

# GraphCom: A multidimensional measure of graphic complexity applied to 131 written languages

Li-Yun Chang<sup>1</sup> · Yen-Chi Chen<sup>2</sup> · Charles A. Perfetti<sup>3</sup>

Published online: 19 April 2017  
© Psychonomic Society, Inc. 2017

**Abstract** We report a new multidimensional measure of visual complexity (GraphCom) that captures variability in the complexity of graphs within and across writing systems. We applied the measure to 131 written languages, allowing comparisons of complexity and providing a basis for empirical testing of GraphCom. The measure includes four dimensions whose value in capturing the different visual properties of graphs had been demonstrated in prior reading research—(1) perimetric complexity, sensitive to the ratio of a written form to its surrounding white space (Pelli, Burns, Farell, & Moore-Page, 2006); (2) number of disconnected components, sensitive to discontinuity (Gibson, 1969); (3) number of connected points, sensitive to continuity (Lanthier, Risko, Stolz, & Besner, 2009); and (4) number of simple features, sensitive to the strokes that compose graphs (Wu, Zhou, & Shu, 1999). In our analysis of the complexity of 21,550 graphs, we (a) determined the complexity variation across writing systems along each dimension, (b) examined the relationships among complexity patterns within and across writing systems, and (c) compared the dimensions in their abilities to differentiate the graphs from different writing systems, in order to predict human perceptual judgments ( $n = 180$ ) of graphs with varying

complexity. The results from the computational and experimental comparisons showed that GraphCom provides a measure of graphic complexity that exceeds previous measures in its empirical validation. The measure can be universally applied across writing systems, providing a research tool for studies of reading and writing.

**Keywords** Connected points · Disconnected component · Gestalt principles · Graphic complexity · Perimetric complexity · Simple features · Writing systems

The world's writing systems contain graphs that span a wide variety of visual forms. Much of this variety is associated with variable mappings that graphic units can have to linguistic units (abjad, alphabetic, syllabary, alphasyllabary, and morphosyllabary). This mapping variety has been the focus of comparative reading research (e.g., universal grammar of reading, Perfetti, 2003; phonological grain size, Ziegler & Goswami, 2005; orthographic depth, Katz & Frost, 1992; semantic transparency, Wydell, 2012; for reviews, see Frost, 2012; Perfetti & Harris, 2013; Seidenberg, 2011). The actual forms of the graphic units have received less attention. However, the visual forms of graphs—reflecting their visual complexity and discriminability—have the potential to affect the identification of both individual graphs and graph combinations (e.g., single letters and letter combinations in alphabets and abjads, akshara in alphasyllabaries, syllables in syllabaries, and characters in morphosyllabaries; Pelli, Burns, Farell, & Moore-Page, 2006), and thus to affect learning to read.

To study the effects of graphic challenges to learning to read in a universal way, free of biases based on a particular writing system, it is important to have a measure of graphic complexity that is sensitive to the variety of devices used in

✉ Li-Yun Chang  
liyunchang@ntnu.edu.tw

✉ Charles A. Perfetti  
perfetti@pitt.edu

<sup>1</sup> Department of Applied Chinese Language and Culture, National Taiwan Normal University, Taipei, Taiwan

<sup>2</sup> Department of Statistics, University of Washington, Seattle 98195, WA, USA

<sup>3</sup> Learning Research and Development Center, University of Pittsburgh, 3939 O'Hara Street, Room 833, Pittsburgh, Pennsylvania 15260, USA

writing. We report here such a measure, a multidimensional measurement system for quantifying graphic complexity, GraphCom, and its application to 131 written languages. We demonstrate the value of the system in predicting similarity ratings made by speakers of different languages.

In what follows, we first discuss the broader context for descriptions of graphic units, delineating the units that are the object of our study and reviewing previously developed measures of graphic complexity (i.e., perimetric complexity: Pelli et al., 2006; Watson, 2012) and considering perceptual principles in human cognition. We then present the rationale and descriptions of our new multi-dimensional measure, the results of applying the measure to 131 languages, and the performance of the measure in predicting visual similarity judgments.

### Graphic units: Graphs and graphemes

It is common in alphabetic reading research to refer to a grapheme as the basic unit of writing—in particular, one or more letters that map onto a single phoneme. Such a definition lacks both universality (e.g., Chinese characters do not map to phonemes) and also departs from the logic of linguistic descriptions. A definition of grapheme that conforms to linguistic analysis by being parallel to descriptions of phoneme and morpheme is this: a grapheme is a functional unit of writing that abstracts over variations in graphs—allographs; for instance, all the fonts for the letter *b* that exist at in a given language. The unit is functional in that the grapheme is the minimal graphic unit distinguishing two *written* morphemes, thus analogous to the phoneme, which distinguishes two spoken morphemes. For example, in English all letters are graphemes as well as graphs, because all letters distinguish among *written* English morphemes. According to this definition, the functional role of graphemes does not depend on mapping to phonemes, as attested by the contrast between homophonic morphemes such as *buy/bye* and *reel/real*. This technical definition of grapheme also includes nonletter graphemes such as the apostrophe, which distinguishes *teacher's* from *teachers*. What counts as a grapheme is language-dependent even within a writing system. Thus, a capital letter and a lowercase letter seem to be allographs of a single grapheme in English, but probably not in German, where capitalization distinguishes between grammatically derived morphemes (*Wissen* as noun vs. *wissen* as verb). This sense of a grapheme as a distinguisher of written morphemes is more systematic and universal than the commonly used definition in the English language research literature. Thus, it applies to Chinese as well, where a character is a grapheme as well as a morpheme and distinguishes between multimorpheme words.

For purposes of measuring graphic complexity, our view is that the common psychological use of “grapheme,” which



originated in alphabetic research, is too narrow. However, the more universal linguistic definition requires a detailed morphological analysis of each written language, a goal that is beyond the scope of our research. These issues of the definition of grapheme have led us to focus instead on the minimal unit of the graph, a written form that can be combined with other graphs to form graphemes (in any sense of grapheme). These graphs are readily recognized by literate users of a language as basic writing units—several thousand characters in Chinese, 26 letters in English, 33 letters in Russian, and 46 kana in the Japanese syllabaries—that are combined to produce written language, whatever their mapping. Most important is that a metric based on writing graphs (rather than graphemes) can be applied to any written language according to the goals of a researcher. For an English example, the complexity measure applied to the letters S and H separately can also be applied to the combination SH if one wants to measure “grapheme” complexity.

### Writing graphs, as a culture product, are different from other visual categories

Every writing graph (henceforth, simply “graphs”) is a basic, two-dimensional visual form that participates alone or in combination in coding a linguistic unit (e.g., phoneme, syllable, or morpheme). In the information they convey, these graphic forms are different from other visual categories such as natural scenes, objects, and faces, but similar to line drawings (Changizi, Zhang, & Shimojo, 2006). Scenes carry more complex information about color, texture, shading, illumination, and occlusion (Sayim & Cavanagh, 2011). Objects, similar to scenes, provide more information about three-dimensional space, depth, and texture than line drawings. Faces, although they are composed with fewer elements than scenes and objects, are still more complex than line drawings because faces are usually seen from many viewpoints. In contrast, line drawings are simple. Because their complexity varies along fewer visual dimensions, indices that are useful for natural visual categories—for example, entropy on information, Fourier analysis on spatial-frequency, JPEG compression for size of an image (see Chikhman, Bondarko, Danilova, Goluzina, & Shelepin, 2012)—are not applicable.

Although graphs and line drawings share the general properties of two-dimensional simplicity, graphs become differentiated from line drawings with the earliest emergence of literacy contexts. Letters, the graphs used in alphabetic systems for examples, are differentiated from line drawings by children by the age of three (Levin & Bus, 2003; Robins & Treiman, 2009). For English-speaking young children, preschoolers 3–5 years of age who are preliterate have some understanding that a written word represents a specific spoken word, differing in this way from a drawing (Treiman,

**Table 1** Comparison of two graphs in terms of their visual complexity

	English Graph (Alphabet) 	Thai Graph (Alphasyllabary) 
Perimetric Complexity (PC)	13	13
Number of Disconnected Components (DC)	1	2
Number of Connected Points (CP)	3	2
Number of Simple Features (Strokes) (SF)	4	4

Hompluem, Gordon, Decker, & Markson, 2016). Moreover, these children are sensitive to the visual spatial layout of their own writing system, as compared to foreign writing systems (Treiman, Mulqueeny, & Kessler, 2014). When they are asked what writing is, these young children are more likely to choose sets of graphs from their own language (i.e., English) as instances of writing than graphs in other languages (e.g., Chinese characters; Lavine, 1977). These observations point to a categorical importance of graphs, as they becomes a functionally distinct perceptual objects in learning to read.

### Previous measures of graphic complexity






One well-defined and well-attested dimension for quantifying graphs is *perimetric complexity* (Pelli, Bums, Farell, & Moore-Page, 2006): the ratio of the square of the sum of the inside and outside perimeters to the product of  $4\pi$  and the area of the foreground (Pelli et al., 2006; Watson, 2012; see Tables 1 and 2 for examples, and the Method section for the algebraic expression). More informally, perimetric complexity captures the density of the written marks (“black ink”) relative to the background space in which they are located. Perimetric complexity has some valuable characteristics. First, it is objective, quantitative, and size-invariant. Thus, its values are not affected by font size. Second, it is empirically tested and correlates well with other subjective measures, such as pattern

goodness and information load (for a discussion, see Jiang, Shim, & Makovski, 2008). Third, it is computerized, and the algorithm can be used for binary-code (black-and-white) images (Watson, 2012), making it a tool that is general across visual categories.

Pelli et al. (2006) applied perimetric complexity to a range of graphs and demonstrated that perimetric complexity is inversely proportional to graph identification efficiency. Specifically, they sampled graphs across a wide range of written languages (i.e., Arabic, Armenian, Chinese, Devanagari, English, and Hebrew) and different fonts (e.g., Bookman, Courier, Helvetica, Kustler, and Sloan). They asked participants (ranging from 3 to 68 years of age) to look at a briefly displayed graph and then to identify it from a list of graphs in the given language. Graphic complexity was negatively correlated with human identification efficiency. Given the reliability and validity of perimetric complexity, it became a useful measure for controlling the complexity of stimuli in studies on learning to read (e.g., Liu, Chen, & Wang, 2016; Wang, McBride-Chang, & Chan, 2014; Yin & McBride, 2015).

Research on the relation between visual complexity and learning to read across writing systems suggests that learning to read more visually complex first language (L1) may require higher visual skills and may, in turn, strengthen such skills. In particular, a series of studies by Nag and colleagues provided evidence that visual skills required for reading Indian languages tend to be relatively high as compared with alphabetic

**Table 2** Five graphs with complexity values using GraphCom, the measurement system with four dimensions

Writing system	Abjad	Alphabet	Syllabary	Alphasyllabary	Morphosyllabary
Written language	Hebrew	Russian	Cree	Telugu	Chinese
Example graph					
PC	5.16	7.83	12.04	18.06	20.85
DC	1	1	3	3	1
CP	1	1	3	2	14
SF	2	2	6	5	9

PC = perimetric complexity, DC = number of disconnected components, CP = number of connected points, SF = number of simple features

languages (e.g., Nag, 2008; Nag & Snowling, 2011; Nag, Snowling, Quinlan, & Hulme, 2014; Nag, Treiman, & Snowling, 2010). These high demands come from the large number of graphs in these “extensive” written languages (Nag, 2007, 2014) and impose a strong influence on the pace of learning to read (Nag, Caravolas, & Snowling, 2011). It is possible that meeting the higher learning demands imposed by visually complex writing systems leads to improved visual skills: in a cross-writing-system study (McBride-Chang et al., 2011), children learning to read traditional Chinese outperformed age-matched kindergarteners who were learning to read less complex languages (Hebrew and Spanish) in a visual–spatial processing task. Similarly, in a comparison of 8- to 14-year-old readers of Chinese and Greek, controlling for reading experience, Chinese readers of all ages outperformed their age-matched Greek counterparts on visual-spatial processing (Demetriou et al., 2005). Collectively, these findings underscore the importance of visual complexity of graphs regarding their roles in impacting learning to read across writing systems.

### Complexity characteristics in different writing systems

Not all characteristics of graphs found to be important in reading research are captured by perimetric complexity. There are many examples of two graphs that share the same value in perimetric complexity, while differing substantially in other ways. For instance, in Table 1, perimetric complexity quantifies both the graph <w> (an English letter) and the graph <๐> (a Thai letter) as 13; however, the two graphs have salient visual differences in their numbers of disconnected components (i.e., <w> has one component, and <๐> has two components), and thus also in their numbers of connected points; <w> has three points, each composed by two lines, and <๐> has two connected points, each composed by one circle and one line.

Variation in disconnected components is typical for alphasyllabaries, and variation in numbers of connected points is typical for alphabets. The numbers of connected points in letters of the Roman alphabet used in English (e.g., line terminations in <R>) are the features most critical in letter identification (Fiset et al., 2008). In alphasyllabaries, letters featuring disjointed components (e.g., the Thai letter <๐>) are highly associated with visual confusion in early literacy (Winkel, 2010). However, it is unclear whether the number of connected points, an important factor in the recognition of alphabetic writing, also affects letter identification in an alphasyllabary; similarly, we do not know whether the number of disconnected components, a salient measure in an alphasyllabary, plays a role in early alphabetic literacy.

In the morphosyllabic system for Chinese languages, the number of strokes (usually defined as a one-time movement of pen) has been long used as a complexity index with demonstrated psychological reality. For instance, Su and Samuels (2010) report that, in a Chinese character recognition task, response latencies to characters increased with the number of strokes for Chinese-speaking second-graders. In a study of Japanese kanji, Tamaoka and Kiyama (2013) found that both lexical decision times and naming depended on the number of strokes as well as on kanji frequency. Chinese character reading studies have also examined (and experimentally controlled) the number of strokes in both simplified (Wu, Zhou, & Shu, 1999) and traditional (Y. P. Chen, Allport, & Marshall, 1996) Chinese. Although all writing varies in the number of strokes, this measure remains unapplied to any writing system other than Chinese.

To examine complexity characteristics of graphs in different written languages and to examine and compare reading and writing across writing systems, a general, multidimensional measure that can apply to all writing systems is needed.

### Gestalt principles for perceptual organization of graphs

Some of the features highlighted in research (i.e., numbers of disconnected points, connected components, and strokes, respectively) seem to echo principles of the perceptual organization of relations among visual components (proximity, symmetry, convexity, closure, connectedness, and continuation) that were emphasized in Gestalt theory (Koffka, 1935/1963). These principles were proposed as a partial answer to the question of how individual elements group into parts that then group into the larger perceptual object that is separated from other perceptual objects (Ehrenstein, 2008; Spillmann & Ehrenstein, 2004). For example, continuation affords clues to the relationship between simple features (Biederman, 1987), and connectedness is sensitive to information regarding continuity (Lanther, Risko, Stolz, & Besner, 2009). In contrast, discontinuity highlights relations between more complex features.

An emphasis on continuity and discontinuity echoes the criteria for making a well-designed written language suggested by Watt (1983, 1994; see also Treiman & Kessler, 2011). Watt argued that shapes in such written language should be (1) *similar*, or have a degree of homogeneity; (2) *contrasting*, or distinguishable from one to another; (3) *economical*, or easy to perceive and produce; (4) *redundant*; (5) *attractive*; and (6) *expressive*. The systematicity of graph shapes was also emphasized by Treiman and Kessler (2014), who observed that, across writing systems, there is a tendency for graphs to look similar. This similarity may reflect basic principles of learning, one of which is that learners abstract

patterns that hold across a set of graphs and use these patterns to supplement their memory for individual graphs.

Consideration of the different ways in which graphs vary across writing systems led us to develop a new measure that uses these different complexity-related variations, while also building on perimetric complexity. This measure, GraphCom, consists of four dimensions: perimetric complexity, number of disconnected components, number of connected points, and number of simple features (strokes). We applied this visual complexity measure to a large number of written languages, representing all five of the major writing systems. To the best of our knowledge, this is the first attempt to apply multidimensional complexity measures to quantify a larger number of graphs, in order to provide a valid tool to study the visual forms of those graphs.

### The graph complexity measure, GraphCom

GraphCom includes four dimensions of graph measurement. The three dimensions added to perimetric complexity are quantified by the following basic units: A *simple feature*, following Pelli et al.'s (2006) definition, is a discrete element of an image that can be discriminated independently from other features. For example, <T> has two simple features, a vertical segment and a horizontal segment. A *connected point* (or a junction) is an adjoining of at least two features. For example, <T> has one (the junction of the horizontal line and the vertical line) and <F> has two connected points (the junctions of one vertical line with two horizontal lines). A *disconnected component* is a simple feature that is not linked to other features in a set. For example, disconnected components are shown, respectively, in <i> (the dot and the vertical line) and <云> (the horizontal line on the top and the integral component at the bottom). Given these basic definitions, we can describe our four dimensions:

**Perimetric complexity (PC)** PC is the ratio of the squared perimeter of a graph (number of pixels) to the number of background pixels in the graph. Specifically, PC is  $\frac{P^2}{A4\pi}$ , the square of the sum of the inside and outside perimeters of the foreground (P), divided by the foreground area (A), divided by  $4\pi$  (Pelli et al., 2006; Watson, 2012). For example, if upper-case <W> has a 4,656-pixel perimeter and 136,602-square-pixel area, its perimetric complexity is 12.6287 ( $= 4,656 \times 4,656 / 136,602 / 4\pi$ ). This dimension is sensitive to the changes in luminance across space (i.e., spatial frequency) of a graph and its value is invariant to the size of the graph (Grainger, Rey, & Dufau, 2008).

**Number of disconnected components (DC)** DC is defined as a simple feature or a feature that is not linked to other features

in a set. If a given graph is composed of multiple disconnected components, there are *spaces* among these components; for instance, <◊> has four disconnected components created by spaces among the circle and the three dots. This dimension is sensitive to discontinuity information (Gibson, 1969).

**Number of connected points (CP)** CP is a point of contact between features. This dimension is sensitive to information regarding continuity (Lanthier et al., 2009) and provides clues to relations between simple features (Biederman, 1987), counter to the DC dimension. Note that CP is not simply the inverse of DC; for instance, Vai syllables <|> and <◊> have the same number of disconnected components (three), but the number of connected points of <◊> is four (for the diamond) and the number of connected points of <|> is zero.

**Number of simple features (SF)** SF is a discrete element that can be discriminated from others (Pelli et al., 2006); a typical example is a stroke within a Chinese character (Wu, Zhou, & Shu, 1999). Other examples for one simple feature include a line, a dot, a circle, or a curved line. To make the measure size-invariant, length, width, and thickness are not considered properties of features. This dimension is sensitive to the extent to which the graph combines simple features.

Collectively, these four dimensions provide objective, quantitative, and size-invariant estimations of graphic complexity. Table 2 shows how these four dimensions of GraphCom capture different characteristics of five example graphs.

### Method

#### The written languages

For the application of the GraphCom to actual writing, we selected 131 written languages to represent five writing systems<sup>1</sup> (alphabet, 60; abjad, 16; alphasyllabary, 41; syllabary, 11; morphosyllabary, 3), using languages examined in previous cross-writing-system (Changizi & Shimojo, 2005), cross-alphabet (Seymour, Aro, & Erskine, 2003), and cross-script

<sup>1</sup> Writing systems are broad families of written languages, delineated by the linguistic units represented by their graphemes (e.g., abjads: consonants; alphabet: phonemes; alphasyllabaries: consonant–vowel units; syllabaries: syllables; morphosyllabaries: syllable morphemes; Cook & Bassetti, 2005). Among writing scholars, there is some variation in the names for these systems. For example, what we refer to here as *alphasyllabaries*, following Bright (1992), have been named *abugida* by Daniels (1990) to reflect their distinctiveness and to counter the impression that they are a mix of alphabetic and syllabic writing. This five-way classification system is a more accurate reflection of differences between written languages than is the traditional three-way system (e.g., Gelb, 1952). In particular, the three-way system included both alphasyllabaries and consonantal abjads as alphabets. The differences between these systems and alphabets are sufficient for us to adopt the five-way classification. For a penetrating discussion of these issues, see Share and Daniels (2016); for a contrary view of alphasyllabaries as alphabets, see Rimzhim, Katz, and Fowler (2014).

(traditional vs, simplified Chinese; H. C. Chen, Chang, Chiou, Sung, & Chang, 2011) studies. To identify the inventory of graphs and writing system categories for these languages, we followed Changizi and Shimojo (2005), who used *Ager's Omniglot: A Guide to Writing Systems* (Ager, 1998). For the three languages on which Omniglot offers no information, we consulted other sources: H. C. Chen et al. (2011) for the two major scripts of Chinese (i.e., traditional and simplified Chinese), and an official list of 1,006 Japanese kanji by school year (Ministry of Education in Japan, 2015). Finally, for purposes of the complexity measure, we used only the forms of isolated graphs. For most written languages (ignoring handwriting), this is of no consequence. However, in some, especially the akshara of alphasyllabaries, graphs can change shape when they are combined in actual writing: Vowel graphs are reduced to diacritics when conjoined with consonants. These variations, which are important in actual writing, are not captured in our analyses, which defines the graphs of every language in their canonical forms.

### Graphic complexity quantification

We generated images of each of 21,550 graphs using the Processing software ([www.processing.org](http://www.processing.org); Reas & Fry, 2010). Graphs were presented in black Arial font against a  $500 \times 500$  pixel white background. In all, 25% of the selected languages are not supported by the Arial font; for these, an alternative font similar to Arial was adopted. Appendix Table 10 summarizes the detailed information about these 131 written languages. Measures of the four dimensions of the GraphCom were then applied to each of these 21,550 images.

## Results

### Complexity variation along individual dimensions

We describe the complexity of a graph as a set of values along the four dimensions of GraphCom. Figure 1 shows the complexity variations across writing systems as boxplots for each of the four dimensions: perimetric complexity (PC), number of disconnected components (DC), number of connected points (CP), and number of simple features (SF).

To assess the relationships among these dimensions, we correlated the complexity values on each of the four dimensions across the five writing systems, as well as separately for each writing system. Table 3 summarizes the results for the overall correlations, collapsed across writing systems: All correlations are greater than .82 (all  $ps < .001$ ), except for the  $r = .65$  correlation of DC (the number of disconnected components) with CP (the number of connected points). Perimetric complexity shows high correlations with the other dimensions, although a

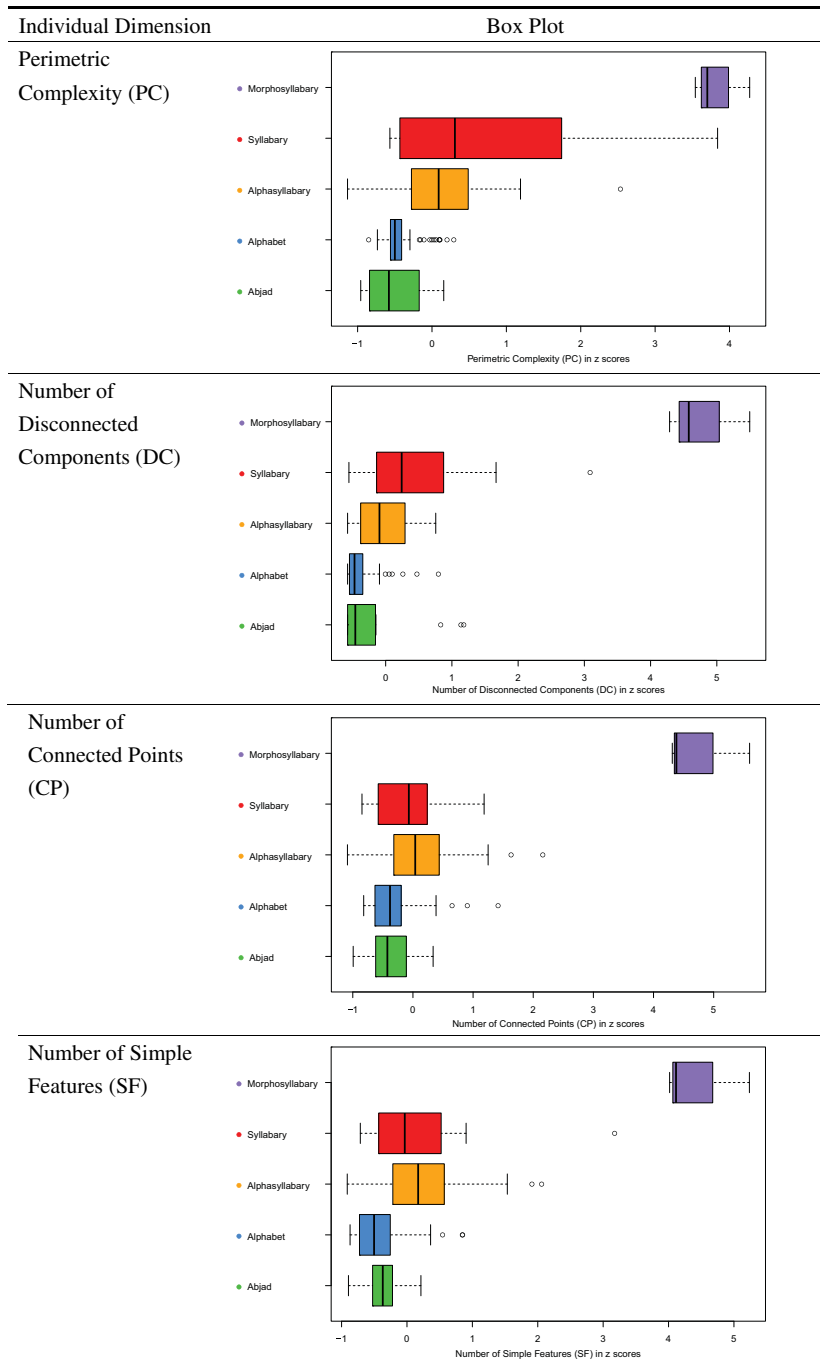
lower correlation with DC, reflecting PC's ability to capture indirectly much of what the other dimensions target specifically. However, the measure with the greatest shared variance is the number of simple features, the building blocks of the graphs. Finally, the correlations show that the number of discontinuous components (DC) is the most distinctive measure, sharing no more than 67% of variance with other measures, and only 42% with the number of connected points. Significantly, not all writing systems showed the same pattern of correlations among the dimensions. These specific writing system differences are discussed in Chang (2015).

### Dimensions differentiate writing system pairs

Next, we determined which dimension is best at differentiating among parent writing systems. If different dimensions play a role in such differentiation, this would support the value of the multidimensional approach. For this analysis, we used the nonparametric Kolmogorov–Smirnov (KS) distance<sup>2</sup> (Stephens, 1974), one of the most commonly used distance measures for comparing two samples. In our case, the two samples correspond to two writing systems. The KS distance, which does not assume a normal distribution, is sensitive to the difference in the cumulative distribution functions of two samples, and thus is suitable for the highly nonnormal distributions of our writing systems on the various dimensions. Our five writing systems yielded ten writing system pairs. For each pair, we calculated the KS distances on each dimension (see Table 4); the dimension responsible for the greatest KS distance was taken as the dimension that is most sensitive to differences between those two writing systems.

Table 5 shows the complexity dimension that maximally differentiates each pair of writing systems. Thus, the alphabet and abjad writing systems are most differentiated by their numbers of disconnected components; alphasyllabaries and morphosyllabaries are most differentiated by their numbers of simple features; and so forth. The number of connected points was not a maximal differentiator for any pair of systems. Interestingly, perimetric complexity, which has been the only dimension used to compare graphic complexity across writing systems in prior research (Pelli et al., 2006), was the most

<sup>2</sup> The KS distance between two samples is the maximal difference between their (empirical) cumulative distribution functions. Given one sample, the cumulative distribution function is the function  $F(x)$  whose value is “the ratio of data points below  $x$ .” For instance, when  $x$  is the median, then  $F(x) = .5$ ; when  $x$  is the first quartile,  $F(x) = .25$ . Given two samples with cumulative distribution functions  $F1(x)$  and  $F2(x)$ , the absolute difference,  $|F1(x) - F2(x)|$ , is a function of  $x$ . The KS distance is defined as the maximal value of  $|F1(x) - F2(x)|$  for all possible values of  $x$ .



**Fig. 1** Boxplots comparing graphic complexity across writing systems, for each dimension. The boxplots indicate, for each writing system

(coded by color), the range of written languages (in terms of Quartile 1, Quartile 2, Quartile 3, and outliers)

reliable differentiator only for the alphasyllabary–alphabet pair. These results suggest that the most effective dimension for differentiating writing system pairs is the number of disconnected components (DC); in Table 5, DC provides maximal differentiation for six of the ten writing system pairs. More generally, the results highlight the value of the multidimensional approach. No single dimension is universally the most effective at distinguishing any two arbitrarily selected writing systems.

### Behavioral validation: Similarity ratings of graph pairs

To provide a behavioral test of GraphCom and its individual dimensions, we had participants with different first language (L1) backgrounds make similarity ratings on pairs of graphs from a single written language. We chose similarity ratings because they represent a paradigm commonly used in visual science and psychology over the past 130 years (for a review,

**Table 3** Correlations of graphic complexity across writing systems

	Perimetric Complexity (PC)	Number of Disconnected Components (DC)	Number of Connected Points (CP)	Number of Simple Features (SF)
PC	1.00			
DC	.82***	1.00		
CP	.89***	.65***	1.00	
SF	.95***	.83***	.93***	1.00

\*\*\*  $p < .001$

see Mueller & Weidemann, 2012). The assumption is that two graphs that are more similar in complexity will be judged more similar than two graphs that are less similar in complexity.

## Method

### Stimuli

To select a language that is representative of its writing system, we identified a centroid for each writing system within a multidimensional complexity space<sup>3</sup> defined by the four dimensions of GraphCom. A centroid is the geometric center of a multidimensional space; in our case, the centroid of a writing system is the location of the unweighted mean of all the written languages within this writing system—that is, the average of their coordinates along the four dimensions. Thus, for each writing system, one language was designated as its *centroid written language*: Hebrew (abjad), Russian (alphabet), Telugu (alphasyllabary), Cree (syllabary), and Chinese (morphosyllabary). The stimuli included graphs from these five centroid written languages.

We created two categories for Chinese, because it contains thousands of characters and more than one type of graphic unit. Basic components (including radicals), which are the functional “building blocks” in the Chinese language (Shen & Ke, 2007), can stand alone as characters and are composed of a small number of strokes (average: 4.52). Compound characters, which are composed from these building blocks, have a large number of strokes (average: 13.21). Thus, we classified all characters as either basic or compound; note that these characters have the same forms in the traditional and simplified Chinese systems.

The division of Chinese into basic and compound types resulted in six groups of graphs based on the centroid written languages. These are ordered, in terms of increasing

complexity, Hebrew, Russian, Cree, Telugu, basic Chinese characters, and compound Chinese characters. For the similarity ratings, graphs were paired within each written language, with graphs in each pair matched on either upper or lower case and, where applicable, vowel or consonant; all graphs in each written language (except for Chinese) were exhaustively used.

We created four stimulus lists, each consisting of six groups of graphs. Each list contained 180 pairs, more than the 120 pairs that are adequate to induce meaningful similarity judgments (Simpson, Mousikou, Montoya, & Defior, 2013). Appendix Table 11 shows the graph pairs for each list; Table 6 provides further information regarding these pairs of graphs.

### Observers

A total of 180 observers participated in this experiment. All reported normal or corrected-to-normal vision. Table 7 presents demographic information about these observers. We chose observers whose first language was among those five languages with the most speakers worldwide (Arabic, English, and Hindi) and for whom the graphs to be judged were not from their first language.

### Procedure

The experiment was carried out via a large crowdsourcing platform, Amazon Mechanical Turk (MTurk). MTurk data have been demonstrated to be indistinguishable from laboratory data in different research fields (e.g., economics: Horton, Rand, & Zeckhauser, 2011; politics: Berinsky, Huber, & Len, 2012; social science: Buhrmester, Kwang, & Gosling, 2011; psycholinguistics: Sprouse, 2011; and psychology: Simcox & Fiez, 2014); to ensure data quality, we also followed the principles for using MTurk (Chandler, Mueller, & Paolacci, 2014) to design our online experiment. Four human intelligence tasks (HITs) for recruiting observers from four writing systems were posted on MTurk’s online recruitment interface. Each HIT had a two hour completion limit. Consent was obtained prior to the experiment; after MTurk volunteers agreed to participate, they were directed via a Web link to any of the four stimuli lists for similarity ratings.

The sequence of tasks was the same for each observer: a similarity rating, a language history questionnaire, a demographic background task, and a translation task (except for the English HIT) for verifying the observer’s L1 backgrounds. After completing the last task, a unique 13-digit code associated with the observer’s responses appeared on the screen automatically, along with debriefing information. The observer was instructed to report the code to MTurk to obtain monetary compensation. Successful generation of the 13-digit code also indicated that all of the observer’s responses were successfully sent from their local machines to our server. Below, we give a brief introduction for each task.

<sup>3</sup> In constructing this multidimensional space, we calculated, for each written language, the mean score on each dimension. Because the scale varies among dimensions (min–max range and interquartile range [Q3 – Q1] for each dimension: perimetric complexity: 1–76, Q3 – Q1 = 22; number of disconnected components: 1–18, Q3 – Q1 = 4; number of connected points: 0–37, Q3 – Q1 = 8; number of simple features: 1–32, Q3 – Q1 = 8), we transformed the resulting means to within-dimension *ss* scores. This standardization procedure addressed the issue of scale difference, allowing for comparisons of complexity on the same scale.



**Table 4** KS distances between two given writing systems for each dimension

	Perimetric Complexity (PC)						Number of Disconnected Components (DC)				
	AL	AS	AJ	SY	MS		AL	AS	AJ	SY	MS
AL	–					AL	–				
AS	.20	–				AS	.19	–			
AJ	.10	.14	–			AJ	.27	.32	–		
SY	.23	.06	.18	–		SY	.38	.27	.32	–	
MS	.18	.08	.13	.14	–	MS	.28	.24	.25	.19	–
	Number of Connected Points (CP)						Number of Simple Features (SF)				
	AL	AS	AJ	SY	MS		AL	AS	AJ	SY	MS
AL	–					AL	–				
AS	.06	–				AS	.16	–			
AJ	.07	.06	–			AJ	.09	.11	–		
SY	.06	.09	.08	–		SY	.12	.09	.12	–	
MS	.05	.06	.07	.07	–	MS	.24	.34	.27	.31	–

AL = Alphabet; AS = Alphasyllabary; AJ = Abjad; SY = Syllabary; MS = Morphosyllabary

**Similarity rating task** This task was designed to tap variability in observers' judgments of visual similarity. Each trial began with a black fixation cross appearing for 300 ms, followed by a pair of graphs appearing for up to 5,000 ms, followed by a blank for 1,000 ms. The observer saw a pair of graphs appear at the center of the screen and the heading "1 = very different 2 = mainly different 3 = mainly similar 4 = very similar" at the bottom of the screen. The observers were instructed to rate how visually similar the two graphs were by pressing one of four keys on their alphanumeric keypad (not the numeric keypad) to indicate the rated similarity. Once the observers had responded, the screen moved on to the next trial.

After instructions, the observers were given 12 demonstration trials with explicit statements on the degree of similarity, 12 practice trials without feedback, and 180 experimental trials; the ordering of grapheme pairs was randomized. Table 8 shows trials in the demonstration, practice, and experimental phases. Responses and response time were recorded. This task took approximately 15 minutes to complete.

**Language questionnaire** The language questionnaire (Tokowicz, Michael, & Kroll, 2004) was used to assess participants' language-learning experiences both quantitatively (e.g., rating general language learning skill) and qualitatively (e.g., comments about language-learning experience). Observers were encouraged to give their best answers to the questions, without any time limit.

**Demographic background questionnaire** The demographic background questionnaire was developed to learn more about observers' educational, cultural, and health statuses (e.g., visual and hearing problems) and their surroundings during participation in this study. The responses on visual and hearing

questions were used to filter data quality. We imposed no time limit to complete this survey.

**Translation task** The translation task was developed to filter the data for quality. This task consisted of 20 English words chosen from the instructions for this experiment. Observers saw one word at a time and were asked to type the first translation to their L1 that came to mind within 12 seconds; timing was determined in a pilot study. Observers who failed to provide translations in a written language consistent with their reported L1 were excluded from the analysis.

## Results

*Each dimension played a prominent role in predicting human similarity ratings.*

To test the effects of complexity on the perceptual judgments, we used a mixed-effects modeling approach, which is well-suited to assess the effects of both items (graphs) and subjects (observers) (Baayen, Davidson, & Bates, 2008). We assumed that two graphs that were more similar in complexity, as measured in GaphCom, would be judged more visually similar than two graphs that were less similar in complexity. Accordingly, we expected that a model that used all four dimensions of graphic complexity would provide the best fit to the human similarity ratings. We thus tested alternative models by means of a backward elimination procedure, which ensured that any joint predictive capability<sup>4</sup> of the dimensions could be observed (Burnham & Anderson, 2003). We first tested the full

<sup>4</sup> Backward elimination (or a back-out procedure) has an advantage over forward and stepwise selections. Forward selection and stepwise methods can fail to identify predictive models based on joint contributions of variables, because if the variables don't predict well individually, they will never enter the model. Because the backward method starts with everything in the model, joint predictive capability will be seen.

**Table 5** Dimensions that maximally differentiate writing system pairs

	Alphabet	Alphasyllabary	Abjad	Syllabary	Morphosyllabary
Alphabet	–				
Alphasyllabary	PC	–			
Abjad	DC	DC	–		
Syllabary	DC	DC	DC	–	
Morphosyllabary	DC	SF	SF	SF	–

PC = perimetric complexity, DC = number of disconnected components, SF = number of simple features

model containing all the predictors (the four complexity dimensions); then we constructed a second model that removed one of the predictors, to test whether removing one predictor would reduce the predictive performance. If so, that was evidence that the predictor should remain in the model.

Our predictors were the absolute differences in similarity ratings between two graphs in each of the complexity dimensions (i.e., PC, DC, CP, and SP). We performed a series of model comparisons with Laplace estimation using the *lmer()* function of the *lme4* package (Bates, Maechler, & Dai, 2010) to fit the models, and the likelihood ratio test (Lehmann, 1986) to determine model performance. The *lme4* model formulae used to fit each model are displayed in Appendix A.

The full mixed-effects model (FULL) included fixed effects of the four predictors and crossed random effects for subjects and items. The additional four models had fixed effects for three predictors (one predictor removed for each model) and crossed random effects for subjects and items. Thus, four model comparisons were carried out. Table 9 summarizes the model tests in terms of the Akaike information criterion (AIC) and Bayesian information criterion (BIC), two common criteria for model selection, as well as the chi-square values (and associated degrees of freedom) for the likelihood ratio test. A lower AIC/BIC indicates a better-fitting model (Wasserman, 2006). As is shown in Table 9, both AIC and BIC suggested that the FULL model scored best on these criteria. Similarly, for all likelihood ratio tests, the FULL model showed significant advantages over any reduced model ( $p$  values below .001): [FULL without PC vs. FULL,  $\chi^2(8) = 372.88$ ; FULL without DC vs. FULL,  $\chi^2(8) = 138.44$ ; FULL without CP vs. FULL,  $\chi^2(8) = 390.08$ ; FULL without SF vs. FULL,  $\chi^2(8) = 558.83$ ]. These tests indicate that removing any one of the predictor dimensions made the model significantly worse in accounting for variance in the data. This

suggests that each dimension played a role in accounting for observers' judgments of similarity.

## Discussion

The multidimensional measure of graphic complexity, GraphCom, is a useful tool for assessing visual complexity in any writing system. Its dimensions are grounded in basic perceptual factors—the number of simple visual features (lines, curves, and dots), the number of connected points, and discontinuities in the configural form. These dimensions are added to perimetric complexity, a proven measure that captures overall configurational complexity (Pelli et al., 2006). We applied GraphCom to 131 written languages across the world's five major writing systems, demonstrating that this measurement system surpassed previous measures in predicting human perceptual judgments. Importantly for research, GraphCom can be applied to any of the many other written languages beyond our sample of 131.

The value of GraphCom is supported by several results. First, it resulted in an ordering of complexity among the 131 languages that aligns with informal observations of these languages. Thus, Chinese written in its traditional script is measured as the most complex written language, more than the simplified Chinese script. At the other end of the scale, abjads and alphabets show similar low levels of complexity, and are distinguished primarily by their number of discontinuous elements. Of course, these alignments are to be expected to some extent, because we developed GraphCom measures to reflect properties of real writing. Thus, the ordering of the written languages is not a validation, but a demonstration that the measure produces sensible outcomes.

**Table 6** Characteristics and numbers of graph pairs for each written language in similarity rating (per list; four lists in total)

Writing system	Abjad	Alphabetic	Syllabary	Alphasyllabary	Morphosyllabary
Written language	Hebrew	Russian	Cree	Telugu	Chinese
Number of graphs in the written language	32	Upper: 33 Lower: 33	80	Vowels:35 Consonant:35	Basic: 242 Compound: 5,600+
Number of pairs per list	32	33	40	35	20 20

**Table 7** Demographic information for the participating observers ( $n = 180$  in total)

First Language (L1)	Estimated L1 Speakers	Sample Size	Gender	Mean Age ( $SD$ )	Mean Number of Learned Language ( $SD$ )
Arabic	295 million	60	29 males	27.42 (4.86)	2.53 (0.93)
English	339 million	60	22 males	26.37 (5.64)	1.67 (0.86)
Hindi	260 million	60	42 males	27.75 (5.67)	2.53 (0.72)

Number of learned languages include first language;  $SD$  = standard deviation; Estimates of L1 speakers were retrieved from Wikipedia: <https://goo.gl/ILTtaOu>

More interesting are the results concerning the individual dimensions. Perimetric complexity, the most commonly used measure for capturing configurational complexity of graphs in prior research, may not be suitable in some situations. When we applied each dimension to pairs of writing systems, using the nonparametric KS distance measure, we found that perimetric complexity was not the best differentiator among writing systems. It was the most successful differentiator only for separating alphabetic from alphasyllabary languages. The number of disconnected components was generally the most important distinguisher of writing systems.

Also relevant are the results of a modeling study that simulated graph learning across hundreds of languages (Chang, Plaut, & Perfetti, 2016). In the learning model, each dimension of the GraphCom was found to uniquely account for the training times the model needed to reach mastery. Indeed, perimetric complexity was the weakest predictor in the graph learning simulation; the number of simple features was the strongest one.

The most direct validation of the measure comes from its prediction of human perceptual similarity judgments. In fitting the perceptual judgment data to regression models, we found that all dimensions contributed to explaining the data. Removing any one dimension score from the model significantly reduced the model’s ability to predict visual similarity judgments.

We emphasize that these dimensions are not independent, and indeed they are highly inter-correlated when the data are collapsed across writing systems to allow a correlation based on 21,550 graphs. Writing systems differ in how they use the visual, graphic characteristics that are measured by GraphCom dimensions (Chang, 2015). In alphabets, connected points (or line junctions such as <L>, <T>, and <Y>) are especially important in letter identification (Lanthier et al., 2009; Szwed, Cohen, Qiao, & Dehaene, 2009). This importance reflects the relatively small number of graphs needed in most alphabetic languages. This allows the re-use of a small set of simple features that can be combined at junctions to

**Table 8** Examples for graph pairs at different phases in the similarity rating task

Trial	Suppose to be “Very different”	Suppose to be “Very similar”
Demonstration phase		
Practice phase		
Experimental phase		

**Table 9** A summary for four model comparisons, using dimensions to predict human similarity ratings ( $n = 180$ )

The FULL model	AIC	BIC			
Model with PC, DC, CP, SF	71,452	71,606			
Model comparison	AIC	BIC	$\chi^2$	df	p
FULL without PC vs. FULL	71,809	71,898	372.88	8	<.001
FULL without DC vs. FULL	71,575	71,664	138.44	8	<.001
FULL without CP vs. FULL	71,826	71,916	390.07	8	<.001
FULL without SF vs. FULL	71,995	72,084	558.83	8	<.001

form unique graphs. In contrast, morphosyllabic writing (Chinese and Japanese kanji) requires a very large number of graphs to code syllable morphemes. As the number of graphs increases, recombining features through connected points becomes impossible; instead additional graphs must add more simple features that also create more connected points and discontinuous components. Overall, graphic complexity is largely driven by the number of graphs that is needed in a written language. Collapsed over all 131 orthographies in our study, the number of graphs is highly correlated with the GraphCom measure of written language complexity,  $r = .78$  ( $p < .001$ ). This correlation is governed by how the written language manages the mapping of graphs to linguistic units in spoken language, because the writing system largely determines the number of graphs required (for a discussion, see Perfetti & Harris, 2013; Perfetti & Verhoeven, *in press*).

Although the neural basis of visual perception is beyond the scope of our study, it seems relevant to consider the relation between the properties of the graphs developed for written language and the properties of human vision. Hubel and Wiesel (1962, 1965) established that the receptive field of the cats' visual system included line, curvature, and edge detectors and computations that estimate their numbers. Primate visual systems have layered receptive fields that selectively respond to specific dimensions—for instance, V1 neurons to orientations; V2 neurons to corners; or V4 neurons to linear gratings, colors, angles, and curves—and computational abilities that operate across these layers (Van Essen, Anderson, & Felleman, 1992; for more recent work, see Coen-Cagli & Schwartz, 2013; Grill-Spector & Malach, 2004; Troncoso, Macknik, & Martinez-Conde, 2011).

Details aside, it is reasonable to suggest that the development of writing graphs has become aligned with human vision capabilities within other constraints, especially time and effort in graph production (Changizi & Shimojo, 2005). The three added dimensions (beyond perimetric complexity) of GraphCom seem to align with basic detection functions (simple features) and computational capabilities (connected points, discontinuous components) of human vision. Perimetric complexity seems to indirectly capture most of these detection and computational capabilities. Contributing substantially to

perimetric complexity's measures of inside and outside perimeters are graphemes' simple features and their junctions. Indeed, the number of simple features and the number of connected points together account for over 88% of the variance in perimetric complexity.

Finally, we note the practical value of GraphCom as a research tool. Researchers can access the dimension-specific complexity values of the 21,550 graphs from 131 written languages in the graphic complexity database, available at [https://dl.dropboxusercontent.com/u/28768192/GraphemeAll/GraphDataset\\_131\\_languages.zip](https://dl.dropboxusercontent.com/u/28768192/GraphemeAll/GraphDataset_131_languages.zip). The database can be used in various applications, depending on the research goals. For example, within a single language, graphic complexity measures can be applied to the graphs a child encounters in reading instruction; across languages, graphic complexity in one language can be compared with those of another. For some research aims, specific complexity dimensions can be applied to within-language and between-languages comparisons; for other aims, researchers can create composite scores at the level of individual graphs, the language using them, or the writing system to which the language belongs. More generally, data at the graph, grapheme, written language, or writing system level can be useful for a wide range of applications from comparative writing studies to learning to read to models of graph processing; in short, for studies that take account of visual factors in written language.

## Summary and conclusion

We introduced GraphCom, a multidimensional measurement system for quantifying the visual complexity of graphs across the world's writing systems. Starting with perimetric complexity, a well-validated single measure of complexity, GraphCom adds three dimensions that reflect the ways that graph forms differ in their composition over simple features, their connection points, and their discontinuities. These four dimensions were validated by their abilities to predict human perceptual judgments on graphs that varied in complexity as measured by GraphCom. As a tool for research, the GraphCom measures are available online for 131 written languages and 21,550 graphs. In addition, its measures are defined precisely, in order to allow application to any of the world's writing systems. This provides a practical research tool for constructing studies of perception and orthographic learning by children and adults, and also cross-language studies of reading and writing.

**Author note** This work was supported by the National Science Foundation (Grant #SBE-0836012) through Pittsburgh Science of Learning Center (PSLC) and "Aim for the Top University Project" of the National Taiwan Normal University and the Ministry of Education, Taiwan, R.O.C. The authors thank Adrian Maries and members in the Perfetti Lab at the University of Pittsburgh for their assistance with various tasks, and all observers for their participation. Moreover, the authors acknowledge the insightful comments of David Share and other, anonymous reviewers.

## Appendix

**Table 10** Detailed information for 131 written languages

Languages by Writing System	Font for Creating Graph Images	GI	PC Mean	DC Mean	CP Mean	SF Mean
Alphabet						
Albanian (Elbasan)	Arial	40	7.73	1.08	1.45	2.53
Albanian (Todhri)	Arial	53	7.18	1.13	2.08	2.92
Armenian (Eastern)	Arial	38	7.43	1.03	1.55	2.63
Asomtavruli	Arial	38	7.53	1.00	2.00	2.97
Avestan	Ahuramazda	54	9.83	1.00	2.11	3.52
Bassa	Arial	30	7.66	1.02	1.43	2.68
Belarusian	Arial	32	7.35	1.17	1.83	2.74
Bosnian	Arial	30	7.59	1.37	1.52	2.63
Bulgaria	Arial	30	7.40	1.03	2.05	2.85
Celtiberian	Arial	28	6.59	1.11	2.61	3.39
Cyrillic (Abkhaz)	Arial	56	9.41	1.33	2.68	3.68
Danish	Arial	29	7.25	1.05	1.62	2.38
Deseret	Code 2000	38	6.65	1.00	1.71	2.61
Dutch	Arial	26	6.85	1.04	1.44	2.25
English	Arial	26	6.85	1.04	1.44	2.25
Enochian	Enochian	22	7.41	1.14	1.59	2.68
Finnish	Arial	28	7.00	1.20	1.41	2.34
Fraser	Arial	40	7.82	1.00	1.68	2.43
French	Arial	26	6.85	1.04	1.44	2.23
Glagolitic	Arial	42	10.13	1.02	4.81	5.07
Gothic (Wulfila)	Alphabetum Unicode	25	7.11	1.08	1.16	2.36
Greek	Arial	24	7.09	1.06	1.43	2.27
German	Arial	26	6.85	1.04	1.44	2.25
Hungarian Runes	Arial	46	9.09	1.00	2.85	3.70
Icelandic	Arial	32	7.11	1.28	1.47	2.45
Italian	Arial	21	6.74	1.02	1.45	2.17
Kazakh	Arial	42	7.31	1.11	1.94	2.85
Korean (Hangeul)	MS Mincho	40	14.71	1.38	2.15	3.40
Kyrgyz	Arial	36	7.48	1.11	2.04	2.89
Latin (ancient)	Alphabetum Unicode	21	6.12	1.00	1.86	2.71
Latin (modern)	Arial	41	8.41	1.49	1.61	2.71
Macedonian	Arial	31	7.27	1.08	1.89	2.81
Marsiliana	Arial	26	9.97	1.00	2.27	2.88
Mkhedruli	BPG Glaho	38	7.80	1.00	1.32	2.39
Mongolian	Arial	35	7.49	1.11	2.01	2.87
Montenegrin	Arial	33	7.65	1.39	1.47	2.64
N’Ko	JG Nko	27	5.50	1.00	2.11	2.96
Norwegian	Arial	29	7.25	1.05	1.62	2.38
Nuskhuri	BPG Nino Khutsuri U	38	7.12	1.00	3.97	5.08
Old Church Slavonic	Arial	45	8.42	1.21	2.26	3.24
Old Permic (Abur)	Arial	38	9.16	1.05	2.39	3.42
Pahawh Hmong	Naadaa	166	11.05	1.80	2.16	4.02
Pollard Miao	Ahmao	85	7.19	1.61	1.31	2.87
Portuguese	Arial	26	6.85	1.04	1.44	2.25

**Table 10** (continued)

Languages by Writing System	Font for Creating Graph Images	GI	PC Mean	DC Mean	CP Mean	SF Mean
Romanian	Arial	31	6.96	1.19	1.45	2.34
Runic (Danish Futhark)	Code 2000	16	6.55	1.00	1.88	2.75
Runic (Elder Futhark)	Code 2000	24	7.20	1.04	2.50	3.25
Russian	Arial	33	7.51	1.12	2.05	2.89
Santali (OICemet')	Arial	30	10.27	1.07	2.43	3.40
Serbian	Arial	30	7.34	1.02	2.02	2.83
Somali (Osmanya)	MPH 2B Damase	30	11.52	1.00	1.47	2.63
Sorang Sompeng	Arial	24	10.55	1.00	3.13	4.25
Spanish	Arial	27	6.93	1.07	1.48	2.31
Swedish	Arial	29	7.14	1.19	1.47	2.40
Tajik	Arial	35	7.48	1.14	1.94	2.90
Theban	Theban	25	10.49	1.12	3.56	4.56
Ukrainian	Arial	33	7.16	1.11	1.89	2.79
Varang Kshiti	Arial	30	6.62	1.00	2.27	3.23
Yupik	Arial	44	7.84	1.25	2.19	3.18
Zhuyin Fuhao	DFKai-SB	37	10.51	1.11	2.35	3.51
<b>Abjad</b>						
Ancient Berber (Vertical)	Tamalout Standard Unicode	25	9.64	2.00	2.20	4.00
Arabic	Arial	28	8.78	1.82	1.36	3.07
Aramaic (Early)	Aramaic Early Br Rkb	22	6.31	1.00	2.32	2.91
Hebrew	Arial	22	4.90	1.07	0.89	1.96
MiddlePersian (Pahlavi)	Arial	22	4.95	1.00	1.64	2.82
Nabataean	Arial	22	5.99	1.09	1.59	2.68
Neo Tifinagh	Hapax Berbère	33	9.82	1.18	2.12	3.15
Parthian	Arial	22	5.37	1.05	1.64	2.82
Pashto	Arial	40	9.14	2.03	1.43	3.35
Phoenician	MPH 2B Damase	22	7.55	1.00	2.32	2.95
Psalter	Arial	21	4.94	1.00	1.38	2.29
Sabaeen Minean	Sabaen44	29	6.09	1.00	2.83	3.62
Samaritan	Samaritan	22	9.02	1.00	3.05	3.82
South Arabian	Arial	28	7.95	1.00	2.46	3.18
Syriac	Estrangelo Edessa	22	5.74	1.09	1.55	2.64
Tifinagh	MPH 2B Damase	33	10.81	1.24	1.97	3.09
<b>Alphasyllabary</b>						
'Phags-pa	BabelStone Phags-pa Book	41	9.87	1.00	4.44	5.17
Ahom	Ahom	45	11.00	1.51	2.04	3.42
Amharic	GF Zemen Unicode	282	7.47	1.03	2.74	3.50
Balinese	JG Aksara Bali	84	23.32	1.64	2.56	4.13
Batak	Arial	32	5.19	1.41	0.72	2.09
Bengali	Akaash Normal	57	14.60	1.21	4.26	5.51
Brahmi	Brahmi TTF	52	4.89	1.12	1.56	2.67
Buhid (Mangyan)	Arial	48	8.03	1.46	3.29	4.60
Burmese	Myanmar1	62	13.72	1.53	2.27	3.68
Dehong	Arial	30	4.00	1.03	2.43	3.03
Devanagari	Sanskrit 2003	62	9.41	1.03	2.98	4.27
Dives Akuru	Arial	46	10.45	1.15	1.70	3.09
Ethiopic (Ge'ez)	Code 2000	234	7.63	1.00	2.56	3.32
Gujarati	Shruti	64	9.23	1.28	1.47	2.81
Gurmukhi	Anmol Uni	60	11.81	1.22	3.32	4.68

**Table 10** (continued)

Languages by Writing System	Font for Creating Graph Images	GI	PC Mean	DC Mean	CP Mean	SF Mean
Hanuno'o(Mangyan)	Arial	48	11.05	1.48	2.52	4.13
Hindi	Sanskrit 2003	66	9.25	1.14	2.91	4.27
Inuktitut	Aboriginal Serif Regular	112	7.65	1.61	1.28	2.88
Kannada	Tunga	50	12.55	1.42	2.40	3.84
Kharosthi	MPH 2B Damase	39	8.57	1.05	1.33	2.44
Khmer	Khmer OS	130	10.42	1.44	6.02	7.12
Lao	Saysettha Web	78	13.40	1.63	2.90	4.71
Lepcha_Rong	JG Lepcha	77	9.13	1.06	2.71	3.74
Limbu	MPH 2B Damase	45	8.60	1.16	1.98	3.16
Malayalam	ML-NILA01	69	14.13	1.13	1.97	3.64
Manipuri	Akaash Normal	57	11.98	1.19	3.21	4.42
Marathi	Sanskrit 2003	65	8.53	1.29	2.94	4.22
Meroitic (non-hieroglyphic)	Arial	23	6.89	1.30	2.30	3.48
Oriya	Raghu Oriya	66	16.25	1.11	2.30	3.27
Redjang (Kaganga)	Arial	36	6.12	1.17	1.83	2.97
Sindhi	Bahij Nassim-Regular	51	9.86	1.31	3.37	4.63
Sinhala	Potha	71	14.71	1.51	3.20	3.93
Soyombo	JG Soyombo	86	10.79	1.76	5.16	6.87
Syoti-Nagri	Arial	38	10.49	1.11	3.37	4.79
Tagalog	Tagalog Stylized	45	14.87	1.78	1.53	3.02
Tagbanwa	Arial	42	12.18	1.64	2.07	3.93
Tamil	Code 2000	47	14.58	1.15	3.19	4.68
Telugu	NATS	70	11.41	1.33	2.68	4.10
Thaana	Free Serif	49	6.20	1.71	1.63	3.41
Thai	Angsana New	102	14.88	1.68	4.54	6.24
Tibetan	Arial	34	11.79	1.00	3.44	4.38
<b>Syllabary</b>						
Carrier Dene	Code 2000	195	10.27	1.22	2.93	4.14
Cherokee	Aboriginal Sans	85	7.07	1.01	1.87	2.86
Cree (Woodland)	Aboriginal Serif	80	7.00	1.53	1.33	2.68
Cypriot	Alphabetum Unicode	55	11.60	1.87	1.55	3.58
Japanese (Hiragana)	MS Mincho	48	23.32	1.29	2.75	4.19
Japanese (Katakana)	MS Mincho	48	16.06	1.38	1.56	2.96
Kpelle	JG Kpelle A	86	22.26	3.14	4.43	9.01
LinearB	Penuturesu	71	30.20	2.31	2.89	5.17
Ndjuka'	Arial	57	8.34	1.04	2.39	3.05
Ojibwe	Aboriginal Serif	88	7.04	1.48	1.11	2.43
Vai	Dukor	208	13.07	1.82	2.88	4.86
<b>Morphosyllabary</b>						
Chinese (Simplified)	DFKai-SB	6097	29.47	4.01	9.54	10.60
Chinese(Traditional)	DFKai-SB	6097	32.47	4.55	11.64	12.50
Japanese (Kanji)	DFKai-SB	1006	28.62	3.84	9.65	10.43

*Note.* GI = graph inventory, PC = perimetric complexity, DC = number of disconnected components, CP = number of connected points; SF = number of simple features

## Appendix

Table 11 Graph pairs per list used in the similarity rating

	List 1		List 2		List 3		List 4	
Hebrew	א	צ	ב	א	ס	א	א	ץ
	ת	ב	ב	ג	ב	ה	ר	ב
	ב	ד	ב	ב	ת	ב	ב	ך
	פ	ג	ג	ן	ג	ז	ק	ג
	ד	ך	ש	ד	ב	ד	ד	ח
	ם	ה	ה	פ	ה	ש	כ	ה
	ו	נ	ת	ו	ל	ו	ו	א
	ו	ז	ז	י	ז	ג	ן	ז
	ח	מ	ה	ח	ב	ח	ח	ת
	כ	ט	ט	ס	ט	ע	ם	ט
	י	ש	ז	י	ץ	י	י	נ
	כ	כ	כ	ץ	כ	פ	ש	כ
	כ	ט	ל	כ	א	כ	כ	מ
	ך	ך	ך	ץ	ך	ן	ד	ך
	ל	ר	נ	ל	ש	ל	ל	צ
	ש	מ	מ	ע	מ	כ	כ	מ
	ם	ב	מ	ם	ח	ם	ם	ט
	ג	נ	נ	ר	נ	י	פ	נ
	ן	ז	ך	ו	נ	ן	ן	ו
	ק	ס	ס	כ	ס	פ	ת	ס
	ע	ל	צ	ע	ר	ע	ע	ש
	פ	פ	פ	כ	פ	ך	ע	פ
	פ	ב	ו	פ	ט	פ	פ	ב
	ה	ך	ך	פ	ך	ד	ג	ך
	צ	א	ט	צ	ם	צ	צ	ל
	ו	ץ	ץ	ת	ץ	ק	י	ץ
	ק	ח	ד	ק	צ	ק	ק	ה
	י	ר	ר	ך	ר	ו	ז	ר
	ש	ש	ח	ש	מ	ש	ש	ס
	ע	ש	ש	ק	ש	ת	פ	ש



Table 11 (continued)

	л	л	к	л	г	л	л	л
	о	л	л	д	л	к	г	л
Russian	А	В	н	а	Б	А	а	ж
	Т	Б	б	ъ	Б	Ф	о	б
	В	З	ю	в	Д	В	в	ш
	С	Г	г	р	Г	Ь	р	г
	Д	Л	ц	д	Н	Д	д	п
	Ё	Е	е	г	Е	И	с	е
	Ё	Ш	в	ё	В	Ё	ё	ю
	К	Ж	ж	щ	Ж	М	э	ж
	З	У	ш	з	М	З	з	х
	М	И	и	о	И	Э	щ	и
	Й	Ы	ч	й	Ю	Й	й	л
	И	К	к	м	К	Й	г	к
	Л	А	х	л	Ш	Л	л	ч
	Е	М	м	э	Ц	Щ	й	м
	Н	П	ы	н	З	Н	н	б
	Ф	О	о	и	О	С	е	о
	П	Д	л	п	У	П	п	н
	Ь	Р	р	е	Р	О	ф	р
	С	Ч	з	с	А	С	с	я
	Г	Т	т	с	Т	К	ь	т
	У	Х	б	у	П	У	у	в
	Э	Ф	ф	ь	Ф	Р	ё	ф
	Х	Б	я	х	Я	Х	х	а
	Й	Ц	ц	ё	Ц	Ъ	к	ц
	Ч	Н	п	ч	Х	Ч	ч	у
	Щ	Ш	ш	й	Ш	Е	м	ш
	Щ	Ц	ж	щ	Л	Щ	щ	ы
	Р	Ъ	ь	к	Ъ	Г	т	ь
	Ы	Ю	д	ы	Ж	Ы	ы	ц
	Ъ	Ь	ь	т	Ь	Т	и	ь
	Э	Я	у	э	Ч	Э	э	з
	О	Ю	ю	ф	Ю	Ё	ь	ю
	Я	Ж	а	я	Ы	Я	я	д
Cree	Δ	▷	◁	▷	Δ	◁	<	Δ
	Δ	Δ	◁	Δ	▷	◁	▷	Δ



**Table 11** (continued)

ఇ ఓ	గ ర	బ గ	ఐ ఇ
ఉ ఈ	ఝ ఘ	ఘ ఘ	ఈ ఋ
ఉ ఊ	జ న	ష జ	అః ఉ
ఈ ఊ	గ చ	చ హ	ఉః ఙ
ఋ ఔ	చ డ	చ చ	ఉః ఋ
ఋ ఋ	ఞ జ	జ ళ	ఋ ఉ
౞ ఐ	ఝ డ	య ఝ	ఏ ౞
౞ ఙ	ధ ఞ	ఞ గ	౞ ఈ
ఎ ఏ	ట ల	ద ట	అం ఎ
ఇ ఏ	భ ర	ర ర	ఏ ౞
ఐ ఎ	డ త	శ డ	ఒ ఐ
ఆ ఒ	జ డ	డ ఝ	ఒ అ
ఓ ఒ	ణ బ	న ణ	ఎ ఓ
ఋ ఔ	ళ త	త ఞ	ఔ అ
అం అః	థ ప	ర థ	ఔ అం
౞ అః	ర ద	ద జ	అః ఋ
కా క్	ధ థ	డ ధ	కా కా
కి కి	క న	న వ	కి కా
కీ క్	ప చ	న ప	కొ కి
కూ కు	ఘ ఘ	ఫ ఖ	కు కం
కూ క్	బ ట	ల బ	కు కూ
కం కృ	ఖ భ	భ చ	కృ కః

Table 11 (continued)

	కౌ	కృ	మ	ష	ట	మ	కై	కౌ
	కృ	కృ	హ	య	య	ఘ	కృ	కౌ
	కృ	కౌ	ర	ద	ప	ర	కౌ	కృ
	క	క	వ	ల	ల	ధ	క	క
	క	క	వ	య	ణ	వ	క	క
	కృ	కై	జ	శ	శ	ఇ	కై	కృ
	కౌ	క	ష	స	మ	ష	కౌ	కౌ
	క	క	ప	స	స	క	క	క
	కౌ	కై	హ	మ	ధ	హ	క	కౌ
	క	క	ఇ	ళ	ళ	జ	క	కృ
	క	క	ఱ	ణ	త	ఱ	కృ	క
Simple Chinese	七 本 口 刀 尸 子 目 失 目 甲 土 丰 且 白 王 占 杏 用 里 吏	乜 禾 日 力 尺 子 早 矢 自 曰 王 井 日 百 生 舌 杏 甬 甲 更	夫 矢 中 天 井 丁 子 曰 于 九 舌 田 甲 甘 早 史 李 更 冉 里	矢 天 史 大 卅 了 丈 白 手 子 古 由 中 耳 卓 吏 季 申 再 重	了 九 于 己 木 壬 天 矢 中 王 日 廿 另 禾 止 生 甲 申 果 弗	十 几 干 巳 本 士 大 天 申 工 有 世 吊 木 正 圭 申 里 更 夷	大 丁 人 土 巳 干 夫 夕 曰 日 斤 止 友 未 吏 皿 申 由 甲 且	夫 十 八 工 巴 牛 失 歹 田 目 斥 丘 反 朱 串 血 曰 曲 更 早

**Table 11** (continued)

Complex Chinese	僮 憾 濡 瞭 徼 褐 漕 醅 跛 椀 黑 察 愚 紊 娶 幕 寒 斐 粟 喜	撞 撼 燭 瞭 徼 褐 潭 酪 歧 棕 熏 蔡 蕙 索 最 暮 塞 裴 粟 嘉	僧 揪 檄 暗 梢 情 棒 嗑 跟 媒 覃 熏 窠 菟 煦 纂 莠 寅 亮 堂	憎 揪 激 暗 稍 惜 椿 磕 跟 媒 覃 墨 裹 冤 照 纂 萎 富 毫 掌	懊 搞 濯 蟠 睦 植 棹 蚧 鞘 勤 署 窘 寞 奢 梨 笙 莫 髡 奠 竟	懊 搞 濯 蟠 睦 植 楸 楮 蚧 鞘 勤 署 窘 寞 奢 替 焚 筐 募 髡 尊 章	漫 塘 晴 喋 褊 填 梧 跌 擒 鄙 覃 寥 惠 眷 耆 管 管 壁 罩 恙	慢 塘 晴 牒 褊 埧 楮 楮 踈 擒 敲 置 廖 愚 智 屠 篙 萱 擘 署 羔
-----------------	--	--	--	--	--	--	--	---

### Appendix A *lme4* (Bates, Maechler, & Dai, 2010) model formulae for fitting models and comparing the four models

#### Fitting five models

```
Full <- lmer(Response ~ 1 + dPC * dDC * dCP * dSF +
(1|Subject) + (1|Item), data = rating)
```

```
FullwoPC <- lmer(Response ~ 1 + dDC * dCP * dSF +
(1|Subject) + (1|Item), data = rating)
```

```
FullwoDC <- lmer(Response ~ 1 + dPC * dCP * dSF +
(1|Subject) + (1|Item), data = rating)
```

```
FullwoCP <- lmer(Response ~ 1 + dPC * dDC * dSF +
(1|Subject) + (1|Item), data = rating)
```

```
FullwoSF <- lmer(Response ~ 1 + dPC * dDC * dCP +
(1|Subject) + (1|Item), data = rating)
```

#### Comparing the full model with the other four models, respectively

```
anova(Full, FullwoPC)
```

```
anova(Full, FullwoDC)
```

```
anova(Full, FullwoCP)
```

```
anova(Full, FullwoSF)
```

1 denotes the intercept, \* denotes an interaction plus main effects between two predictors, and (x|a) denotes random effects of x by a.

#### References

- Ager, S. (1998). Omniglot: A guide to writing systems. In *Encyclopedia omniglot*. Retrieved from [www.omniglot.com](http://www.omniglot.com)
- 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. doi:10.1016/j.jmla.2007.12.005
- Bates, D., Maechler, M., & Dai, B. (2010). *lme4*: Linear mixed-effects models using S4 classes (R package version 0.999375-37) [Computer software manual]. Retrieved from <http://lme4.rforge.r-project.org>
- Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis*, *20*, 351–368. doi:10.1093/pan/mpr057
- Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*, 115–147. doi:10.1037/0033-295X.94.2.115
- Bright, W. (Ed.). (1992). *International encyclopedia of linguistics*. New York: Oxford University Press.
- Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, *6*, 3–5. doi:10.1177/1745691610393980

- Burnham, K. P., & Anderson, D. R. (2003). *Model selection and multimodel inference: A practical information-theoretic approach*. New York: Springer.
- Chandler, J., Mueller, P., & Paolacci, G. (2014). Nonnaïveté among Amazon Mechanical Turk workers: Consequences and solutions for behavioral researchers. *Behavior Research Methods*, *46*, 112–130. doi:10.3758/s13428-013-0365-7
- Chang, L. Y. (2015). *Visual orthographic variation and learning to read across writing systems* (Unpublished doctoral dissertation). University of Pittsburgh, Pennsylvania. <http://d-scholarship.pitt.edu/23959/>
- Chang, L. Y., Plaut, D. C., & Perfetti, C. A. (2016). Visual complexity in orthographic learning: Modeling learning across writing system variations. *Scientific Studies of Reading*, *20*, 64–85.
- Changizi, M. A., & Shimojo, S. (2005). Character complexity and redundancy in writing systems over human history. *Proceedings of the Royal Society B*, *272*, 267–275.
- Changizi, M. A., Zhang, Q., Ye, H., & Shimojo, S. (2006). The structures of letters and symbols throughout human history are selected to match those found in objects in natural scenes. *American Naturalist*, *167*, E117–E139.
- Chen, Y. P., Allport, D. A., & Marshall, J. C. (1996). What are the functional orthographic units in Chinese word recognition: The stroke or the Stroke pattern? *Quarterly Journal of Experimental Psychology*, *49*, 1024–1043.
- Chen, H. C., Chang, L. Y., Chiou, Y. S., Sung, Y. T., & Chang, K. E. (2011). Construction of Chinese orthographic database for Chinese character instruction. *Bulletin of Educational Psychology*, *43*, 269–290.
- Chikhman, V., Bondarko, V., Danilova, M., Goluzina, A., & Shelepin, Y. (2012). Complexity of images: Experimental and computational estimates compared. *Perception*, *41*, 631–647. doi:10.1068/p6987
- Coen-Cagli, R., & Schwartz, O. (2013). The impact on midlevel vision of statistically optimal divisive normalization in V1. *Journal of Vision*, *13*(8), 1–20. doi:10.1167/13.8.13
- Cook, V., & Bassetti, B. (2005). An introduction to researching second language writing systems. In V. Cook & B. Bassetti (Eds.), *Second language writing systems* (pp. 1–67). Clevedon: Multilingual Matters.
- Daniels, P. T. (1990). Fundamentals of grammatology. *Journal of the American Oriental Society*, *110*, 727–731.
- Demetriou, A., Kui, Z. X., Spandoudis, G., Christou, C., Kyriakides, L., & Platsidou, M. (2005). The architecture, dynamics, and development of mental processing: Greek, Chinese, or universal? *Intelligence*, *33*, 109–141.
- Ehrenstein, W. H. (2008). Gestalt psychology. In *Encyclopedia of neuroscience* (pp. 1721–1724). New York: Springer.
- Fiset, D., Blais, C., Ethier-Majcher, C., Arguin, M., Bub, D., & Gosselin, F. (2008). Features for identification of uppercase and lowercase letters. *Psychological Science*, *19*, 1161–1168.
- Frost, R. (2012). Towards a universal model of reading. *Behavioral and Brain Sciences*, *35*, 263–279. doi:10.1017/S0140525X11001841
- Gelb, I. J. (1952). *A study of writing*. Chicago: University of Chicago Press.
- Gibson, E. J. (1969). *Principles of perceptual learning and development*. New York: Meredith.
- Grainger, J., Rey, A., & Dufau, S. (2008). Letter perception: From pixels to pandemonium. *Trends in Cognitive Sciences*, *12*, 381–387. doi:10.1016/j.tics.2008.06.006
- Grill-Spector, K., & Malach, R. (2004). The human visual cortex. *Annual Review of Neuroscience*, *27*, 649–677. doi:10.1146/annurev.neuro.27.070203.144220
- Horton, J. J., Rand, D. G., & Zeckhauser, R. J. (2011). The online laboratory: Conducting experiments in a real labor market. *Experimental Economics*, *14*, 3990425.
- Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, *160*, 251–260.
- Hubel, D. H., & Wiesel, T. N. (1965). Receptive fields and functional architecture in two non-striate visual areas (18 and 19) of the cat. *Journal of Neurophysiology*, *28*, 229–289.
- Jiang, Y. V., Shim, W. M., & Makovski, T. (2008). Visual working memory for line orientations and face identities. *Perception & Psychophysics*, *70*, 1581–1591. doi:10.3758/PP.70.8.1581
- Katz, L., & Frost, R. (1992). Reading in different orthographies: The orthographic depth hypothesis. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 67–84). Amsterdam: North-Holland.
- Koffka, K. (1963). *Principles of Gestalt psychology*. New York: Harcourt, Brace & World (Original work published 1935).
- Lanthier, S. N., Risko, E. F., Stolz, J. A., & Besner, D. (2009). Not all visual features are created equal: Early processing in letter and word recognition. *Psychonomic Bulletin & Review*, *16*, 67–73. doi:10.3758/PBR.16.1.67
- Lavine, L. O. (1977). Differentiation of letterlike forms in prereading children. *Developmental Psychology*, *13*, 89–94. doi:10.1037/0012-1649.13.2.89
- Lehmann, E. L. (1986). *Testing statistical hypotheses* (2nd ed.). New York: Springer.
- 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. doi:10.1037/0012-1649.39.5.891
- Liu, D., Chen, X., & Wang, Y. (2016). The impact of visual-spatial attention on reading and spelling in Chinese children. *Reading and Writing*, *29*, 1435–1447. doi:10.1007/s11145-016-9644-x
- McBride-Chang, C., Zhou, Y., Cho, J.-R., Aram, D., Levin, I., & Tolchinsky, L. (2011). Visual spatial skill: A consequence of learning to read? *Journal of Experimental Child Psychology*, *109*, 256–262. doi:10.1016/j.jecp.2010.12.003
- Ministry of Education in Japan. (2015). *Official list of kyōiku kanji by grade*. Retrieved from the Ministry of Education, Culture, Sports, Science and Technology–Japan website, [www.mext.go.jp/a\\_menu/shotou/new-cs/youryou/syo/koku/001.htm](http://www.mext.go.jp/a_menu/shotou/new-cs/youryou/syo/koku/001.htm)
- Mueller, S. T., & Weidemann, C. T. (2012). Alphabetic letter identification: Effects of perceivability, similarity, and bias. *Acta Psychologica*, *139*, 19–37. doi:10.1016/j.actpsy.2011.09.014
- Nag, S. (2007). Early reading in Kannada: The pace of acquisition of orthographic knowledge and phonemic awareness. *Journal of Research in Reading*, *30*, 7–22.
- Nag, S. (2008). *Kannada vocabulary test*. Bangalore: Promise Foundation.
- Nag, S. (2014). Alphabetism and the science of reading: From the perspective of the akshara languages. *Frontiers in Psychology*, *5*, 866. doi:10.3389/fpsyg.2014.00866
- Nag, S., Caravolas, M., & Snowling, M. J. (2011). Beyond alphabetic processes: Literacy and its acquisition in the alphasyllabic languages. *Reading and Writing*, *24*, 615–622.
- Nag, S., & Snowling, M. J. (2011). Cognitive profiles of poor readers of Kannada. *Reading and Writing*, *24*, 657–676.
- Nag, S., Snowling, M., Quinlan, P., & Hulme, C. (2014). Child and symbol factors in learning to read a visually complex writing systems. *Scientific Studies of Reading*, *18*, 309–324. doi:10.1080/10888438.2014.892489
- Nag, S., Treiman, R., & Snowling, M. J. (2010). Learning to spell in an alphasyllabary: The case of Kannada. *Writing Systems Research*, *2*, 41–52.
- Pelli, D. G., Burns, C. W., Farell, B., & Moore-Page, D. C. (2006). Feature detection and letter identification. *Vision Research*, *46*, 4646–4674. doi:10.1016/j.visres.2006.04.023

- Perfetti, C. A. (2003). The universal grammar of reading. *Scientific Studies of Reading*, 7, 3–24.
- Perfetti, C. A., & Harris, L. N. (2013). Reading universals are modulated by language and writing system. *Language Learning and Development*, 9, 296–316. doi:10.1080/15475441.2013.813828
- Perfetti, C. A., & Verhoeven, L. (in press). *Learning to read across languages and writing systems*. Cambridge, UK: Cambridge University Press.
- Reas, C., & Fry, B. (2010). *Getting started with processing*. Sebastopol: O'Reilly.
- Rimzhim, A., Katz, L., & Fowler, C. A. (2014). Brahmi-derived orthographies are typologically aksharic but functionally predominantly alphabetic. *Writing Systems Research*, 6, 41–53. doi:10.1080/17586801.2013.855618
- Robins, S., & Treiman, R. (2009). Learning about writing begins informally. In D. Abram & O. Korat (Eds.), *Literacy development and enhancement across orthographies and cultures* (pp. 17–29). New York: Springer.
- Sayim, B., & Cavanagh, P. (2011). What line drawings reveal about the visual brain. *Frontiers in Human Neuroscience*, 5(118), 1–4. doi:10.3389/fnhum.2011.00118
- Seidenberg, M. S. (2011). Reading in different writing systems: One architecture, multiple solutions. In P. D. McCardle, B. Miller, J. R. Lee, & O. J. L. Tzeng (Eds.), *Dyslexia across languages: Orthography and the brain–gene–behavior link* (pp. 146–168). New York: Paul Brookes.
- Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, 94, 143–174.
- Share, D. L., & Daniels, P. T. (2016). Aksharas, alphasyllabaries, abugidas, alphabets and orthographic depth: Reflections on Rimzhim, Katz and Folwer (2014). *Writing System Research*, 8, 17–31. doi:10.1080/17586801.2015.1016395
- Shen, H. H., & Ke, C. (2007). Radical awareness and word acquisition among nonnative learners of Chinese. *Modern Language Journal*, 91, 97–111.
- Simcox, T., & Fiez, J. (2014). Collecting response times using Amazon Mechanical Turk and Adobe Flash. *Behavior Research Methods*, 46, 95–111. doi:10.3758/s13428-013-0345-y
- Simpson, I. C., Mousikou, P., Montoya, J. M., & Defior, S. (2013). A letter visual-similarity matrix for Latin-based alphabets. *Behavior Research Methods*, 45, 431–439.
- Spillmann, L., & Ehrenstein, W. H. (2004). Gestalt factors in the visual neurosciences. In L. Chalupa & J. S. Werner (Eds.), *The visual neurosciences* (pp. 1573–1589). Cambridge: MIT Press.
- Sprouse, J. (2011). A validation of Amazon Mechanical Turk for the collection of acceptability judgments in linguistic theory. *Behavior Research Methods*, 43, 155–167. doi:10.3758/s13428-010-0039-7
- Stephens, M. A. (1974). EDF statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association*, 69, 730–737.
- Su, Y.-F., & Samuels, S. J. (2010). Developmental changes in character-complexity and word-length effects when reading Chinese script. *Reading and Writing*, 23, 1085–1108. doi:10.1007/s11145-009-9197-3
- Szwed, M., Cohen, L., Qiao, E., & Dehaene, S. (2009). The role of invariant line junctions in object and visual word recognition. *Vision Research*, 49, 718–725. doi:10.1016/j.visres.2009.01.003
- Tamaoka, K., & Kiyama, S. (2013). The effects of visual complexity for Japanese kanji processing with high and low frequencies. *Reading and Writing*, 26, 205–223. doi:10.1007/s11145-012-9363-x
- Tokowicz, N., Michael, E., & Kroll, J. F. (2004). The roles of study abroad experience and working memory capacity in the types of errors made during translation. *Bilingualism: Language and Cognition*, 7, 255–272.
- Treiman, R., Hompluem, L., Gordon, J., Decker, K., & Markson, L. (2016). Young children's knowledge of the symbolic nature of writing. *Child Development*, 87, 583–592. doi:10.1111/cdev.12478
- Treiman, R., & Kessler, B. (2011). Similarities among the shapes of writing and their effects on learning. *Written Language and Literacy*, 14, 39–57.
- Treiman, R., & Kessler, B. (2014). *How children learn to write words*. New York: Oxford University Press.
- Treiman, R., Mulqueeny, K., & Kessler, B. (2014). Young children's knowledge about the spatial layout of writing. *Writing System Research*, 7, 235–244. doi:10.1080/17586801.2014.924386
- Troncoso, X. G., Macknik, S. L., & Martinez-Conde, S. (2011). *Visual prosthetics* (G. Dagnelie, Ed.). Boston, MA: Springer. doi:10.1007/978-1-4419-0754-7
- Van Essen, D. C., Anderson, C. H., & Felleman, D. J. (1992). Information processing in the primate visual system: An integrated systems perspective. *Science*, 255, 419–423. doi:10.1126/science.1734518
- Wang, Y., McBride-Chang, C., & Chan, S. (2014). Correlates of Chinese kindergarteners' word reading and writing: The unique role of copying skills? *Reading and Writing*, 27, 1281–1302. doi:10.1111/1467-9817.12016
- Wasserman, L. (2006). *All of statistics: A concise course in statistical inference* (pp. 218–222). New York: Springer.
- Watson, A. B. (2012). Perimetric complexity of binary digital images: Notes on calculation and relation to visual complexity. *Mathematica Journal*, 14, 1–41. doi:10.3888/tmj.14-5
- Watt, W. C. (1983). Grade der Systemhaftigkeit: Zur Homogenität der Alphabetschrift [Degrees of systematicity: On the homogeneity of the alphabetic script]. *Zeitschrift für Semiotik*, 5, 371–399.
- Watt, W. C. (1994). Curves as angles. In *Writing systems and cognition—Perspectives from psychology, physiology, linguistics, and semiotics* (pp. 215–246). Amsterdam, The Netherlands: Springer.
- Winkel, H. (2010). Spelling development in Thai children. *Journal of Cognitive Science*, 11, 7–35.
- Wu, N., Zhou, X., & Shu, H. (1999). Sublexical processing in reading Chinese: A development study. *Language and Cognitive Processes*, 14, 503–524.
- Wydell, T. N. (2012). Cross-cultural/linguistic differences in the prevalence of developmental dyslexia and the hypothesis of granularity and transparency. In T. N. Wydell & L. Fern-Pollak (Eds.), *Dyslexia: A comprehensive and international approach* (pp. 1–14). Rijeka: InTech.
- Yin, L., & McBride, C. (2015). Chinese kindergarteners learn to read characters analytically. *Psychological Science*, 26, 424–432.
- Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, 131, 3–29. doi:10.1037/0033-2909.131.1.3