Modeling the Visual Similarity of Chinese Characters Across Expertise Groups

Zhuqian Zhou[1], James E. Corter[1]

Teachers College, Columbia University

Abstract

This study investigated how well similarity models of Chinese characters developed in previous research could be used to model human judgments across different levels of proficiency in Chinese. The behavioral data collected from the three groups of participants confirmed the superiority of and preference for configurations over components in experts' perceptions. In contrast, Chinese learners' and novices' criteria for similarity judgments were less clear, as indicated by the low proportion of variance that could be accounted for by extended tree analysis of their group judgments. We discuss computational challenges in modeling human perception and judgments about Chinese characters and propose future directions for research, including the potential use of statistical and machine learning techniques with larger datasets for improved model development.

Keywords: Chinese characters; computational modeling; extended tree analysis; perceptual expertise; visual similarity

Introduction

Learning of a novel visual category arises through perceptual experience. But increasing experience, and expertise, often leads to the grouping and re-grouping of perceptual elements. The trajectory from novices to experts, therefore, may manifest in qualitative shifts in the perception of the object being studied (Chase and Simon, 1973; Vogt & Magnussen, 2007; Slovic, 2016).

Language learning is no exception. The study of visual perception of a language's basic characters becomes even more important when it comes to nonalphabetic languages like Chinese whose writing system possesses thousands of characters. Thus, understanding how people's perception of Chinese characters varies developmentally with increasing expertise in the Chinese language should not only help Chinese language educators to devise effective instructional strategies, but also inform visual cognition research, especially research focused on the debate on holistic versus analytical processing as a mark of perceptual expertise for word recognition, in contrast to face perception (Chen & Yeh, 2015; Moret-Tatay et al., 2020; Ventura, & Cruz, 2023; Wong et al., 2012).

In the dynamic process of visual perception of Chinese characters, being able to recognize a character and to discriminate it from its confusingly similar neighbors is basic to Chinese perceptual expertise (Woodrome & Johnson, 2009). Discrimination in particular relies on making effective and accurate similarity judgments in the character pool (Ashby & Perrin, 1988; Tversky, 1977). Therefore, in order to study the perception of Chinese characters, investigating their visual similarity for language learners can provide telling information as to recognition processes and visual learning for complex stimuli. Previous research concerning Chinese character recognition followed this general logic.

To investigate human judgments of the visual similarity of Chinese characters, previous research adopted either direct or indirect assessments of similarity that were in line with common measures of psychological similarity in various domains (Medin et al., 1993). Direct assessments had participants rate the degree of similarity of each pair of characters on a Likert-type scale (Yencken and Baldwin, 2006) while indirect assessments derived similarity data from behavioral tasks, such as sorting cards with Chinese characters into piles according to similarity (Rosenberg, 1975; Yeh & Li, 2002) and discriminating confusing character pairs under time pressure (Yang & Wang, 2018).

While different behavioral measures of character similarity provide generally consistent results, there has not yet been a consensus on how the visual similarity of Chinese characters should be modeled computationally. The only common ground so far has been to specify the visual features of the characters. Visual features are vital because there has been an established consensus in cognition and vision research that people recognize patterns and letters through feature detection (Coates et al., 2019; Geyer & DeWald, 1973; Gibson, 1969). Even when Pelli et al. (2006) made a serious effort to disprove the feature detection theory, strong results from a series of experiments they conducted (including experiments on Chinese character identification) compelled them to reaffirm it. Historically, various lists of features have been proposed to account for English letter recognition (Geyer, 1970); Gibson, 1969; Laughery, 1971; Coates et al., 2019) with pros and cons for each. The definitive set of visual features of English letters has yet to be established. Meanwhile, there is even less consensus about the visual features of Chinese characters.

Current computational models of Chinese character similarity can be classified into two groups by how each character was decomposed and coded. One type of model requires an explicit experimenter-specified description of the visual features of a character, that recognizes the hierarchical structure of Chinese ideographs. Figure 1 demonstrates this hierarchy—a Chinese character is composed of one or more radical components; a component is further composed of one or more basic strokes—continuous marks in writing. Therefore, one *a priori* way to code Chinese characters is to specify what/how many components and/or strokes are in a character (Yang, 1998; Yencken & Baldwin, 2008). Moreover, different relative positions of components in a character lead to different configurations of the character. 明 in Figure 1 is of a left-right configuration type, because its components 日 and 月 are on the left and right side

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

respectively. Other common configurations include the above-below configuration (e.g. 吞 *swallow* = 天 *sky* + 口 *mouth*) and the boxed configuration (e.g. 囚 *prison* = 人 *person* + 口 *mouth*). Thus, apart from components and strokes, the configuration of a character is also a commonly used human-coded feature (Yeh & Li, 2002; Yeh et al., 2003).

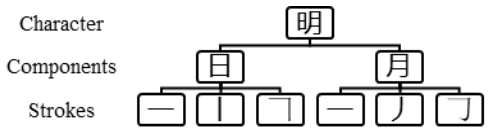

Figure 1: The hierarchical structure of Chinese character 明

The other type of computational models of Chinese character similarity stems from computer vision research and relies on low-level features machine-extracted directly from the images of characters. Lepage (2014) simulated a Chinese character analogy task using this approach. The four groups of feature vectors were calculated based on a 16 × 16 binary image of each character, with each pixel's value of either one or zero indicating whether it was a black pixel or a white one. The four groups of feature vectors included:

1) features based on random positions: the number of black pixels at 1/*n* pixel positions randomly selected from the binary image (named as "pix-*n*"), and the number of black pixels at (*n*-1)/*n* pixel positions randomly selected from the binary image (named as "pyx-*n*"),

2) features of rows and columns: the number of black pixels on each of the 16 rows (named as "lin"), and on each of the 16 columns (named as "col"), or of rows and columns combined (named as "lincol"),

3) features inspired by classic image processing algorithms: the number of pixels with given contexts defined by possible patterns of a 3 × 3 window with the pixel of interest in the center (named as "nei"), and the number of pixels with the same gray level for its neighborhood defined by the possible number of black pixels in a 3 × 3 window with the pixel of interest in the center (named as "ngr"), and

4) features inspired by classical image data structure: the number of black pixels in each quarter of each level of the binary image until level *n* (named as "qua-*n*", see Figure 2). When *n* is zero, qua-*n* is a one-dimensional feature vector with one value equal to the total number of black pixels on the image. When *n* is one, qua-*n* is a five-dimension feature vector with one value equal to the total number of black pixels on the image and the rest four values equal to the number of black pixels on the four evenly divided parts of the image. The character image was cut in quarters by cutting the rows and columns in halves simultaneously.

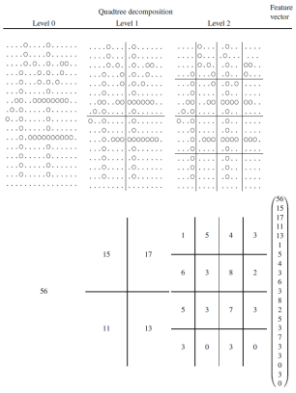

Figure 2: The "qua-2" feature vector of 伴 from Lepage (2014, p. 48)

But in general, the adequacy of particular visual features of Chinese characters for modeling their visual similarity has been little investigated. Thus, the present research seeks to compare the abovementioned visual features of Chinese characters that have been proposed by computational psychologists and computer scientists against behavioral data collected from three groups of adults—Chinese native readers, Chinese second-language learners, and Chinese novices who did not know any Chinese. Differences in model performance across different groups of participants were analyzed and implications are discussed.

Methods

Materials

We used 16 characters (i.e. 啄, 曉, 偌, 排, 售, 曇, 惹, 菲, 問, 曆, 匿, 痱, 重, 爾, 幾, 事) in the present study, a subset of characters employed in previous research of human judgments of Chinese character similarity (Yeh & Li, 2002). The characters were deliberately chosen to provide varied information for subjective similarity judgments. As shown in Table 1, three basic configurations of elements (i.e. left-right, above-below, and boxed) and four specific components (i.e. 口, 日, 若, 非) were factorially varied to select 12 characters. That is, each of the twelve characters corresponded to a unique configuration-component combination. Four additional characters (the last four in Table 1) were added as foils. The font type of Chinese characters used in the current study was Microsoft YaHei.

Table 1: Sixteen Chinese characters and their corresponding configurations and key components.

Stimulus	Configuration	Component
啄	Left-Right	口
曉	Left-Right	日

偌	Left-Right	若
排	Left-Right	非
售	Above-Below	口
曇	Above-Below	日
惹	Above-Below	若
菲	Above-Below	非
問	Boxed	口
曆	Boxed	日
匿	Boxed	若
痱	Boxed	非
重	Foil	Foil
爾	Foil	Foil
幾	Foil	Foil
事	Foil	Foil

Participants

Visual similarity judgments were collected from 18 Chinese novices who did not know any Chinese, 17 Chinese second-language learners, and seven native Chinese readers. The participants were recruited via the Amazon Mechanical Turk (MTurk) marketplace. All of the 42 participants were age 18 or older. Each participant received one US dollar (USD) for approximately ten minutes of work. The goal was to have at least 15 valid responses from each proficiency group via a single channel. However, due to difficulty recruiting participants from MTurk who were Chinese native readers, we collected another 15 responses from Chinese native readers via *wjx.cn*, the Chinese equivalent of MTurk. These participants were also above the age of 18. Each received five Chinese Yuan (CNY) for their participation. In total, 57 valid responses were collected and submitted for subsequent analysis.

Procedure

The entire procedure was delivered online via Qualtrics. After answering seven survey questions on their age, gender, education level, and proficiency level in Chinese, participants completed four rounds of sorting tasks by giving responses to the four subsequent questions. Each round requested the participant to sort the 16 characters into two, three, four, or five piles "according to their Visual Similarity". Participants accomplished the sorting task by dragging the randomly sequenced characters listed vertically on the left of the screen and dropping them into the corresponding number of boxes on the right of the screen. This drag-and-drop form of interaction was implemented using the "Pick, group, and rank" question tool on Qualtrics. The last question on the survey was an open-ended question asking participants to "explain why you made the above groupings". This question was not mandatory, however.

(Dis)similarity Matrices of Participants

For each participant, a 16 × 16 matrix was built, indicating how many times the character on the corresponding row and the character on the corresponding column were dragged to the same box. The highest score possible in a cell for each participant was four because there were four rounds of sorting. The lowest score possible could be zero if the pair of characters were never put into the same box by the participant. For each proficiency group, we summed up individual matrices and subtracted the score of each cell from the maximum possible score (i.e. the summed value in diagonal cells) to get the dissimilarity matrix for later clustering analysis.

Feature Vectors and Dissimilarity Matrices of Models

We deployed and compared multiple visual similarity models of Chinese characters that differed in what feature vectors were used to represent the set of characters in the study. Five models were based on *a priori* human-coded features, i.e. 1) type of configuration, 2) key component, 3) number of total strokes, 4) number of horizontal strokes, and 5) number of vertical strokes, and 14 models were based on machine-extracted features—*pix-3, pix-5, pix-7, pyx-3, pyx-5, pyx-7, lin, col, lincol, nei, ngr, qua-2, qua-3,* and *qua-4* (see the Introduction for their definitions and descriptions).

The dissimilarity matrix for a particular model was computed by calculating the Euclidean distance between the feature vectors of a pair of characters (Yang, 1998). We chose Euclidian distance rather than cosine similarity because the overall level (i.e., how many black pixels on a dimension) matters when all characters are of the same size.

Results

Evaluating Model Fit

We evaluated how well a model performed in two ways—by comparing individual model performance against human performance and by comparing the performance of two sets of models in predicting human performance.

First, we assessed to what extent similarity measures of characters resulting from a model reproduced actual similarity data from the participants. This was done by calculating the rank-order correlation coefficient using Kendall's τ (Galili, 2015) between a model distance matrix and each proficiency group's dissimilarity matrix. The result is that none of the five high-level human-defined features show significant correlations with human similarity judgments. In contrast, almost all distance matrices generated by the low-level machine-coded features, except features *col* and *ngr*, significantly correlate with Chinese novices' dissimilarities (see Table 2).

Table 2: Correlation coefficients (Kendall's τ) between the dissimilarity matrices for the three participant groups and the distance matrices generated by low-level machine-extracted features.

		Native Readers	Learners	Novices
RP	pix-3	-0.033	.000	.157*
	pix-5	-0.02	-0.012	.156*
	pix-7	-0.027	-0.028	.136*
	pyx-3	0.008	-0.028	.157*
	pyx-5	-0.001	-0.039	.152*
	pyx-7	0.003	-0.026	.174**
RC	lin	-0.065	0.021	.152*
	col	0.119	-0.046	0.089
	lincol	0.01	-0.006	.156*
DS	qua-2	0.1	-0.038	.257**
	qua-3	0.1	-0.04	.274**
	qua-4	0.099	-0.042	.276**
PA	nei	0.074	-0.015	.255**
	ngr	0.072	-0.036	0.093

*$p < .05$, **$p < .01$

Note. RP: features based on randomly sampled positions; RC: features of rows and columns; DS: features inspired by classical image data structure; PA: features inspired by classical image processing algorithms

Second, we compared two general models—based on human-coded features and machine-extracted features—by conducting multiple regressions predicting the observed dissimilarity matrix (MRM; Lichstein, 2007) of each participant group by a linear combination of the model-derived distance matrices generated by features in each set.

These regression analyses generally confirmed the results from the simple correlation analyses. First, none of the Chinese character dissimilarity matrices produced by the three participant groups could be significantly predicted by the linear combination of five distance matrices generated by high-level human-defined features of Chinese characters. In contrast, the linear combination of 14 distance matrices generated by the low-level machine-extracted features significantly predicted dissimilarity matrices from Chinese novices (R^2 = .320, p < .01) with *pyx-3* and *nei* being significant predictors (see Table 3).

Table 3: Multiple regressions between dissimilarity matrices of participant groups and distance matrices generated by low-level machine-extracted features.

	Native Readers	Learners	Novices
		Estimate	
Intercept	40.843	74.234	6.743
pix-3	1.288	1.272	0.256
pix-5	-0.743	-0.806	-0.760
pix-7	-0.323	-0.115	0.427
pyx-3	1.646	-0.698	1.804*
pyx-5	-2.034	-2.742	-2.062
pyx-7	0.219	3.001	0.718
col	-0.122	-0.136	0.043
lin	0.241	-0.160	0.241
lincol	-0.195	0.290	-0.336
nei	0.143	-0.369	-1.033
ngr	-0.063	1.132	3.301
qua-2	-0.063	-0.835	-2.132
qua-3	0.234	0.001	0.371**
qua-4	0.098	-0.062	0.052
R^2	.178	.079	.320**

*$p < .05$, **$p < .01$

Extended Tree Analysis

Because the machine-extracted and human-coded features account for only a modest proportion of the variance in the three participant groups' sorting data, it can be concluded that other features underlie the assessment of visual similarity for Chinese characters. Thus, we undertook exploratory analyses of the similarity-sort data, submitting the summed and transformed dissimilarity matrix for each group for analysis by the EXTREE program (Corter & Tversky, 1986). The EXTREE model can identify both hierarchical or nested feature structures, like any hierarchical clustering algorithm, but also identifies additional features that "cut across" the tree hierarchy, allowing the visual display of more general overlapping cluster structures (Corter, 2023). EXTREE represents non-nested feature sets by adding "marked" or labeled extra features to the branches of an additive tree to represent features that cannot be accommodated by the tree structure per se. Such a non-hierarchical clustering method is needed here where the stimulus set is constructed in a type of factorial design, creating "crossed" features. Note that the EXTREE algorithm first estimates the best-fitting additive tree for the dissimilarity data. The additive tree generalizes the ultrametric tree fit by most hierarchical clustering algorithms by including the estimation of stimulus-specific weights for the "unique features" of each object (Sattath & Tversky, 1987; Corter, 1996).

The EXTREE analyses yielded extended tree diagrams of data from the three participant groups. These extended tree solutions provide a direct visual display of how people with different proficiency of Chinese perceive the similarity of Chinese characters. The description and discussion of the three tree diagrams below are organized as follows: first, we walk readers through the tree diagram branch by branch ignoring added features; second, we point out the added marked features and suggest what these features might mean; third, we compare the three trees for the three expertise groups and summarize their characteristics.

Figure 3: The extended tree of Chinese native readers

Figure 3 displayed the extended tree diagram produced by the EXTREE program to account for Chinese native readers' perceived similarity of Chinese characters. There were three main branches. The first branch contains all four left-right characters—啄, 曉, 偌, and 排. The second branch contained two boxed characters—問 and 曆—nested under the same sub-branch and character 曇 and two foils—爾 and 幾—under another sub-branch. The third branch again contained two boxed characters—匿 and 痱—nested under the same sub-branch and three above-below characters—售, 菲, and 惹—and two foils—重 and 事—under another sub-branch. The plain additive tree accounted for 74.0% of the variance (R^2) in the original data.

On top of the tree structure, eight marked features labeled as C [排, 痱, 菲], D [排, 痱, 菲, 售], E [曉, 曇, 問, 曆, 爾, 幾], H [曉, 曇, 問, 曆, 爾, 幾, 啄], I [啄, 偌, 排], N [曆, 匿, 痱], O [曆, 匿, 痱, 問], and U [曇, 售, 惹, 菲] on the tree in Figure 3 increased the R^2 to 95.9%. Characters sharing Feature C were all characters in the study containing Component 非. Adding 售 to the Feature C set led to the set of characters with Feature D. This set seemed to reflect a visual feature containing a long vertical line and a couple of short horizontal lines perpendicularly attached to the vertical one (e.g. 丰 or 韭). Features E and H seem to represent crossed strokes, with E emphasizing orthogonal crosses, e.g.＋, and H emphasizing tilted crosses, e.g.×. Feature I, N, O, and U point unambiguously to configurations. All three characters with Feature I are left-right characters. All three characters with Feature N and all four with Feature O are boxed characters. All four characters with Feature U are above-below characters. Although the above-below and boxed configurations were not perfectly reflected in the plain additive tree, the extended features fully recovered the configuration dimension in data collected from Chinese native readers, suggesting that they are strongly influenced by the configuration information in their similarity sorts.

Figure 4 displays the extended tree of Chinese learners. There were three main branches. (Note: the third branch split into two sub-branches right after the first split that gave birth to the three main branches, so it appears to be four main branches in Figure 4.) The first branch contained all characters with Component 日 or 非. The second branch contained one left-right character with Component 口, one above-below character with Component 若, and two foils. The third branch contained the rest two characters with Component 口, the rest two characters with Component 若, and the rest two foils. The R^2 increased from 42.6% to 66.5% when eight features were added to the tree by the EXTREE program—C [排, 事, 菲, 曆, 痱], D [惹, 問], E [菲, 重], H [問, 啄, 惹], I [爾, 曉, 排, 曇], N [菲, 幾], O [匿, 重], and U [爾, 排]. Feature C seemed to capture characters with open parallel lines (e.g. ≡) although 曆 is somehow an exception. Feature D seemed related to a rectangle at the center of the character. Feature E could be about symmetry. Feature I could be related to crosses of strokes (e.g. ＋ and ×). However, Feature H, N, O, and U could not easily be interpreted.

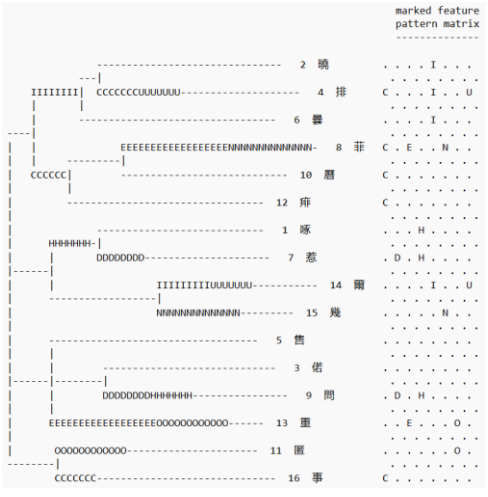

Figure 4: The extended tree of Chinese learners

In the novices' extended tree (see Figure 5), there were two main branches with the first branch containing all four left-right characters, three of the four above-below characters, two of the four boxed characters, and two of the four foils. The other six characters—two with Component 日, one with Component 若, and three foils—are in the second branch. The R^2 increased from 43.8% to 68.0% after eight features

were added to the extended tree by EXTREE—C [售, 事, 曆], D [售, 事, 菲, 痱], E [售, 事, 曇, 曆, 匿, 重], H [痱, 事], I [幾, 爾], N [重, 排, 菲, 痱], O [問, 排], and U [問, 匿]. Feature C, D, E, H, N, and O seemed related to visual elements—Component 非, 扌or 卜, and ≡ that appeared in Feature C and D found in native readers' data and Feature C found in learners' data. Feature I seemed related to stroke crosses—＋and ×—that appeared in Feature E and H found in native readers' data and Feature I found in learners' data. Feature U highlighted the boxed configuration of the two characters.

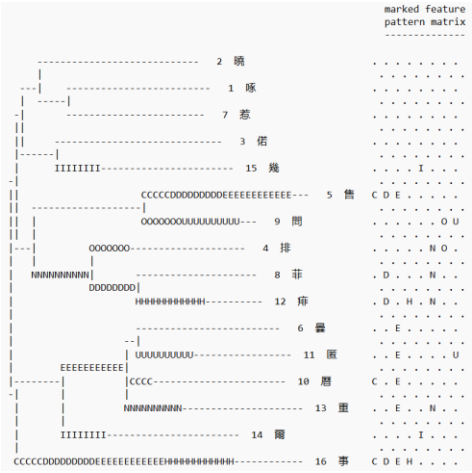

Figure 5: The extended tree of Chinese novices

Comparing the three extended threes with each other, we found that their tree structures had little in common. The number of main branches of the tree diagrams dropped from four for native readers, to three for learners, and to two for novices. While the four branches for native readers corresponded roughly to the three configurations plus one set of foils, the tree structures of the other two proficiency groups showed no priority of configurations or components as their classification criteria. There were only two sub-branches (clusters) that were present in more than one tree, i.e. the 爾-幾 branch in the trees of native readers and of learners and the 啄-惹 branch in the trees of learners and of novices. Trees of native readers and novices shared no sub-branches at all. Such distinctiveness among the three proficiency groups was also quantified by the weak correlations between each two dissimilarity matrices: τ(native, learner) = .06, τ(native, novice)=.10, τ(learner, novice) = -.03. None of these correlations are statistically significant. Nevertheless, despite the heterogeneity among the three groups, we also discovered some visual elements attended to by all three groups, such as 扌, 卜, ≡, ＋, and ×.

Discussion

This study utilizes computational modeling to explore the perception of Chinese characters by people with differing levels of proficiency in the language. We focused on evaluating how well feature models of Chinese character similarity proposed in previous research could predict human judgment of the characters' visual similarity, as measured by a sorting task. The visual features of Chinese characters suggested *post hoc* by EXTREE analysis of the behavioral data were also discussed and compared among the language proficiency groups.

The correlation and regression analyses show that low-level machine-extracted features significantly predict the visual similarity judgments of Chinese characters by Chinese novices, but not the judgments of Chinese second-language learners or native readers. The fact that the significantly predictive low-level features include features based on pixels selected at random suggests that Chinese novices who do not know any Chinese may perceive the characters based mainly on elementary low-level perceptual routines for object recognition and shape perception.

The *post hoc* features emerging in the EXTREE diagrams in the forms of the tree structure and extra marked features suggest that the type of configuration, one of the proposed high-level human-defined features evaluated here, is a primary criterion for judging the similarity of Chinese characters among Chinese native readers. This finding confirms and extends findings in the literature on expertise concerning the influence of high-level abstract and relational features (e.g., configurations) over low-level features (e.g., components) in experts' perception (Chase and Simon, 1973; Gobet, 2005; Vogt & Magnussen, 2007; Slovic, 2016; Yeh & Li, 2002; Yeh et al., 2003).

Although configuration is suggested by the EXTREE solution to be an important high-level feature characterizing Chinese native readers, the correlation and regression analysis do not show a statistically significant effect of configuration, which suggests that a more nuanced model of native readers' perception of Chinese characters may be needed. Furthermore, how Chinese learners classify Chinese characters cannot be explained by either the high-level or the low-level features.

Future research might attempt to expand the set of potential features, either by exploring features manifested by EXTREE *post hoc*, or via automatic search. We note that previous efforts for identifying visual features for Chinese character recognition have not implemented machine learning algorithms trained on large datasets. To discover features with stronger predicting power, we plan to train new models against large open datasets of Chinese characters. With the human-defined and machine-extracted features in this paper and machine-learned features in future research, we hope that progress can be made to better understand expertise development in identifying, reading, and writing Chinese characters.

References

Ashby, F. G., & Perrin, N. A. (1988). Toward a unified theory of similarity and recognition. *Psychological review, 95*(1), 124.

Chase, W. G., & Simon, H. A. (1973). *Perception in chess. Cognitive psychology*, *4*(1), 55-81.

Chen, Y. C., & Yeh, S. L. (2015). Binding radicals in Chinese character recognition: Evidence from repetition blindness. *Journal of Memory and Language, 78*, 47-63.

Coates, D. R., Bernard, J. B., & Chung, S. T. (2019). Feature contingencies when reading letter strings. *Vision research, 156*, 84-95.

Corter, J. E. (1996). *Tree models of similarity and association* (Vol. 7). Sage.

Corter, J. E. (2023). Clustering approaches. In R.J. Tierney, F. Rizvi, & K. Erkican, (Eds.), *International Encyclopedia of Education (Fourth Edition)*, 627-636. Elsevier. https://doi.org/10.1016/B978-0-12-818630-5.10072-7.

Corter, J. E., & Tversky, A. (1986). Extended similarity trees. *Psychometrika, 51*(3), 429-451.

Galili, T. (2015). dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics, 31*(22), 3718-3720.

Geyer, L. H., & DeWald, C. G. (1973). Feature lists and confusion matrices. *Perception & Psychophysics, 14*(3), 471-482.

Gibson, E. J. (1969). *Principles of perceptual learning and development*. New York: Appleton-Century-Crofts.

Hsiao, J. H. W., & Cheung, K. (2011). The modulation of word type frequency on hemispheric lateralization of visual word recognition: Evidence from modeling Chinese character recognition. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 33, No. 33).

Lepage, Y. (2014). Analogies between binary images: Application to Chinese characters. In *Computational Approaches to Analogical Reasoning: Current Trends* (pp. 25-57). Berlin, Heidelberg: Springer Berlin Heidelberg.

Lichstein, J. W. (2007). Multiple regression on distance matrices: a multivariate spatial analysis tool. *Plant Ecology, 188*, 117-131.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological review, 100*(2), 254.

Moret-Tatay, C., Fortea, I. B., & Sevilla, M. D. G. (2020). Challenges and insights for the visual system: Are face and word recognition two sides of the same coin?. *Journal of Neurolinguistics, 56*, 100941.

Pelli, D. G., Burns, C. W., Farell, B., & Moore-Page, D. C. (2006). Feature detection and letter identification. *Vision research, 46*(28), 4646-4674.

Rosenberg, S. & Kim, M. (1975). The method of sorting as a data-gathering procedure in multivariate research. *Multivariate Behavioral Research, 10*, 489-502.

Sattath, S., & Tversky, A. (1987). On the relation between common and distinctive feature models. *Psychological Review, 94*(1), 16.

Slovic, P. (2016). *The perception of risk*. Routledge.

Tversky, A. (1977). Features of similarity. *Psychological review, 84*(4), 327.

Ventura, P., & Cruz, F. (2023). Domain Specificity vs. Domain Generality: The Case of Faces and Words. *Vision, 8*(1), 1.

Vogt, S., & Magnussen, S. (2007). Expertise in pictorial perception: Eye-movement patterns and visual memory in artists and laymen. *Perception, 36*(1), 91-100.

Wong, A. C. N., Bukach, C. M., Hsiao, J., Greenspon, E., Ahern, E., Duan, Y., & Lui, K. F. (2012). Holistic processing as a hallmark of perceptual expertise for nonface categories including Chinese characters. *Journal of Vision, 12*(13), 7-7.

Woodrome, S. E., & Johnson, K. E. (2009). The role of visual discrimination in the learning-to-read process. *Reading and Writing, 22*, 117-131.

Yang, R., & Wang, W. S. Y. (2018). Categorical perception of Chinese characters by simplified and traditional Chinese readers. *Reading and Writing, 31*, 1133-1154.

Yang, Y. Y. (1998). Adaptive recognition of Chinese characters: imitation of psychological process in machine recognition. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 28*(3), 253-265.

Yeh, S. L., & Li, J. L. (2002). Role of structure and component in judgments of visual similarity of Chinese characters. *Journal of Experimental Psychology: Human Perception and Performance, 28*(4), 933.

Yeh, S. L., Li, J. L., Takeuchi, T., Sun, V., & Liu, W. R. (2003). The role of learning experience on the perceptual organization of Chinese characters. *Visual Cognition, 10*(6), 729-764.

Yencken, L., & Baldwin, T. (2006). Modelling the orthographic neighbourhood for Japanese kanji. In *International Conference on Computer Processing of Oriental Languages* (pp. 321-332). Berlin, Heidelberg: Springer Berlin Heidelberg.

Yencken, L., & Baldwin, T. (2008). Measuring and predicting orthographic associations: Modelling the similarity of Japanese kanji. In *Proceedings of the 22nd International Conference on Computational Linguistics (Coling 2008)* (pp. 1041-1048).