The perception of simplified and traditional Chinese characters in the eye of simplified and traditional Chinese readers

Tianyin Liu (kanalty@hku.hk)
Janet Hui-wen Hsiao (jhsiao@hku.hk)

Department of Psychology, University of Hong Kong
604 Knowles Building, Pokfulam Road, Hong Kong SAR

Abstract
Expertise in Chinese character recognition is marked by analytic/reduced holistic processing (Hsiao & Cottrell, 2009), which depends mainly on readers’ writing rather than reading experience (Tso, Au, & Hsiao, 2011). Here we examined whether simplified and traditional Chinese readers process characters differently in terms of holistic processing. When processing characters that are distinctive in the simplified and traditional scripts, we found that simplified Chinese readers were more analytic than traditional Chinese readers in perceiving simplified characters; this effect depended on their writing rather than reading/copying performance. In contrast, the two groups did not differ in holistic processing of traditional characters, regardless of their performance difference in writing/reading traditional characters. When processing characters that are shared in the two scripts, simplified Chinese readers were also more analytic than traditional Chinese readers. These results suggest that simplified Chinese readers may have developed better analytic processing skills than traditional Chinese readers from experiences with simplified characters, and these skills are transferrable to the processing of shared and even traditional characters.

Keywords: Chinese character recognition, holistic processing, reading, writing

Introduction
Chinese characters are the basic writing units in Chinese orthography. They consist of strokes packed into a square configuration, in contrast to words in most alphabetic languages, which are linear combinations of letters. It has been suggested that Chinese characters have many visual similarities with faces (McKeough et al., 2008). For example, both faces and Chinese characters have a homogenous shape, are recognized at the individual level, and learnt in an upright orientation. However, expertise in face recognition is marked by holistic processing, that is, to perceive features of an object as a whole instead of various parts (Gauthier & Tanaka, 2002); in contrast, expertise in recognizing Chinese characters is marked by reduced holistic processing (Hsiao & Cottrell, 2009). This effect may be due to expert Chinese readers’ knowledge about Chinese orthography. Chinese characters are composed of strokes, which combine to form over a thousand different stroke patterns in Chinese orthography (Hsiao & Shillcock, 2006), and stroke patterns are the smallest functional units in Chinese character recognition (Chen, Allport, & Marshall, 1996). For expert Chinese readers, when recognizing Chinese characters, they may be more sensitive to the internal constituent components and have better ability to ignore some unimportant configural information for recognition, such as exact distances between features (Ge, Wang, McKeough, & Lee, 2006) compared with novices (Ho, Ng, & Ng., 2003; Hsiao & Cottrell, 2009). Consequently, expert readers may process Chinese characters less holistically than novices.

There are currently two Chinese writing systems in use in Chinese speaking regions, namely simplified and traditional Chinese. The simplified script was created during the writing reform initiated by the central government of the People’s Republic of China in the 1960s for easing the learning process. Today the majority of Chinese speaking regions including Mainland China, Singapore, and Malaysia use the simplified script, while Hong Kong and Taiwan continue to use the traditional script. The simplification process did not apply to all characters; among the most frequently used 3,500 characters, around 40% were simplified, which have approximately 22.5% fewer strokes than the traditional counterparts (Gao & Kao, 2002); the remaining 60% remained the same in two scripts.

The effects of simplifying the script have aroused some discussion since last decade. For instance, while simplified characters were designed to ease the learning process, many researchers (e.g., Chen, 1999) believe that pure reduction of strokes without standardization of principles may make the characters harder to learn: on one hand, reducing the stroke numbers may make the characters more legible and easier to write for beginners; on the other hand, up to a certain point, characters may become less distinguishable due to high visual similarity among one another as readers’ lexicon size expands (Chen, 1999). Consistent with this speculation, McBride-Chang et al. (2005) recently found that visual skills of children who learned simplified characters were significantly better than those of Hong Kong children who learned the traditional script. Peng, Winett, and Wang’s (2010) ERP data were also consistent with the finding that simplified character readers have greater visual discrimination skills than traditional character readers in perceiving Chinese characters.

Thus, it is possible that the simplification has significantly increased visual similarity among characters in the Chinese lexicon. Simplified characters may differ from one another in fewer strokes than traditional characters. As the ability to identify these diagnostic features is important for recognition (e.g., Oliva & Schyns, 1997), reading simplified characters may involve more analytic/reduced holistic processing than reading traditional characters. Here we aim
to examine whether native traditional and simplified Chinese readers process Chinese characters differently due to the differences in their scripts. We first examine their perception of characters in either the simplified or the traditional form; we predict that simplified Chinese readers will perceive simplified characters less holistically than traditional readers due to their expertise with the simplified script, and vice versa in the perception of traditional characters. In addition to the characters that have different visual forms in the two scripts, around 60% of the most frequently used characters have the same form in the two scripts, and these shared characters provide us a unique opportunity to test the transfer effect of reading simplified/traditional characters. Because both reader groups are experts in reading shared characters, if the two groups differ in the way they perceive/process the shared characters, it will suggest a transfer effect from their experience with the simplified or traditional scripts.

Tso, Hsiao, and Au (2011) recently examined how writing experience influences holistic processing in Chinese character recognition. They recruited proficient Chinese readers who were skilled in both writing and reading (Writers), and those who had limited writing experience (Limited-Writers). They found that Writers perceived Chinese characters less holistically than Limited-Writers, and holistic processing effect was dependent on writing rather than reading performance. Although simplified Chinese readers may still be able to read traditional characters through their similarity with simplified characters or context information, they generally do not know how to write them (and vice versa for traditional Chinese readers). Thus, similar to Limited-Writers, they may perceive characters in their unfamiliar script more holistically, and this effect may depend on their writing rather than reading performance. To verify this hypothesis, we also measure participants’ reading and writing performance in the two scripts and examine whether their (reduced) holistic processing can be predicted by these measures. This study is also the first to investigate holistic processing effects in the two Chinese scripts across two groups of readers.

Methods

Participants

24 native simplified Chinese readers (5 males and 19 females) from Mainland China and 24 native traditional Chinese readers (9 males and 15 females) from Hong Kong participated in the study. They were all skilled writers in their own script: all Mainland participants had passed the Chinese test of National Entrance Examination to college, and all Hong Kong participants had passed the Chinese test of Hong Kong Advanced Level Examination. They were all students at University of Hong Kong; all simplified Chinese readers had resided in Hong Kong for less than 1 year (average length of stay was 9.35 months) by the time they were recruited and had limited exposure to traditional Chinese before. Both groups had similar education background (average years of education, Mainland = 15.46, SE = .37; Hong Kong = 15.38, SE = .44) and similar age (Mainland average = 22.25, SE = .65; Hong Kong average = 22.42, SE = .81). All of them had normal or corrected-to-normal vision and were right-handed as measured by the Edinburgh Handedness Inventory (Oldfield, 1971).

Holistic processing

The complete composite paradigm was used to examine holistic processing effects (Gauthier & Bukach, 2007). The experiment procedure was adopted from Hsiao and Cottrell (2009; Fig. 1). In each trial, participants were presented with a pair of Chinese characters simultaneously, and told to attend to only half of each character and judge whether they were the same or different. In congruent trials, the attended and irrelevant halves of the characters led to the same response (i.e., both were the same or different); in incongruent trials, they led to different responses. The level of holistic processing was assessed by the performance difference between the congruent and incongruent trials.

Materials

The materials consisted of 480 pairs of low to medium frequency (Kwan, 2001) Chinese characters in Ming font, divided equally into three script types: 160 pairs were simplified characters; 160 pairs were the corresponding traditional version of the simplified characters, i.e., having same meanings and pronunciations and differing in orthography; the remaining 160 pairs were characters shared between the two scripts (i.e., shared characters). Characters of different script types were matched in relative character frequency, and the traditional characters were significantly more complex than the simplified ones ($t(159) = 6.17, p < .01$). In each script type, half of the characters had a top-bottom (TB) configuration, and the other half were left-right (LR) structured, and two groups were matched in complexity and frequency. The 80 character pairs in each script type and character configuration combination were further divided into the four conditions in the complete composite paradigm, with 20 pairs in each condition shown in Fig. 1. Each character could be divided into two components, horizontally for TB and vertically for LR characters. In either character configuration condition, the attended halves were matched across congruent and incongruent trials (see Fig. 1a for an illustration), and character frequency and visual complexity were also matched across congruent and incongruent trials.

Design

The design had three within-subject variables: congruency (congruent vs. incongruent), character

---

1. Note that the official written languages used in Hong Kong are English and traditional Chinese, and the official language for instruction at University of Hong Kong is English.

2. We used both TB and LR characters to counterbalance possible influence from character structure; the LR structure is the most dominant structure in Chinese orthography, followed by the TB structure (see, e.g., Hsiao & Shillcock, 2006).
configuration (TB vs. LR), and script type (simplified vs. traditional vs. shared); and a between-subject variable: group (simplified vs. traditional Chinese readers). The dependent variable was discrimination sensitivity measured by A', a bias-free nonparametric measure of sensitivity.

Procedure All characters were shown in low contrast to avoid ceiling effects. In each trial, participants were presented with a central fixation cross for 1000 ms, followed by a symbol indicating which half of the character (top or bottom for TB characters; left or right for LR characters) they should attend to. They were then presented with a pair of characters above and below the initial fixation respectively for 500 ms, followed by a mask (Fig. 1b). Both characters were about 2.5° of visual angle away from the center, each occupying around 1.5° of visual angle. Participants performed a same-or-different judgment task as quickly and accurately as possible with a response box; their accuracy was collected. There were six blocks of test; each block had 80 trials; characters with different configurations or in different script types were presented in different blocks. The sequence of blocks was counterbalanced across participants. A practice session with characters not used in the materials was given before the test.

![Fig. 1: Illustration of stimulus pairs in the complete composite paradigm (a) and trial sequences (b). In (a), the attended components are the bottom halves shaded in grey.](https://example.com/fig1.png)

Reading & Writing Performance Tasks assessing participants’ reading and writing abilities were adopted from Tso et al. (2011). Participants’ reading ability was assessed by a character naming task, in which they named the characters in their mother tongue (i.e., Putonghua for Mainland China participants and Cantonese for Hong Kong participants). Their writing ability was assessed by a character copying and a word dictation task.

Character Copying Participants copied 60 characters (20 shared, 20 simplified, and 20 traditional) as quickly and as accurately as possible. The characters were randomly selected from those used in the character naming task; half of them had a TB structure whereas the other half had a LR configuration. All characters were of medium to high frequency and were matched in relative frequency across script types. The traditional characters had significantly more strokes than their simplified counterparts (t(19) = 8.26, p < .01). In each trial, after a 500 ms central fixation, participants were presented with a character occupying 1.5° of visual angle at the center of the screen, and they were asked to read it out in front of a microphone. Upon their response, the screen turned blank and the experimenter pressed buttons on a response box to record the accuracy and initiate the next trial.

Character Naming The materials consisted of 120 Chinese characters, half with a TB structure and the other half with a LR configuration. In either configuration, 20 characters were simplified characters, 20 were the corresponding traditional version of the simplified characters, and the remaining 20 were shared characters; these characters were not used in the holistic processing task. All characters were of medium to high frequency (Kwan, 2001); they were matched in relative frequency across the script types. The traditional characters had significantly more strokes than the simplified ones (t(39) = 10.92, p < .01). In each trial, after a 500 ms central fixation, participants were presented with a character occupying 1.5° of visual angle at the center of the screen, and they were asked to read it out in front of a microphone. Upon their response, the screen turned blank and the experimenter pressed buttons on a response box to record the accuracy and initiate the next trial.

Word Dictation Task 40 characters (20 shared and 20 traditional/simplified) were selected from the character naming task. Each character was concatenated with a second character to compose a two-character word, and these words were used here. Words instead of characters were used to avoid the ambiguity due to the many homophone characters in Chinese. All words were of medium to high frequency (Taiwan Ministry of Education, 1997) and were matched in relative word frequency across different script types. Participants listened to the words presented in a female voice in their native language respectively, i.e., Cantonese for Hong Kong participants and Mandarin for Mainland participants. The audio recordings of the words were presented by a computer in a random order, and participants wrote down each word in their own script first and then in the other script, even if they thought the characters were the same in the two scripts. If they did not know how to write a character, they indicated it by putting a cross on the space. In each trial, after the words “get ready” presented on the screen for 500 ms, participants were presented with a stimulus; they then pressed buttons on a response box to indicate whether they knew how to write it or not. After
writing the word in both scripts, they pressed a button to indicate completion and start the next trial. Their accuracy of writing the first character of each word was assessed (to match the character naming task).

**Results**

**Reading and Writing Performances**

**Reading Performance** Participants’ character naming performance was summarized in Table 1. Repeated measures ANOVA revealed a main effect of script type in accuracy \(F(2, 92) = 5.40, p < .01\) and response time (RT) \(F(2, 92) = 5.60, p < .01\); and an interaction between group and script type in accuracy \(F(2, 92) = 14.66, p < .01\) and RT \(F(2, 92) = 15.70, p < .01\). Simplified Chinese readers were more accurate in naming shared \((t(46) = 1.51, p < .05)\; \text{although the difference was only 1\%}\) and simplified characters \((t(46) = 2.87, p < .05)\) than traditional Chinese readers, and traditional Chinese readers were more accurate in naming traditional characters \((t(46) = 3.33, p < .01)\; \text{these differences were not significant in RT)}\).

<table>
<thead>
<tr>
<th>Task</th>
<th>Script</th>
<th>Simplified Chinese Readers</th>
<th>Traditional Chinese Readers</th>
<th>Comparison between two groups</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Acc</td>
<td>RT (ms)</td>
<td>Acc</td>
</tr>
<tr>
<td>Nani</td>
<td>Shared</td>
<td>1.00</td>
<td>251.05</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>Simplified</td>
<td>1.00</td>
<td>251.85</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>Traditional</td>
<td>0.97</td>
<td>286.79</td>
<td>0.99</td>
</tr>
<tr>
<td>Copyan</td>
<td>Shared</td>
<td>0.96</td>
<td>7531.16</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>Simplified</td>
<td>0.98</td>
<td>4815.76</td>
<td>0.94</td>
</tr>
</tbody>
</table>
|                       | Traditional| 0.95 | 12251.36 | 0.94 | 6173.10 | ** **
| Dictate               | Shared     | 0.99 | 0.97   |       |       | *  |
|                       | Simplified | 1.00 | 0.66   |       |       | ** **
|                       | Traditional| 0.20 | 0.99   |       |       |       |

Table 1: Summary of participants reading and writing performance

* \(p < .05\), ** \(p < .01\)

**Writing Performance** Participants’ writing performance was summarized in Table 1. In character copying, repeated-measures ANOVA revealed a main effect of script type in both accuracy \(F(2, 92) = 93.40, p < .01\) and RT \(F(2, 92) = 135.28, p < .01\), and an interaction between group and script type in accuracy \(F(2, 92) = 81.76, p < .01\) and RT \(F(2, 92) = 115.38, p < .01\). Traditional Chinese readers were faster \((t(46) = 7.58, p < .01)\) and more accurate \((t(46) = 9.58, p < .01)\) in copying traditional characters, but less accurate in copying simplified characters \((t(46) = -3.58, p < .05)\) than simplified Chinese readers. In contrast, the two groups did not differ in the accuracy or RT of copying shared characters; this suggests that both group had similar level of copying skills in shared characters. In the dictation task, a main effect of script type \(F(2, 92) = 59.15, p < .01\) and an interaction between script type and group \(F(2, 92) = 171.96, p < .01\) were found. Simplified Chinese readers were more accurate in recalling and writing shared \((t(46) = 2.41, p < .05)\) and simplified characters \((t(46) = 6.13, p < .01)\), but were less accurate in recalling and writing traditional characters \((t(43) = 21.86, p < .01)\) than traditional Chinese readers.

**Holistic Processing**

In A’, repeated measures ANOVA revealed main effects of congruency \((F(1, 46) = 70.40, p < .01)\) and script type \((F(1, 46) = 33.79, p < .01)\). Participants did better in congruent \((M = 98, SE = .00)\) than incongruent trials \((M = .94, SE = .01)\), and in processing LR \((M = .97, SE = .02)\) than TB \((M = .95, SE = .02)\) characters. A significant interaction between congruency and group \((F(1, 46) = 6.60, p < .05)\) indicated that traditional Chinese readers were more holistic than simplified Chinese readers in general. There was also a three-way interaction between script type, congruency, and group \((F(2, 92) = 5.027, p < .01)\), suggesting that the interaction between congruency and group was different across the three script types. To investigate how the two groups differed in processing different script types, we first contrasted their difference in processing simplified vs. traditional characters; we then compared their behavior in processing shared characters to examine how their experience in processing simplified/traditional characters influenced their perception of shared characters.

**Simplified vs. Traditional characters** Repeated-measure ANOVAs revealed main effects of congruency \((F(1, 46) = 65.55, p < .01)\) and character configuration \((F(1, 46) = 33.26, p < .01)\). There was an interaction between congruency and group \((F(1, 46) = 5.30, p < .05)\): simplified Chinese readers perceived both characters less holistically overall; and a three-way interaction between script type, congruency, and group \((F(1, 46) = 5.11, p < .05)\), suggesting the interaction between congruency and group was different between the two scripts.

When we examined their performance in processing the two scripts separately, in processing simplified characters, as predicted, the interaction between congruency and group was significant \((F(1, 46) = 5.74, p < .05)\): simplified Chinese readers processed simplified characters less holistically than traditional Chinese readers (Fig. 2a), possibly due to their expertise with simplified characters. Nevertheless, in processing traditional characters, there was no interaction between group and congruency \((p = .76; \text{Fig. 2b)}\). This suggests that the two groups processed traditional characters with a similar level of holistic processing, regardless of their performance difference in reading and writing traditional characters.

Since the two groups differed in some reading and writing performance measures in processing simplified characters (Table 1), to examine whether the difference in holistic processing of simplified characters was dependent on these measures, we put them as covariates in separate ANCOVA tests. Participants’ reading performance in traditional and shared characters could hardly explain the holistic processing difference, because the interaction was still significant if we put their shared character reading accuracy \((F(1, 46) = 5.62, p < .05)\) or RT \((F(1, 46) = 6.08, p < .05)\), or traditional character reading RT \((F(1, 46) = 5.66, p < .05)\) as a covariate separately, and it was marginal when
traditional character reading accuracy was put as a covariate ($F(1, 46) = 3.62, p = .06$). Similarly, when putting either their simplified character naming accuracy ($F(1, 46) = 4.01, p < .05$) or simplified character copying RT ($F(1, 46) = 5.99, p < .05$) as covariates the interaction between congruency and group was still significant. When we put their simplified character reading RT ($F(1, 46) = 3.23, p = .08$) or copying accuracy ($F(1, 46) = 2.71, p = .11$), the interaction between congruency and group became marginal. Only when we put simplified character dictation accuracy ($F(1, 46) = .627, p = .43$) as a covariate did the interaction become insignificant. Furthermore, participants $A'$ difference between congruent and incongruent trials (the measure of holistic processing) in processing simplified Chinese characters was significantly correlated with their simplified character dictation ($r = -.39, p < .05$) but not with reading or copying performances. These results were consistent with Tso et al.’s (2011) finding that the reduced holistic processing effect in expert Chinese character processing may depend more on writing rather than reading or copying performance.

Fig. 2: Performance of simplified and traditional Chinese readers in the holistic processing task with (a) simplified Chinese characters and (b) traditional Chinese characters.

**Shared characters** A main effect of congruency was found ($F(1, 46) = 54.96, p < .01$). There was an interaction between congruency and group ($F(1, 46) = 5.01, p < .05$): simplified Chinese readers were less holistic than traditional Chinese readers (Fig. 3), even though these characters were shared in the two scripts. Since the two groups also differed in some reading/writing performance measures (Table 1), to examine whether the difference in holistic processing was dependent on these measures, they were put as covariates in separate ANCOVA tests. We found that the interaction between congruency and group was still significant when putting shared character naming accuracy ($F(1, 46) = 4.72, p < .05$) or RT ($F(1, 46) = 5.25, p < .05$), copying accuracy ($F(1, 46) = 4.97, p < .05$), or dictation accuracy ($F(1, 46) = 4.74, p < .05$) as a covariate. When putting shared character copying RT ($F(1, 46) = 3.44, p = .07$), as a covariate, the interaction became marginally significant. However, when putting simplified character copying accuracy as a covariate, the interaction became very marginal ($F(1, 46) = 2.52, p = .12$), and when putting simplified character dictation accuracy as a covariate ($F(1, 46) = .60, p = .43$), the interaction became insignificant. Also, participants’ shared character holistic processing significantly correlated with simplified character dictation ($r = -.36, p < .05$) but not with any other reading/writing performance measures. These results suggested that the difference in holistic processing of shared characters could not be completely accounted for by their performance difference in reading/writing shared characters; rather, it was dependent on their writing performance of simplified characters. Thus, the holistic processing difference was likely due to a transfer effect from simplified Chinese readers’ expertise with simplified characters.

**Figure 3:** Performance of simplified and traditional Chinese readers in the holistic processing task with shared characters.

**Discussion**

Here we examined whether simplified and traditional Chinese readers processed Chinese characters differently in terms of holistic processing, and whether their writing and reading performance measures could uniquely predict these differences. We found that when processing simplified characters, simplified Chinese readers were less holistic than traditional Chinese readers, and the difference was dependent on their word dictation performance rather than reading or copying performances. This finding is consistent with Tso et al.’s (2001) study, which showed a close relationship between writing experience and reduced holistic processing in Chinese character recognition. In contrast, although simplified Chinese readers performed much worse in both reading and writing traditional characters than traditional Chinese readers, their performance in holistic processing of traditional characters did not differ from traditional Chinese readers. This effect may be because processing simplified characters generally requires more analytic processing due to higher visual similarity among characters compared with traditional characters. Thus, simplified Chinese readers may have developed a better analytic processing skill in reading Chinese characters in general, and it could be more easily transferred to reading traditional characters compared with the generalization from traditional to simplified characters in traditional Chinese readers. This speculation is consistent with the recent finding that simplified Chinese readers have
better visual skills than traditional Chinese readers (e.g., McBride-Chang et al., 2005; Peng et al., 2010).

When processing characters that are the same in the two scripts (shared characters), simplified Chinese readers were also less holistic than traditional Chinese readers, even though both groups were experts in processing shared characters. Although simplified Chinese readers had better dictation performance in shared characters, further analysis showed that this difference in holistic processing was not dependent on their writing abilities of shared characters, but on their writing abilities of simplified characters. These findings further support our hypothesis that recognizing simplified characters requires more analytic processing than recognizing traditional characters, and this enhanced analytic processing skill is transferrable to the processing of characters that are shared in both scripts, and even to traditional characters.

Note however that the enhanced analytic processing in the simplified Chinese readers compared with the traditional Chinese readers here may also be accounted for by the difference in Chinese teaching method adopted in Mainland China and Hong Kong. In Hong Kong, children learn to read and write Chinese mainly through rote repetition, whereas in Mainland, children are taught explicitly about character components. As the result, simplified Chinese readers in Mainland may be generally more sensitive to the internal constituent components of characters than traditional Chinese readers in Hong Kong (McBride-Chang et al., 2005). To rule out the influence from the teaching methods in accounting for the current results, future work will examine whether similar effects can be found in traditional Chinese readers in Taiwan, where Children are also taught about character components explicitly.

In conclusion, here we show that expertise in reading and writing simplified Chinese characters equips readers with better analytic processing skills that are transferrable to the processing of shared and even traditional characters.

Acknowledgments
We are grateful to the Research Grant Council of Hong Kong (project code: HKU 745210H to J.H. Hsiao).

References