levin and Bus did in their by-child sorting task by having the adults study several pairs of written and drawn productions of an individual child before making a judgment about a final pair from the same child. In this way, we sought to take advantage of the adult participants' perceptual learning skills to show us whether young Chinese children distinguish writing from drawing.

## Study 1

### *Method*

#### *Children*

The child participants were 109 native Chinese speakers (57 girls and 52 boys) who attended a public kindergarten in a middle-class area of Beijing. Children ranged in age from 2;0 to 6;8. Table 1 shows the number of children in each year group. Children were selected from two classes at each age level.

**Table 1**
Mean proportions of correct responses by adults in Studies 1a and 1b.

| Year group | Number of children | Study 1a | Study 1b |
|---|---|---|---|
| 2 | 23 | .79 | .77 |
| 3[a] | 21 | .94 | .93 |
| 4[b] | 21 | 1.00 | 1.00 |
| 5 | 22 | 1.00 | 1.00 |
| 6 | 22 | 1.00 | 1.00 |

[a] The productions of 1 child in this year group could not be included in Study 1a because there was no one target that the child both wrote and drew.
[b] The productions of 3 children in this year group could not be included in Study 1a because there was no one target that these children both wrote and drew.

For the writing task, the children were asked to write their names, *huǒ* ('fire', 火), *rì* ('sun', 日), and *shuǐ* ('water', 水). We chose these characters because they are fairly common and simple, all containing four strokes. The experimenter began by saying that many people like writing and that they may write individual characters, stories, and other things.[1] The experimenter went on to say that the children's mothers and fathers may write such things as the children's names and shopping lists. The experimenter stated that the children also write and that they have their own ways of doing so. The experimenter said that he or she would like to see how the children wrote. The experimenter told the children that however they chose to write, he or she would like it. The experimenter encouraged the children to try writing all of the targets. If the children insisted on skipping a particular target, the experimenter proceeded to the next one. Children were permitted to write more than one target on the same page or to use separate sheets for each target.

For the drawing task, the children were asked to draw a picture of themselves as well as pictures of sun, fire, and water. The instructions were similar to those for the writing task, suitably modified for the case of drawing. Other aspects of the procedure were also similar.

The experimenter allowed the children to choose the order of the drawing and writing tasks and then the order of targets within a task. Allowing the children a choice in these respects helped to maintain their interest in the tasks. We allowed the children to choose from among four types of paper (grid, yellow, pink, and white) and five implements (black pencil, black pen, yellow crayon, red crayon, and blue crayon) for each task. We selected crayons of these colors because, according to other children and adults we asked, they are appropriate colors for sun, fire, and water.

*Adults*

*Study 1a.* A total of 16 native Chinese-speaking college students were tested individually. After a brief introduction to the writing and drawing tasks the children had performed, the experimenter showed the adults all of the productions from a given child, which served as the practice items for that child, except for one pair that was the child's written and drawn versions of one specific target, the mystery pair. Adult participants were told what each practice item was meant to represent and whether it was intended as drawing or writing. After studying the practice items, the participants were shown the mystery pair, told what it was intended to represent, and asked which production came from which task. Participants were given feedback about whether their judgments of the mystery pairs were right or wrong, and they were told that the top three scorers would receive a prize. For the productions of 2- and 3-year-olds, the participants were asked to explain their reasons for each judgment.

Each adult made judgments about the productions of 105 children during a session of approximately 50 min that included a 10-min break. The order of the children was randomized for each participant. Of the 105 children, 61 drew and wrote all items and so had four potential mystery pairs. In these cases, 4 adults were assigned to each pair. Another 18 children had three such pairs, and in such cases 5 adults were assigned to one pair, 5 adults were assigned to the second pair, and 6 adults were

[1] In the instructions, the experimenter used the term *xiě* for writing (e.g., *xiě chū nǐ de míng zì* 'write your name') and the term *huà* for drawing (e.g., *huà shuǐ* 'draw water'), which are commonly used colloquial Chinese for children.

assigned to the third pair. Another 19 children had two pairs, meaning that 8 adults judged each pair, and 7 children had only one pair, in which case all of the adults judged it.

*Study 1b.* A total of 16 new native Chinese-speaking college students were tested individually. After a brief introduction to the writing and drawing tasks the children had performed, the experimenter showed the adult participants two piles of productions by a given child: one with all of the drawings and the other with all of the writings. Participants were given time to study the piles and, when ready, were asked to judge which one was made under instructions to draw and which one was made under instructions to write. For the productions of 2- and 3-year-olds, the participants were also asked to explain the reasons for their guesses. Participants received feedback about whether their judgments were right or wrong and were told that the top three scorers would receive a prize.

Each adult participant made judgments about the productions of all 109 children during a session of approximately 45 min that included a 10-min break. The order of the children was randomized for each participant.

### *Results and discussion*

We analyzed the correct responses using a multilevel model with children and judges as random factors. The analysis was carried out at the trial level with a binary response variable. We used the software package lme4 (Bates, 2009), selecting a generalized mixed-effects model with a logit link function and using the $p$ values calculated by this package that are based on the Wald statistic. For Study 1a, we included the factors of child sex, age (in months), order of the writing task relative to the drawing task, and whether or not the judge had available the maximum number of practice trials for a child (6) or not. (This last variable was dichotomized because it had a highly non-normal distribution that could not be improved by transformation.) Stimulus was treated as a fixed effect, and we contrasted self/name with each of the other stimuli. The analysis for Study 1b was similar except that stimulus was not included as a variable and the number of productions available to the judge was coded as the maximum possible (8) or not. In all analyses, nonbinary measures were centered at their grand mean.

In both studies, adults did better with the products of older children than with those of younger children ($p < .001$ for both). These effects may be seen in Table 1, which shows the results for each study broken down by year group. For neither study did any variable other than age contribute significantly to the model. Because ceiling effects in children age 4 years or over could have reduced the sensitivity of the analyses, we repeated the analyses with only the 2- and 3-year olds. The same pattern of significant and nonsignificant results was obtained.

The mean proportion of correct judgments on self/name was .86 for the 2- and 3-year-olds in Study 1a, as compared with .90, .84, and .84 for sun, fire, and water, respectively. We might have expected better performance on self/name if Chinese children distinguish between writing and drawing better for their own names than for other targets, as learners of alphabetic writing systems have been suggested to do. However, the contrasts between self/name and the other stimuli were not statistically significant.

A combined analysis of Studies 1a and 1b using the variables of age, study, and their interaction showed a main effect of age ($p < .001$) but no effects involving study. If children made distinctions involving properties that adults do not normally associate with writing versus drawing, such as always using red crayon for writing and blue crayon for drawing, we would have expected adults to perform well in Study 1a but not in Study 1b. The similar results of the two studies suggest that children produced distinctions that made sense to adults.

For each of the 44 2- and 3-year-olds, we used a binomial test to determine whether the proportion of correct responses by the judges across the two studies was significantly greater than the .50 that would be expected by chance. Of the 23 2-year-olds, 17 showed significantly above-chance performance, as did 20 of the 21 3-year-olds. Of the 10 children under age 2;6, 7 performed reliably above the level of chance.

The results suggest that many Chinese children, before their 3rd birthday and sometimes even before 2½ years of age, know that they should do something different when they write and when they

draw and that they make some differences in their productions. Moreover, the differences that children produce are ones that adults associate with writing as compared with drawing. However, we do not know from the results of this study which specific distinctions children make. It is possible that children choose different types of paper and/or different implements for writing and drawing but that they make the same types of marks in the two tasks. Such results would support the claim of Gombert and Fayol (1992) that children under age 3½ or 4 produce writings that "exhibit similar workmanship" to their drawings (p. 32). However, it is also possible that children produce different types of marks in the two tasks. We addressed these possibilities in Study 2 by examining the nature of the children's marks as well as the implements and paper they chose.

The comments of the adult participants in Study 1 provide some hints about the distinctions that young children may make between writing and drawing. Consider Fig. 1, which shows the productions of a boy age 2;0 and a girl age 2;8 who attempted all of the targets and for whom the adults performed relatively well (.88 and .75 for the boy and girl, respectively, in Study 1a; .75 and .69, respectively, in Study 1b). Some of the characteristics that the adults mentioned involved the children's choice of implements and paper. The judges also noticed some differences in workmanship between writing and drawing. For example, the boy's drawings contained fewer straight lines and more closed circles than his writing, and the girl's marks were smaller for writing than for drawing. The judges'

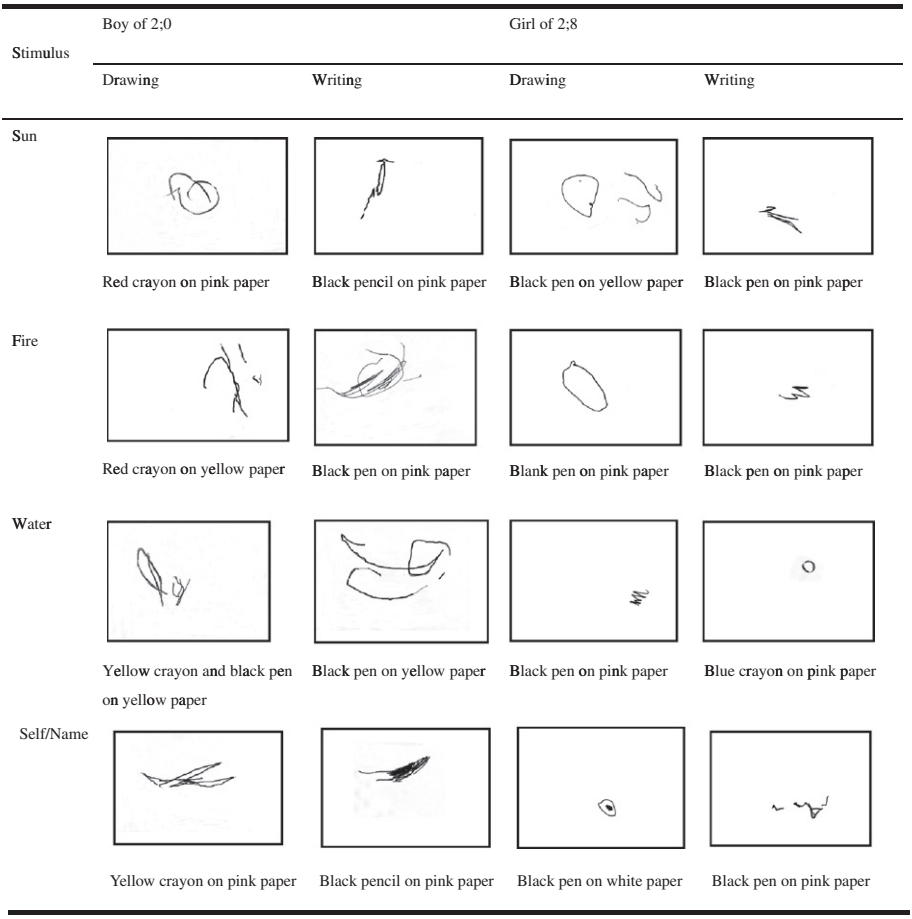


**Fig. 1.** Productions by a boy aged 2;0 and by a girl aged 2;8. (some redrawn to increase contrast).

comments, together with findings from the case studies that were cited earlier and studies with older children (Brenneman, Massey, Machado, & Gelman, 1996), were used to select characteristics for study in Experiment 2. Our selection of characteristics was also based on the properties that writing shares across systems.

## Study 2

### *Method*

A number of characteristics of the productions were examined. For each characteristic for which there was some subjectivity, two Chinese adults served as judges. The judges were from the same population as the Study 1 participants but had not taken part in that study. Unless otherwise mentioned, there were different judges for each characteristic. Interrater reliability for categorical measures was calculated using Cohen's kappa. When the judges disagreed, an additional judge was enlisted and the disagreement was solved through discussion. For continuous variables, the intraclass correlation coefficient was used as a measure of reliability and the final score was the average of the two judges' ratings. In what follows, we describe the specific characteristics that were scored, how they were scored, and the reliability of the scoring.

We measured the recognizability of the productions by showing judges the productions of all children by stimulus mixed randomly across ages. The judges were told what each product was intended to be, for example, a picture of the sun or the character for water. The judges determined whether the product resembled the target to at least some extent or whether it did not resemble the target at all. The reliability for this measure was .97 ($p < .001$).

Productions were scored as involving a single writing implement or involving more than one implement. Judges saw the productions of all children by stimulus mixed randomly across age groups. They were not told which products were intended as drawing and which ones were intended as writing, and they were not told the identity of the targets. Unless otherwise noted, a similar procedure was used for each of the following characteristics to be discussed. The reliability of the single versus multiple implement measure was 1.00 ($p < .001$).

Productions were scored as black only if children used a black pen and/or pencil but no colored implements (reliability = 1.00, $p < .001$). Productions that included color were further scored for whether they included any of yellow, blue, and red, with the same judges making decisions about each of these colors. The reliabilities for these judgments were .87, .90, and .91, respectively ($p < .001$ for all). (The reliabilities were less than perfect because when a very small amount of a color was used, judges sometimes disagreed on which color it was.)

We noted the type of paper that was selected on each trial, including those few cases in which children selected a paper but did not produce anything. We also noted whether each child ever wrote or drew more than one stimulus on a single piece of paper. These measures were objective and did not require judges. Using the videotapes, judges assessed whether or not the child rotated the paper during each production. The rotation measure had a reliability of .96 ($p < .001$). Because the scoring of this and other characteristics that were based on the videotaped record involved judges watching the videos in the sequence in which they were made, the judges in these cases, unlike most of the others, knew whether the child was attempting to draw or write and the identity of the target.

Judges measured the size of each product by finding the smallest rectangle that included the whole production, using two rulers, and calculating the area of the rectangle in squared centimeters (reliability = .95, $p < .001$). The "squarishness" of each product was determined by calculating the ratio of the short side to the longer side, such that a perfect square would receive a value of 1.0 (reliability = .90, $p < .001$). Judges also scored whether the child filled in his or her productions, as when drawing a circular outline for the sun and filling it with a block of color (reliability = .93, $p < .001$).

Density of marks was assessed by having judges compare the written and drawn versions of the same stimulus that were produced by the same child, without knowing which one was drawing and which one was writing, and determining which production had denser marks. Judges were shown pairs of products mixed randomly across age groups, and they made comparisons only when a child

both drew and wrote a given stimulus. Although this procedure necessitated the elimination of some data, we considered it appropriate because it is difficult to judge a single product on its own as having dense or sparse lines. The reliability of the density measure was a modest but significant .64 ($p < .001$). Using a similar procedure, judges determined which product in each pair had more curved marks (reliability = .69, $p < .001$).

To score the consistency of the products within a task, judges were shown two piles for each child: one with all of the drawings and the other with all of the writings. They were not told which pile was drawings and which one was writings. Judges were asked to determine which pile showed more similarity across the products. They were told that they could make this judgment on any basis that made sense to them, including type of paper, type of implement, or nature of marks. The reliability of this measure was .69 ($p < .001$).

We obtained an indication of task difficulty by noting whether the children attempted each item in each task. In addition, judges scored difficulty-related comments from the videotapes. For example, a child who said "I can't write sun" or asked "How should I draw fire?" was coded as having made a difficulty-related comment on that trial. We also coded object-related comments, as when a child mentioned properties that are associated with the target object (e.g., the yellowness of the sun), and writing-related comments, as when a child mentioned properties that are associated with a character (e.g., the strokes in the character for the sun). The same two individuals made judgments about all three types of comments. The reliabilities for difficulty-related, object-related, and writing-related comments were .96, .97, and .94, respectively ($p < .001$ for all).

Finally, judges scored the written productions for correctness. Each character of the children's names were scored individually as correct or incorrect, as were the characters for sun, fire, and water (reliability = .95, $p < .001$).

### *Results and discussion*

We analyzed the results for each characteristic at the trial level, unless otherwise stated, using multilevel analyses with participants as a random factor and age as a fixed factor. When there were individual data points for each task, stimulus, or both, we included task and stimulus as fixed factors. All possible interactions involving age, task, and stimulus were included as well. In most of the analyses that included stimulus as a factor, we contrasted self/name with each of the other stimuli. For continuous variables, we used the languageR program (Baayen, 2009) to estimate $p$ values using posterior distributions for the model parameters obtained by Markov chain Monte Carlo sampling. Because sex and order of the writing task relative to the drawing task did not produce any significant effects in Study 1, and because we wished to avoid the risk of overparameterization, we did not include these factors in the Study 2 analyses.

The first columns of data in Table 2 show the proportions of products that were judged to resemble the target to at least some extent. The multilevel analysis showed that there were more such products in drawing than in writing ($p = .002$) and more for older children than for younger children ($p < .001$). An additional effect, not shown in Table 2, is that children's productions of self/name were more recognizable than their productions of fire (.63 for self/name vs. .50 for fire, $p = .036$). This held true for both writing and drawing, as shown by the lack of interaction with task.

Even though the 2- and 3-year-olds' writings did not much resemble the conventional characters, the results of Study 1 suggest that they had some of the characteristics that are conventionally associated with writing. One such characteristic is that a character is usually written with a single implement, whereas several implements may be used to draw a picture. Table 2 shows the proportions of productions for each year group that were made with a single implement. There was a main effect of task ($p = .022$), such that children were more likely to use a single implement to write than to draw. The main effect of age ($p = .029$) arose because older children were less likely to confine themselves to one instrument than were younger children. There was also a three-way interaction involving task, age, and the contrast between water and self ($p = .023$) as well as a main effect of the water/self contrast ($p = .021$). Children were more likely to use multiple implements to draw pictures of water than to draw pictures of themselves, an effect that tended to be larger among the older children. One possible explanation for this result is that older children were influenced by traditional Chinese line

**Table 2**
Proportions of writings and drawings with various characteristics.

| Yeargroup | Resemble target to at least some extent | | Produced with single implement | | Produced wholly or in part with colored crayon | | Grid paper chosen | | Paper rotated during production | | More than one stimulus ever produced per page | | Filled-in forms produced | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Writing | Drawing | Writing | Drawing | Writing | Drawing | Writing | Drawing | Writing | Drawing | Writing | Drawing | Writing | Drawing |
| 2 | .15 | .27 | .98 | .89 | .43 | .69 | .16 | .12 | .01 | .05 | .00 | .00 | .00 | .01 |
| 3 | .27 | .48 | 1.00 | .74 | .57 | .67 | .15 | .02 | .06 | .17 | .30 | .10 | .00 | .00 |
| 4 | .75 | .86 | .98 | .66 | .27 | .55 | .09 | .00 | .00 | .34 | .47 | .00 | .00 | .22 |
| 5 | .89 | .97 | .99 | .53 | .15 | .72 | .17 | .03 | .05 | .36 | .77 | .00 | .02 | .23 |
| 6 | .99 | 1.00 | .98 | .57 | .14 | .63 | .04 | .08 | .00 | .41 | .59 | .02 | .00 | .42 |

drawings, which often feature human figures and which are drawn with a single implement, a black pen. The most important outcome for current purposes, however, is that differences among stimuli were virtually nonexistent in the writing task, where even the 2- and 3-year-olds almost always used a single implement both when writing their names and when writing other words.

Table 2 shows the proportions of productions in which colored crayons were used wholly or in part. Children were more likely to include color in their drawings than in their writings ($p < .001$), and younger children were generally more likely to use color than were older children ($p = .013$). There were differences among items especially in drawing, such that children were less likely to use colored crayons when drawing pictures of themselves than when drawing pictures of sun, fire, or water (.50 vs. .71). This led to significant interactions between task and the contrasts between self/name and each of the other items ($p < .032$) as well as a significant main effect of each contrast ($p < .001$). Differences among items in the drawing task tended to increase with age, resulting in a significant three-way interaction involving task, age, and the contrast between self and fire ($p = .02$) as well as significant two-way interactions involving age and the contrasts between self and fire ($p < .001$) and between self and water ($p = .007$). The differences in color use among items in the drawing task may reflect children's experience with traditional black-on-white Chinese line drawings, as mentioned earlier, in which human figures are a salient element. The most important point for current purposes, however, is that differences among items in the use of color were minimal in the writing task. Regardless of whether they were writing their own names or other characters, children as young as 2 and 3 years of age were less likely to use color than they were in drawing.

When children did use one or more colored crayons in writing or drawing, which color(s) did they choose? To address this question, we looked at those products that included color and calculated the proportion of these that included each specific color. The proportion of colored drawings that included blue was .73 for water as compared with .18 for the other targets. There was a significant association between target identity and the use of blue in drawing ($p < .001$ by Fisher's exact test, pooling across ages). (We used Fisher's exact test rather than multilevel analyses in analyzing children's use of colors because, especially for writing, relatively few productions included each specific color.) If children used pictorial elements in their writing, we would expect them to use blue crayon more often when writing water than when writing the other items. However, this difference was not statistically significant (.48 vs. .31). Red was more popular in drawings of sun and fire than in drawings of water and self (.73 vs. .33, $p < .001$). In writing, however, red was not significantly more popular for sun and fire than for water and self (.49 vs. .33). Children were more likely to use yellow when drawing the sun than when drawing the other targets. A little more than half of the drawings of the sun included yellow, more than for the other items (.51 vs. .28, $p < .001$). In writing, however, the proportion of products that included yellow was not higher for sun than for the other items (.17 vs. .28). Thus, although the children tended to choose object-appropriate colors for drawing, they did not generally prefer such colors for writing even when they did write with a colored implement.

Table 2 shows the proportions of trials on which children selected grid paper, which is associated with writing in China. The means for writing and drawing (.12 and .04, respectively) both were significantly less than the .25 that would be expected if children chose randomly among the four available types of paper ($p < .001$ by chi-square tests). However, grid paper was significantly more popular for writing than for drawing according to a multilevel analysis ($p = .011$). There was also a main effect of age ($p = .028$), such that older children used grid paper less often overall than younger children.

As Table 2 shows, children were more likely to rotate their paper while drawing than while writing ($p < .001$). Older children showed more rotation than younger children ($p = .004$). The interaction between age and task was not statistically significant.

Table 2 shows the proportions of children who produced more than one item on a single piece of paper at least once while performing a task. This was more likely to occur in the writing task than in the drawing task ($p < .001$). There was a significant interaction between task and age ($p = .039$), such that the difference between writing and drawing emerged starting at around 3 years of age.

The proportions of filled-in productions by children in each year group is shown in the rightmost columns of data in Table 2. Filling in was more common in drawing than in writing ($p = .006$), and it was also more common among older children than among younger children ($p < .001$). A separate analysis of drawings using the factors of stimulus and age showed that filled-in drawings were more

**Table 3**
Mean values of area and squarishness for writings and drawings.

| Year group | Area (log $cm^2$) | | | | Squarishness (shorter side/longer side) | | | |
|---|---|---|---|---|---|---|---|---|
| | Writing | | Drawing | | Writing | | Drawing | |
| | Name | Sun, fire, water | Self | Sun, fire, water | Name | Sun, fire, water | Self | Sun, fire, water |
| 2 | 1.38 | 1.57 | 1.89 | 1.87 | .69 | .70 | .70 | .72 |
| 3 | 1.35 | 1.21 | 1.55 | 1.70 | .71 | .67 | .68 | .67 |
| 4 | 1.53 | 1.13 | 2.01 | 2.00 | .72 | .76 | .71 | .63 |
| 5 | 1.22 | 0.80 | 2.04 | 1.81 | .73 | .74 | .77 | .64 |
| 6 | 1.41 | 0.57 | 2.10 | 2.03 | .41 | .70 | .70 | .75 |

common for fire than for self ($p = .042$) and that filled-in drawings of the sun showed a greater increase with age than filled-in drawings of self ($p = .002$).[2] Most important for current purposes is that these patterns were not observed in the writing task, where children hardly ever filled in their written names or other characters.

Table 3 provides information about the sizes of the children's writings and drawings. We used a log transformation in the analyses of product size since the untransformed variable was highly skewed. There was a main effect of task ($p < .001$), such that children's drawings were larger than their writings. There was also an interaction of task and age ($p = .007$). The difference in size between writing and drawing was not statistically significant for the 2-year-olds but was significant for each of the older groups. We also found three-way interactions involving task and age and the contrasts between self and sun and between self and water ($p = .042$ and $p < .001$, respectively) as well as a two-way interaction involving task and the contrast between water and self ($p = .012$). These interactions arose because, although there were no significant differences in the sizes of the drawings as a function of stimulus or age, some such differences were found in the writings. As the results in Table 3 show, children's names did not decrease in size with age, whereas other written forms did get smaller. This outcome reflects the fact that older children often wrote or attempted to write all of the characters in their names, and most of the children's names contained three characters.

Information about the squarishness of children's writings and drawings of name/self and the other items is provided in Table 3. Before 6 years of age, children's written renditions of their names and of other targets were similar in squarishness to their drawings. The 6-year-olds, in contrast, produced written names that were substantially less square than their other writings and drawings. This occurred because, as mentioned above, the older children often wrote all of the characters in their names. The statistical analyses showed main effects of task, the contrast between self and water, and the contrast between self and sun ($p < .001$ for all). The contrast between self and water interacted with task ($p < .001$), as did the contrast self and fire ($p = .005$). Age interacted with task ($p < .001$), the contrast between self and water ($p < .001$), and the contrast between sun and water ($p = .009$), and there were three-way interactions involving age, task, and the contrasts between self and water and between self and fire ($p < .001$ for both).

Table 4 shows the proportion of cases for each year group in which children's writings were judged to be denser and more curved than the corresponding drawings. A main effect of age was found for both variables ($ps = .015$ and .009, respectively). The 2-year-olds did not show a significant difference in density between writing and drawing according to a binomial test. For children in each of the older groups, however, writings were reliably less dense than drawings, with values significantly less than .50 ($p < .01$). Children's writings were less curved than their drawings, a difference that was statistically reliable even for the 2-year-olds ($p < .001$) and that increased with age. There were no significant differences between name/self and the other items for either density or curvedness.

[2] In a multilevel analysis that included the factors of task, stimulus, and age, the absence of filling in for three of the four items in the writing task led to an extremely large coefficient and an accompanying large standard error for the task effect. This is a well-known problem with the use of the Wald statistic for logit models (e.g., Menard, 2002).

**Table 4**
Proportions of cases in which writing was judged to be more dense, more curved, and more similar across productions than drawing.

| Year group | Density | Curvedness | Similarity across productions |
|---|---|---|---|
| 2 | .45 | .31 | .70 |
| 3 | .36 | .26 | .74 |
| 4 | .12 | .02 | .94 |
| 5 | .12 | .03 | .85 |
| 6 | .07 | .00 | .85 |

Also shown in Table 4 are the results involving similarity across productions. Children's writings were judged to be more similar to one another than their drawings 81% of the time, pooling across age groups, significantly more than the 50% that would be expected by chance ($p < .001$ by a binomial test). Although the results in Table 4 show a tendency for similarity to increase with age, this tendency was not significant according to a logistic regression.

Children attempted the requested items 99% of the time in the drawing task but only 76% of the time in the writing task. A multilevel analysis using the factors of task, age, and stimulus yielded a main effect of task ($p = .012$) but no other significant effects. Children's greater reluctance to write than to draw was also evident in the fact that difficulty-related comments were more common in the writing task, where they occurred on 13% of trials, than in the drawing task, where they occurred on 5% of trials. The data for this and other types of comments were generally too sparse for multilevel analyses involving the factors of task, age, and stimulus and their interactions. However, analyses using Fisher's exact tests showed that difficulty-related comments were more common in writing than in drawing ($p < .001$). Object-related comments, conversely, were more common in drawing than in writing (10% vs. 2%, $p < .001$). Writing-related comments were rare, occurring only in the writing condition (2% vs. 0%, $p = .004$).

The results presented so far show that as early as 2 years of age, children made some distinctions between writing and drawings in both the implements and papers they used and the types of marks they made. From an early age, children often used a single black implement to write. They were more likely to use multiple implements and colored implements to draw. Children used grid paper more often in writing than in drawing, and by 3 years of age their writing took up less of the page than did their drawing. The marks of writing were less dense and less curved than those of drawing, and drawing was characterized by more filled-in outlines and more paper rotation. In addition, children appeared to consider writing to be more difficult than drawing.

For the 2- and 3-year-olds, distinctions between writing and drawing were not more evident for name/self than for other targets. Differences emerged only at older ages when names, which consist of several characters, began to be more rectangular and to take up more space than single-character words. To further compare the results for names and other targets, we examined the proportions of characters that were written correctly. As the results in Table 5 show, children age 4 years or over did better on the characters from their names. This was true even though the characters in the children's names contained, on average, more strokes than the other characters. The 2- and 3-year-olds, however, never wrote either names or non-name characters correctly. A multilevel analysis of these data using the factors of age and character type (name vs. other) showed significant effects of age and character type and of their interaction ($p < .001$ for all).

**Table 5**
Mean proportions of characters that children produced correctly in writing of name and writing of other characters.

| Year group | Name | Sun, water, fire |
|---|---|---|
| 2 | .00 | .00 |
| 3 | .00 | .00 |
| 4 | .25 | .08 |
| 5 | .53 | .36 |
| 6 | .96 | .44 |

## General discussion

The current study had two main goals: to determine whether young Chinese children produce distinctions between writing and drawing and, if so, to examine quantitatively the distinctions they make. Several previous studies carried out in Western countries suggest that most children under 3½ or 4 years of age produce indistinguishable marks when they write and when they draw (Bader & Hildebrand, 1991; Gombert & Fayol, 1992; Noyer & Baldy, 2005). Other researchers have suggested that younger children make some distinctions (Baghban, 1984; Martens, 1996; Rowe, 2008; Schickedanz, 1990), but that evidence comes largely from studies using informal methods. We used more rigorous methods, assessing the ability of adults who were not present during the process of production to differentiate writings from drawings in Study 1 and measuring specific characteristics of the productions in Study 2. We found that even 2- and 3- year-olds made some distinctions between writing and drawing. There were distinctions in the marks themselves, in implement and paper use, and in the comments children made during the process of production.

The distinctions that Chinese 2- and 3-year-olds made between writing and drawing are similar to some of the distinctions that have been mentioned but rarely quantified in previous studies of learners of alphabetic writing systems. Among the apparent similarities between two groups are that writings were more likely to be made with a single black implement and less likely to consist of filled-in outlines. Also, writings were sparser and smaller than drawings and more likely to be similar to one another. These differences reflect characteristics of writing that hold across cultures and writing systems. Indeed, some investigators have suggested that children's early writings show graphic features that are common to all writing systems but not features that are specific to the particular system to which the children are exposed (Puranik & Lonigan, 2010; Tolchinsky, 2003). However, some properties of children's writings that were observed here, in particular the squarishness and angularity, may reflect properties of Chinese characters that do not hold universally. Moreover, children's greater use of grid paper for writing than for drawing appears to reflect aspects of Chinese culture. Comparisons of the early writings of Chinese children with those of children from other societies will be needed to determine whether universal features of writing emerge earlier than language- or culture-specific features or whether, as we suspect, features of the different types sometimes emerge contemporaneously.

Some previous studies suggest that learners of alphabetic writing systems sometimes incorporate pictorial features into their writing by, for example, writing the word for tomato with a red implement or writing the word for a large person with many letters (e.g., Levin & Tolchinsky Landsmann, 1989; Luria, 1929/1978; Tolchinsky-Landsmann & Levin, 1985). We found little evidence for this hypothesis in the case of color. Although the children tended to use object-appropriate colors in drawing, they did not do so consistently in writing. Additional research will be needed to replicate the current findings with learners of Chinese and to determine whether previous reports of use of pictorial strategies by learners of alphabetic writing systems, some of which are anecdotal, can be replicated. In any case, the fact that Chinese children in this study showed little tendency to write words in the same colors in which they drew them demonstrates a way these children distinguished drawing from writing.

Because some Chinese characters have a degree of iconicity, one might expect that Chinese children would be slower to learn about the distinction between writing and drawing than learners of alphabets and that Chinese children would use pictorial features in their writing more often than learners of alphabets. Comparisons across studies and cultures must be made with caution, but the data do not support these ideas. This probably reflects the fact that even the most iconic Chinese characters are much less iconic than most pictures. Indeed, Chinese parents do not seem to make much use of iconicity when teaching their youngsters to write characters (Lin et al., 2009).

As discussed in the Introduction, studies with learners of alphabetic writing systems suggest that the children's names play a special role in the development of writing. In the current study, Chinese children age 4 years or over produced more correct renditions of the characters in their own names than of other characters. In this respect, our results replicate the findings of several studies with learners of alphabetic writing systems (Levin et al., 2005; Tolchinsky-Landsmann & Levin, 1985). Chinese children under 4 years of age never wrote the characters of their names correctly, and their written

names were no more advanced than their written versions of other targets in that names were no more angular, no more likely to be produced with a single black implement, and so on. These findings, which are similar to those of a previous study with Hebrew children (Tolchinsky-Landsmann & Levin, 1985), suggest that at first differentiation between writing and drawing is no more advanced for children's names than for other words. After several years, names gain an advantage. This pattern may hold for learners of Chinese as well as for learners of alphabetic writing systems.

The current study went beyond previous work by looking at children who were younger than those in previous studies of writing, by looking at learners of Chinese, and by quantifying the characteristics of children's writing in ways that have not been done before. However, a limitation of the study is that each child wrote and drew only four targets. This was the most that many of the 2-year-olds could handle in a single session. However, future studies could include more than one session.

Despite the limitations of the study, the results have some potentially important implications for our understanding of literacy development. Studies have shown that the quality of a young child's writing attempts is a good predictor of that child's later literacy skills. In the studies reviewed in the National Early Literacy Panel's (2008) report, many of which were carried out in the United States, writing quality was usually assessed in kindergarten, which children typically begin after their 5th birthday or soon before kindergarten entry. If understanding that writing is different from drawing is an important foundation stone for the development of literacy, then meaningful differences among children might appear as early as 2 and 3 years of age. In future work, it will be important to determine whether the ability to produce distinctions between writing and drawing at these early ages helps to predict later literacy skills.

Our results also shed light on the general issue of how cultural tools are transmitted from one generation to the next. This transmission is often informal. Previous studies outside the domain of writing have shown that children imitate the purposeful actions of others. They imitate even actions that yield no obvious or interesting results, even actions whose functions they do not understand (e.g., Lyons, Young, & Keil, 2007; Yang, Sidman, & Bushnell, 2010). Our findings suggest that the same conclusions apply to writing. The 2- and 3-year-olds studied here attempted to reproduce some of the actions they had seen other people perform while writing, such as choosing a black pencil. They attempted to reproduce some of the results, such as forming marks that are sparse and small. Children did this even though they probably had little understanding of how the marks of writing function to represent meaning. Children learn a good deal about writing before they are formally taught to produce it, and studies of their written productions can shed light on the learning of literacy and on learning more generally.

## Acknowledgments

This research was supported by a grant from the National Natural Science Foundation of China (30800305). We thank Yan Hong, Wang Jingwen, Fu Huiqin, and the members of the Reading and Language Lab for their assistance and for comments on an earlier draft of the manuscript.

## References

Baayen, R. H. (2009). *LanguageR: Data sets and functions with Analyzing Linguistic Data: A Practical Introduction to Statistics.* Retrieved from http://www.CRAN.R-project.org/package=languageR.

Bader, L. A., & Hildebrand, V. (1991). An exploratory study of three to five year olds' responses on the Bader Reading and Language Inventory to determine developmental stages of emerging literacy. *Early Child Development and Care, 77*, 3–95.

Baghban, M. (1984). *Our daughter learns to read and write: A case study from birth to three*. Newark, DE: International Reading Association.

Bates, D. (2009). *Computational methods for mixed models.* Retrieved from http://www.cran.r-project.org/web/packages/lme4/vignettes/Theory.pdf.

Brenneman, K., Massey, C., Machado, S. F., & Gelman, R. (1996). Young children's plans differ for writing and drawing. *Cognitive Development, 11*, 397–419.

Chan, L., & Louie, L. (1992). Developmental trend of Chinese preschool children in drawing and writing. *Journal of Research in Childhood Education, 6*, 93–99.

Ferreiro, E., & Teberosky, A. (1982). *Literacy before schooling*. Portsmouth, NH: Heinemann.

Gombert, J. E., & Fayol, M. (1992). Writing in preliterate children. *Learning & Instruction, 2*, 23–41.

Grinnell, P. C., & Burris, N. A. (1983). Drawing and writing: The emerging graphic communication process. *Topics in Learning & Learning Disabilities, 3*, 21–32.
Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science*. Cambridge, MA: MIT Press.
Lavine, L. O. (1977). Differentiation of letterlike forms in prereading children. *Developmental Psychology, 13*, 89–94.
Levin, I., Both-de Vries, A., Aram, D., & Bus, A. (2005). Writing starts with own name writing: From scribbling to conventional spelling in Israeli and Dutch children. *Applied Psycholinguistics, 26*, 463–478.
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*, 891–905.
Levin, I., & Tolchinsky Landsmann, L. (1989). Becoming literate: Referential and phonetic strategies in early reading and writing. *International Journal of Behavioral Development, 12*, 369–384.
Lin, D., McBride-Chang, C., Aram, D., Levin, I., Cheung, R. Y. M., Chow, Y. Y. Y., et al (2009). Maternal mediation of writing in Chinese children. *Language and Cognitive Processes, 24*, 1286–1311.
Luk, G., & Bialystok, E. (2005). How iconic are Chinese characters? *Bilingualism: Language and Cognition, 8*, 79–83.
Luria, A. R. (1978). The development of writing in the child. In M. Cole (Ed.), *The selected writings of A.R. Luria* (pp. 145–194). White Plains, NY: M.E. Sharpe [Original work published 1929].
Lyons, D. E., Young, A. G., & Keil, F. C. (2007). The hidden structure of overimitation. *Proceedings of the National Academy of Sciences of the United States of America, 104*, 19751–19756.
Martens, P. (1996). *I already know how to read: A child's view of literacy*. Portsmouth, NH: Heinemann.
Menard, S. (2002). *Applied logistic regression analysis* (2nd ed.). Thousand Oaks, CA: Sage.
National Early Literacy Panel (2008). *Developing early literacy: Report of the National Early Literacy Panel*. Washington, DC: National Institute for Literacy.
Noyer, M., & Baldy, R. (2005). Influence de la complexité des formes imagées et langagières d'un référent et rôle du prénom dans l'acquisition de l'écriture entre trois et six ans [Influence of the complexity of imaged and linguistic forms of the referent and role of the first name in the acquisition of writing between 3 and 6 years]. *Archives de Psychologie, 71*, 199–215.
Pollo, T. C., Kessler, B., & Treiman, R. (2009). Statistical patterns in children's early writing. *Journal of Experimental Child Psychology, 104*, 410–426.
Puranik, C. S., & Lonigan, C. J. (2010). From scribbles to scrabble: Preschool children's developing knowledge of written language. *Reading and Writing,* doi:10.1007/s11145-009-9220-8.
Rowe, D. W. (2008). Social contracts for writing: Negotiating shared understandings about text in the preschool years. *Reading Research Quarterly, 43*, 66–95.
Schickedanz, J. A. (1990). *Adam's righting revolutions: One child's literacy development from infancy through Grade One*. Portsmouth, NH: Heinemann.
Shu, H., Chen, X., Anderson, R. C., Wu, N., & Xuan, Y. (2003). Properties of school Chinese: Implications for learning to read. *Child Development, 74*, 27–47.
Tolchinsky, L. (2003). *The cradle of culture and what children know about writing and numbers before being taught*. Mahwah, NJ: Lawrence Erlbaum.
Tolchinsky-Landsmann, L., & Levin, I. (1985). Writing in preschoolers: An age-related analysis. *Applied Psycholinguistics, 6*, 319–339.
Treiman, R., Kessler, B., & Bourassa, D. (2001). Children's own names influence their spelling. *Applied Psycholinguistics, 22*, 555–570.
Tunmer, W. E., Herriman, M. L., & Nesdale, A. R. (1988). Metalinguistic abilities and beginning reading. *Reading Research Quarterly, 23*, 134–158.
Yang, D., Sidman, J., & Bushnell, E. W. (2010). Beyond the information given: Infants' transfer of actions learned through imitation. *Journal of Experimental Child Psychology, 106*, 62–81.