Differential effects of script system acquisition and social immersion experience on face perception: Evidence from event-related brain potentials

Dissertation

Zur Erlangung des akademischen Grades Doktor rerum naturalium (Dr. rer. nat.)

im Fach Psychologie

Eingereicht an der
Lebenswissenschaftlichen Fakultät der Humboldt-Universität zu Berlin

von

M. A. Xiaoli Ma

Präsidentin der Humboldt-Universität zu Berlin

Prof. Dr. Julia von Blumenthal

Dekan der Lebenswissenschaftlichen Fakultät

Prof. Dr. Dr. Christian Ulrichs

Gutachter/in: 1. Prof. a. D. Dr. Werner Sommer

2. Prof. Dr. Rasha Abdel Rahman

3. Prof. Dr. Daniel Brandeis

Tag der mündlichen Prüfung: 08.12.2022

Table of Contents

Acknowledgment

During my PhD study, there are so many people I would like to thank for their support. First and foremost, I am extremely grateful to my supervisor, Prof. Dr. Werner Sommer for his invaluable supervision, continuous support, and incredible patience during my PhD study. I have benefited greatly from your immense knowledge and plentiful experience. I would also like to thank Prof. Dr. Urs Maurer for his thoughtful and detailed feedback on my studies. Additionally, I would like to express gratitude to Prof. Dr. Xiaohua Cao for his generous support of my data collection in China.

Secondly, I would like to thank many interns and students who provided help with the data collection of my studies. Special thanks to two master's students, Mrs. Xinran Li and Miss. Jing Kang for their great help with full enthusiasm on my studies as my teammates and words of support and patience as my friends during my PhD life. I would also like to thank my lab colleagues for their technical support.

Last but not least, I would like to express my deepest gratitude to my family, especially my parents and spouse. They gave me unconditional support and continuous encouragement throughout my PhD life. In addition, I would also like to thank my unborn child who gave me high spirits and motivation to complete my PhD thesis in the last stage of my PhD study.

Zusammenfassung

Menschen sind Experten in der Wahrnehmung und Erkennung menschlicher Gesichter, was stark durch die Erfahrungen des täglichen Lebens geprägt ist. Die menschliche Expertise bei der Wahrnehmung von Gesichtern wird auf eine konfigurative Verarbeitung (konfigurative Prozesse) zurückgeführt. In elektrophysiologischen Studien wurde ein neuronaler Marker für Gesichtsexpertise identifiziert - die okzipito-temporale N170-Komponente des Ereigniskorrelierten Hirnpotenzials (EKP). Vergrößerungen der N170-Komponente wurden mit der Wahrnehmungsexpertise für Stimuli und ihrer konfiguralen Verarbeitung in Verbindung gebracht. Interessanterweise fand ich in einer Reihe neuerer kulturübergreifender ERP-Studien aus westlichen Labors ein konsistentes Muster, dass ostasiatische (vor allem chinesische und japanische) Erwachsene eine größere N170-Aktivierung auf das Gesicht zeigen als kaukasische Erwachsene. Obwohl diese Beobachtungen nicht schlüssig sind, bietet ihre Konsistenz über die Studien hinweg einen wichtigen Anhaltspunkt, um die Auswirkungen der zugrundeliegenden Faktoren auf die Gesichtswahrnehmung zu klären. Bei einer gründlichen Durchsicht dieser kulturübergreifenden Studien wurden zwei potenzielle Faktoren identifiziert, die die erworbene Erfahrung betonen: der Erwerb des Schriftsystems und die Erfahrung der sozialen Immersion. In Bezug auf den Faktor Schriftsystem habe ich eine Schriftsystemhypothese vorgeschlagen. Sie geht davon aus, dass die sehr unterschiedlichen visuellen Anforderungen beim Lesen von alphabetischen Schriften, die von westlichen Kaukasiern häufig erworben werden, und logographischen Schriften, die von Ostasiaten häufig beherrscht werden, die visuelle Gesichtsverarbeitung unterschiedlich beeinflussen, was zu den berichteten Unterschieden in der N170 zwischen westlichen Kaukasiern und Ostasiaten führt. Andererseits wurde eine soziale Immersionshypothese in Bezug auf den sozialen Erfahrungsfaktor vorgeschlagen, weil in diesen kulturübergreifenden Studien die Ostasiaten Ausländer waren, die wahrscheinlich einer relativ großen Anzahl neuer Gesichter in einem

ungewohnten sozialen Umfeld ausgesetzt waren, während die Kaukasier Einheimische in der Region waren, in der die Studien durchgeführt wurden. Es wird angenommen, dass die erhebliche Zunahme der Exposition durch unbekannte Gesichter während des Eintauchens in ein neues soziales und/oder ethnisches Umfeld das System zur Verarbeitung von Gesichtern stimulieren oder sogar überaktivieren könnte, was zu erhöhten N170-Amplituden bei Gesichtern führt.

In meiner Dissertation soll daher untersucht werden, ob die durch die N170-Komponente indizierte Gesichtsexpertise durch den Erwerb des Schriftsystems und/oder soziale Immersionserfahrung moduliert wird. Genauer gesagt, wurden in dieser Arbeit drei Fragen untersucht: 1) Inwieweit erklärt die Schriftsystemhypothese bzw. die Hypothese der sozialen Immersion die in früheren kulturübergreifenden Studien beobachteten Unterschiede in der N170-Gesichtswahrnehmung zwischen ostasiatischen und kaukasischen Erwachsenen? 2) Inwieweit beeinflusst das Schriftsystem die Gesichtswahrnehmung bei Leseanfängern und Leseexperten und 3) sind diese Effekte auf die Gesichter-N170 mit konfiguraler Verarbeitung assoziiert? Um diese Fragen zu beantworten, werden zwei kulturübergreifende ERP-Studien durchgeführt.

Die erste Studie (Studie 1) untersucht die Hypothese des Schriftsystems und der sozialen Immersion, indem sie die beobachteten N170-Unterschiede aus früheren kulturübergreifenden Studien bei jungen Erwachsenen überprüft. In zwei getrennten Experimenten in Hongkong und Berlin wurden EKPs in einer Reihe von One-back Aufgaben aufgezeichnet, wobei Gesichter von chinesischen und kaukasischen Personen in aufrechter und invertierter Orientierung sowie Kritzelbilder als Stimuli verwendet wurden. Im Hongkong Experiment wurden 18 chinesische Ortsansässige und 18 ausländische Gaststudenten, die keine logografische Schrift gelernt hatten, getestet. Im Berliner Experiment wurden 32 Langzeit- und 29 Kurzzeit-Einwohner Berlins, die alle Deutsche waren und keine logografische Schrift lesen konnten, sowie 32 ausländische chinesische Besucher getestet. In beiden Experimenten zeigten Ausländer signifikant größere N170-Amplituden bei Gesichtern, unabhängig von der ethnischen Zugehörigkeit der Stimuli, als Einheimische. Darüber hinaus unterschieden sich die N170-Amplituden von deutschen Berlinern, die erst vor kurzem nach Berlin gekommen waren, nicht von denen von Langzeit-Berlinern. Insgesamt unterstützen die Ergebnisse der Studie 1 die Hypothese der sozialen Immersion, die darauf hindeutet, dass die intensive Konfrontation mit neuen Gesichtern anderer Ethnien während des Eintauchens in eine fremde Kultur die neuronale Reaktion auf Gesichter verstärken könnte. Darüber hinaus wurden die beobachteten Gruppenunterschiede nicht durch die Gesichtsinversion moduliert, was darauf hindeutet, dass die konfigurale Gesichtsverarbeitung nicht durch die Erfahrung der sozialen Immersion beeinflusst wurde.

Die Studie 1 kann jedoch aus zwei Gründen unterschiedliche Effekte des Schriftsystems bei der Gesichtswahrnehmung nicht völlig ausschließen. Erstens sind viele Erwachsene, vor allem mit höherer Bildung, nach dem ersten Leseerwerb mit zusätzlichen Schriftsystemen vertraut, was die Effekte möglicherweise verwischt. Zweitens hat sich gezeigt, dass sich die Schriftkenntnis während des frühen Leseerwerbs, insbesondere in den ersten beiden Jahren des Lesetrainings, dramatisch entwickelt, was darauf hindeutet, dass die Auswirkungen der Schriftkenntnis auf die Gesichtswahrnehmung bei frühen Lesern und erwachsenen Experten unterschiedlich sein könnte. Daher wurde eine Folgestudie (Studie 2) durchgeführt, um die Schriftsystemhypothese bei Kindern zu untersuchen, die etwa ein Jahr lang formelle Leseanweisungen (hauptsächlich) nur in ihrem eigenen muttersprachlichen Schriftsystem erhalten hatten. In Studie 2 wurden zwei getrennte Experimente in Jinhua, China, und Berlin, Deutschland, durchgeführt. Dabei wurden EKPs in einer Reihe von One-back-Aufgaben mit naturalistischen Gesichtern, schwarz-weißen Mooney-Gesichtern und Kritzelbildern sowie in einer Adaptationsaufgabe mit Paaren von Gesichtern, die entweder identisch waren oder sich

in den räumlichen Beziehungen zweiter-Ordnung der Gesichtsmerkmale unterschieden, aufgezeichnet. 28 deutsche und 27 chinesische Kinder der zweiten Schul-Klasse im Alter von 7-8 Jahren wurden untersucht. Die Ergebnisse zeigten größere N170-Amplituden auf naturalistischen Gesichtern bei chinesischen Kindern als deutschen Kindern, was die Skript-System-Hypothese stützt, wonach das umfangreiche Training mit dem hochkomplexen logographischen Skript-System, das chinesische Kinder erlernt haben, die neurokognitive Verarbeitung von Gesichtern tatsächlich verbessert. Darüber hinaus wurde ein ähnlicher Gruppenunterschied auch für Mooney-Gesichter beobachtet, nicht aber für den Adaptationseffekt von räumlichen Beziehungen zweiter Ordnung, was darauf hindeutet, dass das Schriftsystem eine spezifische Art der konfiguralen Verarbeitung - holistische Verarbeitung - beeinflusst, nicht aber die andere Art - Sensibilität für räumliche Beziehungen der zweiter-Ordnung.

Zusammenfassend ist die vorliegende Arbeit die erste, die zeigt, dass die Gesichtswahrnehmung, die durch die N170 des EKP indiziert wird, durch die Faktoren Erwerb des Schriftsystems und soziale Immersionserfahrung beeinflusst wird, was auf eine hohe Plastizität des Gesichtsverarbeitungssystems des Gehirns hindeutet. Im Vergleich zu den Effekten sozialer Immersionserfahrung sind die Effekte des Schriftsystemerwerbs jedoch durch das Stadium der Leseentwicklung eingeschränkt und spiegeln sich in der konfigurativen Verarbeitung wider, was neue Einblicke in das Verständnis der Rolle und der Natur verschiedener Aspekte der Erfahrung in der Entwicklung der Expertise für die Gesichtsverarbeitung über die Lebensspanne ermöglicht.

Schlüsselwörter: Gesichtswahrnehmung; Wahrnehmungsexpertise; konfigurale Verarbeitung; holistische Verarbeitung; Beziehungen der zweiter-Ordnung; N170; EKPs; Lesen; Schriftsystem; Chinesisch; Deutsch; soziale Immersion

Abstract

Humans are experts at perceiving and recognizing human faces, which is heavily shaped by daily life exposure experience. The human expertise in face perception has been attributed to configural processing. Electrophysiological studies identified a neural marker of face expertise - the occipito-temporal N170 component of the event-related brain potential (ERP). Enhancements of N170 responses have been linked to perceptual expertise with stimuli and configural processing. Interestingly, I found that several recent cross-cultural ERP studies conducted in Western labs manifest a consistent pattern in which East Asian (mainly Chinese and Japanese) adults show a larger face-elicited N170 than Caucasian adults. Although this observation is suggestive and inconclusive, its consistency across studies provides an important clue in elucidating the effects of underlying factors on face perception. Two potential factors emphasizing the acquired experience, the script system acquisition, and the social immersion experience, were identified by a thorough review of these cross-cultural studies. In regard to the script system factor, I proposed a script system hypothesis. It assumes that highly different visual demands involved in reading alphabetic scripts commonly acquired by Western Caucasians versus logographic script frequently mastered by East Asians would differentially influence visual face processing, resulting in the reported differences in face N170 between Western Caucasians and East-Asians. On the other hand, a social immersion hypothesis regarding the social experience factor was proposed because in these cross-cultural studies the East-Asians were foreigners who likely had been exposed to a relatively larger number of new faces in an unfamiliar social environment whereas Caucasians were locals in the region where the studies were conducted. It presumes that the substantial increase of unfamiliar face exposure during immersion into a new social and/or ethnic environment might stimulate or even over-activate the face processing system, resulting in enhanced N170 amplitudes to faces.

Hence, my thesis aims to investigate if the face expertise indexed by the N170 component is modulated by script system acquisition and/or social immersion experience. More specifically, the present thesis addressed three questions: 1) to what extent does the script system hypothesis vs. social immersion hypothesis account for the observed face N170 differences between East-Asian and Caucasian adults in previous cross-cultural studies; 2) to what extent does the script system influence face perception in early readers and expert readers, and 3) whether and how the effects in the face N170 are associated with configural processing. To answer these questions two cross-cultural ERP studies are conducted.

The first study (Study 1) investigates the script system and social immersion hypothesis by reassessing the observed N170 differences from previous cross-cultural studies in young adults. Two separate experiments were conducted in Hong Kong and Berlin by recording ERPs in a series of one-back tasks, using faces of Chinese and Caucasian individuals in upright and inverted orientation and doodle drawings as stimuli. In the Hong Kong experiment, 18 local Chinese residents and 18 foreign guest students who had not learned logographic script were tested. In the Berlin experiment, 32 long-term and 29 short-term Berlin residents who were local Germans and could not read logographic script, and 32 foreign Chinese students were tested. In both experiments, foreigners showed significantly larger N170 amplitudes to faces, regardless of ethnicity, than locals. Furthermore, short-term Berlin residents who moved to Berlin recently did not differ in N170 amplitude from long-term Berlin residents who had lived in Berlin for a long time. Together, the findings of Study 1 support the social immersion hypothesis, indicating that the extensive confrontation with novel other-ethnicity faces during immersion in a foreign culture might enhance the neural response to faces. In addition, the observed group differences were not modulated by the face inversion, suggesting that the configural face processing was not influenced by the social immersion experience.

However, Study 1 cannot completely rule out differential effects of the script system in face perception due to two reasons. First, many adults, especially those with higher education have substantial experience with additional script systems after the initial reading acquisition, possibly blurring the effects. Second, it has been demonstrated that script expertise develops dramatically during early reading acquisition, especially in the first two years of reading training, suggesting the effects of script reading experience on face perception might be different in early readers and adult expert readers. Therefore, a follow-up study (Study 2) was conducted to investigate the script system hypothesis in children who had received about one year of formal reading instructions in only their native script system. In study 2, two separate experiments were conducted in Jinhua, China, and Berlin, Germany by recording ERPs in a series of one-back tasks with naturalistic faces, two-tone Mooney faces, and doodles, and in an adaptation task with pairs of faces that were either identical or differed in their second-order spatial relations of facial features. Twenty-eight German and 27 Chinese second-grade children (mainly) aged 7-8 years were assessed. The results revealed that Chinese children showed larger N170 amplitudes to naturalistic faces than German children, supporting the script system hypothesis in which extensive training with the highly complex logographic script system learned by Chinese children may indeed enhance the neurocognitive processing of faces. In addition, a similar group difference was also observed for Mooney faces but not for the adaptation effect of second-order spatial relations, suggesting that the script system affects one specific type of configural processing – holistic processing – but not the other type – sensitivity to second-order spatial relations.

Overall, the present thesis is the first to show face perception indexed by the N170 of ERP is influenced by factors the script system acquisition and social immersion experience, suggesting a high plasticity of the brain’s face processing system. Nevertheless, as compared to the effects of the social immersion experience, the effects of script system acquisition are

constrained by the stage of reading development and reflected in configural processing, providing novel insights into understanding the role and nature of different aspects of experience in the development of expert face processing across the lifespan.

Keywords: face perception; perceptual expertise; configural processing; holistic processing; second-order relations; N170; ERPs; reading; script system; Chinese; German; social immersion

Chapter 1 General Introduction

1.1 Overview of Face Perception

Human faces are unique social stimuli containing rich information about the people around us, such as identity, emotions, age, gender, ethnic background, and many other attributes about an individual which can be derived from a face. In recent decades, the field of face perception has attracted substantial attention.

1.1.1 Underlying Processing: Configural Processing

Human beings show remarkable face recognition abilities and can rapidly discriminate highly similar faces at a glance. It has been commonly demonstrated that the high efficiency in discriminating and recognizing faces in human beings is attributed to the use of configural processing strategies, which is typically defined as the ability to process forms that involve perceiving relations among the features of a compound stimulus (Maurer et al., 2002). It is commonly contrasted with analytic processing. Nevertheless, in the face perception literature the term "configural processing" has sometimes been used interchangeably with the term "holistic processing" (Tanaka & Gordon, et al., 2011). In the present thesis, I adopted a three-level definition of configural processing made in the review by Maurer and colleagues as a framework. The configural face perception is not a unitary construct, it has been suggested to include three types of processing (see Maurer et al., 2002 for a review): i) sensitivity to first-order relations, in which basic relationships between features are taken into account (e.g. detecting a face because its features conform to a standard arrangement in which two eyes are located above, and a nose is in turn located above a mouth, etc.); ii) holistic processing, in which features are bound together into a "Gestalt"; and iii) sensitivity to second-order relations, in which metric distances between features (e.g., the distance between the eyes) are perceived and used for discrimination. They can be measured or manipulated separately based on different experimental manipulations (see Figure 1). The first-order relations processing can be measured by schematic faces (see Figure 1A) that maintain general spatial configurations of faces but do not have recognizable individual facial features. Thus, schematic faces are suggested to predominantly employ configural processing by allowing the retrieval of first-order relational information (Latinus & Taylor, 2006). The holistic processing can be detected by two-tone Mooney faces (Mooney, 1957, see Figure 1B) that include neither individually recognizable facial components nor first-order configural arrangement of facial features. Therefore, processing Mooney faces as faces have been demonstrated to primarily involve holistic processing (Latinus & Taylor, 2005, 2006). To demonstrate the role of second-order relations in face perception, the distances among facial features are manipulated which contains the very little effect on individual facial features (e.g., decreasing the distance between the eyes by moving them closer together, see Figure 1C, Mondloch et al., 2002).

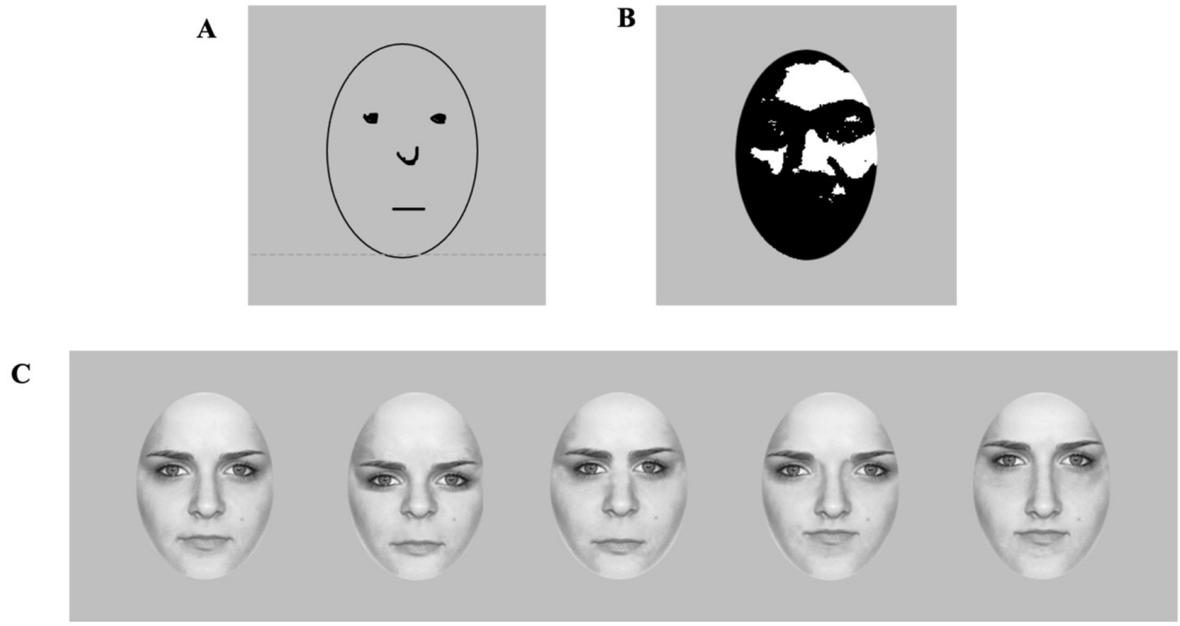


Figure 1. Different stimuli manipulations for separability of three types of configural processing. (A) Sensitivity to first-order relations: A schematic face without individual facial features but retaining first-order relational information, which can be recognized by allowing the retrieval of first-order relational information. (B) Holistic processing: A two-tone Mooney

face contains neither individually recognizable facial features nor the first-order configural arrangement of facial components, which can be recognized as a face due to the involvement of holistic processing. (C) Sensitivity to second-order relations: faces that are manipulated in the spatial relations between facial features, such as decreasing distances between eyes, are used to investigate the role of second-order relational information in face processing.

The most compelling empirical evidence for the configural processing of faces is the face inversion effect where the perception and recognition performance decrease when faces are shown upside down and this effect is much larger for faces than for non-face objects (Yin, 1969). The effect of face inversion is extremely robust, and has been observed in a variety of tasks (e.g., matching or old/new recognition tasks) and stimuli conditions (familiar or unfamiliar faces) (see Rossion, 2008 for a review). Since the low-level properties (e.g., luminance, spatial frequency, contrast) in inverted faces are the same as in upright faces, the impaired performance for the inverted face cannot be due to stimulus properties per se but to differences in the cognitive processes exerted in the recognition of upright as compared to inverted faces. However, whether the inversion of faces leads to a qualitative or quantitative change in face processing is still under debate. According to the qualitative view, researchers argue that the processing of inverted faces is qualitatively different from the processing of upright faces, namely inversion disrupts performance disproportionately in the processing of different facial information (e.g., configural vs. featural information; see Rossion et al, 2008 for a review). By contrast, the quantitative view assumes that upright and inverted faces are processed in the same way, but that configural information is extracted more efficiently in the former than in the later (Willenbockel et al., 2010; Yovel & Duchaine, 2006; Yovel & Kanwisher, 2004). Although the mechanism underlying the inversion is left unspecified, it is no doubt that upside-down face orientation disrupts configural processing and has been demonstrated to interfere with all three types of configural processing (Civile et al., 2016;

Mondloch et al., 2002; Robbins & McKone, 2003; Tanaka & Farah, 1991; Taubert et al., 2011). Hence, face inversion is a robust and reliable paradigm to measure the configural processing of faces.

1.1.2 Underlying Electrophysiological Neural Mechanism: N170

In recent years, a growing number of studies used a variety of neuroimaging methods to investigate the neural mechanisms underlying face perception. For example, studies using high spatial resolution functional magnetic imaging resonance (fMRI) have revealed several visual regions that respond preferentially to faces (e.g., the so-called fusiform face area [FFA]; Haxby et al., 2000; Kanwisher et al., 1997). Due to the high temporal resolution of noninvasive electrophysiological techniques, mostly event-related potentials (ERPs) provide great contributions to elucidating the underlying neural processes for face perception. A large number of ERP studies have revealed several ERP components that are related to different stages in face processing (e.g., Bentin et al., 1996; Pourtois et al., 2005; Rossion, Dricot, et al., 2000; see Schweinberger & Neumann, 2016 for a review). The most commonly studied face-related ERP component is the negative-going N170, which is occurred around 170 ms after stimulus onset at occipitotemporal electrodes (e.g., Bentin et al., 1996; Bötzel et al., 1995, see Figure 2). Although the N170 is considered to be a face-sensitive ERP component, it does not mean that this component is elicited exclusively in response to faces. The N170 corresponds to the visual N1 component that can be triggered by any visual stimulus. However, the N1 is usually larger in response to faces around 170 ms after stimulus onset than in response to other visual stimuli (e.g., Bentin et al., 1996; Cao et al., 2015; Rossion et al., 2003, see Figure 2). Furthermore, source localization studies indicated that the N170 to faces are generated by face-sensitive brain regions (e.g., FFA) located in occipitotemporal areas (e.g., Bötzel et al., 1995; Rossion et al., 2003). Thus, the N170 is more commonly used in face perception research than the N1 because of its peak latency and its scalp topography corresponding to face-sensitive brain areas. In short, in the present thesis, I adopt the widely used term N170 in the face processing literature for both face and non-face stimuli.

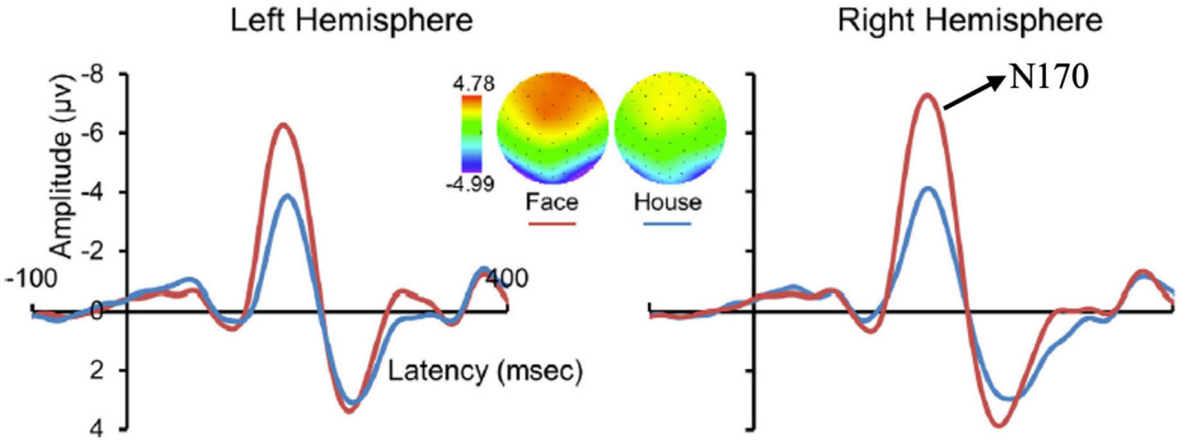


Figure 2. The N170 ERP component recorded to pictures of faces and houses (figure adapted from Cao et al., 2015). Left and right panel: Grand-average waveforms for both faces and houses between -100 and 400 ms following stimulus onset from the left and right occipito-temporal sites separately. Middle panel: topographies showing the bilateral negative component (N170) at occipito-temporal sites for both faces and houses. The N170 is larger for faces than for houses. Note that the negative is plotted upward here.

The N170 has been suggested to be an electrophysiological marker for the structural encoding stage of face perception (e.g., Bentin et al., 1996; Eimer, 2000); in this stage visual representations of facial structures are generated prior to face recognition according to the influential face processing model proposed by Bruce and Young (1986). Specifically, the N170 has been demonstrated to reflect the face detection stage that involves basic-level categorization of the stimulus as a face (e.g., Bentin et al., 1996; George et al., 2005; Rousselet et al., 2004). For example, the N170 is larger in response to upright two-tone Mooney images that are generally perceived as faces than those in inverted orientation that cannot be recognized as faces (George et al., 2005; Jeffreys, 1993). However, recent observations measuring ERP responses to repeated individual faces indicate that N170 amplitudes are reduced when a face is preceded by the same face as compared to a different preceding face, suggesting individual face encoding stage also occurs in the N170 time window (e.g., Cao et al., 2015; Jacques et al., 2007; Li et al., 2019; Zimmer & Kovács, 2011; see Rossion & Jacques, 2011 for a review). Thus, both face detection and coding of individual faces are suggested to take place within the N170 time window.

The N170 is also considered a neural signature for configural processing of faces. One of the most robust evidence is that the N170 component shows a face inversion effect (see Figure 3), in which the N170 is significantly delayed by about 10 ms and can also be larger in response to inverted faces as compared to upright faces (e.g., Bentin et al., 1996, 2007; Civile et al., 2018; Itier et al., 2006; Rossion, Gauthier, et al., 2000). Furthermore, the N170 has been demonstrated to be linked to all three types of configural processing. Schematic faces were found to elicit a comparable N170 as naturalistic faces, suggesting that this component is sensitive to the first-order configural arrangement of facial features (Eimer, 2011; Latinus & Taylor, 2006; Sagiv & Bentin, 2001). The link between holistic face processing and the N170 is reported by evidence that upright Mooney faces triggering a face-like N170 (Eimer, 2011; Latinus & Taylor, 2005, 2006); furthermore, the N170 in response to inverted Mooney faces was larger in trials where stimuli were detected as faces (George et al., 2005). In contrast to the robust evidence for the associations between the N170 and first-order relations and holistic face processing, whether the N170 is sensitive to second-order relations is still in debate and the corresponding findings are mixed. Some studies found that the N170 is not sensitive to manipulations that change the second-order relations (e.g., Halit et al., 2000; Mercure et al., 2008). On the other hand, other studies indicate that the N170 is associated with the processing of second-order relations (e.g., Caharel et al., 2009; Itz et al., 2014; Kaufmann & Schweinberger, 2008; Scott & Nelson, 2006; Vakli et al., 2014; Zimmer & Kovács, 2011). The discrepancy across these studies might be due to the manipulations of stimuli and/or paradigms used. For example, some experiments found that the N170 is larger for faces of caricatures that exaggerate metric distances of facial features from each other (i.e., second-order relations) in comparison to veridical faces, suggesting that the N170 is modulated by the manipulations of second-order spatial information (Itz et al., 2016; Kaufmann & Schweinberger, 2008; Schulz et al., 2012). Moreover, studies using the more sensitive neural adaptation paradigm also support the sensitivity to second-order relations in the N170 (Scott & Nelson, 2006; Vakli et al., 2014). In sum, suitable stimuli and paradigms should be considered when investigating how exactly the N170 is linked to configural face processing.

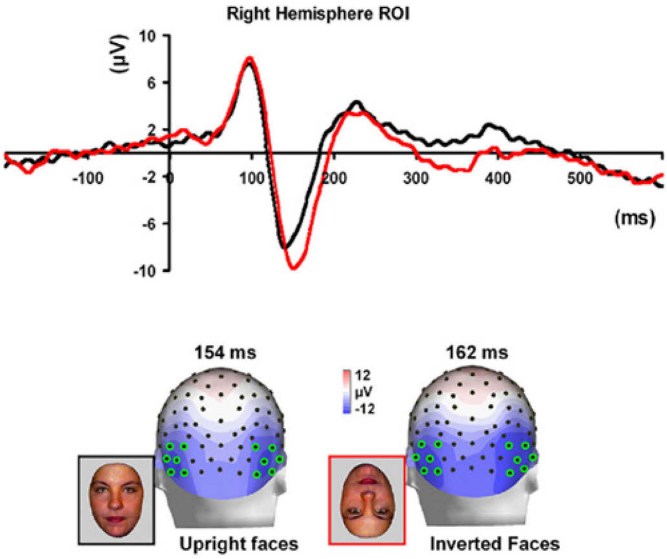


Figure 3. The face inversion effect on the right N170 component for young adults (figure adapted from Alonso Prieto et al., 2011). The N170 is delayed and larger in amplitude to inverted faces than to upright faces. Note that the negative is plotted downward here.

1.1.3 Perceptual Expertise Hypothesis

A fundamental question in the field of face perception is whether face-selective processes should be considered domain-specific mechanisms. With respect to these questions, two distinct hypotheses were proposed and supported by a large body of research (Gauthier & Bukach, 2007; Kanwisher, 2000; McKone et al., 2007; Young & Burton, 2018). One view, the modularity hypothesis, argues that a specialized visual system in the human brain developed selectively and exclusively for processing faces (Kanwisher, 2000). This hypothesis is supported by evidence where the perceptual processing of faces is underpinned by brain systems distinct from those supporting the perception of non-face objects (e.g., Eimer, 2000; Hoehl & Peykarjou, 2012; McKone & Robbins, 2011). By contrast, the perceptual expertise hypothesis posits that the part of our visual system that responds most strongly to faces may respond more broadly to stimuli for which a great deal of experience and exposure has occurred (Young & Burton, 2018). This hypothesis is supported by evidence that the ability to recognize objects of expertise (e.g., cars, birds) does not differ from the ability to recognize faces (e.g., Bukach et al., 2006; Gauthier et al., 1999; Tarr & Gauthier, 2000). To date, the empirical evidence does not conclusively allow to reject either of these hypotheses, in other words, the debate about whether the brain systems involved in face processing are domain-specific or more domain-general is ongoing. Nevertheless, this debate is not the focus of the present thesis, and the perceptual expertise hypothesis would be adopted as the basic theoretical framework of the present dissertation.

The perceptual expertise hypothesis has been confirmed by substantial behavioural (e.g., Gauthier & Tarr, 1997; Searston & Tangen, 2017a, 2017b; Tarr & Cheng, 2003) and neuroimaging (e.g., Gauthier et al., 1999, 2000; McGugin et al., 2015; Tarr & Gauthier, 2000) evidence. Furthermore, two markers of expertise have been identified: the behavioral level - configural processing (e.g., Bukach et al., 2006), and the neural level - N170 (e.g., Eimer et al., 2011). In the following sections, I will review how these two markers of expertise are influenced by expertise effects.

1.1.3.1 Expertise Effects on Configural Processing

The perceptual expertise hypothesis argues that expertise in face perception implies that faces are distinguished and recognized at a subordinate (i.e., individual) level rather than at a basic level as used in most other visual objects (e.g., cars, birds, and houses; Tarr & Gauthier, 2000). Therefore, configural processing develops as an optimal way to meet the demand for perceptual expertise of faces because individual level processing of faces involves subtle analysis of configural information (Carey & Diamond, 1977; Diamond & Carey, 1986). For example, some evidence indicates that face identification is associated with configural face processing (e.g., Richler et al., 2011). A number of studies indicate that non-face objects of expertise are processed more configurally as compared to objects of non-expertise, suggesting that the acquired perceptual expertise is directly associated with enhanced configural processing (Bukach et al., 2006; Busey & Vanderkolk, 2005; Curby et al., 2009; Gauthier et al., 2003). For example, participants received short-term perceptual training to individuate a created laboratory visual stimulus category "Greeble", showing face-like inversion effects for Greebles, and the effects also led to changes in face-sensitive brain areas (e.g., FFA) as compared to novices (e.g., Gauthier & Tarr, 1997; Tarr & Gauthier, 2000). In addition, studies on written words for which people have long-term exposure experience in real life found that literate adults showed configural processing of words, and processed the words from their familiar script system more configurally than those from an unfamiliar script system (Chen et al., 2013; Wong et al., 2011). In sum, the configural processing is a robust behavioral marker of expertise, which can be developed with substantial perceptual experience not only for faces but also for non-face objects of expertise.

1.1.3.2 Expertise Effects on the N170

The N170 ERP component has been suggested to be associated with perceptual expertise (Eimer, 2011; Rossion & Jacques, 2011). This claim is supported by several lines of evidence. Firstly, participants show a more pronounced N170 to stimuli with which they have extensive experience and expertise compared with stimuli they encounter less frequently (e.g., Busey & Vanderkolk, 2005; Gauthier et al., 2000; Rossion et al., 2002, 2004; Tanaka & Curran, 2001). For instance, Tanaka and Curran (2001) found that the N170 amplitude was enhanced when bird experts categorized objects within their area of expertise (i.e., birds) as compared to other domains (e.g., dogs). Also, short-term perceptual training with specific stimuli (e.g., birds, cars) increases the N170 to those trained stimulus categories (Devillez et al., 2019; Scott et al., 2008; Scott & Nelson, 2006). Secondly, non-face objects of expertise show face-like N170 inversion effects (e.g., Busey & Vanderkolk, 2005; Rossion, Gauthier, et al., 2000), suggesting that the N170 is also linked with configural processing of non-face objects of expertise. For example, participants who received intensive perceptual training with upright greebles showed a larger and delayed N170 in response to inverted as compared to upright greebles (Rossion, Gauthier, et al., 2000). Real-world stimuli, such as Chinese characters also elicit face-like N170 inversion effects (Wang et al., 2011). Thirdly, some studies demonstrated a neural competition in the N170 between faces and objects of visual expertise by using a concurrent stimulation paradigm (e.g., Fan et al., 2015; Rossion et al., 2004, 2007), suggesting that perceiving and recognizing faces and other non-face objects to which people are familiar may share similar neural mechanisms in the visual systems at the early stage of visual processing. For instance, Fan et al. (2015) showed a decrease in the amplitude of the N170 elicited by lateral face stimulation when concurrently processing identifiable Chinese characters presented at the center screen. In sum, the face-selective mechanisms reflected by the N170 component should not be considered domain-specific and merely driven by face processing, but instead, as the face-sensitive mechanism that can also be employed in the processing of non-face objects of expertise.

1.2 The Role of Experience on Face Perception

1.2.1 The Interplay of Nature and Nurture on Face Perception

Nature versus nurture is an extensive (and often futile) debate in several fields of science. Specifically, with respect to face perception, whether it is an innate human ability or an acquired skill is a matter of debate. Nevertheless, like so many nature or nurture debates, a simple binary view of the role of nature versus nurture in face perception is both problematic and unnecessary. Thus, it is easy to accept that there will be a complex interplay between two sources of face perception (Honeycutt, 2019; Leopold & Rhodes, 2010). A large number of evidence have demonstrated that both nature and nurture play a crucial role in face perception (see Park et al., 2009 for a review).

Strong evidence for a genetic influence on face perception has been provided by twin studies (Shakeshaft & Plomin, 2015; Wilmer et al., 2010; Zhu et al., 2010), indicating the substantial heritability of face processing (> 60%). For example, Wilmer et al. (2010) found that the performance of monozygotic twins across face recognition tests was more strongly correlated ($r = .70$) than that of dizygotic twins ($r = .29$). In addition, an fMRI study using twin participants found that the patterns of cortical responses to faces were more similar in monozygotic twins than dizygotic twins, suggesting that the genetics plays a role in both neural function and neuroanatomy (Polk et al., 2007). Further evidence demonstrated that visually inexperienced human newborns and even fetuses show a preference for faces over non-face visual objects with equal complexity (Macchi Cassia et al., 2012; Mondloch et al., 1999; Johnson & Morton, 1991; Reid et al., 2017; Valenza et al., 1996), and the newborns' ability of face recognition was diminished by face inversion (Turati et al., 2004, 2006). In sum, these findings indicate that infants might be born with an innate representation of the structural form of a face, suggesting the contribution of innateness to face perception (Deen et al., 2017; Morton & Johnson, 1991).

Although the above-mentioned findings indicate that face perception is largely determined by genetics, numerous studies have demonstrated that the face perception system retains a great deal of experience-contingent plasticity. It is well known that people show better performance in recognition of their own-race faces than of other-race faces—a robust phenomenon termed "other-race effect" (Malpass & Kravitz, 1969; see Meissner & Brigham, 2001 for a review). This effect indicates the plasticity of the face perception system because it does not depend on race per se, but rather on the cultural experience. Especially, the other-race effect appears already in the first year of life (Anzures et al., 2010, 2013; Balas et al., 2011; Kelly et al., 2007, 2009), and also in adolescence (Golarai et al., 2020; Goodman et al., 2007; Pezdek et al., 2003; Sangrigoli et al., 2005). Moreover, previous research confirms that proficiency in other-race face recognition can develop following relatively brief visual exposure to photographs or videos of other-race faces in children (Anzures et al., 2012; Sangrigoli & de Schonen, 2004) and adults (Hills et al., 2020; Hills & Lewis, 2006; Rhodes et al., 2009; Tanaka & Pierce, 2009). These findings suggest that face processing is regulated by experience across the lifespan. In addition, the neural organizations underlying face processing have also been demonstrated to be influenced by experience, for instance, the brain regions (e.g., FFA) known to be involved in processing faces that can respond to other objects (e.g., birds or cars) with sufficient experience (Gauthier et al., 2000).

To summarize, the current evidence suggests that humans are genetically wired for face perception, but that experience plays a crucial role in fine-tuning the mechanisms underlying face processing and pruning them to the specific demands of one’s environment. In the present thesis, it will be highlighted how experience affects face perception during development or even in adulthood.

1.2.2 An Interesting Observation of Cultural Effects on the Face-Sensitive N170 Component

To what extent naturally occurring environmental variations, such as cultural or ethnic diversity, influences perceptual processes in visual perception has been investigated in the past decades (Nisbett & Masuda, 2003; Nisbett & Miyamoto, 2005). Numerous recent ERP studies have been conducted to understand a cultural-specific effect, i.e., the other-race effect. However, the findings on the sensitivity of race in the face-elicited N170 component were mixed. A number of studies failed to report a sensitivity of the N170 to the race of faces (e.g., Caldara et al., 2003, 2004; Chen et al., 2013; Herzmann et al., 2011; Vizioli et al., 2010, 2011; Wiese et al., 2009). On the other hand, although the race sensitivity in the N170 time window was revealed in some studies, the results are also mixed. Some studies showed that the N170 amplitudes in response to own-race faces is smaller than to other-race faces (e.g., Caharel et al., 2011; Gajewski et al., 2008; Herrmann et al., 2007; Stahl et al., 2008, 2010; Walker et al., 2008; Wiese et al., 2014). However, other studies found that own-race faces elicited a larger N170 than other-race faces (Ito & Urland, 2005; Senholzi & Ito, 2013). The heterogeneity of the N170 race effects might be due to methodological differences across previous studies such as task demand (Senholzi & Ito, 2012; Wiese et al., 2013). In short, these mixed findings cannot provide direct and robust evidence for the modulation of N170 by cultural experience.

Nevertheless, some cross-cultural ERP studies (see Table 1) attracted our attention and interest because they indicate differences between participant groups of different ethnic origins. In these studies, a consistent observation although statistically inconclusive that East-Asian participants show larger N170 amplitudes to faces than Caucasian participants. The consistent cultural differences demonstrated above in the face-elicited N170 responses provide new insights into the effects of underlying factors on face perception. Therefore, it would be meaningful to investigate the mechanisms underlying such differences, specifically, to probe possible theoretical accounts for the causes underlying the cultural differences. At least two accounts emphasizing differential acquired experience across groups were proposed in the present thesis: the script system hypothesis and the social immersion hypothesis. In the following sections, I will present each hypothesis and the corresponding theoretical and empirical background of each.

Table 1 N170 amplitudes in East Asian and Caucasian participants in previous studies

Study	Participants		N170 amp. (μV)		
	E. Asian	Cauc.	E. Asian	Cauc.	Diff.
Dering et al. (2013)[a]	25 (Chin./Japan.)	25 British	-6.0	-3.5	-2.5[*]
Gajewski et al. (2008)[a]	9 Chin., 1 Japan., 2 Kor., 2 Vietn.	14 Germ.	-12.1	-9.3	-2.8[#]
Herzmann et al. (2011)[a]	18 Chin. 3 Japan. 2 Kor., 1 Vietn., 1 Philippino	32 Americ.	-1.4	-0.6	-0.8[n.s.]
Vizioli et al. (2010)[c]	12 Chin.	12 British	-4.4	-3.6	-0.8[n.s.]
Vizioli et al. (2011)[b]	15 Chin.	15 British	-4.9	-4.8	-0.1[n.s.]
Wiese et al. (2014)[a]	17 Chin., 1 Japan. 1 Kor.	20 Germ.	-5.3	-3.5	-1.5[n.s.]

[a] N170 amplitudes were measured at the peak of the grand-average waveforms in figures.

[b] Mean N170 amplitudes were provided in the publication.

[c] Mean N170 amplitudes were provided in the publication.

Diff.: the group difference between East Asian and Caucasian participants (E.Asian minus Cauc.)

[*] The group difference was statistically significant in the studies ($p < 0.05$).

[n.s.] The group difference was not statistically significant in the studies ($p > 0.05$).

[#] The group difference was not analyzed in the studies.

E. Asian = East Asian; Cauc. = Caucasian; Chin. = Chinese; Japan. = Japanese; Kor. = Korean; Vietn. = Vietnamese; Germ. = German; Americ. = American

1.2.2.1 Accounts of the Cultural Effects on the Face N170

1.2.2.1.1 A Cultural-Specific Account: The Script System Hypothesis

Members of East-Asian cultures often require to master a logographic script system, which is much more complex than the alphabetic script systems used in Western cultures. Therefore, acquiring script systems differing in their demands on visual processing might differentially affect face processing, and in turn, result in cross-cultural differences in the face N170. This assumption is termed here the script system hypothesis. This hypothesis is proposed based on a number of theoretical considerations and empirical findings.

The neuronal recycling hypothesis

Reading and writing is human cultural inventions that a majority of people are engaged in during their daily lives. Nevertheless, reading and writing has been invented less than 6000 years ago, which is too recent for dedicated neural networks to evolve specifically in the service of reading. According to the neuronal recycling hypothesis proposed by Dehaene and Cohen (2007), reading "recycles" preexisting cortical networks evolved for different but similar functions (e.g., face perception) to create its own brain networks. This hypothesis is supported by numerous studies investigating the broad effects of reading acquisition on the cognitive (e.g., Dehaene et al., 2015; Huettig et al., 2018; van Paridon et al., 2021) and neural (e.g., Dehaene et al., 2010; Hervais-Adelman et al., 2019; López-Barroso et al., 2020; Szwed et al., 2012) processes that underlie the acquisition of this evolutionarily recent skill.

Recent researchers have begun to focus on the effects of reading acquisition in the processing of human faces because of considerable commonalities between visual word and face processing. Becoming a fluent expert reader takes years of practice through formal schooling (Horn & Manis, 1987; Samuels et al., 1978), and reading-related skills perhaps continue to improve with practice even in adulthood (Ziegler & Goswami, 2005). The final stage of reading development is that individuals can read words fluently as well as automatically (Ehri, 2005; Nathan & Stanovich, 1991), Thus, skilled readers obtain expertise in word recognition based on massive training, similar to acquiring face expertise (e.g., Maurer et al., 2008; McCandliss et al., 2003; Wong et al., 2012). Recent evidence has revealed similarities with respect to the underlying processes and mechanisms for face and word recognition. Firstly, both word and face recognition involves configural encoding of spatial relations between features (e.g., Maurer et al., 2002; Ventura et al., 2019; Wong et al., 2019). Secondly, with regard to the neural underpinnings of face and word processing, two anatomically neighboring sites in the ventral occipital-temporal area, the “visual word form area (VWFA)” and the “fusiform face area (FFA)”, respectively, are held to be selective for these stimuli (Baker et al., 2007; Golarai et al., 2021; Nestor et al., 2013; Pinel et al., 2015; Tarr & Gauthier, 2000). Moreover, both areas have been associated with a pronounced negative-going N170 component in the ERP, elicited by both words and faces (Brem et al., 2006; Henson, 2003; Herrmann et al., 2005). These similarities suggest that the development of faces and word processing might not be independent but co-dependent.

Although a number of empirical studies indicate that face perception is influenced by reading acquisition (e.g., Carreiras et al., 2009; Dehaene et al., 2010; Dehaene-Lambertz et al., 2018; Hervais-Adelman et al., 2019), there is an ongoing debate whether the training of visual skills in reading will lead to a positive transfer to faces or, instead, “destructive competition” with faces (Dehaene et al., 2010; Rossion & Lochy, 2021; van Paridon et al., 2021). With respect to the initial formulation of the neuronal-recycling hypothesis, learning to read involves “destructive competition” during the recycling of the neural networks that were formerly specialized for the processing of other visual categories, such as faces (Dehaene & Cohen, 2007). This notion is supported by several lines of research (e.g., Dehaene et al., 2010; Dundas et al., 2014; see Behrmann & Plaut, 2020; Dehaene et al., 2015 for a review). Firstly, an influential fMRI study compared Portuguese and Brazilian literate, ex-illiterate and illiterate adults. The results indicated that in the left fusiform gyrus the neural responses to words increased whereas to faces they decreased with increasing levels of literacy (Dehaene et al., 2010). Subsequently, Pegado et al. (2014) tested the same participants using EEG technique. The findings indicated that both left-lateralization for words and right-lateralization for faces were increased in proportion to the participants' reading ability, suggesting a competition between words and faces for the cortical space in high-level visual areas (Pegado et al., 2014). Furthermore, a behavioral study from the same team demonstrated that illiterates were more holistic in processing faces than literates, suggesting that literacy induces a reduction of holistic face processing (Ventura et al., 2013). Additionally, in support of the destructive competition hypothesis a number of studies on children suggest that the increase of print knowledge in children is associated with reduced specialized activation for faces in the left fusiform area (Cantlon et al., 2011; Centanni et al., 2018; Dundas et al., 2014). This line of evidence suggests that the competition occurs in typical reading development.

However, a growing number of studies are not in favor of the destructive competition hypothesis but support the converse possibility that neuronal recycling improves related older abilities rather than impaired (van Paridon et al., 2021; see Rossion & Lochy, 2021 for a review). A recent fMRI study found that literacy enhanced neural responses for both words and faces in the left fusiform gyrus when comparing literate adults with illiterate adults (Hervais-Adelman et al., 2019), which is contrary to the findings of Dehaene et al. (2010). A subsequent behavioral study found a positive correlation between literacy scores and face recognition in a memory task by comparing literate adults with illiterate adults (van Paridon et al., 2021). With EEG recordings, Li et al. (2013) found that 5-6 years old children who had better word reading abilities showed larger right-lateralization of the N170 for faces, namely reading acquisition increased the right-lateralization of face processing. Moreover, contrary to Ventura et al. (2013), Cao et al. (2019) showed that Chinese literate adults had greater sensitivity to spatial configurations of faces as compared to illiterates, suggesting that reading acquisition improves the configural processing of faces. These findings suggest that face recognition is enhanced rather than impaired by the acquisition of reading skills, which is not consistent with the destructive competition hypothesis. Arguably, this inconsistency might be due to the different script systems that participants learned. The acquisition of an alphasyllabic script (e.g., Devanagari script) in Hervais-Adelman et al. (2019) and van Paridon et al. (2021), and the logographic Chinese script in Cao et al. (2019) and Li et al. (2013) are more complex at the visuo-spatial level than the alphabetic script (e.g., Portuguese script and English) that participants acquired in studies supporting the destructive competition hypothesis (e.g., Dehaene et al., 2010, Dundas et al., 2014; Ventura et al., 2013). However, recent studies testing alphabetic script readers also failed to support the destructive competition hypothesis, suggesting no absolute costs of reading acquisition for face processing (Dehaene-Lambertz et al., 2018; Kühn et al., 2021; Lochy et al., 2019). Nevertheless, these findings do not strongly support the improvement of face processing with literacy acquisition as found in studies on more complex script systems (e.g., Cao et al., 2019; van Paridon et al., 2021) since they are confounded by the increase of general cognitive abilities with age.

Overall, the current evidence supports the neuronal recycling hypothesis and demonstrates an influence of reading acquisition on face processing. However, whether the consequence of reading acquisition on face processing is an improvement or impairment is still in debate. Furthermore, whether the variability across script systems modulates the impact of reading acquisition on face processing is also unclear and needs to be investigated. In the following part, I will focus on the comparisons of alphabetic and logographic scripts in terms of a variety of characteristics.

The comparisons of alphabetic and logographic script

A script system is a method of visually representing spoken language using written letters or characters. There are a number of script systems that vary greatly with respect to the appearance, number of characters and size of denoted units (DeFrancis, 1989). All script systems use a script made up of basic units called graphemes. According to how the graphemes map onto units of language, there are three basic writing systems: alphabetic, logographic, and syllabic (Gelb, 1952). Here, I focus on the comparisons between alphabetic and logographic script systems in terms of several differences in their characteristics. Firstly, the alphabetic script system (e.g., English, German) recodes spoken language at the level of individual sounds, whereas the logographic script system (e.g., Chinese, Japanese) represents word meanings without recourse to spoken language (Li & Bi, 2022). More specifically, an alphabetic script system in which each grapheme corresponds to a phoneme, such as English, German and French, represents sound mainly at the phoneme level. Thus, alphabetic script readers can pronounce a written word even if they do not know its meaning. By contrast, in logographic scripts such as Chinese and Japanese kanji each grapheme corresponds to a syllabic morpheme and is mainly encoded at the morpheme level. Although about 66% of frequently used logographic characters include one component that represents the pronunciation of the characters to a certain extent (phonetic component), the phonetic component often does not have the same pronunciation as the whole character (Hu et al., 2013). Accordingly, the pronunciation of the character cannot be derived from parts of the character as in (transparent) alphabetic script systems, and thus logographic script readers must memorize the pronunciation of each character (Li et al., 2022).

Processing the visual forms of graphemes is critical for reading acquisition (Perfetti & Hart, 2002). The differences in terms of the visual demands of grapheme processing across script systems have recently received attention (Chang et al., 2016). In particular, the visual demands of grapheme processing across script systems vary in terms of the size of grapheme inventory and the corresponding visual complexity of configurations (Chang et al., 2016). For alphabetic scripts, the grapheme inventory is made up of a small number of letters (average number of graphemes: 30-40 letters counting both upper and lower case). By contrast, the grapheme inventory of logographic scripts contains a much larger number of characters (the average number of graphemes is greater than 3000 characters). Logographic scripts have much more complex configurations of graphemes than alphabetic scripts. Firstly, the visual form of graphemes in logographic scripts is square shaped and has high spatial flexibility inside, differing from the linear, left-right spatial configural of alphabetic words (Chen & Kao, 2002). Secondly, each logographic character contains an average of 10 strokes (Huang & Hanley, 1995), which is larger than strokes per letter (average of 2.5 strokes) in alphabetic words (Changizi & Shimojo, 2005). Moreover, each logographic character is composed of radicals that are usually packed into a complex spatial pattern in which they may appear at different locations and vary in shape and orientation with respect to the others (Mo et al., 2015). For example, radicals of Chinese characters can be arranged in left-right (e.g., [咪]), top-down (e.g., [吕]), or inside–outside (e.g., [回]) configurations.

Considering the set sizes of graphemic elements and the complexity of spatial configurations, logographic and alphabetic scripts might differ strongly in their demands on visual processing and memory. This claim is supported by evidence showing the effects of complexity of visual information on visual memory for objects (e.g., Alvarez & Cavanagh, 2004; Xu & Chun, 2006). More direct evidence where Chinese readers have better visual memory skills since Chinese reading requires memorizing thousands of visually complex characters as compared to alphabetic script reading (see Flaherty, 2003 for a review). Therefore, it is likely to assume that these differences across script systems might give rise to script-specific perceptual and cognitive processing mechanisms during the reading acquisition in a particular writing system. This assumption is in line with the system accommodation hypothesis in which the script system imposes constraints on processing that the brain must accommodate (Perfetti et al., 2007). Indeed, a number of empirical studies found that the right hemispheric visual regions (e.g., right fusiform gyrus) are involved more in processing of Chinese script than that alphabetic script (Bolger et al., 2005; Cao et al., 2010; Ji et al., 2019; Tan et al., 2005). This difference might be due to the more extensive spatial relationships between sub-character components in Chinese as compared to the more linear alphabetic letters. The spatial analysis required in reading Chinese relies strongly on low spatial frequencies defining the radical relationships, which are encoded in right-hemispheric visual regions (Perfetti et al., 2007). By contrast, alphabetic reading mainly relies on high spatial frequencies; for example, the difference between *deer* and *dear* is coded in high spatial frequencies involving grapheme-phoneme conversion; high spatial frequencies are decoded primarily by the left hemisphere (Hsiao & Lam, 2013; Y. Liu & Perfetti, 2003).

In sum, the variability across script systems impacts reading acquisition and its neural mechanisms. According to the neuronal recycling hypothesis, reading acquisition “recycle” older visual functions (Dehaene & Cohen, 2007); thus, differences in script systems might affect non-verbal visual processing.

Effects of script system on non-verbal visual processing

Reading visually more complex scripts has been suggested to require stronger visual skills and may, in turn, strengthen them, as compared to reading relatively less complex scripts (Chang et al., 2016). Evidence for this mainly comes from cross-cultural comparisons. Mann (1985) found that Japanese second graders were better at memorizing abstract visual stimuli than their American peers. Demetriou et al. (2005) found that Chinese children aged 8 years old outperformed Greek children of the same age in visual spatial reasoning. McBride-Chang et al. (2011) reported that children who learned to read Chinese performed better on a standardized visuospatial relationship task than children who learned to read Spanish. Furthermore, a similar difference in a similar visual spatial task was found by comparing Chinese with English dyslexic readers (McBride-Chang et al., 2013). These findings indirectly suggest that learning to read a visually complex script may boost children's visual processing, establishing the importance that visual complexity has on reading development.

The above-mentioned evidence has demonstrated the effects of reading acquisition on non-verbal visual processing; however, whether the effects induced by script system acquisition may be transferred to processing of human faces with sufficient complexity is still unclear. In the following, I will discuss the possible effects of script systems on face processing according to the comparisons of face and scripts.

The comparisons of scripts and faces

The resemblance with respect to visual characteristics is particularly striking between logographic scripts (e.g., Chinese characters) and faces. Firstly, similar to faces, Chinese characters are represented graphically at the individual level (i.e., their identities). Moreover, a face expert can recognize a large number of faces regardless of their facial expressions, similar to a Chinese literate needing to recognize more than 3000 characters regardless of font. Importantly, Chinese characters are suggested to induce similar perceptual expertise effects as faces do (Hsiao & Cottrell, 2009; Tso et al., 2020).

One robust hallmark of face expertise, configural processing, has been suggested to be also involved in processing of Chinese characters (Wong et al., 2012). Chinese characters also contain three types of configural information (i.e., first-order and second-order relations, information that requires holistic processing) which is analogous to those used in processing of faces (D. Maurer et al., 2002; Perfetti & Harris, 2013). For Chinese characters, first-order relational information refers to the relative locations of character components (i.e., radicals) within the horizontal or vertical character structures (e.g., 音-昱; 部-陪). The second-order relational information for Chinese characters is more discriminating at the stroke level, in which modifying the relative length and distances between strokes may result in different radicals/characters (e.g., 工-千；日-曰). In addition, Chinese characters are holistically represented as a whole perceptual unit involving no independent, intermediate representations of radicals (Mo et al., 2015).

In sum, processing Chinese characters relies on various types of configural information, which shows the importance of configural processing in Chinese character recognition. This claim is supported by a growing basis of evidence. Skilled readers of Chinese characters tended to sort Chinese characters based on similarities in global spatial relationships among character components, while novice readers tended to sort Chinese characters according to character components (Yeh et al., 2003; Yeh & Li, 2002). Recent studies consistently found that expert Chinese readers show face-like configural processing when they recognize Chinese characters based on a variety of paradigms (e.g., inversion effect, composite effect) and methods (e.g., behavioural, neuroimaging) frequently used in face research (Chen et al., 2013; Kao et al., 2010; Wang et al., 2011; Wong et al., 2019). For instance, Wang et al. (2011) found that inversion increased and delayed N170 similarly for faces and Chinese characters, suggesting similar neural mechanisms underlying configural processing of Chinese characters and faces. These findings suggest that configural processing is also a marker of perceptual expertise for visual Chinese characters.

According to these similarities between Chinese characters and faces, it is reasonable to assume that learning Chinese script is likely to influence face processing to a larger extent than learning an alphabetic script. More specifically, the substantial training with visually complex logographic Chinese characters might tune the visual system towards configural processing and have a positive transfer to face processing.

However, the script system might influence face processing differentially in different stages of reading acquisition. This notion is supported by evidence that the developmental ERP studies suggest an inverted "U" model of script expertise effects on the N170; perceptual learning appears to be critically important during the first two or three years of reading acquisition and then gradually declines as expertise consolidates (e.g., Brem et al., 2009; Cao et al., 2011; Maurer et al., 2011). For example, Cao et al. (2011) found that the N170 amplitude to Chinese characters in 2nd grade children showed the largest amplitude among all age groups, but thereafter N170 amplitude declined with increasing age. Thus, it is plausible that the effects of reading acquisition on face processing may appear in early readers and then decline and might even disappear in adult expert readers, resulting in differential effects of script system in different stages of reading acquisition. In addition, the developmental pattern of configural processing of Chinese characters is also suggested to be inverted-U shaped: an increase in configural processing due to initial visual experience in early stage of reading acquisition, followed by enhanced sensitivity to local components as a result of writing experience in later stage of reading acquisition (Liao & Hsiao, 2021). In support of this idea, it has been found that Chinese first graders (who were learning to read Chinese but had limited writing experience) showed increased configural processing in Chinese-character recognition compared with non-Chinese first graders who did not receive the local Chinese curriculum; moreover, the configural processing in Chinese children was reduced as they reached higher grades (Tso et al., 2012). Therefore, it is plausible to assume that the training with visually complex logographic Chinese characters in a configural vision might have a positive transfer to configural face processing for early readers but not for expert readers.

1.2.2.1.2 An Environmental-Specific Account: The Social Immersion Hypothesis

Through a thorough review of the cross-cultural studies listed in Table 1, all the East-Asians were foreigners who had lived in the regions for months to years where the studies were conducted, whereas all the Caucasians were mostly residents. Individuals typically engage in a larger variety of activities, related to their occupation and social life, bringing with them a greater variety of social encounters. Oruc et al. (2019) reported that faces have a prominent presence in the everyday visual experience of adults with an average of 255 faces per day. Therefore, the East-Asians tested in the studies listed in Table 1 probably had been exposed to a relatively large number of new faces in an unfamiliar and other-ethnic social environment as compared to the Caucasians. Arguably, the substantial increase of unfamiliar face exposure in East-Asian foreigners, especially in a predominant other-ethnicity environment may stimulate the face processing system and result in enhanced N170 amplitudes to faces. Thus, the cultural differences in the face N170 between East-Asians and Caucasians likely stem from increased experience with a larger number of new faces during the staying in the new social environment for East-Asians as compared to Caucasians. This claim is termed the social immersion hypothesis. In the following, I will present the theoretical and empirical evidence for this hypothesis.

According to the perceptual expertise hypothesis, increasing experience with a specific visual category could produce and/or improve the perceptual expertise for the stimulus category (Gauthier & Tarr, 1997; Tanaka & Curran, 2001). For instance, Tanaka et al. (2005) trained adult participants to discriminate birds and owls within a short-term period. The results indicated that the discrimination of these trained stimuli was improved after training; moreover, the improvements generalized to novel exemplars within the trained category and new, unfamiliar species of birds or owls. Two follow-up ERP studies (Scott et al., 2006, 2008) not only replicated the findings of Tanaka et al. (2005) at the behavioral level, but also found similar training effects at the neural level where the N170 was increased to the trained stimuli, subsequent transfer to untrained stimuli from learned categories and to new stimuli belonging to a novel but structurally related categories. Recent studies provide direct evidence for training effects on face perception (e.g., Corrow et al., 2019; Davies-Thompson et al., 2017). For example, the perceptual sensitivity for faces in individuals suffering from developmental prosopagnosia was improved through several months of intensive perceptual training in face perception, and the improvement generalized to untrained faces (e.g., Davies-Thompson et al., 2017; Corrow et al., 2019). In addition, training of face cognition is also beneficial for unimpaired adults in which the speed of face processing was enhanced after training in easy (speed) tasks although face memory was not affected by training with more difficult face memory tasks (Dolzycka et al., 2014).

In sum, previous findings in adults suggest the remarkable plasticity of adults' visual systems. Therefore, the substantial exposure experience of novel faces in East-Asian adults in the studies listed in Table 1 is likely to fine-tune their face-related visual systems, thus resulting in enhanced face-elicited N170 as compared to Caucasian adults. Nevertheless, for these East-Asian participants, the geographical region where the study was conducted constituted not only a new social environment but also a different ethnic environment. Thus, two types of immersion experience may have to be distinguished: immersion into a different ethnic social environment (new-ethnicity immersion hypothesis) and/or immersion into a new social context (global new social immersion hypothesis).

According to the face space theory proposed by Valentine (1991), the substantial increase of exposure experience to other-race faces may be a critical factor for the social immersion effects on the observed cultural differences in the face N170. Valentine (1991) proposed a multidimensional face space framework in which different faces are encoded as points in an n-dimensional space based on the dimensions used to discriminate between them, and these dimensions are shaped by experience, optimizing our ability to discriminate between faces typically encountered in our environment. This model is largely considered a robust framework for face processing in general. Within this framework, an average face, representing the origin of the multidimensional space, is the outcome of all the faces encountered in life, weighted by the frequency and recency of the encounter. The face space model has been corroborated by a large amount of empirical evidence (e.g., Hills & Lewis, 2006; Jaquet et al., 2007, 2008; Leopold et al., 2001; Loffler et al., 2005; Papesh & Goldinger, 2010). In accordance with the face space model, experience with other-race faces should tune a person's face space; subsequently improving the recognition of other-race faces, and reducing the other-race effect. Indeed, studies on other-race effects suggest that people can significantly improve their recognition performance with other-race faces consequently to naturalistic exposure (e.g., Hancock & Rhodes, 2008; Rhodes et al., 2009; Tanaka et al., 2004) and laboratory intensive perceptual training (e.g., Goldstein & Chance, 1985; Hills & Lewis, 2006; Lebrecht et al., 2009; Tanaka et al., 2013; Tanaka & Pierce, 2009). Furthermore, some studies suggest that the short-term experience in a new ethnic environment influences the neural correlates of face processing (Derntl et al., 2009; Elfenbein & Ambady, 2002; Gajewski et al., 2008). For example, after about one to two years spent in a new ethnic environment, emotion recognition performance was improved and the correlated neural activity was also increased (Derntl et al., 2009; Elfenbein & Ambady, 2002).

Overall, the above findings suggest that increasing one's exposure experience to other-race faces tune one's face perception system, resulting in improvements in recognition of other-race faces. However, whether such improvements are generalized to own-race faces are not clarified in these studies. To my knowledge, only one study showed that the duration of residence in Europe for East-Asian participants was positively correlated with the latency of the N170 to both own and other-race faces, implying that same and other-race faces are subject to similar perceptual learning processes (Gajewski et al. 2008). It is consistent with the claim that the differences between own-race and other-races faces are not qualitative; instead, the same underlying processes and mechanisms might be involved in processing both races of faces (e.g., DeGutis et al., 2013; Wiese et al., 2009; Zhou et al., 2021). Thus, the substantial increase of unfamiliar face exposure in a different predominant ethnicity may stimulate the face processing system in general, resulting in enhanced N170 amplitudes to faces regardless of race.

1.2.2.1.3 Underlying Perceptual Processes for the Cultural Effects on the Face N170

Since the N170 component is considered as a neural signature of configural processing (e.g., Eimer et al., 2011), the consistent cultural effects in the face N170 might reflect differences in configural face perception across groups. Nevertheless, whether this perceptual difference is culture-specific or modulated by the social environmental factor should be clarified.

The observed cultural effects in face N170 might reflect a culture-specific effect in configural processing. This claim is consistent with previous cross-cultural studies indicating more general cultural differences in perceptual styles (e.g., see Nisbett & Masuda, 2003; Nisbett & Miyamoto, 2005 for a review). The relevant cross-cultural studies have found systematic differences in perceptual styles more generally between Westerners (residing in America and Europe) and East-Asians (residing in China and Japan) in a variety of visual stimuli (e.g., objects or scenes, Chua et al., 2005; Masuda & Nisbett, 2001, 2006) and methods (e.g., eye-tracking or neuroimaging; Chua et al., 2005; Lao et al., 2013; Masuda et al., 2014). These findings indicate that Westerners tend to focus on focal information (analytic perception strategy) whereas East-Asian observers preferably attend to visual stimuli as a whole (holistic perception strategy). Importantly, recent studies have demonstrated that similar cultural differences in perceptual processes can be observed in face perception. A series of eye-tracking studies showed that Western Caucasian observers consistently fixated on the eye region, and partially the mouth, whereas East Asian observers fixated more on the central region of the face when processing faces (Blais et al., 2008; Jack et al., 2009; Kelly, Jack, et al., 2011; Kita et al., 2010; Miellet et al., 2013; Rodger et al., 2010). Such scanning biases even appear to arise early in development, from 7 years old (Kelly, Liu, et al., 2011). The cultural difference in face scanning strategies was assumed to be in line with previous cultural differences in perceptual styles between Westerners and East-Asians, suggesting a tendency towards analytic processing in Westerners with greater attention paid to the facial features and a bias towards holistic processing in East-Asians with a stronger focus on one central point of the face.

However, the above notion was not supported by a very recent study where the holistic processing operationally defined by the inversion effect was not modulated by the face scanning strategies (Zhong et al., 2021). The inconsistency might be due to the definition of holistic processing. In particular, in the face perception literature, the definition of holistic processing varies considerably between authors and is measured by a variety of paradigms (Maurer et al., 2002; Piepers & Robbins, 2012a; Richler et al., 2012). According to the review by Maurer et al. (2002), holistic processing refers to a process that glues together individual features into overall gestalt and it has been considered as one of several types of configural processing. However, it has been demonstrated that all three types of configural processing (i.e., sensitivity to first-order relations, holistic processing and sensitivity to second-order relations are disrupted by the inversion of faces (Maurer et al., 2002). Thus, it should be investigated to what extent is each type of configural processing influenced by cultural experience. This issue has been addressed in only one study where Japanese were more likely to use overall resemblance for choosing a prototypical face (holistic processing) and showed better performance in identifying second-order relations of features as compared to Americans (Miyamoto et al., 2011). This finding indicates that East-Asian shows superiority in two types of configural processing (i.e., holistic processing and second-order relations) over Westerners.

To summarize, although the consensus indicates that that East-Asians have a bias for holistic processing whereas Westerners are more analytic in processing visual stimuli (e.g., objects, scenes), the cultural effects on perceptual styles for faces are still unclear. Thus, whether an enhanced N170 in the East Asian population could relate to a preferentially holistic vision in general needs to be addressed.

1.3 The Aims of the Present Thesis

With respect to the interesting observation of cultural differences between East-Asian and Caucasian adults in the face N170 (see Table 1), I proposed two possible factors emphasizing the acquired experience, one is the cultural-specific experience—the acquisition of script system, the other is the environmental-specific experience—social immersion experience. Specifically, two hypotheses in regard to these two factors, the script system hypothesis and the social immersion hypothesis, were proposed in the present thesis. Therefore, the overall aim of this thesis is to investigate the script system hypothesis and social immersion hypothesis, namely whether and how the script system acquisition and social immersion experience influence face perception. More specifically, I addressed the following questions in the present thesis:

(1) To what extent do the script system hypothesis and/or social immersion hypothesis account for the interesting observation of cultural differences in the face N170 between East-Asian and Caucasian adults across a number of cross-cultural studies?

(2) To what extent does the script system experience affect face processing reflected in the N170 in early and expert readers?

(3) To what extent do the observed effects in the face-elicited N170 imply the effects in configural face processing?

To address these questions, I will conduct two multi-group ERP studies. The first ERP study (Study 1) aims to reassess the observation that East-Asian adults showed larger face N170 than Caucasian adults and elucidate to what extent the script system hypothesis and/or social immersion hypothesis account for the observed differences. In addition, Study 1 investigates whether the observed differences on the N170 reflect differences in configural face processing. To address these goals, two separate ERP experiments will be conducted in Hong Kong and Berlin by recording ERPs in a series of one-back tasks, using pictures of faces of Chinese and Caucasian individuals in upright and inverted orientation and doodle drawings as stimuli. Experiment 1 of this study will be conducted in an East-Asian society (Hong Kong) by recruiting local Hong Kong Chinese (LC) participants and non-Chinese (NC) participants; the NC participants could not read any logographic script and have recently come to Hong Kong from other countries. Experiment 2 will be conducted in a Western society (Berlin, Germany) by recruiting long-term Berlin residents (LB) who are local residents in Berlin, short-term Berlin residents (SB) who have recently come to Berlin from other cities in Germany, and short-term Chinese visitors (SC) who have recently come to Germany from China. All Berlin residents are native German speakers and without any logographic scripts reading experience. The script system hypothesis would be supported by larger face-elicited N170 amplitudes in Chinese participants in both Experiment 1 and Experiment 2. On the other hand, the general social immersion hypothesis would be supported by larger face-N170 amplitudes in participants that had moved into their novel social environments only a short time ago (Experiment 1: non-Chinese foreigners; Experiment 2: short-term Chinese and short-term German participants). Experiment 2 could further differentiate between the global novel social immersion and the new-ethnicity immersion hypothesis by comparing the two German groups and two short-term groups respectively. In addition, if an observed group difference in N170 reflects differences in configural processing, it should be more pronounced for upright than for inverted faces.

The second study (Study 2) aims to investigate whether the script system influence face perception in early readers who only have one-year of formal reading instruction with their native script. Furthermore, this study also investigated the extent to which two types of configural processing (holistic processing vs. sensitivity to second-order relations) are affected by the script system. Two separate ERP experiments will be conducted in Jinhua, China, and in Berlin by recording ERPs in a series of one-back tasks with naturalistic faces, two-tone Mooney faces and doodles, and on an adaptation task with pairs of faces that were either identical or differed in their second-order spatial relations of facial features. According to the script system hypothesis, it is expected that Chinese children would show larger N170 responses to naturalistic faces than German children. If the enhanced N170 specifically reflects a superiority in holistic processing, a similar group difference in N170 should be observed for Mooney faces, which have to be processed holistically as there are no separable facial features (Mooney, 1957). Besides, if the sensitivity to second-order relations assessed by the adaptation effect, plays a role in the acquired group differences for naturalistic faces, a stronger N170 adaptation effect would be expected in Chinese children.

Chapter 2 The Present Studies

2.1 Study 1: Neural Sensitivity to Faces is Increased by Immersion into a Novel Ethnic Environment: Evidence from ERPs

Abstract

Previous reports suggest that East-Asians may show larger face-elicited N170 components in the ERP as compared to Caucasian participants. Since the N170 can be modulated by perceptual expertise, such group differences may be accounted for by differential experience, for example, with logographic versus alphabetic scripts (script system hypothesis) or by exposure to abundant novel faces during the immersion into a new social and/or ethnic environment (social immersion hypothesis). We conducted experiments in Hong Kong and Berlin, recording ERPs in a series of one-back tasks, using same- and other-ethnicity face stimuli in upright and inverted orientation and doodle stimuli. In Hong Kong we tested local Chinese residents and foreign guest students who could not read the logographic script; in Berlin we tested German residents who could not read the logographic script and foreign Chinese visitors. In both experiments, we found significantly larger N170 amplitudes to faces, regardless of ethnicity, in the foreign than in the local groups. Moreover, this effect did not depend on stimulus orientation, suggesting that the N170 group differences do not reflect differences in configural visual processing. A group of short-term German residents in Berlin did not differ in N170 amplitude from long-term residents. Together, these findings indicate that the extensive confrontation with novel other-ethnicity faces during immersion in a foreign culture may enhance the neural response to faces, reflecting the short-term plasticity of the underlying neural system.

Keywords: face perception, event-related potentials, N170, configural processing, social experience

2.1.1 Introduction

How we process visual information is influenced by previous experience. Experience with certain stimulus classes improves the ability to detect, discriminate, and identify these stimuli (e.g., Sigman & Gilbert, 2000; Yi et al., 2006) – an effect termed perceptual learning, leading to perceptual expertise for certain visual objects. Arguably the most important domain of objects for which most individuals have acquired visual expertise is human faces. A robust demonstration of expertise in face perception is the disproportionally strong decrement in performance following picture-plane inversion of faces as compared to inverting other objects (Diamond & Carey, 1986; Rossion & Curran, 2010; Yin, 1969). This face inversion effect is usually explained by assuming that inversion specifically disrupts the processing of configural information (spatial relations between facial features) which is especially important for face perception (see Maurer et al., 2002 for a review). Perceptual training studies demonstrated that the face inversion effect increased after training with upright faces and decreased after training with inverted faces (Ashworth et al., 2008; Hussain et al., 2009).

Valuable tools for investigating the neural correlates of face processing and perceptual expertise are the electroencephalogram (EEG) and event-related potential (ERP). ERP studies revealed a negative-going component peaking at about 170 ms after face presentation onset, termed N170 and suggested to reflect the structural encoding of faces (Eimer, 2011; Rossion & Jacques, 2011). More specifically, the N170 is considered to index configural processing (Eimer et al., 2011; Sagiv & Bentin, 2001; Zion-Golumbic & Bentin, 2007), most prominently evidenced by its sensitivity to face inversion, which increases both N170 amplitude and latency (Caharel et al., 2013; Jacques & Rossion, 2010; Rossion, Gauthier, et al., 2000; Sadeh & Yovel, 2010).

ERP studies have reported effects of perceptual expertise on the N170 component. Long-term perceptual expertise about certain objects is associated with larger N170 amplitudes in response to objects of a participant's area of expertise (e.g., dogs, birds, fingerprints, or cars) as compared to objects outside of this area (Busey & Vanderkolk, 2005; Gauthier et al., 2003; Tanaka & Curran, 2001). Other studies showed that short-term perceptual training with laboratory-created stimulus categories (e.g., greebles) increased N170 amplitudes in response to these stimuli (Rossion et al., 2002, 2004). Face-like N170 inversion effects were also found in domains of non-face objects of expertise (e.g., words) (Dering et al., 2013; Wang et al., 2011). Since the N170 response is frequently modulated by perceptual experience, culture-specific perspectives have attracted interest.

It is well-known that faces of different ethnicity are recognized with greater difficulty than faces of one's own ethnicity (e.g., Byatt & Rhodes, 2004; Walker & Tanaka, 2003). This other-race effect is commonly accounted for by the relative lack of perceptual experience with other-race as compared to same-race faces (e.g., Rhodes et al., 2009; Tanaka et al., 2004). Specifically, greater experience with own-race faces may cause them to be processed in a configural way, whereas other-race faces may process in a more feature-based manner because of low levels of expertise (Michel et al., 2006; Rhodes et al., 1989; Tanaka et al., 2004). Some researchers argued the role of social cognitive factors in other-race effects. In particular, the ingroup/outgroup hypothesis holds that when individuals encounter a face, they first assess whether it belongs to an in-group (i.e., same race) or an out-group (i.e., other race) member (see Hugenberg et al., 2010; Young et al., 2012 for a review). For instance, Hugenberg and Corneille (2009) found that ingroup faces (Western students from the same university) were processed more configurally than outgroup ones (Western students at a different university), suggesting a role of the ingroup/outgroup status in configural processing and supporting the importance of social cognitive aspects in accounting for the other-race effect. The other ongoing debate concerns whether the N170 is sensitive to face race. A number of studies found other-race effects in the N170 but with inconsistent direction, some showing larger N170 to same-race than to other-races faces (e.g., Herrmann et al., 2007; Ito & Urland, 2005) whereas others show the opposite pattern (Gajewski et al., 2008; Stahl et al., 2008; Wiese et al., 2014). However, other studies failed to report the sensitivity to race on the N170 (e.g., Caldara et al., 2003, 2004; Tanaka & Pierce, 2009). The heterogeneity of the N170 race effects might be due to methodological differences across previous studies such as task demand (Senholzi & Ito, 2012; Wiese et al., 2013).

Some studies indicate differences between participant groups of different ethnic origins. Face-elicited N170 in adults was significantly larger in East-Asian than in Caucasian participants in the study of Dering et al. (2013); in other studies (Gajewski et al., 2008; Herzmann et al., 2011; Vizioli et al., 2010; Wiese et al., 2014), similar patterns of (statistically non-significant) larger mean face-elicited N170 amplitudes in East-Asians than Caucasians were observed. In children, Ma et al. (2022) found significantly larger face-elicited N170 responses in Chinese as compared to German early readers. The authors suggested that this difference may be due to the differential demands placed on the visual and mnemonic processing in different script systems learned by the Children after one year of formal reading training. This script system hypothesis is based on the neuronal recycling hypothesis proposed by Dehaene and Cohen (2007), that reading “recycles” preexisting cortical networks that evolved for different but similar functions (e.g., face perception) and creates its own brain networks. Numerous empirical studies have demonstrated the effects of reading acquisition on face perception (e.g., Dehaene et al., 2010, 2015; Hervais-Adelman et al., 2019; Szwed et al., 2011; van Paridon et al., 2021). For example, some ERP studies indicated a dependent relationship between the face N170 and word N170, especially concerning their co-dependent trajectory of development during reading acquisition (e.g., Dundas et al., 2014; Li et al., 2013). Furthermore, previous evidence that readers of visually more complex scripts (e.g., logographic Chinese) tend to demonstrate superior visual skills compared to readers of less complex alphabetic scripts (e.g., Demetriou et al., 2005; Huang & Hanley, 2005; McBride-Change et al., 2011). For example, Demetriou et al. (2005) found that Chinese children outperformed Greek school children on tests of global (holistic) visuospatial processing. Therefore, it may be plausible that differential training of visual perception and memory required for becoming a skilled Chinese reader as compared to an alphabetic script reader might differentially impact non-verbal visual processing, manifesting also in processing other visual stimuli of sufficient complexity, such as faces. Given that the resemblance with respect to visual characteristics is particularly striking between logographic scripts (e.g., Chinese characters) and faces (Liu et al., 2009; Ma et al., 2022), the script system hypothesis assumes that learning Chinese script during reading acquisition is likely to transfer to and influence face processing to a larger extent than learning an alphabetic script.

In the adult studies indicating group differences in N170, the East-Asians were mainly Chinese and Japanese, who are required to master logographic script systems, which are visually more complex than alphabetical scripts used by Western Caucasians (Chang et al., 2016). Therefore, differences in face-elicited N170 between East-Asian adults and Caucasian adults might be contributed to their differences in visual and mnemonic training experience required for becoming a skilled logographic script reader as compared to alphabetic script readers. Specifically, the script system hypothesis assumes that learning a logographic script system may have generalized neurocognitive consequences by strengthening holistic visual processing to visual stimuli of sufficient complexity, such as faces. This notion is supported by the finding of Ma et al. (2022) that Chinese children also showed an enhanced N170 to Mooney faces which have to be processed holistically as there are no separable facial features (Mooney, 1957), as compared to German children, indicating that the N170 differences specifically related to the superiority of Chinese children in holistic face processing. Since the N170 is held to be a neural signature of holistic processing (e.g., Eimer et al., 2011), the enhanced face-N170 in East-Asian adults relative to Caucasian adults might also reflect a relative dominance of holistic processing in East-Asians.

The evidence in favor of more holistic processing in East-Asians has been reported in a large number of cross-cultural studies where Westerners tend to be more independent and more analytic, while East-Asians tend to be more interdependent and holistic (e.g., Markus & Kitayama, 1991; Nisbett et al., 2001; Varnum et al., 2010). In particular, similar cultural differences were found in face perception areas. For instance, Miyamoto et al. (2011) found that Japanese were more likely to use overall resemblance rather than feature-matching for choosing a prototypical face as compared to Americans. Furthermore, a series of eye-tracking studies reported that East-Asians tended to maintain fixation on the nose or central area, (i.e., a holistic pattern) whereas Westerners tended to focus on the eyes and lips of faces (i.e., an analytic pattern) (Blais et al., 2008; Miellet et al., 2013; Rodger et al., 2010), consistent with cultural differences in face processing strategies. Unfortunately, findings of culture-specific eye movement patterns have not been found in other studies on adult participants (Chuk et al., 2017; Or et al., 2015; Rayner et al., 2007) and the holistic face processing measured by inversion effects is not modulated by the fixation patterns on faces (Zhong et al., 2021). Hence, it is unclear, whether an enhanced N170 in the East-Asian adult population would indeed relate to preferentially holistic vision in face processing.

As mentioned above, the N170 can be modulated by short-term training experience. For example, increased exposure to a specific visual stimulus category leads to a significant increase in N170 amplitude to this stimulus category (Scott et al., 2006, 2008), thus the apparently enhanced face N170 in East-Asian participants in the adult studies might be due to increased exposure experience with faces as compared to Caucasians. Notably, the East-Asian participants of the above-mentioned cross-cultural studies had been foreigners and hence immersed for some time into the (Western) society where the study was conducted; in contrast, the Caucasian participants likely had been locals of the host country.

Adults typically engage in a larger variety of activities, related to their occupation and social life, bringing with them, a greater variety of social encounters. Thus, Oruc et al. (2019) reported that faces have a prominent presence in the every-day visual experience of adults with an average of 255 faces per day. Therefore, the East-Asians in the cross-cultural adult studies probably had been exposed to a relatively large number of new faces in an unfamiliar and other-ethnic social environment as compared to the Caucasians. According to face space theory (Valentine, 1991), new faces may alter the representational space of the face recognition system where faces are coded on multiple dimensions in reference to a prototype, representing the average of the faces that a person has encountered. For example, visual training on the dimensions of other-race faces led to a reduction of the other-race effect (Hills et al., 2005). Possibly, the substantial increase of unfamiliar face exposure in foreigners, especially in a different predominant ethnicity may stimulate the face processing system, resulting in enhanced N170 amplitudes to faces. We will refer to this alternative explanation for the increased N170 in the East-Asian participants in previous studies as the social immersion hypothesis.

The social immersion hypothesis is supported by evidence that after about one to two years spent in a new ethnic environment emotion recognition performance improves and the correlated neural activity increases, indicating successful acclimatization and adaptation (Derntl et al., 2009; Elfenbein & Ambady, 2002). More specifically, Gajewski et al. (2008) found that the duration of residence in Europe for East-Asian participants was positively correlated with the latency of the N170 to both own and other-race faces, suggesting that the short-term experience in a new ethnic environment may influence the neural correlates of face processing.

For the participants with increased face-N170 amplitudes in the above-mentioned adult studies, the geographical region where the study was conducted constituted not only a new social environment but also a different ethnic environment. Thus, two types of immersion experience may have to be distinguished: immersion into a different ethnic social environment (new-ethnicity immersion hypothesis) and/or immersion into a new social context (global new social immersion hypothesis).

The present study

We examined the face-elicited N170 in East-Asians and Caucasians in a cross-cultural study by conducting two ERP experiments, aiming to assess differences in face-elicited N170 amplitudes between these groups. More importantly, we wanted to identify the factor(s) underlying such a difference, if confirmed, namely the role of the acquired script system and/or the experienced social immersion. Experiment 1 was conducted in an East-Asian society (Hong Kong) by recruiting local Hong Kong Chinese (LC) participants and non-Chinese (NC) participants who could not read any logographic scripts and had recently come to Hong Kong from other countries. Experiment 2 was conducted in a Western society (Berlin, Germany) by recruiting long-term Berlin residents (LB) who were native German speakers without logographic script reading experience and local residents in Berlin, short-term Berlin residents (SB) who were native German speakers without logographic script reading experience and had recently come to Berlin from other cities in Germany and short-term Chinese visitors (SC) who had recently come to Germany from China. During EEG recordings in both experiments, participants completed a series of one-back tasks using pictures of upright as well as inverted Asian and Caucasian faces and doodle stimuli (complex black and white line-drawings). This task allows to monitor participants' attention and is very easy for most participants (at least in the upright condition). Since the task is very easy, we expect ceiling effects in both groups. Nevertheless, reaction time-based group differences may be expected if there are ability differences between the groups. We used both East-Asian and Caucasian faces as stimuli in order to control for other-ethnicity stimulus effects, which were not the research question of the present study. Inverted faces were used to examine the interpretation of possible group differences in the N170 in terms of holistic processing. When faces are inverted, it is difficult to process them holistically or configurationally (Maurer et al. 2002). If an observed group difference in N170 reflects differences in holistic processing, it should be more pronounced for upright than for inverted faces. The control condition with complex abstract patterns without clear relational/configural information (i.e., doodles) serves to examine whether the group differences in N170 amplitudes are specific to faces.

In summary, the script system hypothesis would be supported by larger face-elicited N170 amplitudes in Chinese participants in both Experiment 1 and Experiment 2. The general social immersion hypothesis, however, would be supported by larger face-N170 amplitudes in participants that had moved into their novel social environments only a short time ago (Experiment 1: non-Chinese foreigners; Experiment 2: short-term Chinese and short-term German participants). In addition, Experiment 2 could further differentiate between the global novel social immersion and the new-ethnicity immersion hypothesis by comparing the two German groups and two short-term groups respectively.

Apart from the N170 we also assessed the visual P1 component. The P1 component is an early occipital positive component peaking at approximately 100 ms following stimulus onset and is thought to originate from the extrastriate brain regions (e.g., Clark et al., 1994; Di Russo et al., 2002). It is sensitive to low-level visual stimulus properties, such as size, luminance or contrast but is also modulated by attention (see Hillyard et al., 1998 for a review). Thus, this component allowed us to assess contributions of low-level processing and attention to any group differences in the following N170. In addition, some studies of face perception suggest that P1 is face-sensitive and larger in response to faces than to other objects (Herrmann et al., 2005; Itier & Taylor, 2004a). Also, larger P1 amplitudes have been reported to inverted than to upright faces, suggesting that P1 might reflect holistic face processing (Itier & Taylor, 2002, 2004b). Therefore, the present study also examined whether the group would be observed in P1.

2.1.2 Experiment 1: Hong Kong

2.1.2.1 Method

Participants

Participants were 38 university students living in Hong Kong at the time of the experiment. Prior to commencing the experiment, all participants provided informed consent, completed the Edinburgh Handedness Inventory and a questionnaire regarding demographic information. Two participants had to be excluded because of poor data quality. Of the remaining *N* = 36 participants, 18 were local Hong Kong Chinese (LC group) [mean age = 20.44 years (*SD* = 3.06); 10 females; 16 right-handed] and 18 were non-Chinese visitors (NC group) who had lived in Hong Kong for an average of 10 months (*SD* = 10.66) [mean age = 19.83 years (*SD* = 1.67); 11 females; 14 right-handed]. Non-Chinese participants were alphabetic script (Latin/Cyrillic) readers with no logographic Chinese reading experience and consisted of eight Western Caucasians, four Southeast-Asians, four Eurasians and two Subsaharan Africans. All participants had normal or corrected-to-normal vision. The experiment was approved by The Joint Chinese University of Hong Kong – New Territories East Cluster Clinical Research Ethics Committee (The Joint CUHK-NTEC CREC) in accordance with the Declaration of Helsinki.

Stimuli

Stimuli consisted of 160 faces, 160 Chinese characters and 160 abstract shapes including 80 doodles and 80 polygons. Face stimuli were grayscale frontal view portraits with neutral expressions (provided by Hildebrandt et al., 2010). Faces were of female and male Chinese and Caucasian adults in equal (see Figure 4). All faces were set to identical luminance and contrast levels and fit into a mask-shaped frame, 4.30° × 5.73° in visual angle at the given viewing distance, which removed any external features (i.e., background, hairline, clothes, and ears). Inverted versions of face stimuli were created by picture-plane rotations of the images by 180°. Doodles were random, highly complex combinations of over-lapping black shapes, curves, and lines (see Figure 4). Polygons were simple bold black outlines of different triangles, quadrangles, pentagons, and hexagons. All doodles and shapes subtended a maximal visual angle of 4.77° × 4.77°. As our hypothesis concerns the processing of faces and other nonverbal stimuli, we will not elaborate on the Chinese character tasks.

All stimuli were presented at the center of a white square (11.9° × 11.9°) on a screen with an otherwise dark grey background.

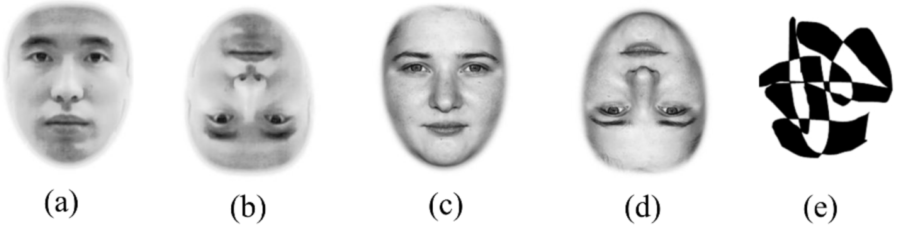


Figure 4. Examples of stimuli used in Experiment 1 and 2: (a) upright Asian faces, (b) inverted Asian faces, (c) upright Caucasian faces, (d) inverted Caucasian faces, and (e) doodles.

Procedure

The experiment was controlled by E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA). Participants were seated in a dimly lit room at 60 cm distance from the computer screen. They performed a simple repetition detection (1-back) task. If the stimulus shown was the same as the preceding stimulus (target trials), they should press the mouse button. If there was no repetition (non-target trials), no action was required. A brief practice session with houses was used to make sure that participants had understood the task. There were eight blocks, one for every stimulus condition (upright Asian faces, upright Caucasian faces, inverted Asian faces, inverted Caucasian faces, upright Chinese characters, inverted Chinese characters, doodles, and polygons). Each block consisted of eighty non-target trials and eight target trials (i.e., repetitions) presented in pseudo-randomized order. In a given block, each image was presented only once except the targets. Further, we balanced inversion of face and character stimuli so that each image was shown upright to half of the participants and upside-down to the other half. Each stimulus condition was presented in a separate block. To control for effects of presentation order, eight sequences of blocks were constructed, where (1) each stimulus condition occupied each possible position and (2) where successive blocks in most cases contained different types of stimuli. The number of participants using each block sequence was matched between groups. In addition, the experiment lasted less than half an hour and included regular breaks in order to counteract fatigue. Each stimulus was presented for 500 ms, followed by an average inter-stimulus interval of 1500 ms (varying between 1250 and 1750 ms). During the inter-stimulus interval, a fixation cross appeared at the center of the screen.

EEG recording and Preprocessing

The EEG was sampled at 500 Hz using a 128-channel geodesic sensor net (Electrical Geodesics Inc.), with 0.1-100 Hz band-pass filter settings. Cz served as the physical reference. Electrode impedances were kept below 50 kΩ throughout recording.

Offline, the EEG data were analyzed with Brain Vision Analyzer 2.0 software and re-calculated to a common average reference. The signals were digitally low-pass filtered at 30 Hz (24 dB/Octave, zerophase-type filter). An automated artifact detection algorithm was run on all channels in each segment. A channel was marked as bad if (1) voltages exceeded ± 100 μV, or (2) absolute differences between two adjacent sample points exceeded 75 µV, or if (3) the voltage range of a given channel within the entire segment exceeded 150 µV. A segment was rejected if it contained any bad channels. Channels that contained more than 20% bad segments across the entire recording were replaced through spline interpolation. We only analyzed non-target trials because there were only 8 target trials per condition and we had no hypotheses about them. The non-target trials were epoched at −100 to 500 ms relative to stimulus onset and baseline-corrected to the average of the pre-stimulus interval. On average the accepted trials in each condition were > 90 %.

Data Analysis

Behavioral data

Performance was measured as response accuracies and reaction times (RTs). Using IBM SPSS Statistics 28, we analyzed these behavioral measures for faces and doodles separately. For faces, we performed ANOVAs with participant group (LC, NC) as a between-subject factor, and stimulus orientation (upright, inverted) and stimulus race (Asian, Caucasian) as within-subject factors. For doodles and polygons, we conducted separate unpaired *t*-tests between the LC and NC group for accuracy and RT.

ERP Data

Topographic ERP analyses

We wanted to ensure that possible group differences in observed amplitudes or latencies of the ERP components can be attributed to differential cognitive processing rather than topographic dissimilarities resulting from inconsistent sources or morphological variation. A global way to determine the time period of significant topographic effects including all experimental conditions is the topographic analysis of variance (TANOVA). TANOVA is a non-parametric randomization test of reference-independent topographic difference measure (Murray et al., 2008). When the EEG data are normalized to unit variance across channels, significant differences between conditions can only be accounted for by differences in source distribution, and not by differences in source strength alone (for details see Koenig & Melie-García, 2009; Lehmann & Skrandies, 1980).

We first computed grand mean ERPs of the non-target trials by participants and stimulus conditions, epoched at −100 to 500 ms relative to stimulus onset. Then the P1 and N170 components were identified from Global Field Power (GFP) waveforms of the grand mean ERPs across participants and stimulus conditions. Next, separate mixed TANOVAs with stimulus type (upright Asian faces, inverted Asian faces, upright Caucasian faces, inverted Caucasian faces, doodles) as a within-subject factor and participant group (LC, NC) as a between-subject factor were applied for a ±30 ms time interval around the corresponding GFP peak of each ERP component in Ragu software (Randomization Graphical User interface; Koenig et al., 2011). We computed 5,000 randomization runs on the analyzed data that have been normalized. An effect will be considered statistically significant if the *p*-value is less than 0.05.

ERP component analyses

Using IBM SPSS Statistics 28, we conducted four-way ANOVAs with a between subject factor participant group (LC, NC), and within-subject factors stimulus orientation (upright, inverted), stimulus race (Asian, Caucasian) and hemisphere (left, right) on amplitudes and latencies of each ERP component to face stimuli. For doodle stimuli, two-way ANOVAs with a between-subject factor participant group (LC, NC) and a within-subject factor hemisphere (left, right) were performed on amplitudes and latencies of each ERP component. Effect sizes are reported as partial eta-square. In addition, Bonferroni correction was applied to account for multiple comparisons in *post-hoc* analysis.

2.1.2.2 Results

Behavioral results

Due to technical problems, behavioral data was only available for 15 out of 18 local Chinese participants and 14 out of 18 non-Chinese participants. For faces, accuracies and RTs were submitted to three-way ANOVAs, which revealed a significant main effect of orientation for both accuracy ($F(1, 27) = 75.97, p < 0.001$, $\eta_p^2 = 0.74$) and RT ($F(1, 27) = 11.99, p = 0.002$, $\eta_p^2 = 0.31$); participants showed higher accuracy and faster responses for upright than for inverted faces (Accuracy: *M* = 0.98 vs. 0.93, *SD* = 0.02 vs. 0.03; RT: *M* = 631.48 vs. 690.20, *SD* = 123.86 vs. 148.79). For doodles and polygons, *t*-tests showed that no significant group difference was observed for both accuracy and RT (*ps* > 0.05).

ERP results

TANOVA

The TANOVAs did not show significant main effects of participant group for P1 (*p* = 0.679) or N170 (*p* = 0.618), nor interactions between participant group and stimulus condition for P1 (*p* = 0.627) or N170 (*p* = 0.142). These results indicate that the topographies of both ERP components to any stimulus type were indistinguishable between the LC and NC group (see Figure 5). Thus, the following analyses were conducted on the assumption of topographic invariance of these components between groups.

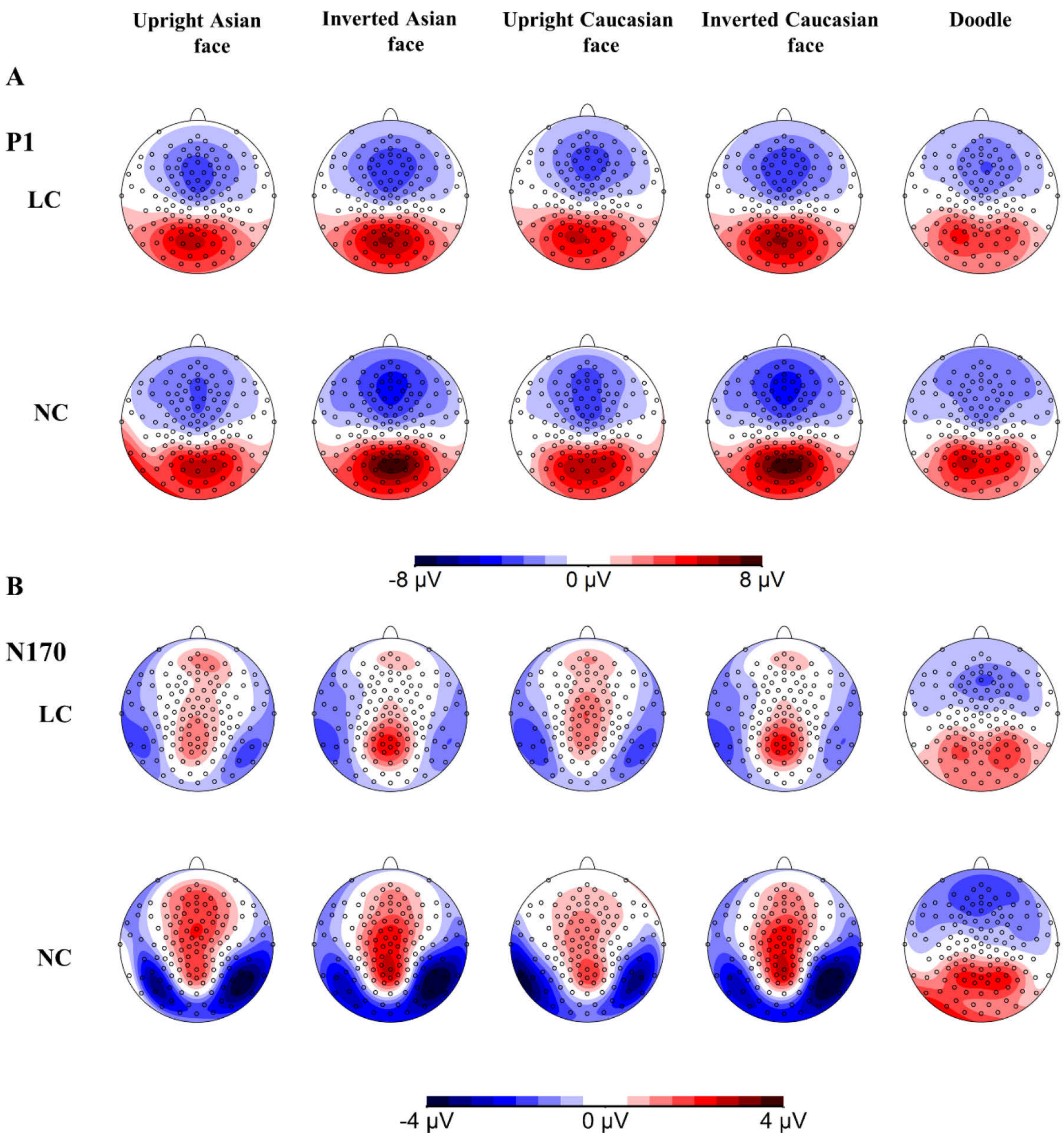


Figure 5. Experiment 1: Topographies of grand-average peak amplitudes of P1 (108 ms, panel A) and N170 (154 ms, panel B) of Chinese local residents (LC) and non-Chinese residents (NC) in Hong Kong elicited by upright Asian faces, inverted Asian faces, upright Caucasian faces, inverted Caucasian faces, and doodles.

ERP component analyses

The N170 was measured at the occipitotemporal electrode sites P7 (left hemisphere) and P8 (right hemisphere), as frequently used in other studies (e.g., Dundas et al., 2014; Eimer et al., 2011; Latinus & Taylor, 2006; Rossion & Jacques, 2008). The P1 was scored at the occipital electrodes O1 (left hemisphere) and O2 (right hemisphere). We analyzed the data from the electrodes of the EGI sensor net that corresponded to these electrodes in the 10-20 system. Thus, adaptive mean amplitudes were averaged within a ±30 ms window centered on the peaks with maximum voltages detected automatically between 110 and 210 ms after stimulus onset for N170 (negative peak) from E58/E96 (P7/P8) sites and between 50 and 150 ms after stimulus onset for P1 (positive peak) from E70/E83 (O1/O2) sites. Peak latencies were measured as the time point of the peak maximum. In addition, ERP latencies in the HK data had to be corrected by a global shift of 18 ms due to the effects of anti-aliasing filters in EGI NA300 at 500 Hz sampling rate (Pegado et al., 2014; Advisory Notice, 29 August 2014, Electrical Geodesics Inc). This correction had no effects on data quality.

P1

The ANOVAs of the P1 to faces (see Figure 6) showed a significant main effect of stimulus orientation for amplitude ($F(1, 34) = 9.83$, $p = 0.004$, $\eta_p^2 = 0.22$) but not for latency ($p > 0.05$), indicating a larger P1 amplitude for inverted than for upright faces ($M = 4.53$ vs. 3.97 µV, $SD = 2.55$ vs. 2.23). The main effect of hemisphere was significant for both amplitude ($F(1, 34) = 9.83$, $p = 0.004$, $\eta_p^2 = 0.22$) and latency ($F(1, 34) = 6.46$, $p = 0.016$, $\eta_p^2 = 0.16$), indicating larger amplitude and longer latency in the left than right hemisphere (Amplitude: $M = 4.50$ vs. 4.00 µV, $SD = 2.50$ vs. 2.26, Latency: $M = 108.99$ vs. 105.75 ms, $SD = 8.13$ vs. 9.62).

For P1 amplitudes and latencies to doodles and polygons (see Figure S1 in Supplementary Material), ANOVAs did not reveal any significant effects (*ps* > 0.05).

N170

For the N170 amplitude to faces (see Figure 6), the ANOVA showed a significant main effect of participant group ($F(1, 34) = 5.52$, $p = 0.025$, $\eta_p^2 = 0.14$), indicating that the NC participants showed the larger face N170 amplitude than the LC participants ($M = -2.04$ vs. -0.42 µV, $SD = 2.41$ vs. 1.62). In addition, there was a significant interaction between participant group, stimulus orientation and hemisphere (*F*(1, 34) = 4.28, p = 0.046, η_p^2 = 0.11). To decompose this interaction, the simple main effect of participant group was analyzed in each hemisphere. The results revealed a significant simple main effect of group in the right hemisphere (*F*(1, 34) = 5.33, p = 0.027, η_p^2 = 0.14), but not in the left hemisphere (*F*(1, 34) = 3.46, p = 0.072, η_p^2 = 0.09). In the right hemisphere the NC group showed a significantly larger face N170 amplitude than LC group (*M* = -2.57 vs. -0.58 µV, *SD* = 3.13 vs. 1.89) whereas in the left hemisphere the NC group showed only numerically (statistically non-significant) larger face N170 amplitudes than the LC group (*M* = -1.50 vs. -0.26 µV, *SD* = 2.09 vs. 1.88). The simple effect of stimulus orientation was also only significant in the right hemisphere (*F*(1, 34) = 4.34, p = 0.045, η_p^2 = 0.11), indicating a larger N170 for inverted than for upright faces (*M* = -1.86 vs. -1.30 µV, *SD* = 3.19 vs. 2.49). The simple two-way interaction between these two factors was not significant for either hemisphere (*ps* > 0.05).

For N170 latency to faces, the ANOVA showed that the main effect of group was not significant (*F*(1, 34) = 1.77, p = 0.192, η_p^2 = 0.05), suggesting the N170 latency to faces was indistinguishable between LC and NC participants (*M* = 159.84 vs. 155.10 ms, *SD* = 10.27 vs. 11.11). There were significant main effects of stimulus orientation (*F*(1, 34) = 73.35, $p < 0.001$, η_p^2 = 0.68), with shorter latency for upright than inverted faces (*M* = 153.57 vs. 161.38 ms, *SD* = 10.36 vs. 11.89), and of stimulus race (*F*(1, 34) = 7.24, p = 0.011, η_p^2 = 0.18), with shorter latency for Caucasian faces than Asian faces (*M* = 156.57 vs. 158.38 ms, *SD* = 11.09 vs. 10.90), and of hemisphere (*F*(1, 34) = 10.35, p = 0.003, η_p^2 = 0.23), with shorter latencies in the right than in the left hemisphere (*M* = 156.08 vs. 159.79 ms, *SD* = 10.64 vs. 11.90). No significant interactions were observed (*p*s > 0.05).

The ANOVAs for doodles did not reveal significant main effects of participant group in amplitude (*F*(1, 34) = 0.00, p = 0.971, η_p^2 = 0.00) or latency (*F*(1, 34) = 0.64, p = 0.430, η_p^2 = 0.02), indicating that these parameters were indistinguishable between LC and NC groups

(Amplitude: *M* = 1.79 vs. 1.77 μV, *SD* = 1.90 vs. 2.11; Latency: *M* = 150.17 vs. 148.89 ms, *SD* = 16.49 vs. 19.56). However, the main effect of hemisphere was significant for latency (*F*(1, 34) = 4.73, $p = 0.037$, $\eta_p^2 = 0.12$) but not for amplitude ($p > 0.05$), suggesting the latency was shorter for the right than for the left hemisphere (*M* = 145.67 vs. 153.39 ms, *SD* = 21.83 vs. 19.61). No interactions approached significance (*p*s > 0.05). The ANOVAs of the polygon-elicited N1 (see Figure S2 in Supplementary Material) did not reveal significant main effects of participant group in amplitude (*F*(1, 34) = 1.40, $p = 0.245$, $\eta_p^2 = 0.04$) or latency (*F*(1, 34) = 0.03, $p = 0.868$, $\eta_p^2 = 0.00$), indicating that these parameters were indistinguishable between groups (amplitude: *M* = 0.19 vs. 1.76 μV, *SD* = -0.59 vs. 2.14; latency: *M* = 149.56 vs. 150.67 ms, *SD* = 16.48 vs. 19.43). In addition, there were not significant interactions (*ps* > 0.05).

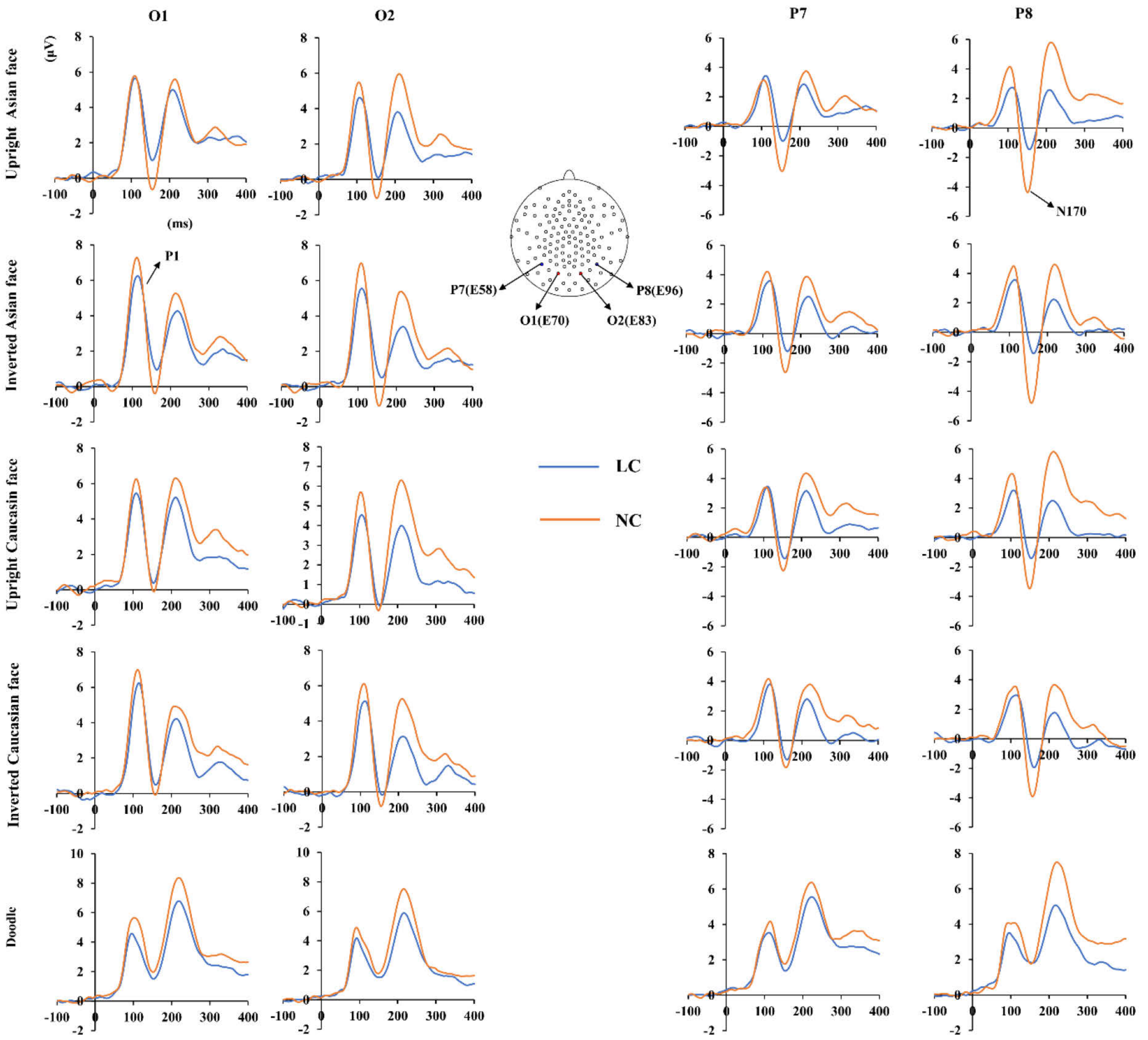


Figure 6. Experiment 1: Grand-average ERP waveforms of Chinese local residents (LC, blue) and non-Chinese residents (NC, orange) in Hong Kong elicited by upright Asian faces, inverted Asian faces, upright Caucasian faces, inverted Caucasian faces, and doodles at two occipital sites showing the P1 amplitude (O1, left hemisphere; O2, right hemisphere) and two occipitotemporal sites showing the N170 amplitude (P7, left hemisphere; P8, right hemisphere).

2.1.2.3 Discussion

Experiment 1 aimed to replicate differences in N170 amplitude between East-Asian and Caucasian participants that were observed in previous reports and to distinguish between the script system hypothesis and the social immersion hypothesis of the effect. The script system hypothesis would predict a replication of larger N170 amplitudes in East-Asian (Hong Kong Chinese) than in non-Chinese participants, whereas the social immersion hypothesis would predict the opposite, that is, a smaller N170 in Chinese than non-Chinese participants.

Experiment 1 showed smaller N170 amplitudes to faces in local Hong Kong Chinese than in non-Chinese participants. Importantly, this difference was not modulated by stimulus race, excluding the contribution of the differences of facial physiognomic information between races of faces. In addition, the absence of similar group differences in P1 makes it unlikely that the observed N170 difference was caused by low-level stimulus properties.

The results strongly argue against the script system hypothesis but support the social immersion hypothesis. The groups showing enhanced face N170 amplitudes in both Experiment 1 and previous studies had been immersed in new social environments. Thus, they likely had been exposed to many new faces, which they needed to perceive, remember and recognize, which might explain their enhanced N170 amplitudes to faces as a consequence of temporary training experience. In addition, the absence of significant group differences for both doodles and polygons in the N170 time window might indicate that the social immersion effect is specific to face processing; however, due to the relatively small sample size, an independent replication would be desirable.

2.1.3 Experiment 2: Berlin

Experiment 1 had provided evidence against the script system account for previous findings from western labs of larger N170 in Chinese than non-Chinese participants, by showing smaller N170 in the Chinese. Hence, Experiment 2 aimed to both confirm the rejection of the script system hypothesis and support the social immersion effect on the face processing system. A limitation of Experiment 1 was that four non-Chinese participants were of Asian descent. Hence, for these participants (who had not learned a logographic script system), the effects of immersion into a new other-ethnicity society may be weaker than for the non-Asians.

Although this might make the finding of larger N170 in this group as a whole even more impressive, it blurs the distinction between the global social immersion and social immersion into other-ethnicity environment. Experiment 2 attempted to make this distinction.

We conducted Experiment 2, closely adhering to the procedure of Experiment 1 in Berlin, Germany. Three groups of participants were tested: two German groups, consisting of long-term Berlin residents (LB) and short-term Berlin residents (SB), and short-term Chinese foreigners (SC). If the short-term immersion experience in a different ethnic environment is a major factor driving the size of the face N170, we expected the SC group to show larger face-elicited N170 amplitudes than the SB group. If the short-term immersion experience in a new social context is the major factor, the SB group should show larger face-elicited N170 amplitudes than the LB group.

2.1.3.1 Method

Participants

Participants were recruited in Berlin, Germany, and fell into three groups: 32 long-term Berlin residents (LB group) who grew up or had lived for at least 2 years in Berlin [mean age = 25.97 years (*SD* = 5.55); 19 females; 32 right-handed], 29 short-term Berlin residents (SB group) who had lived in Berlin for average of 6.54 months (*SD* = 3.48, range: 2.5-14 months) [mean age = 23.86 years (*SD* = 4.79); 18 females; 27 right-handed], and 32 short-term Chinese participants (SC group) who came from mainland China and had lived in Berlin for an average duration of 6.35 months (*SD* = 3.24, range: 2.5-13 months) [mean age = 24.06 years (*SD* = 2.46); 19 females; 32 right-handed]. LB and SB group were native German speakers and did not have logographic Chinese reading experience.

Prior to the experiment, all participants provided informed consent and completed the Edinburgh Handedness Inventory and a questionnaire regarding their demographic information. All participants had normal or corrected-to-normal vision. The experiment was approved by the ethics committee of the Department of Psychology of the Humboldt-University at Berlin in accordance with the Declaration of Helsinki.

Stimuli

Stimuli consisted of faces, Chinese characters, German words and doodles. Except for doodles, all stimuli were presented in both upright and inverted orientation. The polygon condition was omitted in this experiment due to its small and unsystematic ERPs in Experiment 1 (see Figure S1 and S2 in Supplementary Material). Since the Chinese characters and German words do not address the current questions, we will not elaborate on them. The details of face and doodle stimuli are same as in Experiment 1.

Procedure

Presentation 1.0 software controlled the procedure of Experiment 2. It consisted of nine blocks, one for each stimulus condition: upright Asian faces, upright Caucasian faces, inverted Asian faces, inverted Caucasian faces, upright Chinese characters, inverted Chinese characters, upright German words, inverted German words, and doodles. The principles to control the effects of presentation order were the same as in Experiment 1 except there were nine sequences of blocks. The number of participants using each block sequence was closely matched between groups. Other details of the procedure were the same as in Experiment 1.

EEG recording and Preprocessing

EEG was recorded with 39 Ag/AgCl electrodes placed within an elastic cap (Easycap GmbH, Herrsching, Germany) according to the extended International 10-20 System. Two additional electrodes below each eye measured vertical electroocular (EOG) activity. During recordings all channels were referenced to the Cz electrode; AFz served as ground electrode. Impedances were kept < 10 kΩ at a uniform level. Signals were amplified using Brain Amps (Brain products) without additional filter settings during recordings and sampled at 1000 Hz.

Offline EEG data preprocessing was same as in Experiment 1 except that data was digitally high-pass filtered at 0.1 Hz (12dB/Octave, zerophase-type filter) and low-pass filtered at 30 Hz (24 dB/Octave, zerophase-type filter).

Data Analysis

Behavioral data

The Accuracies and RTs for face tasks were submitted to two three-way ANOVAs with a between-subject factor participant group (SC, SB, LB), and within-subject factors stimulus orientation (upright, inverted) and stimulus race (Asian, Caucasian). For doodle task, two-one-way ANOVAs with a between factor participant group (SC, SB, LB) were conducted for accuracy and RT.

ERP data

Topographic ERP analyses

Again, we first checked scalp field homogeneity of ERPs across groups and conditions by TANOVA with a within-subject factor stimulus condition (upright Asian faces, inverted Asian faces, upright Caucasian faces, inverted Caucasian faces, doodles) and a between-subject factor participant group (SC, SB, LB).

ERP component analyses

The four-way ANOVAs with a between-subject factor participant group (SC, SB, LB), and within-subject factors stimulus orientation (upright, inverted), stimulus race (Asian, Caucasian) and hemisphere (left, right) on amplitudes and latencies of each ERP component to face stimuli. For doodle stimuli, two-way ANOVAs with a between-subject factor participant group (SC, SB, LB) and a within-subject factor hemisphere (left, right) were performed on amplitudes and latencies of each ERP component.

In all analyses, effect sizes are reported as partial eta-square. In addition, Bonferroni correction was applied to account for multiple comparisons in *post-hoc* analysis.

2.1.3.2 Results

Behavioral results

ANOVAs showed significant main effects of stimulus orientation for both accuracy ($F(1, 90) = 117.71$, $p < 0.001$, $\eta_p^2 = 0.57$) and RT ($F(1, 90) = 29.39$, $p < 0.001$, $\eta_p^2 = 0.25$), participants across groups showed higher accuracies and faster responses for upright than for inverted faces (Accuracy: $M = 0.98$ vs. 0.94, $SD = 0.02$ vs. 0.04; RT: $M = 618.04$ vs. 673.61 ms, $SD = 106.70$ vs. 111.73). All other main effects and interactions were not significant (*ps* > 0.05).

The ANOVAs on accuracies and RTs in doodle task did not reveal main effects of group (*ps* > 0.05).

ERP results

TANOVA

The TANOVAs showed that neither the main effect of group for P1 ($p = 0.219$) or N170 ($p = 0.239$) nor the interactions between group and stimulus condition for P1 ($p = 0.376$) or N170 ($p = 0.796$) approached significance. Thus, the following component analyses were continued since the absence of group differences on the topographies of ERP components in all stimulus conditions (see Figure 7).

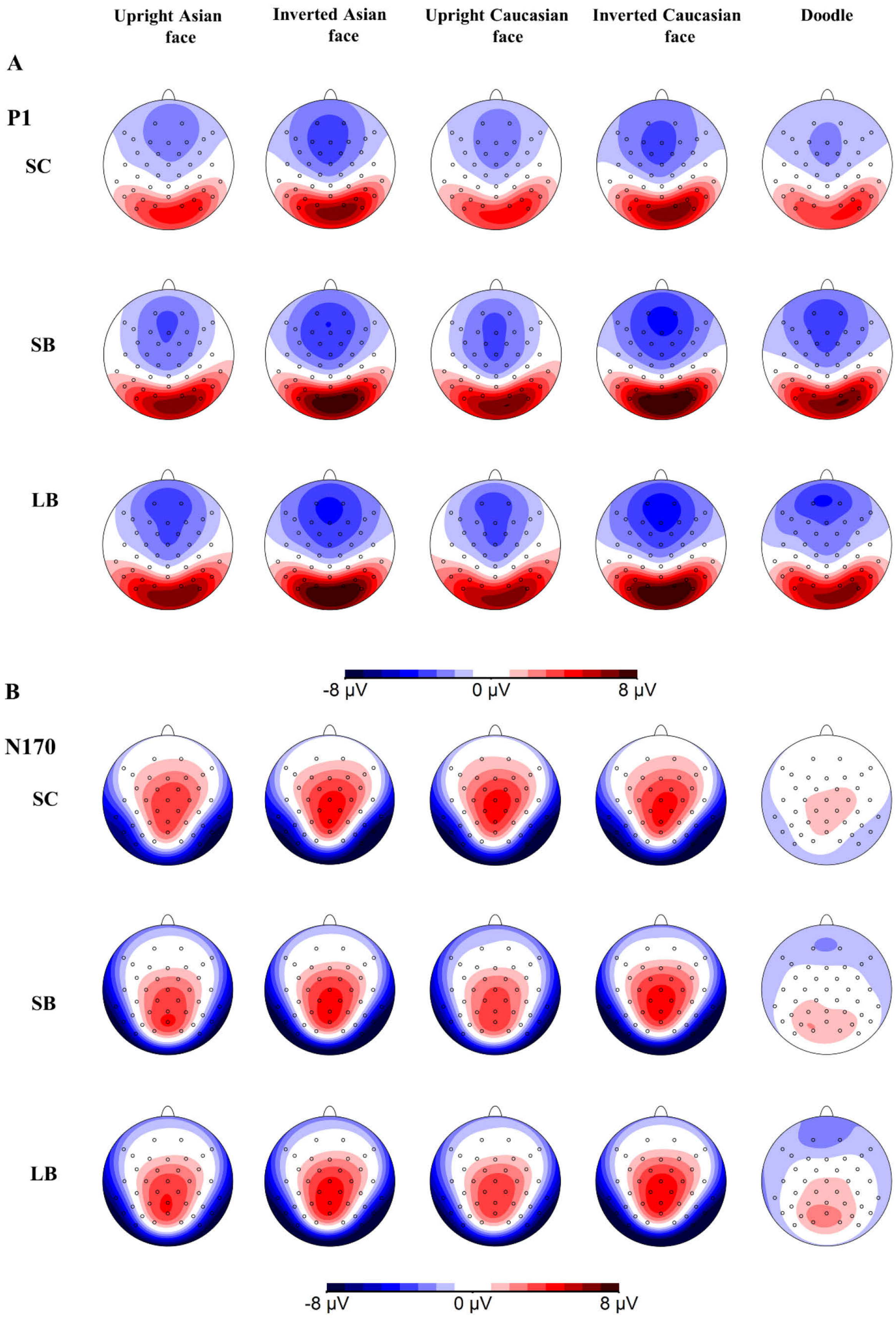

Upright Asian face
Inverted Asian face
Upright Caucasian face
Inverted Caucasian face
Doodle
A
P1
SC
SB
LB
-8 µV
0 µV
8 µV
B
N170
SC
SB
LB
-8 µV
0 µV
8 µV

Figure 7. Experiment 2: Topographies of grand-average peak amplitudes of P1 (103 ms, panel A) and N170 (150 ms, panel B) of Chinese short-term residents (SC), and German short-term and long-term residents (SB vs. LB) in Berlin elicited by upright Asian faces, inverted Asian faces, upright Caucasian faces, inverted Caucasian faces, and doodles.

ERP component analyses

The electrodes for analyses of P1 (O1/O2) and N170 (P7/P8) were chosen as in Experiment 1 and the amplitudes and latencies were derived in the same way.

P1

The ANOVAs results of P1 amplitudes and latencies to faces showed significant main effects of stimulus orientation for both amplitude ($F(1, 90) = 84.06, p < 0.001, \eta_p^2 = 0.48$) and latency ($F(1, 90) = 31.54, p < 0.001, \eta_p^2 = 0.26$). The inverted faces elicited larger amplitudes and longer latencies than upright faces (Amplitude: *M* = 5.85 vs. 4.50 µV, *SD* = 3.79 vs. 3.34; Latency: *M* = 102.48 vs. 96.86 ms, *SD* = 14.53 vs. 14.96). All other main effects or interactions were not significant (*ps* > 0.05).

The ANOVAs performed on P1 amplitudes and latencies to doodles did not yield any significant effects (*ps* > 0.05).

N170

The ANOVA on N170 amplitudes to faces (see Figure 8) indicated a significant main effect of participant group ($F(1, 90) = 3.89, p = 0.024, \eta_p^2 = 0.08$). A Bonferroni corrected *post-hoc* pairwise comparison between SC and SB groups, testing the new-ethnicity immersion hypothesis, showed that the face-elicited N170 amplitude in the SC group was larger than in the SB group (*M* = -2.74 vs. -1.03 µV, *SD* = 2.90 vs. 2.57, *p* = 0.027, Cohen's *d* = 0.62). In contrast, the pairwise comparison between SB and LB group, testing the global new social immersion hypothesis failed significance (*M* = -1.03 vs. -1.50 µV, *SD* = 2.57 vs. 1.89, *p* = 1, Cohen's $d = 0.21$). In addition, without correction the pairwise comparison between the SC and LB groups, combining the two immersion types was significant ($p = 0.049$) and showed a medium effect size (Cohen's $d = 0.51$), but failed significance after Bonferroni correction ($p = 0.148$).

The main effect of stimulus orientation in N170 amplitude was significant ($F(1, 90) = 28.31$, $p < 0.001$, $\eta_p^2 = 0.24$), due to larger N170 amplitudes to inverted as compared to upright faces ($M = -2.05$ vs. -1.51 µV, $SD = 2.75$ vs. 2.48). No other main effects and interactions were significant ($ps > 0.05$).

The ANOVA on N170 latency to faces revealed that the main effect of group was not significant ($F(1, 90) = 0.08$, $p = 0.927$, $\eta_p^2 = 0.02$), suggesting indistinguishable latencies to faces between SC, SB and LB participants ($M = 153.83$ vs. 153.98 vs. 153.02 ms, $SD = 10.36$ vs. 10.42 vs. 11.26). The main effect of orientation was significant ($F(1, 90) = 95.96$, $p < 0.001$, $\eta_p^2 = 0.52$), indicating longer latencies for inverted than for upright faces ($M = 157.34$ vs. 149.85 ms, $SD = 10.14$ vs. 11.88). No other effects and interactions reached significance ($ps > 0.05$).

For N170 to doodles, the ANOVAs revealed a main effect of participant group for amplitudes ($F(1, 90) = 3.20$, $p = 0.045$, $\eta_p^2 = 0.07$), but not for latencies ($F(1, 90) = 0.47$, $p = 0.625$, $\eta_p^2 = 0.01$). The *post-hoc* multiple comparisons with a Bonferroni adjustment showed a more negative N170 amplitude to doodles for SC participants than SB participants ($M = -0.23$ vs. 1.33 µV, $SD = 2.77$ vs. 2.35, $p = 0.049$, Cohen's $d = 0.61$). However, the other two comparisons (SC vs. LB and SB vs. LB) were not significant after Bonferroni correction ($ps > 0.05$). There was a significant main effect of hemisphere for latency ($F(1, 90) = 8.50$, $p = 0.004$, $\eta_p^2 = 0.09$), which was shorter in the right than in the left ($M = 147.54$ vs. 152.53 ms, $SD = 17.57$ vs. 17.57). No significant interactions were observed ($ps > 0.05$).

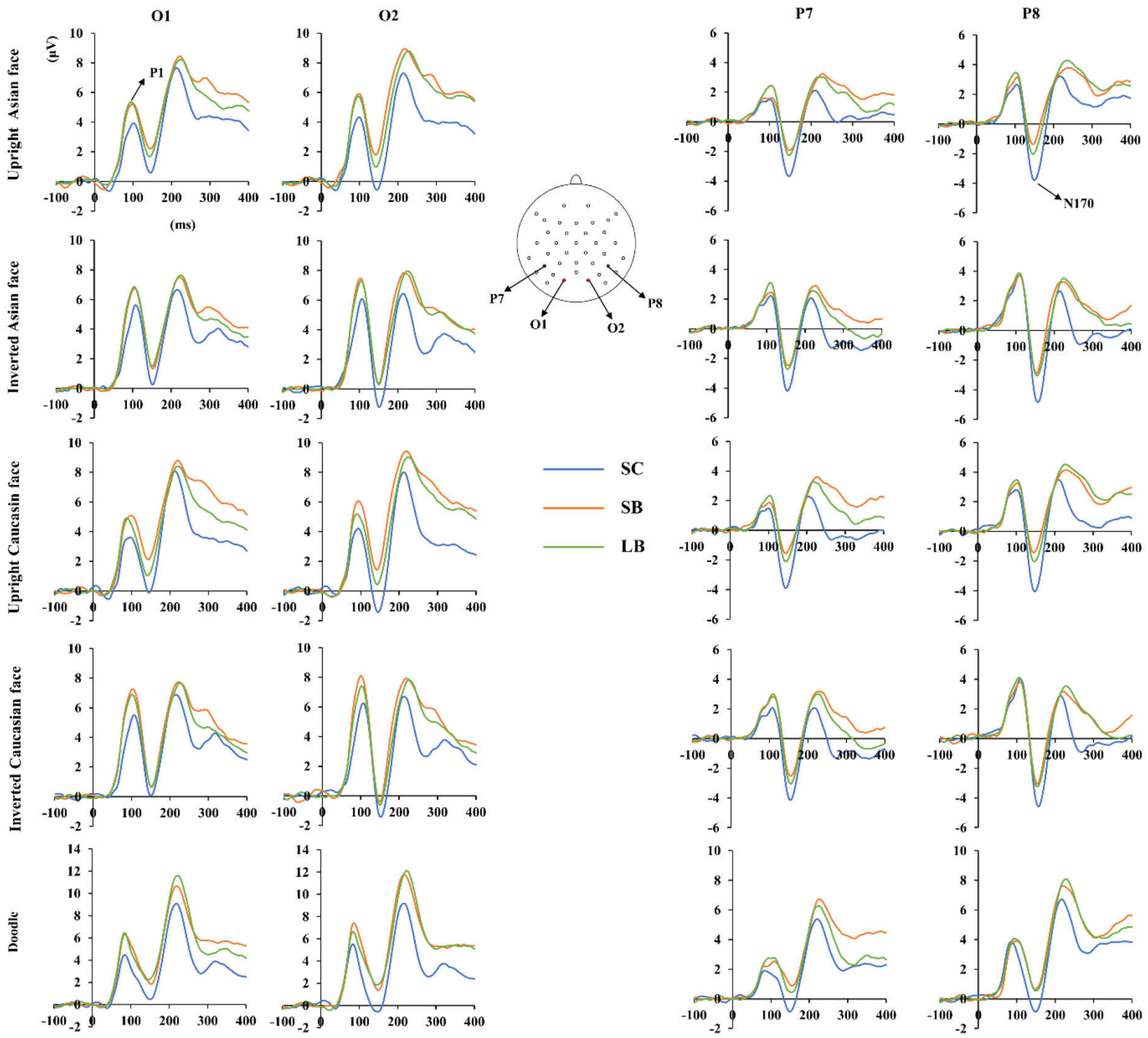


Figure 8. Experiment 2: Grand-average ERP waveforms of Chinese short-term residents (SC, blue), and German short-term (orange) and long-term (green) residents (SB vs. LB) in Berlin elicited by upright Asian faces, inverted Asian faces, upright Caucasian faces, inverted Caucasian faces and doodles at two occipital sites showing the P1 amplitude (O1, left hemisphere; O2, right hemisphere) and two occipitotemporal sites showing the N170 amplitude (P7, left hemisphere; P8, right hemisphere).

2.1.3.3 Discussion

Overall, Experiment 2 demonstrated that the SC participants with short-term immersion experience in a new-ethnicity environment (Germany) show significantly larger face-elicited

N170 amplitudes than SB participants with social immersion experience in Berlin. This finding replicates the group difference between residents and foreigners of Experiment 1 and provides further supporting evidence for the new-ethnicity immersion hypothesis. Meanwhile, Experiment 2 showed that the face N170 was indistinguishable between SB group with short-term immersion experience in Berlin and LB group without any social immersion experience, arguing against the global new social immersion hypothesis. Importantly, as in Experiment 1 the group difference was not related to stimulus race since the group differences did not interact with this factor. In addition, the absence of group differences in P1 rule out contributions of low-level stimulus factors and attention.

The overall analyses of Experiment 1 and 2

To increase the power of the results from two individual experiments, we performed a combined analysis on the face N170 amplitudes across two experiments. Since the N170 group differences in the individual experiments were not modulated by stimulus orientation and stimulus race, data were pooled across these factors. Since the results on N170 amplitude to face stimuli in each experiment support the new-ethnicity immersion hypothesis but together seem to be at variance with the script system hypothesis, we aimed to support this conclusion by a joint ANOVA with between-subject factors group, where groups are coded as locals versus foreigners, and location of lab (HK, Berlin); we also included a within-subject factor hemisphere (left, right). In HK data, the local Chinese and non-Chinese groups were coded as locals and foreigners, respectively. In Berlin data, the short-term Chinese were the foreigner group whereas the local group consisted of the combined short-term Berlin residents and long-term residents. The new-ethnicity immersion hypothesis would be supported by a main effect of group whereas the script system hypothesis would be supported by an interaction between group and location. Indeed, the results revealed a significant main effect of participant group ($F(1, 125) = 10.07$, $p = 0.002$, $\eta_p^2 = 0.08$), confirming larger N170 amplitudes to faces in the foreigner groups than in the local groups across experiments (see Figure 9). Importantly, any interactions involving factors group and location of lab, which might support the script system account, were not significant (*ps* > 0.05).

Overall, across experiments, the observed group differences in face-elicited N170 support an effect of the social immersion experience in a new ethnic social environment but do not align with the script system account.

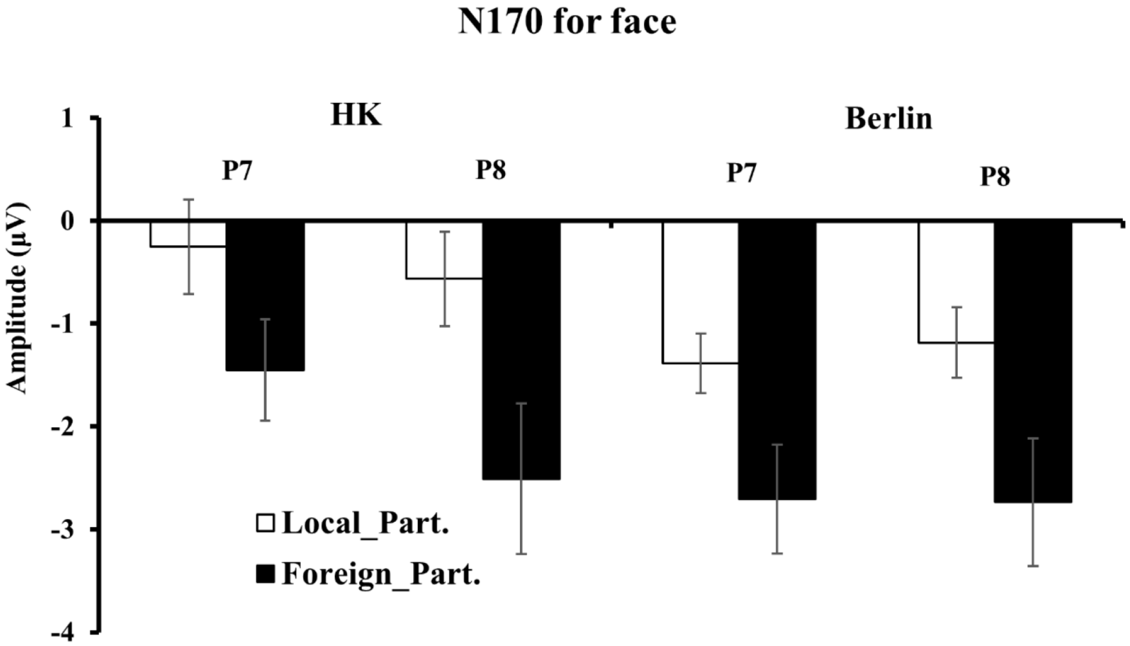


Figure 9. Amplitudes of N170 (P7, left hemisphere; P8, right hemisphere) elicited by faces for the local (white) and foreign (black) participants from HK and Berlin labs. Error bars indicate standard error of the means.

2.1.4 General Discussion

In previous reports of cross-cultural ERP studies conducted in Western labs, we had noticed a consistent yet inconclusive pattern of larger face-elicited N170 amplitudes in East-Asian adults (mainly Chinese) than in Caucasian adults. The pattern was inconclusive because only in one experiment (Dering et al., 2013) it was statistically significant while the numerical (non-significant) differences were consistent across the other studies (e.g., Vizioli et al., 2010;

Wiese et al., 2014). Larger N170 amplitudes in East-Asian participants were supported by a recent cross-cultural study in early Chinese and German readers (Ma et al., 2022). Therefore, the present study aimed to reassess this evidence and elucidate its possible underlying cause. Specifically, the difference might be due to cultural factors, for example due to the differential effects of the acquired script systems (script system hypothesis) or due to short-term immersion experience with a novel social and/or ethnic environment (social immersion hypothesis). Experiment 1, conducted in Hong Kong, revealed smaller face N170 amplitudes in local Chinese participants than in non-Chinese foreign participants. Since this effect is opposite in direction to the effects observed in previous studies, it contradicts the script system hypothesis. However, it is in line with the social immersion account. This was confirmed by Experiment 2, conducted in Berlin, where short-term Chinese visitors showed larger face-elicited N170 amplitudes than German natives. In addition, Experiment 2 allows the further specification that it is not any novel environment that drives the effect but it seems to be the immersion into a novel ethnic environment because long-term German residents and the short-term German residents were very similar in N170 amplitude. Taken together, the present study verifies a significant difference between Chinese and Caucasian participants but not as a consistent cultural difference but depending on whether the ethnic social environment had changed in the recent past.

The script system account of N170 differences

The results from Experiment 1 conducted in HK and the observations from previous cross-cultural ERP studies are contradictory and cannot be consistently explained by the script system hypothesis. According to the script system hypothesis, Chinese participants should have larger N170 to faces than non-Chinese participants in Experiment 1 as in the previous adult studies. However, the smaller face-elicited N170 in Chinese participants compared with non-Chinese participants in the HK lab is at variance with the script system hypothesis. This finding is also inconsistent with the results in children, reported by Ma et al. (2022). Ma et al. has shown that native logographic Chinese readers had larger face-elicited N170 than native alphabetic readers, suggesting the contribution of perceptual experiences with visually complex logographic Chinese characters during reading acquisition in Chinese readers.

The inconsistency of findings might be due to participants' age, that is, adults in the present Experiment 1 versus children in the experiment of Ma et al. (2022). In adults, many ERP studies have consistently reported larger N170 amplitudes in experts to both faces and scripts (Maurer et al., 2005; Rossion et al., 2003), suggesting that the N170 reflects visual expertise for both faces and scripts. Furthermore, previous research indicated that acquisition of script reading involves the "recycling" of the neural system underlying face processing, and might in turn influence the face processing system (e.g., Dehaene et al., 2010). For instance, developmental studies found that the development of face expertise and script expertise reflected in the N170 during reading acquisition were dependent on each other (e.g., Dundas et al, 2014; Li et al., 2013). However, growing evidence suggests an inverted "U" model of script expertise effects on the N170; perceptual learning appears to be critically important during the first two or three years of reading acquisition and then gradually declines as expertise consolidates (e.g., Brem et al., 2009; Cao et al., 2011; Maurer et al., 2011). For example, Cao et al. (2011) found that the N170 amplitude to Chinese characters in 2nd grade children showed the largest amplitude among all age groups, but thereafter declined with increasing age, suggesting that the neural mechanisms for script expertise develop rapidly during the early primary school years. Thus, it is plausible that the effects of reading acquisition on face processing appear in early second-grade readers as found by Ma et al. (2022) and then decline and might even disappear in adult expert readers. In this case, we would not expect any effects of the script system on adult, expert readers and have to resort to alternative accounts for the effects observed in adults.

The social immersion account of N170 differences

As predicted by the social immersion account, our findings indicate that independent of lab location, foreigners showed larger N170 amplitudes to faces than locals. Thus, the enhanced face-elicited N170 in East-Asians as compared to Caucasians at least numerically indicated by previous study results may derive from the social immersion experience of East-Asians. Similar to the East-Asian participants in previous studies (e.g., Dering et al., 2013), our non-Chinese participants in the HK lab and the Chinese participants in the Berlin lab were foreigners to the geographical region and social environment where the study was conducted. During their visit they were probably exposed to many new faces of a different ethnicity as mostly encountered at home, which they needed to encode and recognize. Hence, it is conceivable that a sudden surge in the exposure to novel faces may have enhanced their N170 amplitudes to faces, similar to what has been shown in training studies with novel objects (Rossion et al., 2002, 2004; Scott et al., 2006, 2008). Therefore, although at variance with the script system account, the increase of face N170 in foreigner groups in our study supports the social immersion account.

Importantly, there were no comparable group differences in the P1 – a component, which is sensitive to low-level stimulus properties and attention (see Hillyard et al., 1998 for a review), ruling out contributions of low-level stimulus properties and general attention in the N170 differences for both faces and doodles. As in most experiments, fatigue, practice or habituation effects might be involved as participants are exposed to a multitude of stimuli. Previous studies indicated that short-term adaptation or habituation effects may occur in performance and ERPs even within short experiments (e.g., Maurer, 2008; Schinkel et al., 2014). However, our countermeasures against such time on task effects confounding with the effects of interest should have been adequate, because our main aim was to compare single (e.g., doodles) or at most pairs of conditions (upright vs. inverted) across groups. The paired conditions were (almost) never directly adjacent but separated by other conditions and changed orders across sequences. Moreover, the number of participants using each block sequence was comparable between groups in both studies. Therefore, group comparisons always involved matched positions in the sequences, ruling out short term effects of time on task. Additionally, an empirical indication for the efficiency of our control procedure is the P1 component, which is typically sensitive to adaptation effects (Obrig et al., 2002; Wastell & Kleinman, 1980) but did not show any group differences in the present data. In addition, the TANOVA showed the same topographies of N170 across groups ruling out a contribution of heterogeneity of scalp distributions in the N170 differences.

Additionally, the present study examined whether the observed effects are specific to face perception or transfer to other stimulus domains. This question was pursued by including doodles representing complex visual non-face stimuli in both studies and polygons in Experiment 1, representing simple visual stimuli. In Experiment 1 the absence of group differences in the N170 for both doodles and polygons seems to argue in favor of the face specificity of the observed group effects in N1/N170. However, the findings of Experiment 2 that similar group differences as for the face-elicited N170 were observed for doodles are at variance with the idea that the social immersion effects are limited to faces. Instead, these findings might be accounted for by a more general effect on the neurocognitive generators of the N1/N170 component. Specifically, it might indicate that social immersion into another ethnicity environment influences higher-order structural analysis of complex visual stimuli reflected in the N170/N1 (Rossion et al., 2003; Schendan et al., 1998). However, doodles are meaningless and unfamiliar stimuli for participants and the neurocognitive processes underlying the N1/N170 elicited by these stimuli are not entirely clear, hampering a conclusive interpretation for the group effects in doodle-elicited N170 amplitudes. Future studies should test more familiar non-face stimuli (e.g., cars, houses) to examine the specificity or generality of social immersion effects.

The present findings are consistent with previous research indicating the experience in a new ethnic environment increases the neural actives underlying the processing facial emotion (Derntl et al., 2009; Elfenbein & Ambady, 2002). However, this finding is not in line with the results of Gajewski et al. (2008) of a negative association between the N170 latency to faces and the duration stay of East-Asians in the Western society. This inconsistency might be due to the length of social immersion experience, which on average was 6 months in the present study versus 5.5 years in Gajewski et al. (2008). This explanation needs to be examined in future studies.

For foreigners, the geographical region into which they were immersed constituted not only a different ethnic environment but also a new social environment. Therefore, the social immersion account might hold for a new social immersion without new ethnicity immersion. In order to distinguish these two aspects of social immersion we compared short-term Berlin residents with long-term Berlin residents in Experiment 2; however, the group difference on face N170 amplitudes was not significant. Thus, the enhanced face N170 for foreigners may not be accounted for by the immersion experience in a new social context. Possibly an abundance of new same ethnicity faces is not a comparable challenge to the visual system as an abundance of other ethnicity faces with more different features and configurations as will be discussed next.

Massed exposure to other-race faces as a modulator of face processing?

The other-race effect is a well-known phenomenon whereby individuals discriminate, memorize, or identify other-race faces less effectively than for same-race faces (Byatt & Rhodes, 2004; Meissner & Brigham, 2001; Walker & Tanaka, 2003). However, the present results did not show any race effects on the behavioral or neural level. This is consistent with

Senholzi and Ito (2013) using a similar 1-back task and might be due to low task difficulty since all performance was at or close to ceiling (e.g., mean accuracy for both East-Asian and Caucasian faces was 98%). In addition, since previous findings about race effects in the N170 are mixed, the absence of such an effect in the present study is also not unique. For instance, some researchers argue that the heterogeneity of N170 race effects might be due to methodological differences such as different task demands (Wiese et al., 2013). Nevertheless, the absence of N170 race effect in the present study is at variance with Senholzi and Ito (2013) who observed larger amplitudes for other-race than own-race faces in a 1-back identity repetition detection task. Here, the inconsistency might be due to the face stimuli compared, Caucasian versus East-Asian faces in the present study but Caucasian versus Black faces in Senholzi & Ito (2013). In sum, while the exact reason for the absence of OREs at behavioral and neural levels in the present study remains unclear, this is not uncommon and may initiate more research in order to understand the presence and causes of other-race effects.

On the other hand, the present study suggests that the massed exposure to new other-race faces is critical for enhancing the N170 amplitudes to faces in general. This is evidenced by the results of Experiment 2, indicating no group difference in the N170 between short-term and long-term German residents in Berlin. This view might be accounted for by the face space theory (Valentine, 1991), assuming that the substantial new other-race faces exposure may modify the representation system of face processing, such as influencing the average face weighted by the frequency of the encounter. This is supported by perceptual training studies in which training individuals in facial dimensions of other-race faces improve the performance of other-race faces (Hills et al., 2005). Importantly, there was no modulation of the race on the N170 group differences in the present study, implying that that increased experience to other-race faces in adults affects the neural level underlying the processing of both same- and other-race faces. The conclusion resonates with the problems of improving face recognition abilities in normotypical adults, where even extensive memory training with same-race faces did not lead to a notable improvement (Dolzycka et al., 2014). According to the present findings, other-race faces as training material might provide a better option.

Expertise effects of N170 enhancements

The results of the present study also yield insight into the neural plasticity of expert face processing in adults. Firstly, the increased N170 amplitudes in participants who had social immersion experience in a novel ethnic social environment may reflect their perceptual problems in processing faces. Perceptual problem in face processing is exemplified by the face inversion effect of increases in the latency of the N170, and in many, but not all cases, increases in amplitude as compared to upright faces (Caharel et al., 2013; Jacques & Rossion, 2010; Rossion et al., 2000; Sadeh & Yovel, 2010). However, this suggestion was not supported by the current results, where larger N170 amplitudes in the foreign groups were not accompanied by longer latencies – despite the standard effect of face inversion on both N170 amplitude and latency was found in the present study. By contrast, consistent with previous studies (e.g., Tanaka & Curran, 2001) where the enhanced N170 amplitudes are generally found as a correlate of better visual expertise for certain stimuli, the larger N170 amplitudes in our foreigner groups may reflect their greater expertise in face processing. Thus, the social immersion experience in a novel ethnic social environment improves one’s face expertise reflected by increased N170 amplitudes.

The greater visual expertise reflected by the increased N170 in many cases goes along with more pronounced configural perception (e.g., Gauthier et al., 2003; Rossion et al., 2002). Therefore, the group differences in the N170 may reflect the degree of configural visual processing. To address this issue, we used both upright and inverted faces in the present experiments. We expected that the group difference in the N170 might be smaller for inverted faces than for upright faces or there should be a larger face inversion effect in the N170 in foreign than in local participants. However, the critical interaction between group and orientation was not observed, that is, the observed group differences in the N170 were indistinguishable between upright and inverted faces. Hence, our findings do not provide evidence that the observed N170 increase due to social/ethnic immersion reflects differences in configural visual processing. In addition, the behavioral results also do not support the notion that the group differences in the N170 indicate differences in configural face processing as no group effects were observed. As the behavioral task used in the present was a very easy 1-back task – merely serving the monitoring of attention and controlling for the task difficulty across groups, future studies should investigate the effects of social immersion at the behavioral level by using suitable tasks.

Nevertheless, what remains to be discussed are the exact processes reflected by the group differences in the face-elicited N170 of the current study. As inversion of faces impairs configural, but not featural processing (e.g., Bartlett et al., 2003), the observed N170 differences may indicate the differences in facial feature processing. More specifically, we suggest that participants who are immersing in a new ethnic environment recognize unfamiliar faces in a feature-based process, which relies on the comparison of individual facial features such as the eyes, nose or mouth. This claim is supported by previous studies where the recognition of new faces, compared to old ones, relies relatively more on the processing of featural information (e.g., Lobmaier & Mast, 2007). In addition, perceptual training studies indicate that feature-by-feature strategies may improve unfamiliar face matching (Towler et al., 2017; White et al., 2015). Also, similar group differences as in the face-elicited N170 were observed for the control condition in Exp. 2 —doodles which are complex abstract patterns without clear relational/configural information; this observation, although not significant in Exp. 1, may indicate that the social immersion in a different ethnic environment may influence feature analysis rather than configural visual analysis.

Overall, the present findings suggest that the enhanced face-elicited N170 in foreigners as compared to locals might imply the effects of social immersion in facial feature processing rather than configural processing. This suggestion is not consistent with Ma et al. (2022) where the increased face-elicited N170 in Chinese children as compared to German children indicates a superiority of holistic processing in Chinese script readers. This inconsistency may suggest that there are dissociated underlying face processes affected by the script system acquisition and the social immersion experience. In addition, the current findings do not support claims that East-Asian participants show a general superiority of holistic processing as compared to Caucasian participants (e.g., Blais et al., 2008; Markus & Kitayama, 1991; Miyamoto et al., 2011). Since the evidence for this claim is heterogeneous (see introduction), our findings indicate that it may have to be qualified by the factor social/ethnic immersion.

Limitations and perspectives

The present experiment has some shortcomings that might be dealt with in future studies. Firstly, we examined the effects of immersion experience only for a relatively narrow range of immersion durations. Therefore, it would be interesting to explore the time-course of immersion experience effects. Furthermore, this might be combined with, a more direct proof of the immersion account by using a longitudinal design. Secondly, the sample size of Study 1 was smaller than 20 participants per group, which raises the concern of power. However, the critical results, i.e., the main effect of group ($F(1, 34) = 5.52$, $p = 0.025$, $\eta_p^2 = 0.14$) and the significant interaction between group, stimulus orientation and hemisphere ($F(1, 34) = 4.28$, $p = 0.046$, $\eta_p^2 = 0.11$), showed that the effect sizes estimated by η_p^2 were equal or close to 0.14, corresponding to large effects (Miles & Shevlin, 2001). Thus, the sample size of Study 1 seems to be appropriate to reveal the group differences observed. The third limitation of the present study is the absence of a control group, testing whether social immersion in a foreign country (that is, foreign in language and culture) might influence face processing even if ethnicity of the social environment is constant. Future studies should address this issue by comparing local participants with participants who are foreigners but share the same ethnicity as the locals.

The fourth shortcoming in the present study is the lack of high-level performance data – it would be very interesting to see whether the enhanced N170 during other-ethnicity immersion is accompanied by changes in visual performance or whether the enhanced N170 is merely an epiphenomenon of an overexcited fusiform face area without functional consequences. Finally, our study only focused on the effects of social immersion in Asian-Caucasian settings (i.e., Asians in Caucasian countries and vice versa); thus, the results may not generalize to other settings. For instance, African participants who have immersion experience in Caucasian/Asian countries and vice versa, because of potential differences in associated racial attitudes (e.g., positive and negative social attitudes toward specific ethnicities) and race-related physical differences (e.g., the shape of the eye region seems more similar between African and Caucasian faces than it is between Caucasian and Asian faces). In addition, comparing to Asian-Caucasian setting, the African-Caucasian setting would separate the social immersion account from the script system account and the holistic vs. analytic culture account. Therefore, future studies could test whether the social immersion account applies also to other cultural settings such as African participants with immersion experience in Caucasian/Asian countries and vice versa.

The present findings suggest a fascinating application perspective. Given that immersion in a different ethnic environment activates the face processing system, making it more plastic at least for a while, might provide the chance to use other-race faces as training material to improve face-specific abilities.

2.1.5 Conclusions

In the present study we investigated two possible experience effects on face processing at the neural level, a culture-specific effect – possibly due to exposure to the logographic script, and a more specific effect of immersion into a novel social environment. We found larger N170 amplitudes to faces in participants that were foreigners to the lab locations as compared to local participants; importantly, this difference was independent of the location of residence (East-Asian society vs. Western society). In addition, the observed group difference was indistinguishable between same-and other-race faces. We conclude that immersion into a novel environment of different ethnicity may have an impact on the underlying neural mechanism for faces in general, independent of the ethnicity of stimulus faces. Nevertheless, the N170 group effects were unaffected by stimulus orientation, suggesting the observed N170 increase due to social/ethnic immersion does not reflect differences in configural visual processing. These findings offer novel insights into the effect of visual experience on face processing in times of globalization and increasing international exchange and interactions.

2.1.6 Supplementary Materials

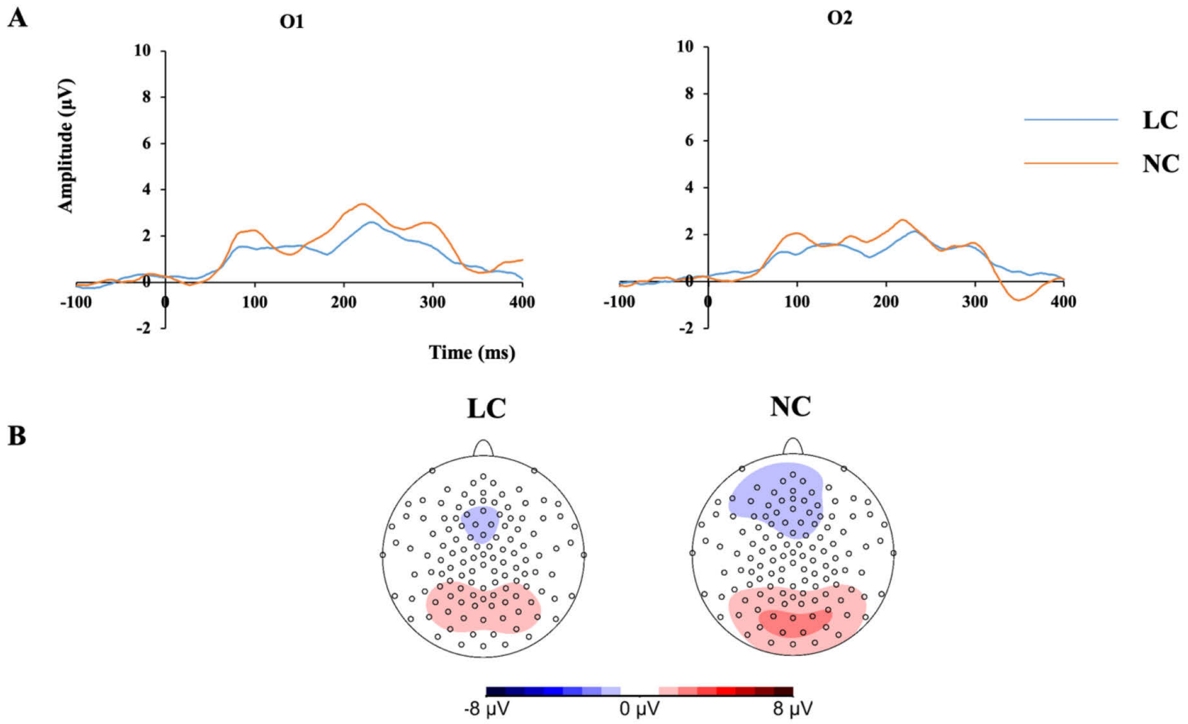


Figure S1. (A) Grand-average ERP waveforms of Chinese local residents (LC, blue) and non-Chinese residents (NC, orange) in Hong Kong elicited by polygons at two occipital sites showing the P1 amplitude (O1, left hemisphere; O2, right hemisphere); (B) Topographies of grand-average peak amplitudes of P1 (108 ms) of LC and NC group in Hong Kong elicited polygons.

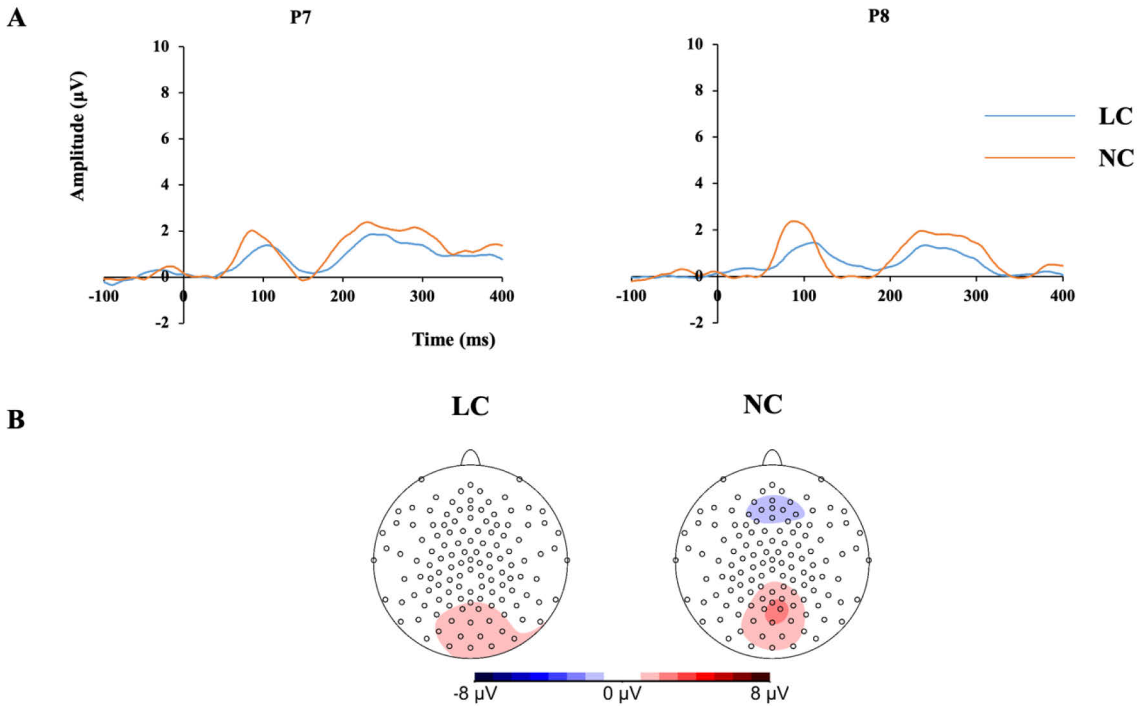


Figure S2. (A) Grand-average ERP waveforms of Chinese local residents (LC, blue) and non-Chinese residents (NC, orange) in Hong Kong elicited by polygons at two occipital sites showing the N170 amplitude (P7, left hemisphere; P8, right hemisphere); (B) Topographies of grand-average peak amplitudes of N170 (154 ms) of LC and NC group in Hong Kong elicited polygons.

2.2 Study 2: Does Learning Different Script Systems Affect Configural Visual Processing? ERP Evidence from Early Readers of Chinese and German

Abstract

Reading is a complex cultural skill requiring considerable training, apparently affecting also the processing of non-linguistic visual stimuli. We examined whether the different visual demands involved in reading different script systems – alphabetic German versus logographic

Chinese script - would differentially influence configural visual processing. Our main dependent measure was the N170 component of the ERP, which is considered as a signature of configural processing. In the present study, German and Chinese children ($N = 28$ vs. 27) who had received about one year of formal instruction in their native script system, worked on a series of one-back tasks with naturalistic faces, two-tone Mooney faces and doodles, and on an adaptation task with pairs of faces were either identical or differed in their second-order relations. Chinese children showed larger N170 amplitudes than German children for naturalistic and Mooney faces, specifically indicating superior holistic processing in Chinese children. In contrast, there was no superiority in Chinese children on the second-order adaptation effect at the N170, providing no evidence for differences in second-order relations processing of facial configurations between the groups. Given the sensitivity of the visual system to reading acquisition, these findings suggest that these group differences in holistic processing might be due to the extensive training with the highly complex logographic script system learned by Chinese children, imposing high demands on higher-order visual perception.

Keywords: face processing, configural processing, ERPs, N170, reading, script system

2.2.1 Introduction

Evidence that cultural experience can alter visual perception has been available for over 100 years (Rivers, 1905), with recent research focusing on differences between Western and East Asian cultures (e.g., Blais et al., 2008; Estéphan et al., 2018; Kuwabara & Smith, 2016; Nisbett & Miyamoto, 2005). Since populations in East-Asian cultures are often required to master a logographic script system, which is much more complex than alphabetic script systems, the differential experience of reading different script systems might contribute to cross-cultural differences. Reading, a relatively recent culturally transmitted skill, must rely on pre-existing visual and cognitive functions, which may be recycled for solving the specific problems posed by reading (Dehaene & Cohen, 2007). Hence, learning to read requires intense exercise of visual functions, evolved for other purposes. Conversely, reading acquisition may induce changes in the processing of non-verbal visual stimuli, especially for faces. However, it is an ongoing debate whether the training of visual skills in reading will involve a positive transfer to other visual domains (faces) or, instead, destructive competition (e.g., Dehaene et al., 2010; Van Paridon et al., 2021). For example, there have been reports about the consequences of reading acquisition on performance in configural face tasks (Cao et al., 2019; Ventura et al., 2013). In the present study we investigated whether acquiring script systems differing in their demands on visual processing will differentially affect configural processing of faces by comparing logographic Chinese and alphabetic German readers.

There are considerable commonalities of written word and face processing. Both kinds of stimuli are members of a large, relatively homogeneous visual object class and require fast and accurate recognition. Skilled readers have gained expertise for word recognition based on massive training, similar to acquiring face expertise (e.g., Maurer et al., 2008; McCandliss et al., 2003; Wong et al., 2012). Both word and face recognition involve configural processing encoding spatial relations between features (e.g., Maurer et al., 2002; Ventura et al., 2019; Wong et al., 2019). Although words and faces seem to compete for representations in the ventral visual cortex (e.g., Cantlon et al., 2011; Centanni et al., 2018; Dehaene et al., 2010), two anatomically neighboring sites in this region, the "visual word form area (VWFA)" and the "fusiform face area (FFA)", respectively, are held to be selective for these stimuli (e.g., Baker et al., 2007; Golarai et al., 2020; Nestor et al., 2013; Pinel et al., 2015; Tarr & Gauthier, 2000). Both areas have been associated with a pronounced negative-going N170 (N1 in some studies) component in the ERP, elicited by both words and faces (e.g., Brem et al., 2006; Henson, 2003; Herrmann et al., 2005) and held to be a neural signature of configural processing (e.g., Eimer et al., 2011; Rossion et al., 2003; Wang et al., 2011).

Logographic Chinese and alphabetic scripts vary greatly in visual complexity as defined by the inventory and complexity of graphemes (Chang et al., 2016). The graphemes correspond to letters and their combinations in alphabetic systems and characters in Chinese script. The set size of the grapheme inventory is much larger for Chinese characters than for alphabetic letters (> 3000 characters vs. 30-40 letters counting both upper and lower case). Chinese characters show a fairly uniform square-shaped outline but are internally organized much more complex (e.g., left-right [咪], top-down [吕], or inside–outside [回]) with up to 50 strokes as compared to the linear, left-right spatial configuration of alphabetic letters within a word (Chen & Kao, 2002). In sum, considering the set sizes of graphemic elements and the complexity of spatial configurations, Chinese and alphabetic scripts differ strongly in their demands on visual processing and memory (e.g., Alvarez & Cavangh, 2004; Xu & Chun, 2006).

Previous studies indicated a stronger association between pure visual skills and logographic Chinese reading as compared to alphabetic script reading (Ho & Bryant, 1999; Huang & Hanley, 1997; Mann, 1985). There is growing evidence that readers of Chinese show advantages in some aspects of visual processing over alphabetic script readers (e.g., Demetriou et al., 2005; Huang & Hanley, 1995; McBride-Chang et al., 2011; Siok et al., 2009) For example, Demetriou et al. (2005) found that Chinese children consistently outperformed Greek schoolchildren in tests of global visuospatial processing. In addition, Yum and Law (2019) found that readers who learned a visually more complex script (e.g., Chinese/Japanese) in early childhood tend to exhibit greater N170 amplitudes to objects than readers of a visually less complex script (e.g., English). Therefore, it may be plausible that differential training of visual perception and memory required for becoming a skilled Chinese reader as compared to an alphabetic script reader might differentially impact non-verbal visual processing, manifesting also in processing other visual stimuli of sufficient complexity, such as faces.

The resemblance between script and face processing is particularly striking for faces and Chinese characters (R. V. Tso et al., 2014; G. Zhou et al., 2012). Similar to faces, Chinese characters are represented graphically at the individual level (i.e., their identities). Although holistic processing of words has been demonstrated in both alphabetic and Chinese expert readers (Ventura et al., 2019; Wong et al., 2011; Wong et al., 2012), we assume that relative to the acquisition of alphabetic script reading, logographic Chinese reading acquisition should involve a relatively greater demand on holistic visual analysis due to the complex spatial configurations of Chinese characters (e.g., Ben-Yehudah et al., 2019; Cao et al., 2015; Mo et al., 2015). Using the complete composite paradigm for face perception (Gauthier & Bukach, 2007), the holistic processing of Chinese characters defined by the obligatory attention to all parts of characters have been demonstrated to be modulated by writing/drawing abilities (Tso et al., 2014, 2021; Tso et al., 2020) which is analogous to the observed pattern in the development of face recognition (Zhou et al., 2012). In line with these findings, it has been suggested that Chinese character reading relies more on low spatial frequencies as compared to alphabetic word reading (Hsiao & Lam, 2013), and that low spatial frequencies support configural processing of faces by conveying global configurational information (e.g., Goffaux et al., 2005; Sergent, 1982). Compared with robustly left-lateralized activation in the ventral occipito-temporal areas in alphabetic word reading, Chinese character reading showed a bilateral activation due to involvement of the right hemisphere (e.g., Bolger et al., 2005). More specifically, Chinese character processing, like face processing, recruits neural systems in the fusiform areas of the right hemisphere, which is sensitive to configural information (e.g., Liu et al., 2009). Because of these similarities between Chinese characters and faces, we expected a greater transfer from reading acquisition to face processing in Chinese than alphabetic reading. Specifically, the substantial training with visually complex logographic Chinese characters might tune the visual system towards configural processing and transfer to face processing, which we term as script system hypothesis.

Some incidental but consistent evidence supports the script system hypothesis. Larger N170 amplitudes to faces were observed in East Asians (mainly Chinese and Japanese) as compared to Western participants (e.g., Gajewski et al., 2008; Herzmann et al., 2011; Vizioli et al., 2010; Wiese et al., 2014). These findings may reflect superiority in readers of logographic Chinese/Japanese kanji characters in configural face processing. Configural processing has been suggested to involve three levels or facets (Maurer et al., 2002): first-order relations (i.e., the arrangement of facial features such as two eyes above the nose and the mouth), holistic processing referred to ‘‘glueing’’ together the facial features into a Gestalt or whole, and the second-order relations or relative metric distances between facial features. A behavioral study found that Japanese excelled over American participants in two kinds of configural processing, holistic perception and sensitivity to second-order relations (Miyamoto et al., 2011). Although suggestive, these studies are inconclusive concerning the script system hypothesis because the participants tested had also learned additional script systems and the studies had not targeted this question.

To test the script system hypothesis, we compared readers of alphabetic German and logographic Chinese in configural processing and, specifically, in terms of its two aspects, both holistic and second-order relation processing (Maurer et al., 2002). We employed a series of one-back tasks and an adaptation task with ERP measures. The one-back tasks used naturalistic faces, Mooney faces (Mooney, 1957), and doodles (complex but meaningless line drawings, see Figure 10). In the adaptation task pairs of adaptor and test faces were either identical or differed in their second-order relations. The tasks were administered to Chinese and German children who had received about one year of formal reading instructions in their native language and acquired basic reading skills, presumably inducing effects in their visual systems (e.g., Cao et al., 2011; Eberhard-Moscicka et al., 2015; Maurer et al., 2006; Walle de Ghelcke et al., 2021). Hence, they had been exposed to different training experiences that might differentially affect configural visual analysis.

Our primary focus was the N170 component in the ERP. According to the script system hypothesis, we expected that Chinese children would show larger N170 responses to naturalistic faces than German children. If the enhanced N170 specifically reflects a superiority in holistic processing, a similar group difference in N170 should be observed for Mooney faces, which have to be processed holistically as there are no separable facial features (Mooney, 1957). Besides, if the sensitivity to second-order relations assessed by adaptation effect, plays a role in the acquired group differences for naturalistic faces, a stronger N170 adaptation effect would be expected in Chinese children. For the control condition with complex abstract patterns without clear relational/configural information, i.e., doodles, we did not expect a group difference.

As a further electrophysiological indicator, we also analysed the P1 component, the occipital positivity preceding the N170. The P1 is considered to reflect basic and early visual processing, for example, low-level physical stimulus properties, such as size, luminance, and contrast, but also spatial attention (see Hillyard et al., 1998 for a review). This component allowed us to assess contributions of low-level processing and attention to any group differences in the following N170.

2.2.2 Method

Participants

The initial sample included 45 German and 37 Chinese children, residing in their native countries. After applying strict exclusion and matching criteria as specified below, the final samples consisted of 28 German and 27 Chinese children (Table 2 for details). Formal reading instruction in both Germany and China begins around the age of six. All German children were native German speakers and received alphabetic German script instruction whereas all Chinese children were native Chinese speakers and received logographic Chinese script instruction with additional instruction in an alphabetic script denoting the pronunciations of Chinese characters called pinyin. Participants were tested around the end of the first grade and the beginning of the second grade. All participants had normal or corrected-to-normal vision. Handedness information was obtained from self-report and/or parent-report using several questions selected from the Edinburgh handedness battery (such as the hand used for writing, drawing, and for using a toothbrush). Prior to the experiment written and verbal informed consent was obtained from the children and their parents. The study was approved by the ethics committee of the departments of Psychology of the Humboldt-University of Berlin and Zhejiang Normal University in accordance with the Declaration of Helsinki.

All participants took tests of fluid intelligence and reading ability. We assessed fluid intelligence with the figural reasoning scale of the Berlin test of fluid and crystallized intelligence for Grades 1-4 (BEFKI-gff; Wilhelm et al., 2014). Each item of the figural reasoning scale consists of a sequence of geometric shapes ordered according to implicit rules. Participants have to infer these rules and choose the next two shapes in the sequence from three alternatives each. Children worked on 11 items of the BEFKI-gff for at most 10 min.

Reading ability in German children was measured with the German 1-min real-word and pseudoword reading fluency test (SLRT-II; Moll & Lander, 2010). In each test, participants have to read aloud as many words as possible from a list of 156 real-words or pseudowords within one minute. In Chinese children reading skills were assessed with the widely applied Chinese character recognition test and the 1-Min reading fluency test (e.g., Lei et al., 2011; Li et al., 2012; Liao et al., 2015). In the former test, children have to name 150 Chinese characters, ordered in increasing difficulty. The reading fluency test requires reading aloud as many two-character Chinese words as possible within one minute from a list of 180. A child is considered as a normal reader if its score in any of these reading tests is not lower than two standard deviations below the mean of the corresponding sample. According to this criterion all of our participants were normal readers (see Table 2). In addition, the reading performance of German children corresponds to the norms (real-word: *M(SD)* = 31.67(16.47); Pseudoword: *M(SD)* = 25.24(9.96)) while the Chinese children's reading performance is comparable to participants in the study by Zhao et al. (2019) where Chinese character recognition in 127 Chinese children in the first half of second-grade was: *M(SD)* = 59.65(23.50)). Hence, the literacy skills for both Chinese and German correspond well with their education level.

Several children were excluded from analysis; the reasons were prematurely terminating the EEG part (*N* = 2), technical problems with EEG data (*N* = 1), excessive movements or problems to pay attention in several experimental blocks (*N* = 2), and performance sensitivity > 2 *SD*s below the sample mean (*N* = 4). Also, to match the groups for age, 11 German children < 7.3 years, 1 German child and 6 Chinese children > 8.1 years were excluded.

Table 2 Descriptive information, reading performance and test statistics of final samples.

	German (N = 28)	Chinese (N = 27)	*t-test* (df = 53)
Age (years)*	7.65 (0.24)	7.76 (0.26)	-1.63
Gender (F/M)	15/13	16/11	/
Handedness (L/R)	4/24	0/27	/
Formal education (months)*	13.56 (1.03)	15.00 (0.60)	-6.31**
Fluid intelligence (items)*	4.96 (1.93)	5.96 (2.49)	-1.67
1-min reading (German words)*	33.86 (20.88)	/	/
1-min reading (pseudo-words)*	25.86 (9.31)	/	/
1-min reading (Chinese words*)**	/	65.04 (19.44)	/
Character recognition (characters)*	/	60.93 (30.59)	/

Note: * *M*(*SD*); ** $p < 0.01$.

Stimuli

We administered five one-back tasks with different stimulus materials and an adaptation task with faces. Stimuli for the one-back task were five sets of 16 grey-scale images, showing standardized naturalistic adult faces, Mooney faces (Mooney, 1957), German words, Chinese characters and doodles (see Figure 10 left panel). Eight naturalistic faces each were of Chinese, and Caucasian adults, in equal numbers female and male with neutral facial expression. The same organization pertained to Mooney faces, generated by increasing the contrast in a separate set of 16 faces until each image was two-tone black and white. The visual angle of naturalistic and Mooney faces was approximately 4.7 ° × 6.3° from a distance of 70 cm. Doodles were black and white line drawings of irregular shapes with visual angle of 5.4 °× 5.6° on average. Since the word conditions do not directly address the current goals, we will not detail them here.

For the adaptation task, 16 grey-scale naturalistic face images of different individuals but standardized in the same way as those for the one-back tasks and also balanced for ethnicity and sex, served as original faces (see Figure 10 right panel). Each original face was manipulated in its second-order relations with Adobe CSC 7 software in four ways: (1) Moving down the eyes by 4 mm (0.5° for the viewing distance of 70 cm) and the mouth by 2 mm (0.2° for the viewing distance of 70 cm), (2) moving the eyes together by 4 mm and moving down the mouth by 2 mm, (3) moving the eyes apart by 4 mm and moving the mouth upward by 2 mm, or (4) moving up the eyes by 4 mm and the mouth by 2 mm. All faces subtended visual angles of 4.7 ° × 6.3°.

Procedure

German and Chinese participants were individually tested by a female experimenter in an EEG laboratory and a separated testing area adjacent to a primary school, respectively. Participants were seated on a comfortable chair 70 cm away from a computer monitor.

Presentation 1.0 and E-prime 3.0 software (Psychology Software Tools, Pittsburgh, PA) controlled the experiments in Germany and China, respectively. In all tasks, stimuli were shown in the center of the screen on a light grey background. The order of the adaptation task and the set of one-back tasks was balanced across participants. The five different one-back tasks were presented about equally often at each possible position.

The series of five one-back tasks was preceded by a practice task with 20 pictures of houses. Each one-back task consisted of 78 trials of a given stimulus category. The first six trials in each task were training items with one immediate repetition, excluded from the final analysis. The 16 stimuli of each category were presented in pseudo-randomized order with the constraint of 8 (11%) immediate repetitions, serving as target trials. Participants were required to press the space button when they saw a repetition immediately. Stimulus duration was 500 ms followed by an interstimulus interval (ISI) that varied randomly from 1250 to 1750 ms (see Figure 10 left panel). In each task a short break was provided after every 26 trials and a longer pause was given at the end.

The adaptation task included 144 trials divided equally into two blocks. On each trial, two images (S1: adaptor face; S2: test face) were presented successively for 500 ms each, separated by a 400-ms interstimulus interval (see Figure 10 right panel). Original and spacing-modified faces were presented equiprobably as S1 or S2. The intertrial interval varied between 1250 and 1750 ms. All combinations of S1 and S2 on the same trials were images of the same individual, with half of the combinations being identical images (adapt condition) and the other half of trials comprising pairs of S1 and S2 stimuli differing in second-order relations (no adapt condition). Participants were instructed to press the space button when the ethnicities of the current face pair and the immediately preceding face pair differed. There were 10% of ethnicity changes in each block. Participants were given practice trials before the adaptation task, and the experimenter made sure that children had understood the task before working at the test trials. A short break was provided after every 18 trials and a longer pause was made between blocks.

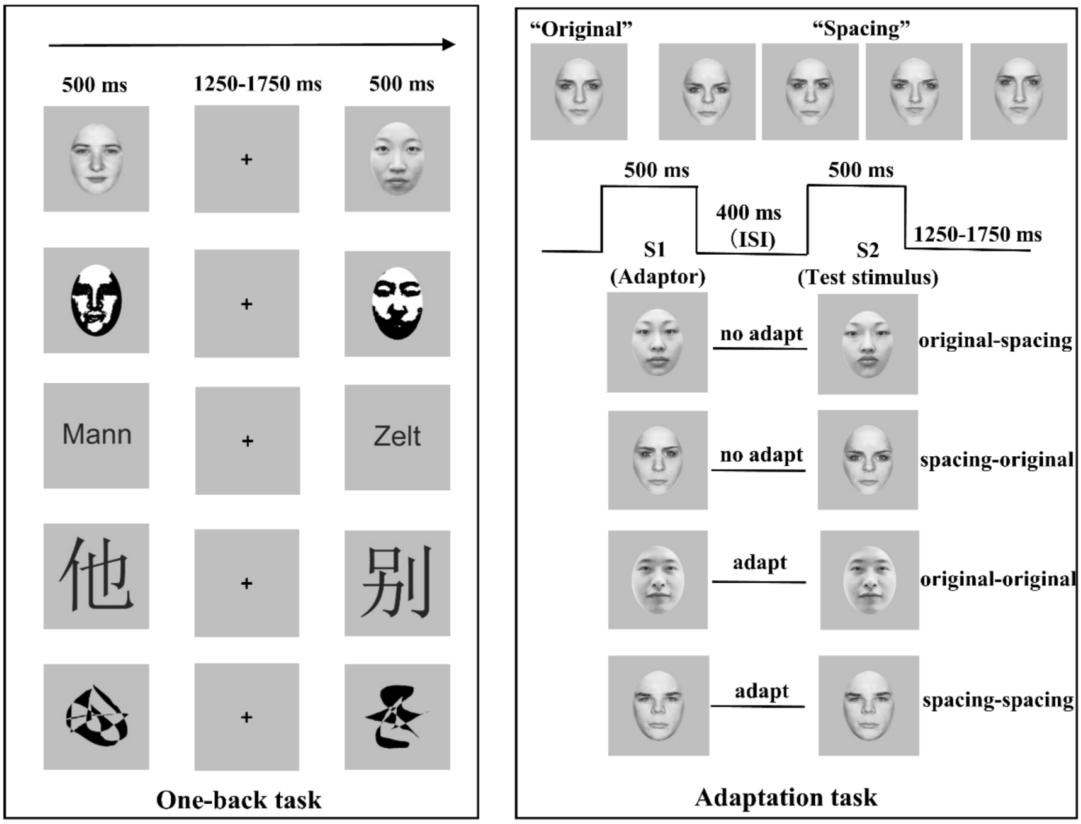


Figure 10. Trial structure and examples of the stimuli for one-back task (left panel) and adaptation task (right panel).

EEG recording and preprocessing

EEG was recorded with 39 Ag/AgCl electrodes placed within an elastic cap (Easycap GmbH, Herrsching, Germany) according to the extended International 10-20 System in Germany and China. Two additional electrodes below each eye measured vertical electroocular (EOG) activity. During recordings all channels were referenced to the Cz electrode; AFz served as the ground electrode. Impedances were kept < 20 kΩ at a uniform level (Ferree et al., 2001). Signals were amplified using Brain Amps (Brain Products) and sampled at 1000 Hz.

Offline, EEG data were re-calculated to common average reference, digitally high-pass filtered at 0.1 Hz (12dB/Octave, zerophase-type filter) and low-pass filtered at 30 Hz (24 dB/Octave, zerophase-type filter), using Brain Vision Analyzer 2.0 software. We analyzed ERPs in all non-target trials, epoched at −100 to 500 ms relative to stimulus onset and baseline-corrected to the average amplitude during the 100-ms pre-stimulus interval. Trials were rejected if (1) voltages exceeded ± 100 μV at any electrode, if (2) the absolute difference between two adjacent sample points exceeded 75 µV, or (3) if the voltage range of a given channel within an epoch exceeded 150 µV. After artifact rejection, the average of all stimulus conditions in each task for each child included at least 15 epochs. A mixed ANOVA with a between-subject factor children group and a within-subject factor stimulus condition revealed that Chinese children (*M* = 50.57, *SD* = 8.43) had higher average number of accepted trials across stimulus conditions than German children (*M* = 44.42, *SD* = 7.98), $F(1,53) = 7.72$, $p < 0.01$, $\eta_p^2 = 0.13$.

Previous studies reported slow-wave activities, especially pronounced in children, overlapping with early ERP components (N1, P2, N2) and affecting their amplitudes and topographies (Dimoska et al., 2007; Johnstone et al., 2003; Johnstone & Barry, 1999). Because our main component of interest, the N170, might be distorted by such slow-wave activity we applied an additional 2 Hz high-pass filter (12dB/Octave, zerophase-type filter) to remove the slow-wave. These “high-pass filtered” ERPs (bandpass 2-30 Hz) will be distinguished from the “broad-band” ERPs before high-pass filtering (bandpass 0.1-30 Hz).

Data analyses

Behavioral analyses

Behavioral performance was assessed as hit rate, false alarm rate, sensitivity, and reaction time (RT). Hit rate was calculated by using the number of target trials, which were correctly judged as targets divided by the total number of target trials. False alarm rate was calculated as the number of non-target trials judged as targets divided by the total number of non-targets. Sensitivity was calculated as hit rate minus false alarm rate. RTs were measured for the correctly judged target trials. However, there were five children (one German, four Chinese) who missed all targets in one or two tasks, resulting in missing RT data. Since these children had always looked at the stimuli during EEG recording as noted by the experimenter, they were still included in the final analyses. The missing RT data in these children were imputed by an R-package titled Multivariate Imputation by Chained Equations (MICE). The MICE algorithm uses RT data from other tasks and other children to predict and impute missing RT data.

We performed separate unpaired *t*-tests with IBM SPSS Statistics 24 to examine differences between German and Chinese children on hit rates, false alarm rates, sensitivity, and reaction times.

EEG analyses

Topographic pattern analysis

Our study aimed to compare amplitude differences of ERP components between German and Chinese children. However, when examining component amplitudes at a few predefined electrodes, possible differences in electrical field topography must be taken into consideration (Lehmann & Skrandies, 1980; C. M. Michel et al., 1999, 2001). Topographic differences provide important additional information about the spatiotemporal dynamics of visual processing, since they may reflect differences in ERP component generator configurations. To address this question, we applied topographic pattern analysis (spatiotemporal segmentation, Brunet et al., 2011) to separate the ERP data into a limited number of topographical map configurations. A map configuration is a period during which the topography of the electrophysiological activity on the scalp remains stable. Any change of the spatial configuration of the electric field on the scalp reveals a difference between configurations of the underlying intracranial sources. Topographic pattern analysis is insensitive to pure amplitude modulations across conditions (i.e., topographies of normalized maps are compared) and independent of the reference electrode (Brunet et al., 2011). In addition, the influence of slow-wave activity was assessed by comparing topographic pattern analyses for broad-band and high-pass filtered ERPs.

Topographic pattern analysis was conducted in the grand mean ERP data (i.e., from 0 to 400 ms post-stimulus onset) for each stimulus condition and group with topographic atomize and agglomerate hierarchical clustering (T-AAHC) implemented in Cartool software (Brunet et al., 2011). A given topography had to be present for at least 30 ms and the maximum correlation between different topographies should not exceed 95% (temporal smoothing, see Brunet et al., 2011). Modified cross-validation (CV) and Krzanowski-Lai (KL) criteria were used to determine the optimal number of maps best explaining the group-averaged data sets across stimulus conditions and groups (Brunet et al., 2011; Murray et al., 2008). In the next step, the dominant maps identified in the group-averaged data were then fitted to the ERPs of each participant using spatial fitting procedures, to quantify their representation across subjects and conditions (Brunet et al., 2011). This procedure allowed us to establish how much of the global variance (GEV) of one condition/group was explained by a given segmentation map.

ERP component analyses

We parameterized the ERP components according to the topographic pattern analysis. Using IBM SPSS Statistics 24, we conducted mixed-design analyses of variance (ANOVA) on ERP component amplitudes and latencies. There were four sets of analyses investigating different questions. Firstly, to investigate the group differences on global configural processing between logographic Chinese and alphabetic German script readers, we performed ANOVAs with a between-subject factor children group (German, Chinese), and within-subject factors hemisphere (left, right) and task (one-back, adaptation) on ERPs to naturalistic faces including faces in the one-back task and adaptors in the adaptation task. Secondly, we performed ANOVAs with a between-subject factor children group (German, Chinese) and a within-subject factor hemisphere (left, right) on ERPs to Mooney faces in one-back task to examine the group differences between two script readers on holistic processing. Thirdly, to examine the group differences between two script readers on the sensitivity to second-order relations during face processing, we conducted ANOVAs with a between-subject factor children group (German, Chinese), and within-subject factors hemisphere (left, right) and adaptation condition (adapt, no adapt) on ERPs to the test faces (S2) in the adaptation task. The sensitivity to second-order relations was measured by the adaptation effect referring to reduced responses for S2 when it is preceded by an identical face with the same second-order relations (adapt condition) relative to when preceded by an identical face but with different second-order relations (no adapt condition). Lastly, we conducted ANOVAs with a between-subject factor children group (German, Chinese) and a within-subject factor hemisphere (left, right) on ERPs to doodles to examine whether any effects observed in the face conditions would also extend to unfamiliar complex visual stimuli that are hard to process configurationally.

2.2.3 Results

Behavioral results

Figure 11 shows the behavioral results for the one-back tasks and adaptation task. T-tests revealed that Chinese children (M = 899.04 ms, SD = 220.70) showed significantly shorter reaction time in naturalistic face one-back task than German children (M = 1012.46 ms, SD = 195.30), $t(53) = 2.02$, $p < 0.05$, Cohen's d = 0.53. In the adaptation task German children (M = 0.71, SD = 0.28) showed significantly higher sensitivity than Chinese children (M = 0.41, SD = 0.40), $t(53) = 2.55$, $p < 0.05$, Cohen's d = 0.67. In contrast, German children (M = 0.07, SD =0.09) showed lower false alarm rate in adaptation task than Chinese children (M = 0.17, SD = 0.20), $t(53) = -2.48$, $p < 0.05$, Cohen's d = 0.63. No other group comparisons reached significance (all $ps > 0.05$).

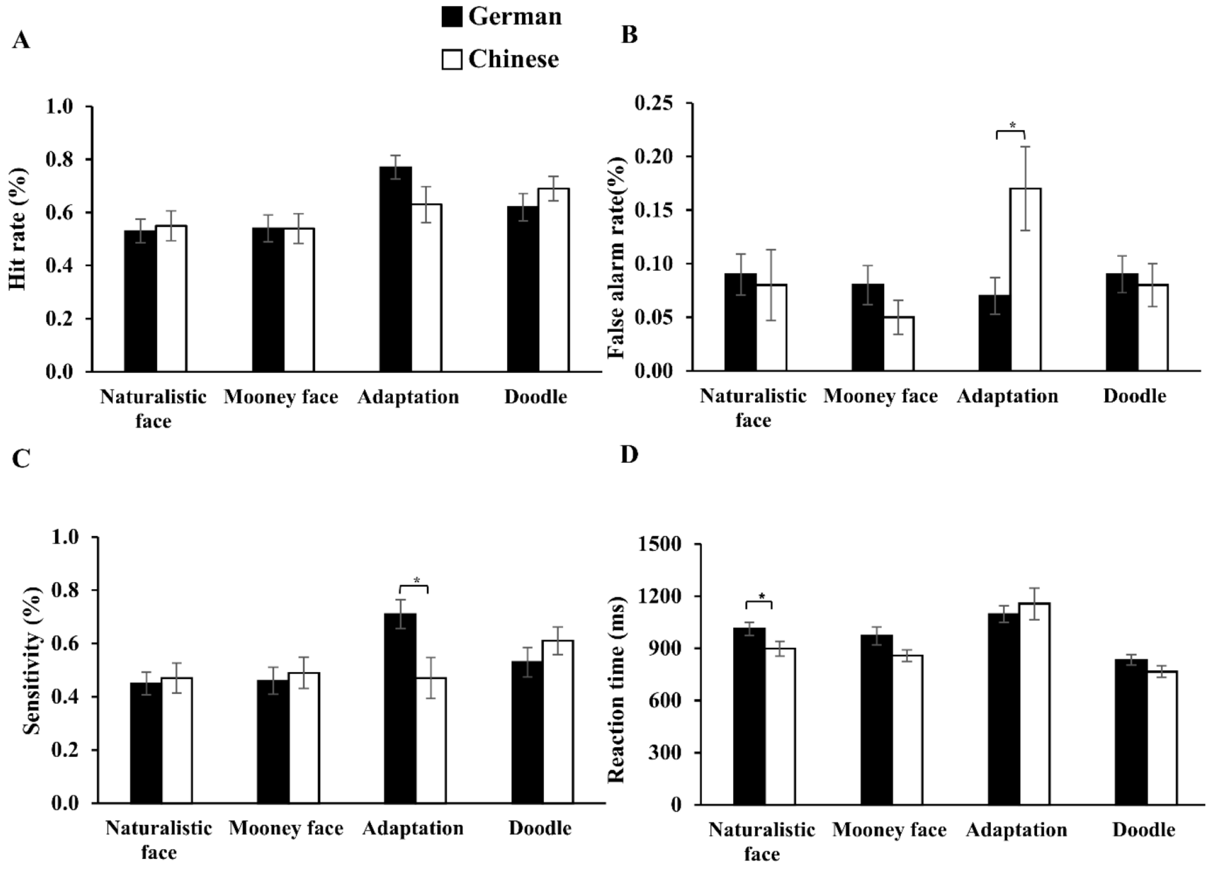


Figure 11. Behavioral results on (A) hit rates, (B) false alarm rates, (C) sensitivity, and (D) reaction times in one-back tasks (naturalistic face, Mooney face, and doodle) and adaptation task. Error bars indicate the standard error of the means. $^{*}p < 0.05$.

Topographic pattern analyses

Figure 12A and 12B illustrate the results of topographic pattern analyses on the broad-band ERP for German and Chinese children in different stimulus conditions. The results revealed that German and Chinese children showed different topographies corresponding to the P1 component in the time window between 70 and 160 ms (Map 6 vs. Map 5), although maps were consistent across all conditions within each group. By contrast, the topographies corresponding to the N170 time range were heterogeneous. Map 3 in the time interval from 160-230 ms, corresponding to N170 was identified in ERPs to naturalistic faces in both tasks and both groups. However, its scalp distribution with positivity in the posterior region is untypical for the N170, and its map strength is quite small. Furthermore, the topographies of

Mooney face in the N170 range in German and Chinese were different (Map 4 vs. Map 7). In the doodle condition both German and Chinese children showed only one unitary but group-specific topography during the whole epoch (Map 6 vs. Map 5). Because of these peculiarities, we eliminated the slow activity overlapping the P1 and N170 by high-pass filtering and reapplied topographic pattern analysis.

As in broad-band ERPs, the results of high-pass filtered ERP (see Figure 12C and 12D) showed different but within-group consistent scalp topographies of P1 for German and Chinese children (Map 6 vs. Map 7) between about 70 and 160 ms. However, topographies in the N170 range were now much more consistent than in broad-band ERPs. Both groups showed the same map (Map 4) between about 160 and 260 ms, corresponding to the N170 for naturalistic faces, Mooney faces and S2-faces in the adaptation task. Besides, this highly consistent map in the N170 range was in line with typical N170 topographies and showed more pronounced strength than that in the broad-band ERPs. In addition, different scalp distributions appeared for the doodle condition between German and Chinese children (Map 3 and 5 vs. Map 4 and 8) between 160 and 260 ms.

The above observations strongly suggest that the slow-wave shift in the broad-band ERP data is responsible for the heterogeneous N170 topographies and their small strengths. Figure S3 in the Supplementary materials shows the grand average waveforms of broad-band ERP and high-pass filtered ERP to illustrate the effect of slow-wave activity on the N170 component. As a consequence, the fitting procedure and statistical analyses were conducted only on high-pass filtered ERPs.

Statistical analysis was performed using ANOVAs of GEV in two time periods. For the first period (70-160 ms), corresponding to the P1 component, we performed ANOVA on GEV with a between-subject factor children group (German, Chinese) and within-subject factors stimulus condition (naturalistic face_one-back, naturalistic face_adaptor, Mooney face, S2_adapt, S2_no adapt, doodle) and map (map 6, map 7). Results showed that the interaction of children group and map was highly significant ($F(1,53) = 51.86, p < 0.001, \eta_p^2 = 0.50$). Post-hoc analysis revealed that German children showed larger GEV for map 6 and smaller GEV for map 7 than Chinese children (all $ps < 0.001$). Also, post-hoc analysis indicated that the GEV for map 6 was larger than for map 7 in German children whereas the GEV for map 7 was larger than for map 6 in Chinese children (all $ps < 0.05$). Thus, the representative map for P1 for German and Chinese children was Map 6 and Map 7, respectively. For the second period (161- 260 ms), corresponding to the N170 component, a mixed ANOVA with a between-subject factor children group (German, Chinese) and within-subject factors stimulus condition (naturalistic face_one-back, naturalistic face_adaptor, Mooney face, S2_adapt, S2_no adapt, doodle) and map (map 3, map 4, map 5, map 8) on GVE was conducted. No significant effect involving the factor group was observed (all $ps > 0.05$). The interaction between stimulus condition and map was significant ($F(1,53) = 24.31, p < 0.001, \eta_p^2 = 0.31$). Post-hoc analysis indicated that the GEV for map 4 (i.e., typical N170 topography) was larger than any other maps across stimulus conditions (all $ps < 0.001$) except the doodle condition (all $ps > 0.05$). Therefore, the dominant map for N170, that is map 4, was indistinguishable for German and Chinese children across stimulus conditions, except for doodles.

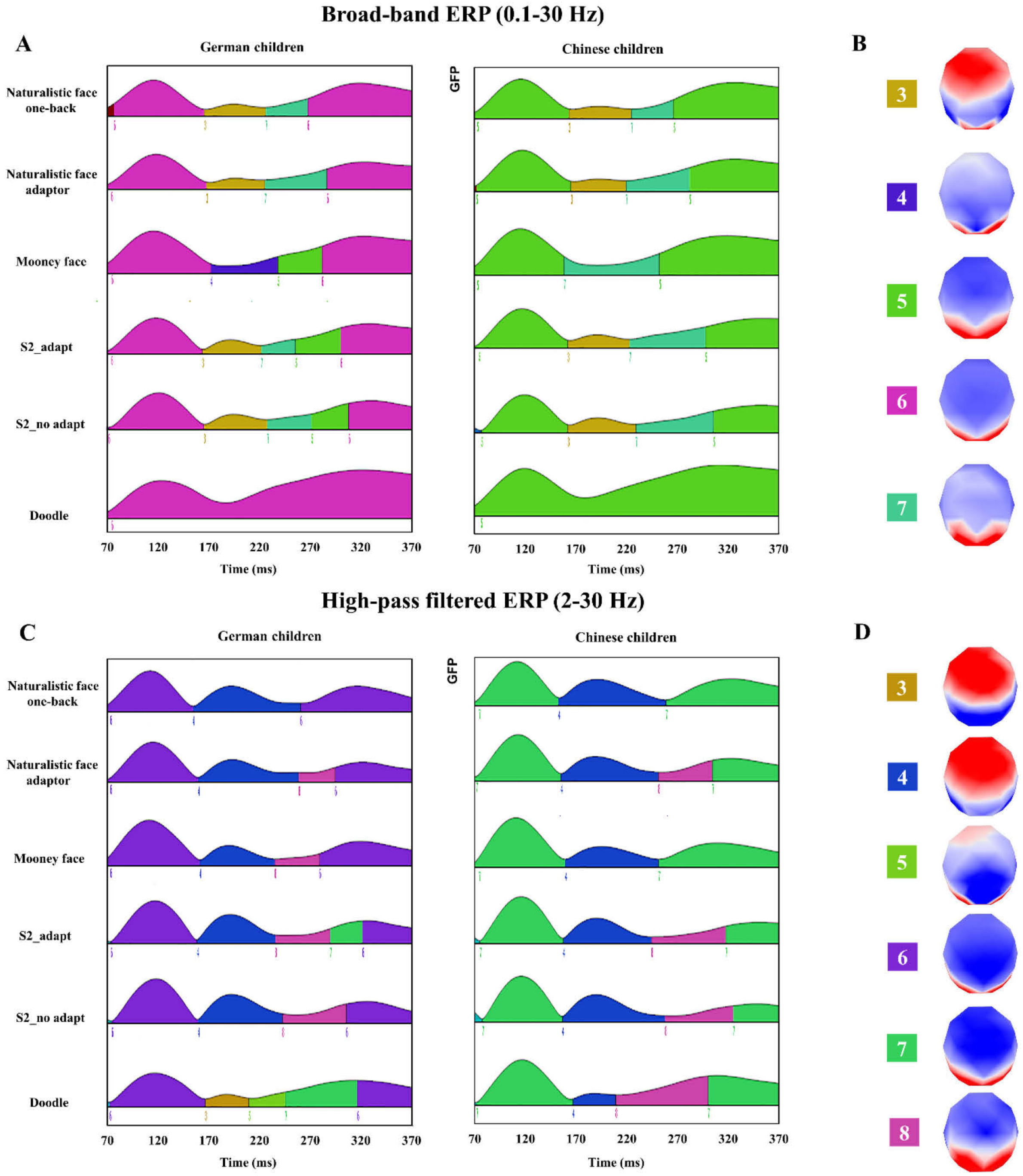


Figure 12. Global field power (GFP) and results of the topographic pattern analysis performed on grand average broad-band ERPs (0.1-30 Hz, A and B) and high-pass filtered ERP (2-30 Hz, C and D). A and C: the temporal distribution of the template map is displayed with different colors for German and Chinese children (left and right sides in panels A and C) between 70 and 370 ms. B and D: the corresponding template maps are displayed with positive voltages

in red and negative voltages in blue. The template map numbers (and their background colors at the left) correspond to the temporal distribution.

ERP component analyses

The topographic pattern analysis revealed different topographies of the P1 component in German and Chinese children (see Figure 12C and 12D) across all stimulus conditions, suggesting different neural generator configurations between groups. A likely source for the difference in ERP topography stems from the differences in head shapes/brain anatomies in Chinese and German children (Ball et al., 2010; Cuffin, 1990; Tang et al., 2018). Because N170 topographies did not differ between groups we can rule out systematic differences in electrode placement between the groups/labs. When surface activities in a certain time interval are distributed differently, as was the case for the P1 time range, amplitude comparisons are close to meaningless. Amplitude differences at a given electrode site might reflect the difference in component strength or topography. In addition, using the maximum across the scalp or global field power is compromised by the unavoidable incomplete coverage of the head surface, again confounding strength and topography. As a consequence of this dilemma, we refrained from further analyzing the P1 components.

Since the German and Chinese children showed indistinguishable topographies in the time period corresponding to the N170 component (see Figure 12C and 12D), we felt justified to compare N170 amplitudes between German and Chinese children at specific electrode sites. However, the N170 was very small for the doodle condition, and peak measures were not reliable in individual participants. Therefore, we parameterized the N170 only for the face conditions. It was measured at the occipitotemporal electrode sites P7 (left hemisphere) and P8 (right hemisphere) frequently used in other studies (e.g., Dundas et al., 2014; Eimer et al., 2011; Latinus & Taylor, 2006; Rossion & Jacques, 2008) and where N170 amplitudes were most prominent in the present data. The N170 amplitude at P7 and P8 was averaged within a ±30 ms window centered on the most negative peak voltage detected automatically between 140 and 300 ms after stimulus onset in each hemisphere. The N170 latency was quantified at the latency of the most negative peak voltage at each hemisphere.

Naturalistic faces

To examine the group differences on global configural processing, we performed the ANOVAs with a between-subject factor group (German, Chinese), and within-subject factors hemisphere (left, right) and task (one-back, adaptation) on N170 amplitudes and latencies to naturalistic faces, including faces of the one-back task and adaptor faces in the adaptation task.

For N170 amplitude (see Figure 13A), the ANOVA revealed a significant main effect of group ($F(1,53) = 5.39$, $p < 0.05$, $\eta_p^2 = 0.09$). Chinese children ($M = -6.84$ µV, $SD = 3.29$) showed larger N170 amplitudes to naturalistic faces, regardless of task, than German children ($M = -5.03$ µV, $SD = 2.43$). The main effect of hemisphere was also significant ($F(1,53) = 7.62$, $p < 0.01$, $\eta_p^2 = 0.13$). The N170 amplitude was larger in the right ($M = -6.52$ µV, $SD = 3.88$) than in the left hemisphere ($M = -5.33$ µV, $SD = 2.86$), independent of group. To make sure that any observed group differences are not stimulus face race-specific or have to be attributed to visual differences between own- and other-race faces, we conducted a further ANOVA with factors group hemisphere, task and race of faces on N170 amplitudes. The results again yielded the main effect of group ($F(1,53) = 5.64$, $p < 0.05$, $\eta_p^2 = 0.10$), which was not modulated by race of faces (all *ps* > 0.05), indicating that the group differences for the face N170 are independent of any stimulus ethnicity effects.

N170 latency yielded a significant interaction between task and group ($F(1,53) = 6.30$, $p < 0.05$, $\eta_p^2 = 0.11$). The post-hoc tests indicated that German children showed significantly longer N170 latency to adaptor faces in the adaptation task ($M = 206.91$ ms, $SD = 22.72$) than Chinese children ($M = 192.64$ ms, $SD = 19.77$, $p < 0.05$). Besides, for German children the N170 latency to adaptor faces in the adaptation task ($M = 206.91$ ms, $SD = 22.72$) was significantly longer than to naturalistic faces in the one-back task (*M* = 194.69 ms, *SD* = 19.77, $p < 0.05$).

Mooney faces

To examine the group differences on holistic processing, we performed ANOVAs with a between-subject factor group (German, Chinese) and a within-subject factor hemisphere (left, right) on N170 amplitudes and latencies to Mooney faces in one-back task. For the N170 amplitude (see Fig 4B), there was a significant main effect of group ($F(1,53) = 4.35$, $p < 0.05$, $\eta_p^2 = 0.08$). Chinese children (*M* = -5.94 µV, *SD* = 3.87) showed significantly larger N170 amplitudes to Mooney faces than German children (*M* = -4.12 µV, *SD* = 2.47). There was no significant effect in N170 latency to Mooney faces (all *ps* > 0.05).

Adaptation effect

The adaptation effect refers to reduced responses to test faces (S2) when preceded by an identical face (adapt condition) relative to the same face but with altered second-order relations (no adapt condition). The adaptation effect was used to assess the sensitivity to second-order relations. Thus, to examine the group differences on the sensitivity to second-order relations during face processing, we conducted ANOVAs with a between-subject factor children group (German, Chinese), and within-subject factors hemisphere (left, right) and adaptation condition (adapt, no adapt) on N170 amplitudes and latencies to test faces in the adaptation task.

We also observed a significant main effect of hemisphere ($F(1,53) = 10.34$, $p < 0.01$, $\eta_p^2 = 0.16$), indicating that the N170 amplitude was more negative in the right (*M* = -7.23 µV, *SD* = 5.31) than in the left (*M* = -5.54 µV, *SD* = 3.00), independent of group. Importantly, no significant effects involving the adaption condition for N170 amplitudes and latencies reached significance (all *ps* > 0.05) although N170 amplitudes to test faces in the adapt condition (*M* =

-6.24 μV, *SD* = 3.30) was numerically smaller than in the no adapt condition (*M* = -6.53 μV, *SD* = 3.18).

However, the absence of adaptation effect might be due to the selection of electrodes. According to the N170 topography for adaptation effect (no adapt minus adapt) in the average ERPs of two groups (see Figure 13C bottom), P9 (left hemisphere) and P10 (right hemisphere) showed the most pronounced adaptation effect. Thus, additional ANOVAs with a between-subject factor group (German, Chinese), and within-subject factors hemisphere (left, right) and adaptation condition (adapt, no adapt) on N170 amplitudes and latencies measured at P9/P10 electrodes was performed.

The results at P9/P10 (see Figure 13C middle panel) showed a significant main effect of adaptation condition on N170 amplitude ($F(1,53) = 5.52$, $p < 0.05$, $\eta_p^2 = 0.09$). The N170 amplitude to test faces in the adapt condition (*M* = -6.16 μV, *SD* = 3.26) was reduced relative to the no-adapt condition (*M* = -6.72 μV, *SD* = 3.60). The main effect of group was significant ($F(1,53) = 7.25$, $p < 0.01$, $\eta_p^2 = 0.120$). German children (*M* = -7.56 μV, *SD* = 3.38) showed larger N170 amplitude to test faces than Chinese children (*M* = -5.28 μV, *SD* = 2.86). A significant main effect of hemisphere was also observed ($F(1,53) = 4.82$, $p < 0.05$, $\eta_p^2 = 0.08$). The N170 amplitude to test faces was larger in right hemisphere (*M* = -6.70 μV, *SD* = 4.30) than left hemisphere (*M* = -5.88 μV, *SD* = 3.30), independent of group. There were no significant effects on N170 latency (all *ps* > 0.05).

Additional planned ANOVAs with within-subject factors hemisphere (left, right) and adaptation condition (adapt, no adapt) on the N170 amplitudes and latencies were conducted for each group. In German children there was a significant main effect of adaptation condition ($F(1,53) = 4.40$, $p < 0.05$, $\eta_p^2 = 0.14$). German children showed reduced N170 amplitudes to test faces in the adapt condition (*M* = -7.11 μV, *SD* = 3.48) as compared to that in the no adapt condition (*M* = -8.00 μV, *SD* = 3.66). In contrast, no significant main effect was observed in

Chinese children ($F(1,53) = 1.26$, $p > 0.05$, $\eta_p^2 = 0.05$) although N170 amplitudes to test faces in the adapt condition ($M = -5.17$ μV, $SD = 2.73$) was numerically smaller than that in the no adapt condition ($M = -5.39$ μV, $SD = 3.08$). For N170 latency, no significant effect was observed (all *ps* > 0.05).

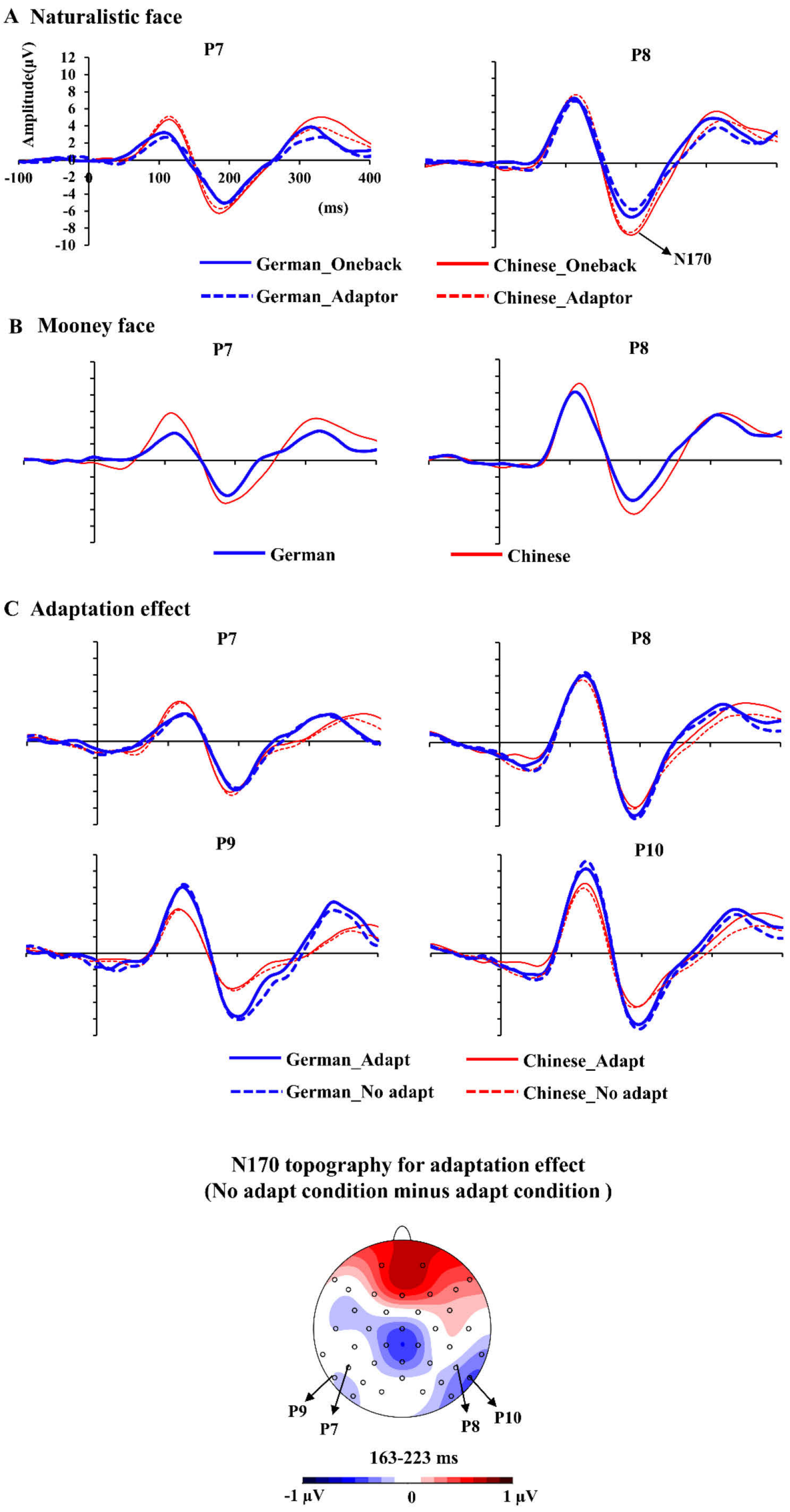

A Naturalistic face
P7
P8
Amplitude(µV)
12
10
8
6
4
2
0
-2
-4
-6
-8
-10
-100
0
100
200
300
400
(ms)
N170
German_Oneback
German_Adaptor
Chinese_Oneback
Chinese_Adaptor
B Mooney face
P7
P8
German
Chinese
C Adaptation effect
P7
P8
P9
P10
German_Adapt
German_No adapt
Chinese_Adapt
Chinese_No adapt
N170 topography for adaptation effect
(No adapt condition minus adapt condition)
P9
P7
P8
P10
163-223 ms
-1 µV
0
1 µV

Figure 13. A and B: the grand average high-pass filtered ERP waveforms for German and Chinese children of Naturalistic faces (A), Mooney faces (B) measured at P7/P8 electrodes. C: the grand average high-pass filtered ERP waveforms for German and Chinese children of S2 faces recorded at P7/P8 (top panel) and P9/P10 (middle panel) respectively; bottom panel: N170 topographies of adaptation effect (no adapt minus adapt) from 163 to 223 ms for the average ERPs of German and Chinese group.

2.2.4 Discussion

Reading acquisition is known to affect the visual processing of non-linguistic visual stimuli. Here we investigated whether these effects would be different after learning script systems with different demands on visual and mnemonic processes, such as logographic Chinese versus alphabetic German. We tested German and Chinese children, who had received about one-year of formal reading education, on different aspects of configural processing of face stimuli, using the N170 component of the ERP as the dependent measure. As compared to German children, Chinese children showed larger N170 amplitudes to naturalistic faces and Mooney faces. In contrast, the adaptation effect to second-order relations at N170 did not differ between the groups. Together, these findings argue for a differential effect of script systems on the holistic processing of non-linguistic visual stimuli, as discussed below.

A holistic processing advantage in Chinese children

The present study showed that the N170 amplitude to both naturalistic faces of the one-back task and the adaptor faces of the adaptation task was larger in Chinese than in German children, indicating that Chinese children show superior configural face processing over German children. Since the findings with naturalistic faces cannot clarify which specific aspect of configural face processing (Maurer et al. 2002) contributes to the observed group difference on N170, the results from the other face tasks have to be considered. For Mooney faces a similar group difference in N170 amplitude was observed as for naturalistic faces, with larger amplitudes in Chinese than in German children. This finding suggests that the group difference in configural processing can be specified to the level of holistic processing. In contrast, in the adaptation task, the N170 adaptation effect to second-order relations was significant in German but not in Chinese children, indicating no superiority of Chinese children in second-order relations processing in faces. We may therefore conclude that the Chinese children's superiority in configural processing as revealed by the N170 to naturalistic faces can be more specifically attributed to the processing holistic aspects rather than second-order relations.

In addition, although previous studies suggested that the facial physiognomic information contributes to the strategies of face processing (e.g., Wang et al., 2015), the present differences between Chinese and German participants cannot be explained by differences of facial physiognomic information between Chinese and Caucasian faces. We used the same stimuli in both experiments, consisting in equal parts of same-and other-race faces. Moreover, the results of the analysis provide direct evidence that the observed group differences were not modulated by face ethnicity.

In general, the observed effects on the N170 component to faces are consistent with previous evidence that the N170 is sensitive to configural processing, including the level of holistic and second-order relations processing (e.g., Eimer et al., 2011; Vakli et al., 2014). One should mention that Chen et al. (2013) found an earlier neurophysiological correlate of holistic processing in the P1 component. This contrasts with the absence of group effects of the P1 component in the present study. Since Chen et al. (2013) had examined holistic processing to Chinese characters in adults, the discrepancy to the present findings indicating holistic processing of faces in children is difficult to account for. Future developmental studies might track the trajectory of neural mechanisms to holistic processing in words and faces.

The script system hypothesis

The present findings support the script system hypothesis we proposed. Enhanced N170 negativity is generally taken to reflect higher visual expertise for certain stimuli (Tanaka & Curran, 2001), which in many cases goes along with more pronounced configural perception (e.g., Gauthier et al., 2003; Rossion et al., 2002). Accordingly, the enhanced N170 in Chinese children suggests that they show more expertise in perceiving visual stimulus configurations. This suggestion is also supported by their performance in the one-back task, where Chinese children were faster than German children without loss of accuracy. Chinese characters have many similarities with faces (Liu et al., 2009) and their complex spatial configurations bias Chinese readers towards holistic processing during reading acquisition as compared to alphabetic readers (e.g., Ben-Yehudah et al., 2019; Demetriou et al., 2005). Thus, the larger N170 to both naturalistic and Mooney faces in Chinese children, indicating more pronounced holistic face processing, might be ascribed to a positive transfer to face processing from their extensive training in holistic visual processing provided by reading acquisition in Chinese.

However, in explaining the group differences in holistic face processing one should also consider an alternative account, that it might be due to the suppression of holistic processing following the reading acquisition of alphabetic languages. Alphabetic readers have been demonstrated to prefer analytic orthographic coding (e.g., Ben-Yehudah et al., 2019). Portuguese illiterates showed more holistic processing of both face and house stimuli than Portuguese literates, suggesting that literacy in alphabetic languages may diminish holistic processing skills (Ventura et al., 2013). Thus, reading acquisition of alphabetic script involving extensive training of analytic processing might induce a generic shift in the ability to deploy analytic visual processing on face processing, resulting in less holistic processing in alphabetic readers than Chinese readers. Although we cannot with certainty distinguish between these two accounts, our results clearly demonstrate a superiority of Chinese children in holistic face processing compared to German children after about one year of formal reading acquisition. Future longitudinal studies starting at preschool age could be conducted to investigate the positive or the negative effects of acquiring different reading systems on holistic visual processing.

The script system account would be supported by correlations between literacy skills and face-elicited ERPs. We calculated Pearson correlations to measure the association of reading skills and N170 amplitudes to faces and Mooney faces in German and Chinese children separately. In general, the correlations were negative, hence in the direction of better reading skills being associated with larger (more negative) N170 amplitudes. However, all correlations were small (all *rs* > -0.30) and not significant in either group (all *ps* > 0.05). It is worth noting, however, that compared to the very small correlations in German children (all *rs* > -0.21), Chinese children showed a close to moderate negative correlation ($r = -.29$) between reading speed (1-minute reading score) and N170 amplitude to Mooney faces in the right hemisphere. Hence, although the observed correlations did not provide conclusive evidence for the script system account, this may be due to the small sample size and it may be premature to ignore the multiple tendencies towards an association. Thus, future studies could further examine the correlations between reading abilities and ERPs with larger samples.

Interestingly, our results in the adaptation task are at variance with a behavioral finding that Chinese and Japanese participants (logographic readers) were better in detecting differences in second-order relations than American participants (alphabetic readers) (Miyamoto et al., 2011). The inconsistency might be due to participants' age, children in the present study versus adults in the experiment of Miyamoto et al. (2011). Developmental studies indicate that the ability to process second-order relations improves at least until age 10 (e.g., Baudouin et al., 2010; Mondloch et al., 2002, 2004). Consequently, German and Chinese children aged 7-8 years, as investigated in the present study may have still immature sensitivity to second-order relations and show different developmental trajectories, leading to contrarian results at different ages. Future developmental studies could track the developmental trajectory of sensitivity to second-order relations in different script readers.

Brain accommodation to specific visual demands of Chinese and alphabetic scripts

We found enhanced N170 responses to faces and Mooney faces over visual occipitotemporal regions in Chinese children as compared to German children, apparently reflecting a flexible accommodation (Perfetti et al., 2007) or recycling (Dehaene & Cohen, 2007) of the visual system to the differential challenges posed by the two script systems studied in the present experiments. According to the accommodation hypothesis, the differences between readers of the two script systems may reflect how the brain optimizes visual processing by accommodating specific properties of the script systems (Perfetti et al., 2007). The neural recycling hypothesis suggested that reading relies on using pre-existing neural systems for vision, which may be employed ("recycled") for the specific problems posed by reading (Dehaene & Cohen, 2007). Pre-existing visual processing skills involved in the processing of other non-verbal visual stimuli are overlearned with extensive training during reading acquisition and, in turn, induce structural and functional changes in the corresponding brain areas (see Dehaene et al., 2015 for a review). Thus, differential visual processing skills required for Chinese and alphabetic reading are accompanied by different neurocognitive accommodation processes, and therefore learning to read different script systems can induce differences in visual processing.

More specifically, the present results indicate that the tuning of the neurocognitive system towards holistic processing is more pronounced by reading acquisition of the Chinese script system than an alphabetic script system. Reading Chinese appears to bias the neurocognitive system towards the holistic spatial analysis of Chinese characters due to their complex and usually square-shaped spatial configurations (e.g., Ben-Yehudah et al., 2019; Kao et al., 2010; Wang et al., 2011). As a result, because Chinese children are massively exposed to Chinese characters during reading acquisition, the accumulation of holistic visual processing expertise might in turn increasingly tune the children's visual system towards holistic processing, resulting in enhanced N170 to both naturalistic and Mooney faces in Chinese children as compared with German children.

Limitations and perspectives

Our study has several potential limitations. First, as in many cross-cultural studies on young participants, it is challenging to ascertain the equivalence of testing situations and equipment. The testing environment was not perfectly matched between German and Chinese children. Our Chinese children were tested in their school where lighting conditions were more difficult to control and less standardized than for the German children who were tested in an EEG lab. However, our group differences in N170 amplitude were condition-specific and not global as one would expect if due to differences in testing condition (e.g., Johannes et al., 1995).

Second, the results of this study are limited by somewhat incomplete matching of the groups. The Chinese group had received significantly longer formal education duration than German children, although the difference was only 1.5 months. On the other hand, the group differences in N170 amplitudes to faces cannot be explained by differences in education duration as there were no significant correlations between education duration and face N170 amplitudes in the German (r = -0.021, p > 0.05) or Chinese (r = 0.116, p > 0.05) group. The socioeconomic background (SES) between German and Chinese groups was also not matched. The Chinese children were recruited from a primary school in a rural area, whereas German children were recruited from an urban area (Berlin). Thus, our Chinese children may have lower SES than our German children. However, participants with higher SES may show better performance in face-related tasks and greater activation in face-related regions than those with lower SES (Noble et al., 2007; Rosen et al., 2018). Therefore, our results point in the opposite direction with Chinese children performing better and showing larger face N170 amplitudes than German children, arguing against an SES account.

Third, as a control for stimulus-unspecific group differences, we used a one-back task with doodles, that is, complex visual stimuli without clear configural information. Since the N170 component for doodles was very small and ambiguous, we refrained from analyzing these results. Therefore, although the second-order relationship task indicates that our group effects are stimulus-specific, future studies should seek rigorous control conditions for non-face-related visual processing.

Fourth, we cross-sectionally compared groups at only one specific time point of development. In a longitudinal study using a composite paradigm, Tso et al. (2012) reported that the holistic processing of Chinese characters decreased across grades one to five. Thus, it needs to be explored whether our findings in 7-8 years old children generalize to other age groups and adults. Hence, a more direct proof of the script system hypothesis should use a longitudinal design, starting at preschool age, where group differences should still be absent.

Fifth, Chinese children in the present study learned the simplified Chinese script. Simplified Chinese script is visually less complex than traditional Chinese script, and the reading acquisition of the two scripts has been demonstrated to impose differential visual demands. Future studies may examine the differences between traditional Chinese readers and alphabetic readers on face processing.

Sixth, the Chinese children had also learned an alphabetic script, that is, Pinyin. It is possible that the reading acquisition of both Pinyin and Chinese script affect holistic face processing more than the effect of one script system alone; however, in the present study we cannot separate the effects of these two factors. Since Pinyin is taught only as an aid to formal Chinese reading (Li & Ping, 2010), its mastery does not constitute formal reading of Chinese itself. Even if the ERP differences observed here would be due to learning Chinese script plus Pinyin versus just alphabetic script, they would still indicate that the reading acquisition experience can shape face processing and, specifically, holistic face processing.

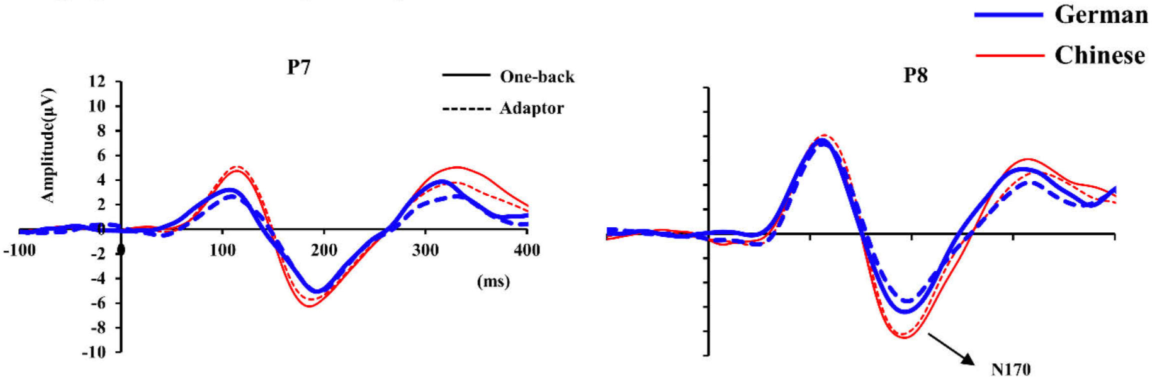
Finally, the visual complexity of the script symbols may not be the only language factor affecting face processing. Across alphabetic languages, spelling-to-sound consistency (i.e., transparency) varies strongly and is a critical factor for reading acquisition (e.g., Ziegler & Goswami, 2005). Thus, future studies could be carried out to examine the effects of orthographic transparency on face processing by comparing children who acquire more or less transparent orthographies, for example, German versus English.

2.2.5 Conclusions

The present study provides the first direct evidence that Chinese children show larger face-related N170 amplitudes than German children. A similar difference in the N170 amplitude to Mooney faces but indistinguishable effects of 2nd order relationship adaptation indicate that the N170 differences specifically related to the superiority of Chinese children in holistic face processing. We suggest that this is due to the specific demands placed on the visual and mnemonic processing in different script systems learned by the Children. Specifically, the present findings support the notion that learning a logographic script system may have generalized neurocognitive consequences by strengthening holistic visual processing to suitable stimuli also in a non-linguistic domain.

2.2.6 Supplementary Materials

A High-pass filtered ERP (2-30 Hz)



B Broad-band ERP (0.1-30 Hz)

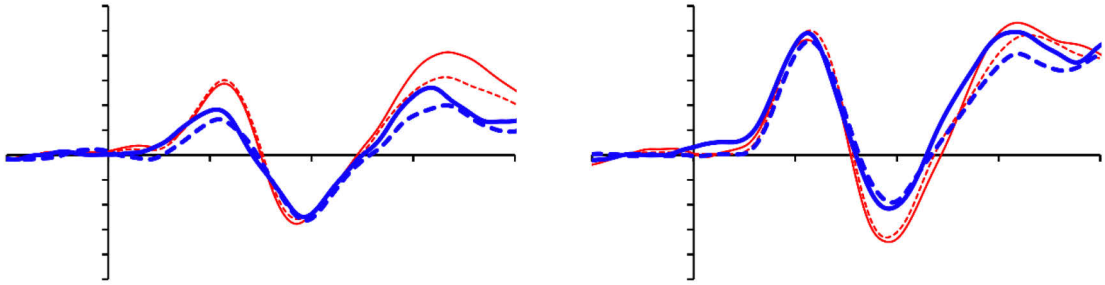


Figure S3. Grand average ERPs to naturalistic faces in different filter settings for German and Chinese children. (A) high-pass filtered ERP (2-30) Hz and (B) Broad-band ERP (0.1-30 Hz).

Chapter 3 General Discussion

The present dissertation is to investigate three main goals: (1) Testing the script system hypothesis and social immersion hypothesis, namely whether and how the script system acquisition and social immersion experience influence face perception; (2) Assessing the effects of script system acquisition on face perception from a developmental perspective, that is, whether and how the script system acquisition influences face perception in early readers and adult expert readers; (3) Investigating whether and how specific configural processes of face perception are influenced by the effects of script reading acquisition and social immersion.

In Study 1, I conducted two ERP experiments in adult participants to examine the script system hypothesis and social immersion hypothesis by reassessing the observed group difference on the face N170 in previous cross-cultural studies. The results indicated that participants who had immersion experience into a new ethnic environment showed larger face N170 than local residents, and this group difference was not dependent on the script system they had acquired. Hence, the findings of Study 1 confirmed the social immersion hypothesis but rejected the script system hypothesis. In addition, this study revealed that the massive exposure experience with other-race faces in a new ethnic social environment is critical for social immersion effects, supporting the new-ethnicity immersion hypothesis. Nevertheless, Study 1 did not show any effects on configural face processing measured by inversion effects. In Study 2, I conducted two ERP experiments on second-grade children participants to examine the effects of the script system on face perception by assessing the group difference between German and Chinese children on the face N170. The results revealed that Chinese children showed larger N170 amplitudes in response to naturalistic faces than German children, supporting the script system hypothesis. In addition, this study found a similar group difference in Mooney faces, suggesting that the script system particularly influences holistic processing. Nevertheless, there was no superiority in Chinese children on the sensitivity to the second-order relations measured by the adaptation effects at the N170, providing no evidence for the effects of script systems on the sensitivity to second-order relations.

In the following, I will first discuss the results in respect of the aims of the present work, and then, implications, limitations, and future perspectives will be considered.

3.1 The Effects of Script System on Face Perception

One of the main goals of the present thesis was to investigate the script system hypothesis, namely whether and how the script system that an individual had acquired influences one's face perception. This goal was addressed in both Study 1 and Study 2. However, the results of these two studies are inconsistent. According to the script system hypothesis, Chinese readers should have larger N170 to faces than non-Chinese readers. However, this hypothesis was not supported by the results of Experiment 1 in Study 1 in which the face-elicited N170 in Chinese participants was smaller than in non-Chinese participants. By contrast, in support of the script system hypothesis the results of Study 2 indicate that the face-elicited N170 was larger in Chinese participants than in German participants.

The inconsistency demonstrated above might be due to the participant's age, that is, adults in Study 1 versus children in Study 2. The script system hypothesis was proposed based on the neural recycling hypothesis in which reading recycles pre-existing visual systems evolved for other stimulus categories with similar functions, which in turn induce changes in the corresponding visual systems (Dehaene & Cohen, 2007; see Dehaene et al., 2015 for a review). Furthermore, recent empirical studies indicate a tight link between face expertise and script expertise during reading acquisition (Cao et al., 2019; Dehaene et al., 2010; Pegado et al., 2014; Ventura et al., 2013). Particularly, the N170 component has been demonstrated to reflect experience-dependent neural changes in visual expertise for both faces and visual scripts (Maurer et al., 2006; Rossion et al., 2003). For example, recent developmental ERP studies found that the development of face expertise and script expertise reflected in the N170 during reading acquisition were not independent but dependent on each other (e.g., Dundas et al, 2014; Li et al., 2013). Thus, the effects of script system on face perception in the N170 might be modulated by the development of script expertise. However, previous studies indicate an inverted “U” developmental trajectory of script expertise reflected by the N170, suggesting perceptual learning of a script appears to be critically important during the first two or three years of reading acquisition and then gradually declines as expertise consolidates (e.g., Brem et al., 2009; Cao et al., 2011; Maurer et al., 2011). For instance, Cao et al. (2011) found that the N170 amplitude to Chinese characters in 2nd grade children was the largest among all age groups but thereafter declined with increasing age, suggesting that the neural mechanisms for script expertise develop rapidly during the early primary school years. Therefore, it is plausible to assume that the effects of script system on face perception would show up in early readers as found in the present Study 2 but declined or disappeared in adult expert readers as found in Study 1.

The observed effects of the script system on face perception in Children reflect a recycling of the visual system in order to meet the differential challenges posed by the two script systems studied in the present experiments, leading to exciting new questions about how the effect of reading acquisition on face perception is modulated by the acquisition of different script systems. In addition, study 2 showed enhanced N170 responses to both faces and Mooney faces in Chinese children as compared to German children, indicating that the tuning of the neurocognitive system toward holistic processing is more pronounced by learning to read a Chinese script system than an alphabetic script system. Thus, this finding suggests that experiencing a visually complex script has a positive transfer to the processing of faces, arguing against the destructive competition view (Dehaene et al., 2010; Dundas et al., 2014; Ventura et al., 2013) in which the consequence of reading acquisition on face processing would be impairment rather than improvement. This suggestion is supported by more direct evidence where consistent negative correlations between reading abilities and face-elicited N170 were found in Study 2. More specifically, although the correlations between reading skills and N170 amplitudes to faces and Mooney faces were small (all $r > -0.30$) and not significant in either children group (all *ps* > 0.05), the negative direction of all correlations indicates that better reading skills are associated with larger (i.e., more negative) N170 amplitudes. In sum, although the observed correlations did not provide conclusive evidence for the script system hypothesis, this may be due to the small sample size and it may be premature to dismiss the multiple, consistently negative tendencies toward an association. Thus, future studies with larger samples should further examine the correlations between reading abilities and ERPs.

3.2 The Effects of Social Immersion Experience on Face Perception

The second main goal of the present thesis was to investigate the social immersion hypothesis, namely whether and how the immersion experience into a new social environment influences one's face perception. As expected by the social immersion hypothesis, the finding of Study 1 indicates that independent of the lab location (Hong Kong vs. Berlin), foreigners showed larger N170 amplitudes to faces than locals. This finding is in line with the observation of group differences in previous cross-cultural studies (see Table 1) that East-Asian participants who were foreigners to the geographical region where the study was conducted showed larger N170 amplitudes to faces than Caucasians who were locals in the same region. Therefore, it is plausible that a surge in exposure to novel faces during immersion into a new environment increases one's N170 response to faces. Like studies showing the effects of short-term intensive training on face perception (e.g., Davies-Thompson et al., 2017; Corrow et al., 2019), the present findings add to the evidence for the role of short-term experience on adults' abilities of processing faces, suggesting the remarkable plasticity in the adult visual system that underlies face perception.

Also, Study 1 distinguished two aspects of social immersion: immersion into a new social environment (global new social immersion hypothesis) vs. immersion into a different ethnic environment (new-ethnicity immersion hypothesis). However, the result indicates that short-term Berlin residents who had mere immersion experience into a new social environment of the same ethnicity (Berlin) did not show significantly different face N170 amplitude from long-term Berlin residents who did not have novel social immersion experience. This finding goes against the global new social immersion hypothesis, ruling out the effects of the immersion experience into a new social environment of the same ethnicity on face perception. Therefore, it is conceivable that massed exposure to other-race faces during immersion into a new environment is critical for enhancing the N170 amplitudes to faces. This claim is consistent with previous studies where short-term staying duration in a new ethnic environment improves the recognition performance of facial emotion and the related neural activities (Derntl et al., 2009; Elfenbein & Ambady, 2002). Furthermore, like other empirical evidence (e.g., Hancock & Rhodes, 2008; Hills & Lewis, 2006; Lebrecht et al., 2009; Rhodes et al., 2009; Tanaka & Pierce, 2009), the present finding supports the face space theory proposed by Valentine (1991) in which substantial new other-race faces exposure may modify the representation system of faces. Nevertheless, the present study extends those evidence by showing no modulation of the race on the N170 group differences, implying that increased experience with other-race faces in adults affects the neural level underlying the processing of both same- and other-race faces. Moreover, the present study also failed to report a sensitivity of N170 to the race of faces. which is consistent with a number of studies where the N170 was not sensitive to the race of faces (e.g., Caldara et al., 2003, 2004; Chen et al., 2013; Herzmann et al., 2011; Vizioli et al., 2010, 2011; Wiese et al., 2009). Therefore, the present finding might imply that the perceptual mechanisms of face processing reflected in N170 are not different for own-and other-race faces.

In general, the observed effects of social immersion experience on face perception yield insight into the plasticity of the perceptual system for processing faces in adults. In contrast with the script system hypothesis, whether the social immersion hypothesis generalizes to children is not addressed in the present thesis. Nevertheless, it is plausible to assume that the effects of social immersion experience can be observed in children as found in adults according to previous findings. Firstly, there is consensus and ample evidence that experience contributes greatly to the development of face processing, suggesting that the face processing system retains a great deal of experience-contingent plasticity throughout development (e.g., Cohen et al., 2007; Pascalis et al., 2009; Rossion & Lochy, 2021). For instance, previous research confirms that proficiency in other-race face recognition can develop following a relatively brief visual exposure to photographs or videos of other-race faces during childhood (Anzures et al., 2012; Sangrigoli & de Schonen, 2004) and adulthood (Goldstein & Chance, 1985; Hills & Lewis, 2006; Rhodes et al., 2009; Tanaka & Pierce, 2009), suggesting that short-term face exposure experience might influence face perception in both children and adults. Secondly, the effects of experience are not comparable across development because of the decrease in the plasticity of the face processing system throughout development (e.g., Macchi Cassia, 2011; McKone et al., 2019). For example, McKone et al. (2019) indicate that the effects of natural social exposure to other-race faces on the improvement of the other-race effects were more pronounced in young children than in adults, suggesting a loss of plasticity in the processing of other-race faces between childhood and adulthood. Thus, it seems reasonable to argue that the social immersion effects on face perception should also appear in children as found in adults in the present study. In the future, questions such as whether the effects of social immersion experience on the face processing system are comparable across developmental stages or variable depending on the time of acquisition, and whether there are interactions between the effects of social immersion experience acquired at different times need to be empirically verified.

3.3 Underlying Perceptual Processes for the Observed Effects on Face Perception

Since the N170 component has been suggested to index configural processing (e.g., Eimer et al., 2011), the present thesis attempted to clarify whether the reported effects of social immersion and script system in face-elicited N170 reflect specific alterations in configural face processing. However, the findings are not consistent across present studies. In Study 1, both upright and inverted faces were used to investigate the observed effects in configural processing of faces. The N170 is considered to index configural processing evidenced by its sensitivity to face inversion, which increases N170 amplitude and latency (Caharel et al., 2013; Jacques & Rossion, 2010; Rossion et al., 2000; Sadeh & Yovel, 2010). Moreover, all three types of configural processing (i.e., sensitivity to first-order relations, holistic processing and sensitivity to second-order relations) are suggested to be disrupted by the inversion of faces (Maurer et al., 2002). Thus, the observed group differences between foreigners and locals in the N170 should be smaller for inverted faces than for upright faces or there should be a larger face inversion effect in the N170 in foreign participants than in local participants. However, the results did not support this prediction—there were no significant interactions involving factors of group and orientation. Hence, the findings of Study 1 do not support the claim that the effects of social immersion in the N170 indexes effects on configural processing of faces. Nevertheless, in Study 2, the group differences in the N170 between Chinese and German children were reported for both naturalistic and Mooney faces. Compared with naturalistic faces, Mooney faces are a good index of holistic processing because it has to be processed holistically as there are no separable facial features (Eimer et al., 2011). Thus, the findings of Study 2 suggest that the group differences in N170 can reflect differences in holistic face processing. By contrast, Study 2 failed to find similar group differences in the N170 adaptation effects for the manipulations of second-order relations of faces, suggesting the group differences in N170 cannot reflect differences in the sensitivity to second-order relations. The inconsistency between the effects of the script system and the social immersion experience in configural processing might suggest that there are dissociated underlying face processes affected by the script system acquisition and the social immersion experience.

What remains to be discussed are the specific underlying processes underlying for the effects of social immersion reflected in the face-elicited N170 of Study 1. As the inversion of faces impairs configural, but not featural processing (e.g., Bartlett et al., 2003), the observed N170 differences in Study 1 may indicate differences in facial feature processing. More specifically, it is plausible to assume that participants who are immersed in a new ethnic environment recognize unfamiliar faces in a feature-based process, which relies on the comparison of individual facial features such as the eyes, nose or mouth. This claim is supported by previous studies where the recognition of new faces, compared to old ones, rely relatively more on the processing of featural information (e.g., Lobmaier & Mast, 2007). Moreover, other supporting evidence from perceptual training studies indicates that feature-by-feature strategies may improve unfamiliar face matching (Towler et al., 2017; White et al., 2015). In sum, the findings of Study 1 might imply that social immersion experience influences facial feature processing rather than configural processing.

The findings of both Study 1 and 2 provide new insights into previous findings that East-Asians show a general bias for holistic processing in processing of visual stimuli as compared to Westerners (e.g., Blais et al., 2008; Markus & Kitayama, 1991; Miyamoto et al., 2011). Specifically, Miyamoto et al. (2011) found that East-Asians adults showed more holistic processing of faces than Caucasian adults. Therefore, the findings of Study 1 with adults are not consistent with Miyamoto et al. (2011) whereas the findings of Study 2 support and extend the previous findings by testing Children. The inconsistency between Study 1 and Miyamoto et al. (2011) might be due to the East-Asians participants tested in the present Study 1, who had social immersion experience in a new social ethnic environment whereas those in Miyamoto et al. (2011) were tested in their own country, suggesting the social immersion experience may obscure the cultural effects on holistic processing. Thus, Study 1 provides new insight into the effects of culture on holistic processing, it may have to be qualified by the factor of social/ethnic immersion. On the other hand, in line with Miyamoto et al. (2011) Study 2 indicates that Chinese children showed superiority of holistic processing over German children, suggesting the role of a potential cultural-specific factor—script system on such cultural effects. More specifically, it suggests that learning to read Chinese improves the holistic processing of faces as compared to alphabetic script reading. This suggestion is consistent with previous findings that holistic processing is also a marker of perceptual expertise for visual Chinese characters as for faces (e.g., Chen et al., 2013; Wong et al., 2019). Thus, extensive training in holistic visual processing provided by reading acquisition in Chinese might lead to a positive transfer to face processing, resulting in an improvement of holistic face processing. In addition, since previous studies indicate an inverted U trend of the developmental pattern of holistic processing for Chinese characters (Liao & Hsiao, 2021; Tso et al., 2012), future studies could track the developmental trajectory of the effects of script system on holistic processing by a development perspective.

3.4 Implications, Limitations, and Future Perspectives

The present thesis is the first report to show that both the cultural-specific factor—script system acquisition and the environmental-specific factor—social immersion experience, influence the early stage of face perception indicated by the N170 ERP component. Furthermore, the present thesis report effects on face-elicited N170 in both children and adults, suggesting that the face processing system in the brain is highly plastic across lifespan. The present findings suggest a fascinating application perspective. For example, the observed effects of social immersion on face processing, namely immersion into a different ethnic environment activates the face processing system, making it more plastic at least for a while, which might provide the chance to use other-race faces as training material to improve face-specific abilities. In addition, since face processing and its neurophysiologic correlates have been suggested to be important markers of social cognition (Lazar et al., 2014; O'Connor et al., 2005, 2007; Schultz, 2005), the present findings might have practical implications for the development of adaptive social behaviour. For instance, Hileman et al. (2011) found that among typically developing children, more negative N170 amplitude to upright faces is associated with less atypical social behaviors. Smaller (less negative) N170 amplitude is thought to reflect less face processing expertise, and this inefficiency in neural processing may result in less fluid and more effortful reciprocal social interactions and therefore more atypical social behavior. Thus, future studies should examine how the script system and social immersion modulate the association between the neural substrate of face processing and the social behavior from an application perspective.

The present thesis has some limitations that might be dealt with in future studies. The first limitation is the lack of behavioral performance data – it would be very interesting to see whether the enhanced N170 is accompanied by corresponding changes in face recognition abilities using formal tests of face processing (e.g., face memory test) or whether the enhanced N170 is merely an epiphenomenon of an overexcited fusiform face area without functional consequences. Thus, future studies should examine to what extent the social immersion and script system experience influence face processing at the behavioral level. The second limitation is the less reliable control condition used in present thesis. In both studies a complex non-face visual stimulus – doodle, was used to investigate whether the observed effects are specific to face perception or transfer to other stimulus domains. However, in Study 1 the results of doodle conditions are inconsistent between experiments and the N170 component for doodles in Study 2 was too small and ambiguous to be measured. Moreover, doodles are meaningless and unfamiliar stimuli for participants and the neurocognitive processes underlying the N170 elicited by it are not entirely clear. Thus, future studies should seek more rigorous and familiar control conditions non-face stimuli (e.g., cars, houses) to examine the specificity or generality of the observed effects in face processing.

A third limitation is that only cross-sectional studies were conducted in the present thesis, in other words, a relatively narrow range of immersion durations in Study 1 and only second-graders in Study 2. Thus, any inferences on the further course of the observed effects cannot be made by the present thesis, in particular whether the observed effects in present thesis generalize to other time points/periods are unclear. The longitudinal studies with several time points allow demonstration of these issues. More specifically, it would be interesting to explore the time course of immersion experience effects, and this might be combined with a more direct proof of the immersion account by using a longitudinal design. In addition, a more direct proof of the script system hypothesis should use a longitudinal design, starting at preschool age, where group differences should still be absent.

3.5 General Conclusions

An interesting observation from a number of cross-cultural studies suggests that East-Asians (mainly Chinese) show a larger face-sensitive N170 of ERP component than Caucasians, it provides an important clue in elucidating the effects of underlying factors on face perception. The present thesis hypothesized two potential factors, a cultural-specific factor—script system acquisition and an environmental factor—social immersion experience, may influence face perception. Two ERP studies that were conducted to investigate the hypotheses indicate that learning to read Chinese script enhanced the face-elicited N170 as compared to alphabetic script reading and immersing into a new ethnic social environment increased one's N170 responses to faces, suggesting face perception is affected by both factors. Nevertheless, the effects of the script system were found in children but not in adults, suggesting the script system effects were constrained by the stage of reading development. By contrast, since the social immersion effects were only investigated in adults, future studies should examine whether these effects generalize to children. Additionally, the present findings reveal that the script system impacts mainly holistic face processing, which is considered to be one aspect of configural face processing, whereas the effects of social immersion were not found in configural face processing, suggesting that there might be dissociated underlying face processes affected by the script system acquisition and the social immersion experience. In sum, the present thesis not only contributed to understanding the role and nature of different aspects of experience in the development of expert face processing across the lifespan but provides new insights into the high plasticity of the face processing system in the brain.

References

Alonso Prieto, E., Caharel, S., Henson, R. N., & Rossion, B. (2011). Early (N170/M170) face-sensitivity despite right lateral occipital brain damage in acquired prosopagnosia. *Frontiers in Human Neuroscience*, *5*. https://doi.org/10.3389/fnhum.2011.00138

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science*, *15*(2), 106–111. https://doi.org/10.1111/j.0963-7214.2004.01502006.x

Anzures, G., Quinn, P. C., Pascalis, O., Slater, A. M., & Lee, K. (2010). Categorization, categorical perception, and asymmetry in infants' representation of face race. *Developmental Science*, *13*(4), 553–564. https://doi.org/10.1111/j.1467-7687.2009.00900.x

Anzures, G., Quinn, P. C., Pascalis, O., Slater, A. M., & Lee, K. (2013). Development of own-race biases. *Visual Cognition*, *21*(9–10), 1165–1182. https://doi.org/10.1080/13506285.2013.821428

Anzures, G., Wheeler, A., Quinn, P. C., Pascalis, O., Slater, A. M., Heron-Delaney, M., Tanaka, J. W., & Lee, K. (2012). Brief daily exposures to Asian females reverses perceptual narrowing for Asian faces in Caucasian infants. *Journal of Experimental Child Psychology*, *112*(4), 484–495. https://doi.org/10.1016/j.jecp.2012.04.005

Ashworth, A. R. S., Vuong, Q. C., Rossion, B., & Tarr, M. J. (2008). Recognizing rotated faces and Greebles: What properties drive the face inversion effect? *Visual Cognition*, *16*(6), 754–784. https://doi.org/10.1080/13506280701381741

Baker, C. I., Liu, J., Wald, L. L., Kwong, K. K., Benner, T., & Kanwisher, N. (2007). Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. *Proceedings of the National Academy of Sciences*, *104*(21), 9087–9092. https://doi.org/10.1073/pnas.0703300104

Balas, B., Westerlund, A., Hung, K., & Nelson III, C. A. (2011). Shape, color and the other-race effect in the infant brain: Shape, color and the other-race effect. *Developmental Science*, *14*(4), 892–900. https://doi.org/10.1111/j.1467-7687.2011.01039.x

Bartlett, J. C. Search J, Abdi H, 2003``What are the routes to face recognition?". Perception of Faces, Objects and Scenes: Analytic and Holistic Processing Eds MP Peterson, G Rhodes (Cambridge, MA: Oxford University Press) pp, 21-52.

Behrmann, M., & Plaut, D. C. (2020). Hemispheric organization for visual object recognition: A theoretical account and empirical evidence. *Perception*, *49*(4), 373–404. https://doi.org/10.1177/0301006619899049

Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience*, *8*(6), 551–565. https://doi.org/10.1162/jocn.1996.8.6.551

Bentin, S., Taylor, M. J., Rousselet, G. A., Itier, R. J., Caldara, R., Schyns, P. G., Jacques, C., & Rossion, B. (2007). Controlling interstimulus perceptual variance does not abolish N170 face sensitivity. Nature Neuroscience, 10(7), 801–802. https://doi.org/10.1038/nn0707-801

Ben-Yehudah, G., Hirshorn, E. A., Simcox, T., Perfetti, C. A., & Fiez, J. A. (2019). Chinese-English bilinguals transfer L1 lexical reading procedures and holistic orthographic coding to L2 English. *Journal of Neurolinguistics*, *50*, 136–148. https://doi.org/10.1016/j.jneuroling.2018.01.002

Blais, C., Jack, R. E., Scheepers, C., Fiset, D., & Caldara, R. (2008). Culture shapes how we look at faces. *PLoS ONE*, *3*(8), e3022. https://doi.org/10.1371/journal.pone.0003022

Bolger, D. J., Perfetti, C. A., & Schneider, W. (2005). Cross-cultural effect on the brain revisited: Universal structures plus writing system variation. *Human Brain Mapping*, *25*(1), 92–104. https://doi.org/10.1002/hbm.20124

Bötzel, K., Schulze, S., & Stodieck, StefanR. G. (1995). Scalp topography and analysis of intracranial sources of face-evoked potentials. *Experimental Brain Research*, *104*(1). https://doi.org/10.1007/BF00229863

Brem, S., Bucher, K., Halder, P., Summers, P., Dietrich, T., Martin, E., & Brandeis, D. (2006). Evidence for developmental changes in the visual word processing network beyond adolescence. *NeuroImage*, *29*(3), 822–837. https://doi.org/10.1016/j.neuroimage.2005.09.023

Brem, S., Halder, P., Bucher, K., Summers, P., Martin, E., & Brandeis, D. (2009). Tuning of the visual word processing system: Distinct developmental ERP and fMRI effects. *Human Brain Mapping*, *30*(6), 1833–1844. https://doi.org/10.1002/hbm.20751

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, *77*(3), 305–327. https://doi.org/10.1111/j.2044-8295.1986.tb02199.x

Brunet, D., Murray, M. M., & Michel, C. M. (2011). Spatiotemporal Analysis of Multichannel EEG: CARTOOL. *Computational Intelligence and Neuroscience*, *2011*, 1–15. https://doi.org/10.1155/2011/813870

Bukach, C. M., Gauthier, I., & Tarr, M. J. (2006). Beyond faces and modularity: The power of an expertise framework. *Trends in Cognitive Sciences*, *10*(4), 159–166. https://doi.org/10.1016/j.tics.2006.02.004

Busey, T. A., & Vanderkolk, J. R. (2005). Behavioral and electrophysiological evidence for configural processing in fingerprint experts. *Vision Research*, *45*(4), 431–448. https://doi.org/10.1016/j.visres.2004.08.021

Byatt, G., & Rhodes, G. (2004). Identification of own-race and other-race faces: Implications for the representation of race in face space. *Psychonomic Bulletin & Review*, *11*(4), 735–741. https://doi.org/10.3758/BF03196628

Caharel, S., d'Arripe, O., Ramon, M., Jacques, C., & Rossion, B. (2009). Early adaptation to repeated unfamiliar faces across viewpoint changes in the right hemisphere: Evidence from the N170 ERP component. *Neuropsychologia*, *47*(3), 639–643. https://doi.org/10.1016/j.neuropsychologia.2008.11.016

Caharel, S., Leleu, A., Bernard, C., Viggiano, M.-P., Lalonde, R., & Rebaï, M. (2013). Early holistic face-like processing of Arcimboldo paintings in the right occipito-temporal cortex: Evidence from the N170 ERP component. *International Journal of Psychophysiology*, *90*(2), 157–164. https://doi.org/10.1016/j.ijpsycho.2013.06.024

Caharel, S., Montalan, B., Fromager, E., Bernard, C., Lalonde, R., & Mohamed, R. (2011). Other-race and inversion effects during the structural encoding stage of face processing in a race categorization task: An event-related brain potential study. *International Journal of Psychophysiology*, *79*(2), 266–271. https://doi.org/10.1016/j.ijpsycho.2010.10.018

Caldara, R., Rossion, B., Bovet, P., & Hauert, C.-A. (2004). Event-related potentials and time course of the 'other-race' face classification advantage. *NeuroReport*, *15*(5). https://journals.lww.com/neuroreport/Fulltext/2004/04090/Event_related_potentials_and_time_course_of_the.34.aspx

Caldara, R., Thut, G., Servoir, P., Michel, C. M., Bovet, P., & Renault, B. (2003). Face versus non-face object perception and the 'other-race' effect: A spatio-temporal event-related potential study. *Clinical Neurophysiology*, *114*(3), 515–528. https://doi.org/10.1016/S1388-2457(02)00407-8

Cantlon, J. F., Pinel, P., Dehaene, S., & Pelphrey, K. A. (2011). Cortical representations of symbols, objects, and faces are pruned back during early childhood. *Cerebral Cortex*, *21*(1), 191–199. https://doi.org/10.1093/cercor/bhq078

Cao, F., Lee, R., Shu, H., Yang, Y., Xu, G., Li, K., & Booth, J. R. (2010). Cultural constraints on brain development: Evidence from a developmental study of visual word processing in mandarin chinese. *Cerebral Cortex (New York, N.Y.: 1991)*, *20*(5), 1223–1233. https://doi.org/10.1093/cercor/bhp186

Cao, X., Li, S., Zhao, J., Lin, S., & Weng, X. (2011). Left-lateralized early neurophysiological response for Chinese characters in young primary school children. *Neuroscience Letters*, *492*(3), 165–169. https://doi.org/10.1016/j.neulet.2011.02.002

Cao, X., Ma, X., & Qi, C. (2015). N170 adaptation effect for repeated faces and words. *Neuroscience*, *294*, 21–28. https://doi.org/10.1016/j.neuroscience.2015.03.009

Cao, X., Yang, Q., Zhong, P., & Chen, C. (2019). The characteristics of face configural effect in illiterates and literates. *Acta Psychologica*, *201*, 102951. https://doi.org/10.1016/j.actpsy.2019.102951

Carey, S., & Diamond, R. (1977). From piecemeal to configurational representation of faces. *Science*, *195*(4275), 312–314. https://doi.org/10.1126/science.831281

Carreiras, M., Seghier, M. L., Baquero, S., Estévez, A., Lozano, A., Devlin, J. T., & Price, C. J. (2009). An anatomical signature for literacy. *Nature*, *461*(7266), 983–986. https://doi.org/10.1038/nature08461

Centanni, T. M., Norton, E. S., Park, A., Beach, S. D., Halverson, K., Ozernov-Palchik, O., Gaab, N., & Gabrieli, J. D. (2018). Early development of letter specialization in left fusiform is associated with better word reading and smaller fusiform face area. *Developmental Science*, *21*(5), e12658. https://doi.org/10.1111/desc.12658

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*(1), 64–85. https://doi.org/10.1080/10888438.2015.1104688

Changizi, M. A., & Shimojo, S. (2005). Character complexity and redundancy in writing systems over human history. *Proceedings of the Royal Society B: Biological Sciences*, *272*(1560), 267–275. https://doi.org/10.1098/rspb.2004.2942

Chen, H., Bukach, C. M., & Wong, A. C.-N. (2013). Early electrophysiological basis of experience-associated holistic processing of chinese characters. *PLoS ONE*, *8*(4), e61221. https://doi.org/10.1371/journal.pone.0061221

Chen, X., & Kao, H. S. R. (2002). Visual-spatial properties and orthographic processing of Chinese characters. Cognitive neuroscience studies of the Chinese language (pp. 175–194). *Hong Kong University Press*.

Chen, Y., Pan, F., Wang, H., Xiao, S., & Zhao, L. (2013). Electrophysiological correlates of processing own- and other-race faces. *Brain Topography*, *26*(4), 606–615. https://doi.org/10.1007/s10548-013-0286-x

Chua, H. F., Boland, J. E., & Nisbett, R. E. (2005). Cultural variation in eye movements during scene perception. *Proceedings of the National Academy of Sciences*, *102*(35), 12629–12633. https://doi.org/10.1073/pnas.0506162102

Chuk, T., Crookes, K., Hayward, W. G., Chan, A. B., & Hsiao, J. H. (2017). Hidden Markov model analysis reveals the advantage of analytic eye movement patterns in face recognition across cultures. *Cognition*, *169*, 102–117. https://doi.org/10.1016/j.cognition.2017.08.003

Civile, C., Elchlepp, H., McLaren, R., Galang, C. M., Lavric, A., & McLaren, I. (2018). The effect of scrambling upright and inverted faces on the N170. *Quarterly Journal of Experimental Psychology*, *71*(11), 2464–2476.

https://doi.org/10.1177/1747021817744455

Civile, C., McLaren, R., & McLaren, I. P. L. (2016). The face inversion effect: roles of first and second-order configural information. *The American Journal of Psychology*, *129*(1), 23–35. https://doi.org/10.5406/amerjpsyc.129.1.0023

Clark, V. P., Fan, S., & Hillyard, S. A. (1994). Identification of early visual evoked potential generators by retinotopic and topographic analyses. *Human Brain Mapping*, *2*(3), 170–187. https://doi.org/10.1002/hbm.460020306

Corrow, S. L., Davies-Thompson, J., Fletcher, K., Hills, C., Corrow, J. C., & Barton, J. J. S. (2019). Training face perception in developmental prosopagnosia through perceptual learning. *Neuropsychologia*, *134*, 107196. https://doi.org/10.1016/j.neuropsychologia.2019.107196

Curby, K. M., Glazek, K., & Gauthier, I. (2009). A visual short-term memory advantage for objects of expertise. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(1), 94–107. https://doi.org/10.1037/0096-1523.35.1.94

Davies-Thompson, J., Fletcher, K., Hills, C., Pancaroglu, R., Corrow, S. L., & Barton, J. J. S. (2017). Perceptual learning of faces: A rehabilitative study of acquired prosopagnosia. *Journal of Cognitive Neuroscience*, *29*(3), 573–591. https://doi.org/10.1162/jocn_a_01063

Deen, B., Richardson, H., Dilks, D. D., Takahashi, A., Keil, B., Wald, L. L., Kanwisher, N., & Saxe, R. (2017). Organization of high-level visual cortex in human infants. *Nature Communications*, *8*(1), 13995. https://doi.org/10.1038/ncomms13995

DeFrancis, J. (1989). Visible speech: The diverse oneness of writing systems. *University of Hawaii Press.*

DeGutis, J., Mercado, R. J., Wilmer, J., & Rosenblatt, A. (2013). Individual differences in holistic processing predict the own-race advantage in recognition memory. *PLoS ONE*, *8*(4), e58253. https://doi.org/10.1371/journal.pone.0058253

Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron*, *56*(2), 384–398. https://doi.org/10.1016/j.neuron.2007.10.004

Dehaene, S., Cohen, L., Morais, J., & Kolinsky, R. (2015). Illiterate to literate: Behavioural and cerebral changes induced by reading acquisition. *Nature Reviews Neuroscience*, *16*(4), 234–244. https://doi.org/10.1038/nrn3924

Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Filho, G. N., Jobert, A., Dehaene-Lambertz, G., Kolinsky, R., Morais, J., & Cohen, L. (2010). How learning to read changes the cortical networks for vision and language. *Science*, *330*(6009), 1359–1364. https://doi.org/10.1126/science.1194140

Dehaene-Lambertz, G., Monzalvo, K., & Dehaene, S. (2018). The emergence of the visual word form: Longitudinal evolution of category-specific ventral visual areas during reading acquisition. *PLOS Biology*, *16*(3), e2004103. https://doi.org/10.1371/journal.pbio.2004103

Demetriou, A., Kui, Z. X., Spanoudis, G., Christou, C., Kyriakides, L., & Platsidou, M. (2005). The architecture, dynamics, and development of mental processing: Greek, Chinese, or Universal? *Intelligence*, *33*(2), 109–141. https://doi.org/10.1016/j.intell.2004.10.003

Dering, B., Hoshino, N., & Theirry, G. (2013). N170 modulation is expertise driven: evidence from word-inversion effects in speakers of different languages. *Neuropsycholo Trend*, 13.

Derntl, B., Habel, U., Robinson, S., Windischberger, C., Kryspin-Exner, I., Gur, R. C., & Moser, E. (2009). Amygdala activation during recognition of emotions in a foreign ethnic group is associated with duration of stay. *Social Neuroscience*, *4*(4), 294–307. https://doi.org/10.1080/17470910802571633

Devillez, H., Mollison, M. V., Hagen, S., Tanaka, J. W., Scott, L. S., & Curran, T. (2019). Color and spatial frequency differentially impact early stages of perceptual expertise training. *Neuropsychologia*, *122*, 62–75. https://doi.org/10.1016/j.neuropsychologia.2018.11.011

Di Russo, F., Martínez, A., Sereno, M. I., Pitzalis, S., & Hillyard, S. A. (2002). Cortical sources of the early components of the visual evoked potential: Cortical Sources of VEP. *Human Brain Mapping*, *15*(2), 95–111. https://doi.org/10.1002/hbm.10010

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, *115*(2), 107–117. https://doi.org/10.1037/0096-3445.115.2.107

Dolzycka, D., Herzmann, G., Sommer, W., & Wilhelm, O. (2014). Can training enhance face cognition abilities in Middle-aged adults? *PLoS ONE*, *9*(3), e90249. https://doi.org/10.1371/journal.pone.0090249

Dundas, E. M., Plaut, D. C., & Behrmann, M. (2014). An ERP investigation of the co-development of hemispheric lateralization of face and word recognition. *Neuropsychologia*, *61*, 315–323. https://doi.org/10.1016/j.neuropsychologia.2014.05.006

Eberhard-Moscicka, A. K., Jost, L. B., Raith, M., & Maurer, U. (2015). Neurocognitive mechanisms of learning to read: Print tuning in beginning readers related to word-reading fluency and semantics but not phonology. *Developmental Science*, *18*(1), 106–118. https://doi.org/10.1111/desc.12189

Ehri, L. C. (2005). Development of sight word reading: phases and findings. In M. J. Snowling & C. Hulme (Eds.), *The Science of Reading: A Handbook* (pp. 135–154). Blackwell Publishing Ltd. https://doi.org/10.1002/9780470757642.ch8

Eimer, M. (2000). The face-specific N170 component reflects late stages in the structural encoding of faces: *NeuroReport*, *11*(10), 2319–2324. https://doi.org/10.1097/00001756-200007140-00050

Eimer, M. (2011). The face-sensitive N170 component of the event-related brain potential. *The Oxford handbook of face perception*, 28, 329-44.

Eimer, M., Gosling, A., Nicholas, S., & Kiss, M. (2011). The N170 component and its links to configural face processing: A rapid neural adaptation study. *Brain Research*, *1376*, 76–87. https://doi.org/10.1016/j.brainres.2010.12.046

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin*, *128*(2), 203–235. https://doi.org/10.1037/0033-2909.128.2.203

Estéphan, A., Fiset, D., Saumure, C., Plouffe-Demers, M.-P., Zhang, Y., Sun, D., & Blais, C. (2018). Time course of cultural differences in spatial frequency use for face identification. *Scientific Reports*, *8*(1), 1816. https://doi.org/10.1038/s41598-018-19971-1

Fan, C., Chen, S., Zhang, L., Qi, Z., Jin, Y., Wang, Q., Luo, Y., Li, H., & Luo, W. (2015). N170 changes reflect competition between faces and identifiable characters during early visual processing. *NeuroImage*, *110*, 32–38. https://doi.org/10.1016/j.neuroimage.2015.01.047

Ferree, T. C., Luu, P., Russell, G. S., & Tucker, D. M. (2001). Scalp electrode impedance, infection risk, and EEG data quality. *Clinical Neurophysiology*, *112*(3), 536–544. https://doi.org/10.1016/S1388-2457(00)00533-2

Flaherty, M. (2003). Sign Language and Chinese Characters on Visual-Spatial Memory: A Literature Review. *Perceptual and Motor Skills*, *97*(3), 797–802. https://doi.org/10.2466/pms.2003.97.3.797

Gajewski, P. D., Schlegel, K., & Stoerig, P. (2008). Effects of human race and face inversion on the N170: A cross-race study. *Journal of Psychophysiology*, *22*(4), 157–165. https://doi.org/10.1027/0269-8803.22.4.157

Gauthier, I., & Bukach, C. (2007). Should we reject the expertise hypothesis? *Cognition*, *103*(2), 322–330. https://doi.org/10.1016/j.cognition.2006.05.003

Gauthier, I., Curran, T., Curby, K. M., & Collins, D. (2003). Perceptual interference supports a non-modular account of face processing. *Nature Neuroscience*, *6*(4), 428–432. https://doi.org/10.1038/nn1029

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, *3*(2), 191–197. https://doi.org/10.1038/72140

Gauthier, I., & Tarr, M. J. (1997). Becoming a “Greeble” expert: exploring mechanisms for face recognition. *Vision Research*, *37*(12), 1673–1682. https://doi.org/10.1016/S0042-6989(96)00286-6

Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform “face area” increases with expertise in recognizing novel objects. *Nature Neuroscience*, *2*(6), 568–573. https://doi.org/10.1038/9224

Gelb, I. J. (1963). A study of writing. *University of Chicago Press.*

George, N., Jemel, B., Fiori, N., Chaby, L., & Renault, B. (2005). Electrophysiological correlates of facial decision: Insights from upright and upside-down Mooney-face perception. *Cognitive Brain Research*, *24*(3), 663–673. https://doi.org/10.1016/j.cogbrainres.2005.03.017

Goffaux, V., Hault, B., Michel, C., Vuong, Q. C., & Rossion, B. (2005). The respective role of low and high spatial frequencies in supporting configural and featural processing of faces. *Perception*, *34*(1), 77–86. https://doi.org/10.1068/p5370

Golarai, G., Ghahremani, D. G., Greenwood, A. C., Gabrieli, J. D. E., & Eberhardt, J. L. (2020). The development of race effects in face processing from childhood through adulthood: Neural and behavioral evidence. *Developmental Science*. https://doi.org/10.1111/desc.13058

Golarai, G., Ghahremani, D. G., Greenwood, A. C., Gabrieli, J. D. E., & Eberhardt, J. L. (2021). The development of race effects in face processing from childhood through adulthood: Neural and behavioral evidence. *Developmental Science*, *24*(3). https://doi.org/10.1111/desc.13058

Goldstein, A. G., & Chance, J. E. (1985). Effects of training on Japanese face recognition: Reduction of the other-race effect. *Bulletin of the Psychonomic Society*, *23*(3), 211–214. https://doi.org/10.3758/BF03329829

Goodman, G. S., Sayfan, L., Lee, J. S., Sandhei, M., Walle-Olsen, A., Magnussen, S., Pezdek, K., & Arredondo, P. (2007). The development of memory for own- and other-race faces. *Journal of Experimental Child Psychology*, *98*(4), 233–242. https://doi.org/10.1016/j.jecp.2007.08.004

Halit, H., de Haan, M., & Johnson, M. H. (2000). Modulation of event-related potentials by prototypical and atypical faces: *NeuroReport*, *11*(9), 1871–1875. https://doi.org/10.1097/00001756-200006260-00014

Hancock, K. J., & Rhodes, G. (2008). Contact, configural coding and the other-race effect in face recognition. *British Journal of Psychology*, *99*(1), 45–56. https://doi.org/10.1348/000712607X199981

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*(6), 223–233. https://doi.org/10.1016/S1364-6613(00)01482-0

Henson, R. N. (2003). Electrophysiological and haemodynamic correlates of face perception, recognition and priming. *Cerebral Cortex*, *13*(7), 793–805. https://doi.org/10.1093/cercor/13.7.793

Herrmann, M. J., Ehlis, A.-C., Muehlberger, A., & Fallgatter, A. J. (2005). Source localization of early stages of face processing. *Brain Topography*, *18*(2), 77–85. https://doi.org/10.1007/s10548-005-0277-7

Herrmann, M. J., Schreppel, T., Jäger, D., Koehler, S., Ehlis, A.-C., & Fallgatter, A. J. (2007). The other-race effect for face perception: An event-related potential study. *Journal of Neural Transmission*, *114*(7), 951–957. https://doi.org/10.1007/s00702-007-0624-9

Hervais-Adelman, A., Kumar, U., Mishra, R. K., Tripathi, V. N., Guleria, A., Singh, J. P., Eisner, F., & Huettig, F. (2019). Learning to read recycles visual cortical networks without destruction. *Science Advances*, *5*(9), eaax0262. https://doi.org/10.1126/sciadv.aax0262

Herzmann, G., Willenbockel, V., Tanaka, J. W., & Curran, T. (2011). The neural correlates of memory encoding and recognition for own-race and other-race faces. *Neuropsychologia*, *49*(11), 3103–3115. https://doi.org/10.1016/j.neuropsychologia.2011.07.019

Hildebrandt, A., Sommer, W., Herzmann, G., & Wilhelm, O. (2010). Structural invariance and age-related performance differences in face cognition. *Psychology and Aging*, *25*(4), 794–810. https://doi.org/10.1037/a0019774

Hileman, C. M., Henderson, H., Mundy, P., Newell, L., & Jaime, M. (2011). Developmental and individual differences on the P1 and N170 ERP components in children with and without autism. *Developmental Neuropsychology*, *36*(2), 214–236. https://doi.org/10.1080/87565641.2010.549870

Hills, P. J., & Lewis, M. B. (2005). Removing the own-race bias in face recognition by shifting attention. *Perception*, 34(S), 206.

Hills, P. J., & Lewis, M. B. (2006). Short article: reducing the own-race bias in face recognition by shifting attention. *Quarterly Journal of Experimental Psychology*, *59*(6), 996–1002. https://doi.org/10.1080/17470210600654750

Hills, P. J., Lowe, L., Hedges, B., & Teixeira, A. R. (2020). The role of extraversion, IQ and contact in the own-ethnicity face recognition bias. *Attention, Perception, & Psychophysics*, *82*(4), 1872–1882. https://doi.org/10.3758/s13414-019-01947-6

Hillyard, S. A., Teder-Sälejärvi, W. A., & Münte, T. F. (1998). Temporal dynamics of early perceptual processing. *Current Opinion in Neurobiology*, *8*(2), 202–210. https://doi.org/10.1016/S0959-4388(98)80141-4

Hoehl, S., & Peykarjou, S. (2012). The early development of face processing—What makes faces special? *Neuroscience Bulletin*, *28*(6), 765–788. https://doi.org/10.1007/s12264-012-1280-0

Honeycutt, H. (2019). Nature and nurture as an enduring tension in the history of psychology. In H. Honeycutt, *Oxford Research Encyclopedia of Psychology*. Oxford University Press. https://doi.org/10.1093/acrefore/9780190236557.013.518

Horn, C. C., & Manis, F. R. (1987). Development of automatic and speeded reading of printed words. *Journal of Experimental Child Psychology*, *44*(1), 92–108. https://doi.org/10.1016/0022-0965(87)90024-5

Hsiao, J. H., & Cottrell, G. W. (2009). Not all visual expertise is holistic, but it may be leftist: The case of chinese character recognition. *Psychological Science*, *20*(4), 455–463. https://doi.org/10.1111/j.1467-9280.2009.02315.x

Hsiao, J. H., & Lam, S. M. (2013). The modulation of visual and task characteristics of a writing system on hemispheric lateralization in visual word recognition-a computational exploration. *Cognitive Science*, *37*(5), 861–890. https://doi.org/10.1111/cogs.12033

Hu, R., Cao, B., & Du, J. (2013). Research on phonetic symbols of phonograms in chinese mandarin. *Journal of Chinese Information Processing*, 27(3), 41-47.

Huang, H. S., & Hanley, J. R. (1995). Phonological awareness and visual skills in learning to read Chinese and English. *Cognition*, *54*(1), 73–98. https://doi.org/10.1016/0010-0277(94)00641-W

Huettig, F., Kolinsky, R., & Lachmann, T. (2018). The culturally co-opted brain: How literacy affects the human mind. *Language, Cognition and Neuroscience*, *33*(3), 275–277. https://doi.org/10.1080/23273798.2018.1425803

Hugenberg, K., Young, S. G., Bernstein, M. J., & Sacco, D. F. (2010). The categorization-individuation model: An integrative account of the other-race recognition deficit. *Psychological Review*, *117*(4), 1168–1187. https://doi.org/10.1037/a0020463

Hussain, Z., Sekuler, A. B., & Bennett, P. J. (2009). Perceptual learning modifies inversion effects for faces and textures. *Vision Research*, *49*(18), 2273–2284. https://doi.org/10.1016/j.visres.2009.06.014

Itier, R. J., Herdman, A. T., George, N., Cheyne, D., & Taylor, M. J. (2006). Inversion and contrast-reversal effects on face processing assessed by MEG. *Brain Research*, *1115*(1), 108–120. https://doi.org/10.1016/j.brainres.2006.07.072

Itier, R. J., & Taylor, M. J. (2002). Inversion and contrast polarity reversal affect both encoding and recognition processes of unfamiliar Faces: A repetition study using ERPs. *NeuroImage*, *15*(2), 353–372. https://doi.org/10.1006/nimg.2001.0982

Itier, R. J., & Taylor, M. J. (2004a). N170 or N1? Spatiotemporal differences between object and face processing using ERPs. *Cerebral Cortex*, *14*(2), 132–142. https://doi.org/10.1093/cercor/bhg111

Itier, R. J., & Taylor, M. J. (2004b). Effects of repetition learning on upright, inverted and contrast-reversed face processing using ERPs. *NeuroImage*, *21*(4), 1518–1532. https://doi.org/10.1016/j.neuroimage.2003.12.016

Ito, T. A., & Urland, G. R. (2005). The influence of processing objectives on the perception of faces: An ERP study of race and gender perception. *Cognitive, Affective, & Behavioral Neuroscience*, *5*(1), 21–36. https://doi.org/10.3758/CABN.5.1.21

Itz, M. L., Schweinberger, S. R., & Kaufmann, J. M. (2016). Effects of caricaturing in shape or color on familiarity fecisions for familiar and unfamiliar Faces. *PLOS ONE*, *11*(2), e0149796. https://doi.org/10.1371/journal.pone.0149796

Itz, M. L., Schweinberger, S. R., Schulz, C., & Kaufmann, J. M. (2014). Neural correlates of facilitations in face learning by selective caricaturing of facial shape or reflectance. *NeuroImage*, *102 Pt 2*, 736–747. https://doi.org/10.1016/j.neuroimage.2014.08.042

Jack, R. E., Blais, C., Scheepers, C., Schyns, P. G., & Caldara, R. (2009). Cultural confusions Show that facial expressions are not universal. *Current Biology*, *19*(18), 1543–1548. https://doi.org/10.1016/j.cub.2009.07.051

Jacques, C., d'Arripe, O., & Rossion, B. (2007). The time course of the inversion effect during individual face discrimination. *Journal of Vision*, *7*(8), 3. https://doi.org/10.1167/7.8.3

Jacques, C., & Rossion, B. (2010). Misaligning face halves increases and delays the N170 specifically for upright faces: Implications for the nature of early face representations. *Brain Research*, *1318*, 96–109. https://doi.org/10.1016/j.brainres.2009.12.070

Jaquet, E., Rhodes, G., & Hayward, W. G. (2007). Opposite aftereffects for Chinese and Caucasian faces are selective for social category information and not just physical face differences. *Quarterly Journal of Experimental Psychology*, *60*(11), 1457–1467. https://doi.org/10.1080/17470210701467870

Jaquet, E., Rhodes, G., & Hayward, W. G. (2008). Race-contingent aftereffects suggest distinct perceptual norms for different race faces. *Visual Cognition*, *16*(6), 734–753. https://doi.org/10.1080/13506280701350647

Jeffreys, D. A. (1993). The influence of stimulus orientation on the vertex positive scalp potential evoked by faces. *Experimental Brain Research*, *96*(1), 163–172. https://doi.org/10.1007/BF00230449

Ji, Y., Liu, J., Zhu, X. Q., Zhao, J., Wang, J., Du, Y. C., & Bi, H. Y. (2019). Atypical N170 lateralization of face and word recognition in Chinese children with autism spectrum disorder. *Journal of Neurolinguistics*, *52*, 100858. https://doi.org/10.1016/j.jneuroling.2019.100858

Johannes, S., Münte, T. F., Heinze, H. J., & Mangun, G. R. (1995). Luminance and spatial attention effects on early visual processing. *Cognitive Brain Research*, *2*(3), 189–205. https://doi.org/10.1016/0926-6410(95)90008-X

Johnson, M.H., Morton, J. (1991). Biology and cognitive development: The case of face recognition. Oxford, England: Blackwell.

Kanwisher, N. (2000). Domain specificity in face perception. *Nature Neuroscience*, *3*(8), 759–763. https://doi.org/10.1038/77664

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face frea: A module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience*, *17*(11), 4302–4311. https://doi.org/10.1523/JNEUROSCI.17-11-04302.1997

Kao, C. H., Chen, D. Y., & Chen, C. C. (2010). The inversion effect in visual word form processing. *Cortex*, *46*(2), 217–230. https://doi.org/10.1016/j.cortex.2009.04.003

Kaufmann, J. M., & Schweinberger, S. R. (2008). Distortions in the brain? ERP effects of caricaturing familiar and unfamiliar faces. *Brain Research*, *1228*, 177–188. https://doi.org/10.1016/j.brainres.2008.06.092

Kelly, D. J., Jack, R. E., Miellet, S., De Luca, E., Foreman, K., & Caldara, R. (2011). Social experience does Not abolish cultural diversity in eye movements. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00095

Kelly, D. J., Liu, S., Lee, K., Quinn, P. C., Pascalis, O., Slater, A. M., & Ge, L. (2009). Development of the other-race effect during infancy: Evidence toward universality? *Journal of Experimental Child Psychology*, *104*(1), 105–114. https://doi.org/10.1016/j.jecp.2009.01.006

Kelly, D. J., Liu, S., Rodger, H., Miellet, S., Ge, L., & Caldara, R. (2011). Developing cultural differences in face processing: Developing cultural differences in face processing. *Developmental Science*, *14*(5), 1176–1184. https://doi.org/10.1111/j.1467-7687.2011.01067.x

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Ge, L., & Pascalis, O. (2007). The other-race effect develops during infancy: evidence of perceptual narrowing. *Psychological Science*, *18*(12), 1084–1089. https://doi.org/10.1111/j.1467-9280.2007.02029.x

Kita, Y., Gunji, A., Sakihara, K., Inagaki, M., Kaga, M., Nakagawa, E., & Hosokawa, T. (2010). Scanning strategies do not modulate face identification: Eye-tracking and near-infrared spectroscopy study. *PLoS ONE*, *5*(6), e11050.

https://doi.org/10.1371/journal.pone.0011050

Koenig, T., Kottlow, M., Stein, M., & Melie-García, L. (2011). Ragu: a free tool for the analysis of EEG and MEG event-related scalp field data using global randomization statistics. *Computational Intelligence and Neuroscience*, *2011*, 1–14. https://doi.org/10.1155/2011/938925

Koenig, T., & Melie-García, L. (2009). Statistical analysis of multichannel scalp field data. In C. M. Michel, T. Koenig, D. Brandeis, L. R. R. Gianotti, & J. Wackermann (Eds.), *Electrical Neuroimaging* (pp. 169–190). Cambridge University Press. https://doi.org/10.1017/CBO9780511596889.009

Kühn, C. D., Wilms, I. L., Dalrymple, K. A., Gerlach, C., & Starrfelt, R. (2021). Face recognition in beginning readers: Investigating the potential relationship between reading and face recognition during the first year of school. *Visual Cognition*, *29*(4), 213–224. https://doi.org/10.1080/13506285.2021.1884151

Kuwabara, M., & Smith, L. B. (2016). Cultural differences in visual object recognition in 3-year-old children. *Journal of Experimental Child Psychology*, *147*, 22–38. https://doi.org/10.1016/j.jecp.2016.02.006

Lao, J., Vizioli, L., & Caldara, R. (2013). Culture modulates the temporal dynamics of global/local processing. *Culture and Brain*, *1*(2–4), 158–174. https://doi.org/10.1007/s40167-013-0012-2

Latinus, M., & Taylor, M. J. (2005). Holistic processing of faces: Learning effects with Mooney Faces. *Journal of Cognitive Neuroscience*, *17*(8), 1316–1327. https://doi.org/10.1162/0898929055002490

Latinus, M., & Taylor, M. J. (2006). Face processing stages: Impact of difficulty and the separation of effects. *Brain Research*, *1123*(1), 179–187. https://doi.org/10.1016/j.brainres.2006.09.031

Lazar, S. M., Evans, D. W., Myers, S. M., Moreno-De Luca, A., & Moore, G. J. (2014). Social cognition and neural substrates of face perception: Implications for neurodevelopmental and neuropsychiatric disorders. *Behavioural Brain Research*, *263*, 1–8. https://doi.org/10.1016/j.bbr.2014.01.010

Lebrecht, S., Pierce, L. J., Tarr, M. J., & Tanaka, J. W. (2009). Perceptual other-race Training reduces implicit racial bias. *PLoS ONE*, *4*(1), e4215. https://doi.org/10.1371/journal.pone.0004215

Lehmann, D., & Skrandies, W. (1980). Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalography and Clinical Neurophysiology*, *48*(6), 609–621. https://doi.org/10.1016/0013-4694(80)90419-8

Lei, L., Pan, J., Liu, H., McBride-Chang, C., Li, H., Zhang, Y., Chen, L., Tardif, T., Liang, W., Zhang, Z., & Shu, H. (2011). Developmental trajectories of reading development and impairment from ages 3 to 8 years in Chinese children: Developmental trajectories and early prediction of reading disability. *Journal of Child Psychology and Psychiatry*, *52*(2), 212–220. https://doi.org/10.1111/j.1469-7610.2010.02311.x

Leopold, D. A., O'Toole, A. J., Vetter, T., & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. *Nature Neuroscience*, *4*(1), 89–94. https://doi.org/10.1038/82947

Leopold, D. A., & Rhodes, G. (2010). A comparative view of face perception. *Journal of Comparative Psychology*, *124*(3), 233–251. https://doi.org/10.1037/a0019460

Li, C., Ma, X., Zhu, C., & Cao, X. (2019). The recovery speed of category sensitive N170 responses to faces and Chinese characters. *Brain Research*, *1723*, 146384. https://doi.org/10.1016/j.brainres.2019.146384

Li, H., Shu, H., McBride-Chang, C., Liu, H., & Peng, H. (2012). Chinese children's character recognition: Visuo-orthographic, phonological processing and morphological skills: CHINESE CHILDREN'S CHARACTER RECOGNITION. *Journal of Research in Reading*, *35*(3), 287–307. https://doi.org/10.1111/j.1467-9817.2010.01460.x

Li, S., Lee, K., Zhao, J., Yang, Z., He, S., & Weng, X. (2013). Neural competition as a developmental process: Early hemispheric specialization for word processing delays specialization for face processing. *Neuropsychologia*, 51(5), 950–959. https://doi.org/10.1016/j.neuropsychologia.2013.02.006

Li, Y., & Bi, H.-Y. (2022). Comparative research on neural dysfunction in children with dyslexia under different writing systems: A meta-analysis study. *Neuroscience & Biobehavioral Reviews*, *137*, 104650. https://doi.org/10.1016/j.neubiorev.2022.104650

Li, X., & Ping, R. (2010). Reconsidering Chinese pinyin teaching in primary schools. *Journal of Hunan First Normal University.*

Liao, C.-H., Deng, C., Hamilton, J., Lee, C. S.-C., Wei, W., & Georgiou, G. K. (2015). The role of rapid naming in reading development and dyslexia in Chinese. *Journal of Experimental Child Psychology*, *130*, 106–122. https://doi.org/10.1016/j.jecp.2014.10.002

Liao, W., & Hsiao, J. (2021). The role of eye movement pattern and global-local information processing abilities in isolated English word reading. In Proceedings of the Annual Meeting of the Cognitive Science Society (Vol. 43, No. 43).

Liu, T., Chuk, T. Y., Yeh, S.-L., & Hsiao, J. H. (2016). Transfer of perceptual expertise: the case of simplified and traditional Chinese character recognition. *Cognitive Science*, *40*(8), 1941–1968. https://doi.org/10.1111/cogs.12307

Liu, Y., & Perfetti, C. A. (2003). The time course of brain activity in reading English and Chinese: An ERP study of Chinese bilinguals. *Human Brain Mapping*, *18*(3), 167–175. https://doi.org/10.1002/hbm.10090

Lobmaier, J. S., & Mast, F. W. (2007). Perception of novel faces: The parts have it! *Perception*, *36*(11), 1660–1673. https://doi.org/10.1068/p5642

Lochy, A., de Heering, A., & Rossion, B. (2019). The non-linear development of the right hemispheric specialization for human face perception. *Neuropsychologia*, *126*, 10–19. https://doi.org/10.1016/j.neuropsychologia.2017.06.029

Loffler, G., Yourganov, G., Wilkinson, F., & Wilson, H. R. (2005). FMRI evidence for the neural representation of faces. *Nature Neuroscience*, *8*(10), 1386–1391. https://doi.org/10.1038/nn1538

López-Barroso, D., Thiebaut de Schotten, M., Morais, J., Kolinsky, R., Braga, L. W., Guerreiro-Tauil, A., Dehaene, S., & Cohen, L. (2020). Impact of literacy on the functional connectivity of vision and language related networks. *NeuroImage*, *213*, 116722. https://doi.org/10.1016/j.neuroimage.2020.116722

Macchi Cassia, V. (2011). Age biases in face processing: The effects of experience across development: Age biases in face processing across development. *British Journal of Psychology*, *102*(4), 816–829. https://doi.org/10.1111/j.2044-8295.2011.02046.x

Macchi Cassia, V., Pisacane, A., & Gava, L. (2012). No own-age bias in 3-year-old children: More evidence for the role of early experience in building face-processing biases. *Journal of Experimental Child Psychology*, *113*(3), 372–382. https://doi.org/10.1016/j.jecp.2012.06.014

Malpass, R. S., & Kravitz, J. (1969). Recognition for faces of own and other race. *Journal of Personality and Social Psychology*, *13*(4), 330–334. https://doi.org/10.1037/h0028434

Mann, V. A. (1985). A cross-linguistic perspective on the relation between temporary memory skills and early reading ability. *Remedial and Special Education*, *6*(6), 37–42. https://doi.org/10.1177/074193258500600607

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, *98*(2), 224–253. https://doi.org/10.1037/0033-295X.98.2.224

Masuda, T., & Nisbett, R. E. (2001). Attending holistically versus analytically: Comparing the context sensitivity of Japanese and Americans. *Journal of Personality and Social Psychology*, *81*(5), 922–934. https://doi.org/10.1037/0022-3514.81.5.922

Masuda, T., & Nisbett, R. E. (2006). Culture and Change Blindness. *Cognitive Science*, *30*(2), 381–399. https://doi.org/10.1207/s15516709cog0000_63

Masuda, T., Russell, M. J., Chen, Y. Y., Hioki, K., & Caplan, J. B. (2014). N400 incongruity effect in an episodic memory task reveals different strategies for handling irrelevant contextual information for Japanese than European Canadians. *Cognitive Neuroscience*, *5*(1), 17–25. https://doi.org/10.1080/17588928.2013.831819

Maurer, D., Grand, R. L., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, *6*(6), 255–260. https://doi.org/10.1016/S1364-6613(02)01903-4

Maurer, U., Rossion, B., & McCandliss, B. D. (2008). Category specificity in early perception: face and word n170 responses differ in both lateralization and habituation properties. *Frontiers in human neuroscience*, 18. https://doi.org/10.3389/neuro.09.018.2008

Maurer, U., Brem, S., Kranz, F., Bucher, K., Benz, R., Halder, P., Steinhausen, H.-C., & Brandeis, D. (2006). Coarse neural tuning for print peaks when children learn to read. *NeuroImage*, *33*(2), 749–758. https://doi.org/10.1016/j.neuroimage.2006.06.025

Maurer, U., Schulz, E., Brem, S., der Mark, S. van, Bucher, K., Martin, E., & Brandeis, D. (2011). The development of print tuning in children with dyslexia: Evidence from longitudinal ERP data supported by fMRI. *NeuroImage*, *57*(3), 714–722. https://doi.org/10.1016/j.neuroimage.2010.10.055

Mcbride-Chang, C., Chow, B. W. Y., Zhong, Y., Burgess, S., & Hayward, W. G. (2005). Chinese character acquisition and visual skills in two Chinese scripts. *Reading and Writing*, *18*(2), 99–128. https://doi.org/10.1007/s11145-004-7343-5

McBride-Chang, C., Shu, H., Chan, W., Wong, T., Wong, A. M.-Y., Zhang, Y., Pan, J., & Chan, P. (2013). Poor readers of Chinese and English: overlap, stability, and longitudinal correlates. *Scientific Studies of Reading*, *17*(1), 57–70. https://doi.org/10.1080/10888438.2012.689787

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*(2), 256–262. https://doi.org/10.1016/j.jecp.2010.12.003

McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: Expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences*, *7*(7), 293–299. https://doi.org/10.1016/S1364-6613(03)00134-7

McGugin, R. W., Van Gulick, A. E., Tamber-Rosenau, B. J., Ross, D. A., & Gauthier, I. (2015). Expertise effects in face-selective areas are robust to clutter and diverted attention, but not to competition. *Cerebral Cortex*, *25*(9), 2610–2622. https://doi.org/10.1093/cercor/bhu060

McKone, E., Kanwisher, N., & Duchaine, B. C. (2007). Can generic expertise explain special processing for faces? *Trends in Cognitive Sciences*, *11*(1), 8–15. https://doi.org/10.1016/j.tics.2006.11.002

McKone, E., & Robbins, R. (2011). *Are faces special?* Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199559053.013.0009

McKone, E., Wan, L., Pidcock, M., Crookes, K., Reynolds, K., Dawel, A., Kidd, E., & Fiorentini, C. (2019). A critical period for faces: Other-race face recognition is improved by childhood but not adult social contact. *Scientific Reports*, *9*(1), 12820. https://doi.org/10.1038/s41598-019-49202-0

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law*, *7*(1), 3–35. https://doi.org/10.1037/1076-8971.7.1.3

Mercure, E., Dick, F., Halit, H., Kaufman, J., & Johnson, M. H. (2008). Differential lateralization for words and faces: category or psychophysics? *Journal of Cognitive Neuroscience*, *20*(11), 2070–2087. https://doi.org/10.1162/jocn.2008.20137

Michel, C. M., Seeck, M., & Landis, T. (1999). Spatiotemporal dynamics of human Cognition. *Physiology*, *14*(5), 206–214. https://doi.org/10.1152/physiologyonline.1999.14.5.206

Michel, C. M., Thut, G., Morand, S., Khateb, A., Pegna, A. J., Grave de Peralta, R., Gonzalez, S., Seeck, M., & Landis, T. (2001). Electric source imaging of human brain functions. *Brain Research Reviews*, *36*(2–3), 108–118. https://doi.org/10.1016/S0165-0173(01)00086-8

Michel, C., Rossion, B., Han, J., Chung, C.-S., & Caldara, R. (2006). Holistic processing is finely tuned for faces of one’s own race. *Psychological Science*, *17*(7), 608–615. https://doi.org/10.1111/j.1467-9280.2006.01752.x

Miellet, S., Vizioli, L., He, L., Zhou, X., & Caldara, R. (2013). Mapping face recognition information use across cultures. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00034

Miles, J., Shevlin, M., 2001. Applying regression and correlation: A guide for students and researchers. Sage, London.

Miyamoto, Y., Yoshikawa, S., & Kitayama, S. (2011). Feature and configuration in face processing: Japanese are more configural than Americans. *Cognitive Science*, *35*(3), 563–574. https://doi.org/10.1111/j.1551-6709.2010.01163.x

Mo, C., Yu, M., Seger, C., & Mo, L. (2015). Holistic neural coding of Chinese character forms in bilateral ventral visual system. *Brain and Language*, *141*, 28–34. https://doi.org/10.1016/j.bandl.2014.11.008

Mondloch, C. J., Le Grand, R., & Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. *Perception*, *31*(5), 553–566. https://doi.org/10.1068/p3339

Mondloch, C. J., Lewis, T. L., Budreau, D. R., Maurer, D., Dannemiller, J. L., Stephens, B. R., & Kleiner-Gathercoal, K. A. (1999). face perception during early infancy. *Psychological Science*, *10*(5), 419–422. https://doi.org/10.1111/1467-9280.00179

Mooney, C. M. (1957). Age in the development of closure ability in children. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *11*(4), 219–226. https://doi.org/10.1037/h0083717

Morton, J., & Johnson, M. H. (1991). CONSPEC and CONLERN: A two-process theory of infant face recognition. *Psychological Review*, *98*(2), 164–181. https://doi.org/10.1037/0033-295X.98.2.164

Murray, M. M., Brunet, D., & Michel, C. M. (2008). Topographic ERP analyses: A step-by-step tutorial review. *Brain Topography*, *20*(4), 249–264. https://doi.org/10.1007/s10548-008-0054-5

Nathan, R. G., & Stanovich, K. E. (1991). The causes and consequences of differences in reading fluency. *Theory Into Practice*, *30*(3), 176–184.

https://doi.org/10.1080/00405849109543498

Nestor, A., Behrmann, M., & Plaut, D. C. (2013). The neural basis of visual word form processing: A multivariate investigation. *Cerebral Cortex*, *23*(7), 1673–1684. https://doi.org/10.1093/cercor/bhs158

Nisbett, R. E., & Masuda, T. (2003). Culture and point of view. *Proceedings of the National Academy of Sciences*, *100*(19), 11163–11170. https://doi.org/10.1073/pnas.1934527100

Nisbett, R. E., & Miyamoto, Y. (2005). The influence of culture: Holistic versus analytic perception. *Trends in Cognitive Sciences*, *9*(10), 467–473. https://doi.org/10.1016/j.tics.2005.08.004

Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review*, *108*(2), 291–310. https://doi.org/10.1037/0033-295X.108.2.291

Noble, K. G., McCandliss, B. D., & Farah, M. J. (2007). Socioeconomic gradients predict individual differences in neurocognitive abilities. *Developmental Science*, *10*(4), 464–480. https://doi.org/10.1111/j.1467-7687.2007.00600.x

Obrig, H., Israel, H., Kohl-Bareis, M., Uludag, K., Wenzel, R., Müller, B., Arnold, G., & Villringer, A. (2002). Habituation of the visually evoked potential and its vascular response: Implications for neurovascular coupling in the healthy adult. *NeuroImage*, *17*(1), 1–18. https://doi.org/10.1006/nimg.2002.1177

O'Connor, K., Hamm, J. P., & Kirk, I. J. (2005). The neurophysiological correlates of face processing in adults and children with Asperger's syndrome. *Brain and Cognition*, *59*(1), 82–95. https://doi.org/10.1016/j.bandc.2005.05.004

O'Connor, K., Hamm, J. P., & Kirk, I. J. (2007). Neurophysiological responses to face, facial regions and objects in adults with Asperger's syndrome: An ERP investigation. *International Journal of Psychophysiology*, *63*(3), 283–293. https://doi.org/10.1016/j.ijpsycho.2006.12.001

Or, C. C.-F., Peterson, M. F., & Eckstein, M. P. (2015). Initial eye movements during face identification are optimal and similar across cultures. *Journal of Vision*, *15*(13), 12. https://doi.org/10.1167/15.13.12

Oruc, I., Balas, B., & Landy, M. S. (2019). Face perception: A brief journey through recent discoveries and current directions. *Vision Research*, *157*, 1–9. https://doi.org/10.1016/j.visres.2019.06.005

Papesh, M. H., & Goldinger, S. D. (2010). A multidimensional scaling analysis of own- and cross-race face spaces. *Cognition*, *116*(2), 283–288. https://doi.org/10.1016/j.cognition.2010.05.001

Park, J., Newman, L. I., & Polk, T. A. (2009). Face processing: The interplay of nature and nurture. *The Neuroscientist*, *15*(5), 445–449. https://doi.org/10.1177/1073858409337742

Pegado, F., Comerlato, E., Ventura, F., Jobert, A., Nakamura, K., Buiatti, M., Ventura, P., Dehaene-Lambertz, G., Kolinsky, R., Morais, J., Braga, L. W., Cohen, L., & Dehaene, S. (2014). Timing the impact of literacy on visual processing. *Proceedings of the National Academy of Sciences*, *111*(49), E5233–E5242. https://doi.org/10.1073/pnas.1417347111

Peng, G., Minett, J. W., & Wang, W. S.-Y. (2010). Cultural background influences the liminal perception of Chinese characters: An ERP study. *Journal of Neurolinguistics*, *23*(4), 416–426. https://doi.org/10.1016/j.jneuroling.2010.03.004

Perfetti, C. A., & Harris, L. N. (2013). Universal reading processes are modulated by language and writing system. *Language Learning and Development*, *9*(4), 296–316. https://doi.org/10.1080/15475441.2013.813828

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. Precursors of functional literacy, 11, 67-86.

Perfetti, C. A., Liu, Y., Fiez, J., Nelson, J., Bolger, D. J., & Tan, L.-H. (2007). Reading in two writing systems: Accommodation and assimilation of the brain's reading network. *Bilingualism: Language and Cognition*, *10*(2), 131–146. https://doi.org/10.1017/S1366728907002891

Pezdek, K., Blandon-Gitlin, I., & Moore, C. (2003). Children's face recognition memory: More evidence for the cross-race effect. *Journal of Applied Psychology*, *88*(4), 760–763. https://doi.org/10.1037/0021-9010.88.4.760

Piepers, D. W., & Robbins, R. A. (2012). A review and clarification of the rerms "holistic," "configural," and "relational" in the face perception literature. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00559

Pinel, P., Lalanne, C., Bourgeron, T., Fauchereau, F., Poupon, C., Artiges, E., Le Bihan, D., Dehaene-Lambertz, G., & Dehaene, S. (2015). Genetic and environmental influences on the visual word form and fusiform face areas. *Cerebral Cortex*, *25*(9), 2478–2493. https://doi.org/10.1093/cercor/bhu048

Polk, T. A., Park, J., Smith, M. R., & Park, D. C. (2007). Nature versus nurture in ventral visual cortex: A functional magnetic resonance imaging study of twins. *Journal of Neuroscience*, *27*(51), 13921–13925. https://doi.org/10.1523/JNEUROSCI.4001-07.2007

Pourtois, G., Thut, G., Grave de Peralta, R., Michel, C., & Vuilleumier, P. (2005). Two electrophysiological stages of spatial orienting towards fearful faces: Early temporoparietal activation preceding gain control in extrastriate visual cortex. *NeuroImage*, *26*(1), 149–163. https://doi.org/10.1016/j.neuroimage.2005.01.015

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. Precursors of functional literacy, 11, 67-86.

Rayner, K., Li, X., Williams, C. C., Cave, K. R., & Well, A. D. (2007). Eye movements during information processing tasks: Individual differences and cultural effects. *Vision Research*, *47*(21), 2714–2726. https://doi.org/10.1016/j.visres.2007.05.007

Reid, V. M., Dunn, K., Young, R. J., Amu, J., Donovan, T., & Reissland, N. (2017). The human fetus preferentially engages with face-like visual stimuli. *Current Biology*, *27*(12), 1825-1828.e3. https://doi.org/10.1016/j.cub.2017.05.044

Rhodes, G., Brake, S., Taylor, K., & Tan, S. (1989). Expertise and configural coding in face recognition. *British Journal of Psychology*, *80*(3), 313–331. https://doi.org/10.1111/j.2044-8295.1989.tb02323.x

Rhodes, G., Ewing, L., Hayward, W. G., Maurer, D., Mondloch, C. J., & Tanaka, J. W. (2009). Contact and other-race effects in configural and component processing of faces. *British Journal of Psychology*, *100*(4), 717–728. https://doi.org/10.1348/000712608X396503

Richler, J. J., Cheung, O. S., & Gauthier, I. (2011). Holistic processing predicts face recognition. *Psychological Science*, *22*(4), 464–471. https://doi.org/10.1177/0956797611401753

Richler, J. J., Palmeri, T. J., & Gauthier, I. (2012). Meanings, mechanisms, and measures of holistic processing. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00553

Rivers, W. H. R. (1905). Observations on the senses of the Todas. *British journal of Psychology*, 1(4), 321.

Robbins, R., & McKone, E. (2003). Can holistic processing be learned for inverted faces? *Cognition*, *88*(1), 79–107. https://doi.org/10.1016/S0010-0277(03)00020-9

Rodger, H., Kelly, D. J., Blais, C., & Caldara, R. (2010). Inverting faces does not abolish cultural diversity in eye movements. *Perception*, *39*(11), 1491–1503. https://doi.org/10.1068/p6750

Rosen, M. L., Sheridan, M. A., Sambrook, K. A., Meltzoff, A. N., & McLaughlin, K. A. (2018). Socioeconomic disparities in academic achievement: A multi-modal investigation of neural mechanisms in children and adolescents. *NeuroImage*, *173*, 298–310. https://doi.org/10.1016/j.neuroimage.2018.02.043

Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta Psychologica*, *128*(2), 274–289. https://doi.org/10.1016/j.actpsy.2008.02.003

Rossion, B., Collins, D., Goffaux, V., & Curran, T. (2007). Long-term expertise with artificial objects increases visual competition with early face categorization Processes. *Journal of Cognitive Neuroscience*, *19*(3), 543–555. https://doi.org/10.1162/jocn.2007.19.3.543

Rossion, B., & Curran, T. (2010). Visual expertise with pictures of cars correlates with RT magnitude of the car inversion effect. *Perception*, *39*(2), 173–183. https://doi.org/10.1068/p6270

Rossion, B., Dricot, L., Devolder, A., Bodart, J.-M., Crommelinck, M., Gelder, B. de, & Zoontjes, R. (2000). Hemispheric asymmetries for whole-based and part-based face processing in the human fusiform gyrus. *Journal of Cognitive Neuroscience*, *12*(5), 793–802. https://doi.org/10.1162/089892900562606

Rossion, B., Gauthier, I., Goffaux, V., Tarr, M. J., & Crommelinck, M. (2002). Expertise training with novel objects leads to left-lateralized facelike electrophysiological responses. *Psychological Science*, *13*(3), 250–257. https://doi.org/10.1111/1467-9280.00446

Rossion, B., Gauthier, I., Tarr, M. J., Despland, P., Bruyer, R., Linotte, S., & Crommelinck, M. (2000). The N170 occipito-temporal component is delayed and enhanced to inverted faces but not to inverted objects: An electrophysiological account of face-specific processes in the human brain. *NeuroReport*, *11*(1).

Rossion, B., & Jacques, C. (2008). Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170. *NeuroImage*, *39*(4), 1959–1979. https://doi.org/10.1016/j.neuroimage.2007.10.011

Rossion, B., & Jacques, C. (2011). The N170: Understanding the time course of face perception in the human brain. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195374148.013.0064

Rossion, B., Joyce, C. A., Cottrell, G. W., & Tarr, M. J. (2003). Early lateralization and orientation tuning for face, word, and object processing in the visual cortex. *NeuroImage*, *20*(3), 1609–1624. https://doi.org/10.1016/j.neuroimage.2003.07.010

Rossion, B., Kung, C.-C., & Tarr, M. J. (2004). Visual expertise with nonface objects leads to competition with the early perceptual processing of faces in the human occipitotemporal cortex. *Proceedings of the National Academy of Sciences*, *101*(40), 14521–14526. https://doi.org/10.1073/pnas.0405613101

Rossion, B., & Lochy, A. (2021). Is human face recognition lateralized to the right hemisphere due to neural competition with left-lateralized visual word recognition? A critical review. *Brain Structure and Function*. https://doi.org/10.1007/s00429-021-02370-0

Rousselet, G. A., Mace, M. J.-M., & Fabre-Thorpe, M. (2004). Animal and human faces in natural scenes: How specific to human faces is the N170 ERP component? *Journal of Vision*, *4*(1), 2–2. https://doi.org/10.1167/4.1.2

Sadeh, B., & Yovel, G. (2010). Why is the N170 enhanced for inverted faces? An ERP competition experiment. *NeuroImage*, *53*(2), 782–789. https://doi.org/10.1016/j.neuroimage.2010.06.029

Sagiv, N., & Bentin, S. (2001). Structural encoding of human and schematic faces: holistic and part-based processes. *Journal of Cognitive Neuroscience*, *13*(7), 937–951. https://doi.org/10.1162/089892901753165854

Samuels, S. J., LaBerge, D., & Bremer, C. D. (1978). Units of word recognition: Evidence for developmental changes. *Journal of Verbal Learning and Verbal Behavior*, *17*(6), 715–720. https://doi.org/10.1016/S0022-5371(78)90433-4

Sangrigoli, S., & de Schonen, S. (2004). Effect of visual experience on face processing: A developmental study of inversion and non-native effects. *Developmental Science*, *7*(1), 74–87. https://doi.org/10.1111/j.1467-7687.2004.00324.x

Sangrigoli, S., Pallier, C., Argenti, A.-M., Ventureyra, V. a. G., & de Schonen, S. (2005). Reversibility of the other-race effect in face recognition during childhood. *Psychological Science*, *16*(6), 440–444. https://doi.org/10.1111/j.0956-7976.2005.01554.x

Schendan, H. E., Ganis, G., & Kutas, M. (1998). Neurophysiological evidence for visual perceptual categorization of words and faces within 150 ms. *Psychophysiology*, *35*(3), 240–251. https://doi.org/10.1111/1469-8986.3530240

Schinkel, S., Ivanova, G., Kurths, J., & Sommer, W. (2014). Modulation of the N170 adaptation profile by higher level factors. *Biological Psychology*, *97*, 27–34. https://doi.org/10.1016/j.biopsycho.2014.01.003

Schultz, R. T. (2005). Developmental deficits in social perception in autism: The role of the amygdala and fusiform face area. *International Journal of Developmental Neuroscience*, *23*(2–3), 125–141. https://doi.org/10.1016/j.ijdevneu.2004.12.012

Schulz, C., Kaufmann, J. M., Walther, L., & Schweinberger, S. R. (2012). Effects of anticaricaturing vs. Caricaturing and their neural correlates elucidate a role of shape for face learning. *Neuropsychologia*, *50*(10), 2426–2434. https://doi.org/10.1016/j.neuropsychologia.2012.06.013

Schweinberger, S. R., & Neumann, M. F. (2016). Repetition effects in human ERPs to faces. *Cortex*, *80*, 141–153. https://doi.org/10.1016/j.cortex.2015.11.001

Scott, L. S., & Nelson, C. A. (2006). Featural and configural face processing in adults and infants: A behavioral and electrophysiological investigation. *Perception*, *35*(8), 1107–1128. https://doi.org/10.1068/p5493

Scott, L. S., Tanaka, J. W., Sheinberg, D. L., & Curran, T. (2006). A reevaluation of the electrophysiological correlates of expert object processing. *Journal of Cognitive Neuroscience*, *18*(9), 1453–1465. https://doi.org/10.1162/jocn.2006.18.9.1453

Scott, L. S., Tanaka, J. W., Sheinberg, D. L., & Curran, T. (2008). The role of category learning in the acquisition and retention of perceptual expertise: A behavioral and neurophysiological study. *Brain Research*, *1210*, 204–215. https://doi.org/10.1016/j.brainres.2008.02.054

Searston, R. A., & Tangen, J. M. (2017a). Expertise with unfamiliar objects is flexible to changes in task but not changes in class. *PLOS ONE*, *12*(6), e0178403. https://doi.org/10.1371/journal.pone.0178403

Searston, R. A., & Tangen, J. M. (2017b). The emergence of perceptual expertise with fingerprints over time. *Journal of Applied Research in Memory and Cognition*, *6*(4), 442–451. https://doi.org/10.1037/h0101814

Senholzi, K. B., & Ito, T. A. (2013). Structural face encoding: How task affects the N170's sensitivity to race. *Social Cognitive and Affective Neuroscience*, *8*(8), 937–942. https://doi.org/10.1093/scan/nss091

Sergent, J. (1982). The cerebral balance of power: Confrontation or cooperation? *Journal of Experimental Psychology: Human Perception and Performance*, *8*(2), 253–272. https://doi.org/10.1037/0096-1523.8.2.253

Shakeshaft, N. G., & Plomin, R. (2015). Genetic specificity of face recognition. *Proceedings of the National Academy of Sciences*, *112*(41), 12887–12892. https://doi.org/10.1073/pnas.1421881112

Sigman, M., & Gilbert, C. D. (2000). Learning to find a shape. *Nature Neuroscience*, *3*(3), 264–269. https://doi.org/10.1038/72979

Siok, W. T., Spinks, J. A., Jin, Z., & Tan, L. H. (2009). Developmental dyslexia is characterized by the co-existence of visuospatial and phonological disorders in Chinese children. *Current Biology: CB*, *19*(19), R890-892. https://doi.org/10.1016/j.cub.2009.08.014

Stahl, J., Wiese, H., & Schweinberger, S. R. (2008). Expertise and own-race bias in face processing: An event-related potential study: *NeuroReport*, *19*(5), 583–587. https://doi.org/10.1097/WNR.0b013e3282f97b4d

Stahl, J., Wiese, H., & Schweinberger, S. R. (2010). Learning task affects ERP-correlates of the own-race bias, but not recognition memory performance. *Neuropsychologia*, *48*(7), 2027–2040. https://doi.org/10.1016/j.neuropsychologia.2010.03.024

Szwed, M., Dehaene, S., Kleinschmidt, A., Eger, E., Valabrègue, R., Amadon, A., & Cohen, L. (2011). Specialization for written words over objects in the visual cortex. *NeuroImage*, *56*(1), 330–344. https://doi.org/10.1016/j.neuroimage.2011.01.073

Szwed, M., Ventura, P., Querido, L., Cohen, L., & Dehaene, S. (2012). Reading acquisition enhances an early visual process of contour integration. *Developmental Science*, *15*(1), 139–149. https://doi.org/10.1111/j.1467-7687.2011.01102.x

Tan, L. H., Spinks, J. A., Eden, G. F., Perfetti, C. A., & Siok, W. T. (2005). Reading depends on writing, in Chinese. *Proceedings of the National Academy of Sciences*, *102*(24), 8781–8785. https://doi.org/10.1073/pnas.0503523102

Tanaka, J. W., & Curran, T. (2001). A neural basis for expert object recognition. *Psychological Science*, *12*(1), 43–47. https://doi.org/10.1111/1467-9280.00308

Tanaka, J. W., Curran, T., & Sheinberg, D. L. (2005). The training and transfer of real-world perceptual expertise. *Psychological Science*, *16*(2), 145–151. https://doi.org/10.1111/j.0956-7976.2005.00795.

Tanaka, J. W., & Farah, M. J. (1991). Second-order relational properties and the inversion effect: Testing a theory of face perception. *Perception & Psychophysics*, *50*(4), 367–372. https://doi.org/10.3758/BF03212229

Tanaka, J. W., & Gordon, I. (2011). Features, configuration, and holistic face processing. The Oxford handbook of face perception, 177-194.

Tanaka, J. W., Heptonstall, B., & Hagen, S. (2013). Perceptual expertise and the plasticity of other-race face recognition. *Visual Cognition*, *21*(9–10), 1183–1201. https://doi.org/10.1080/13506285.2013.826315

Tanaka, J. W., Kiefer, M., & Bukach, C. M. (2004). A holistic account of the own-race effect in face recognition: Evidence from a cross-cultural study. *Cognition*, *93*(1), B1–B9. https://doi.org/10.1016/j.cognition.2003.09.011

Tanaka, J. W., & Pierce, L. J. (2009). The neural plasticity of other-race face recognition. *Cognitive, Affective, & Behavioral Neuroscience*, *9*(1), 122–131. https://doi.org/10.3758/CABN.9.1.122

Tarr, M. J., & Cheng, Y. D. (2003). Learning to see faces and objects. *Trends in Cognitive Sciences*, *7*(1), 23–30. https://doi.org/10.1016/S1364-6613(02)00010-4

Tarr, M. J., & Gauthier, I. (2000). FFA: A flexible fusiform area for subordinate-level visual processing automatized by expertise. *Nature Neuroscience*, *3*(8), 764–769. https://doi.org/10.1038/77666

Taubert, J., Apthorp, D., Aagten-Murphy, D., & Alais, D. (2011). The role of holistic processing in face perception: Evidence from the face inversion effect. *Vision Research*, *51*(11), 1273–1278. https://doi.org/10.1016/j.visres.2011.04.002

Towler, A., Kemp, R. I., & White, D. (2017). Unfamiliar face matching systems in applied settings. *Face processing: systems, disorders and cultural differences. New York: Nova Science Publishing, Inc.*

Tso, R. V., Au, T. K., & Hsiao, J. H. (2014). Perceptual expertise: Can sensorimotor experience change holistic processing and left-side bias? *Psychological Science*, *25*(9), 1757–1767. https://doi.org/10.1177/0956797614541284

Tso, R. V., Chan, R. T., Chan, Y., & Lin, D. (2021). Holistic processing of Chinese characters in college students with dyslexia. *Scientific Reports*, *11*(1), 1973. https://doi.org/10.1038/s41598-021-81553-5

Tso, R. V. Y., Au, T. K. -f., & Hsiao, J. H. -w. (2012). Writing facilitates learning to read in Chinese through reduction of holistic processing: A developmental study. *Journal of Vision*, *12*(9), 530–530. https://doi.org/10.1167/12.9.530

Tso, R. V. Y., Chan, R. T. C., & Hsiao, J. H. (2020). Holistic but with reduced right-hemisphere involvement: The case of dyslexia in Chinese character recognition. *Psychonomic Bulletin & Review*, *27*(3), 553–562. https://doi.org/10.3758/s13423-020-01721-y

Turati, C., Macchi Cassia, V., Simion, F., & Leo, I. (2006). Newborns' face recognition: Role of inner and outer facial features. *Child Development*, *77*(2), 297–311. https://doi.org/10.1111/j.1467-8624.2006.00871.x

Turati, C., Sangrigoli, S., Ruely, J., & Schonen, S. (2004). Evidence of the face inversion effect in 4-month-old infants. *Infancy*, *6*(2), 275–297. https://doi.org/10.1207/s15327078in0602_8

Vakli, P., NÃ©meth, K., Zimmer, M., Schweinberger, S. R., & KovÃ¡cs, G. (2014). Altering second-order configurations reduces the adaptation effects on early face-sensitive event-related potential components. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00426

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *The Quarterly Journal of Experimental Psychology Section A*, *43*(2), 161–204. https://doi.org/10.1080/14640749108400966

Valenza, E., Simion, F., Cassia, V. M., & Umiltà, C. (1996). Face preference at birth. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(4), 892–903. https://doi.org/10.1037/0096-1523.22.4.892

van Paridon, J., Ostarek, M., Arunkumar, M., & Huettig, F. (2021). Does neuronal recycling result in destructive competition? The influence of learning to read on the recognition of faces. *Psychological Science*, *32*(3), 459–465. https://doi.org/10.1177/0956797620971652

Varnum, M. E. W., Grossmann, I., Kitayama, S., & Nisbett, R. E. (2010). The origin of cultural differences in cognition: Evidence for the social orientation hypothesis. *Current Directions in Psychological Science*, *19*(1), 9–13. https://doi.org/10.1177/0963721409359301

Ventura, P., Domingues, M., Ferreira, I., Madeira, M., Martins, A., Neto, M., & Pereira, M. (2019). Holistic word processing is involved in fast parallel reading. *Journal of Expert*, 2, 47-58.

Ventura, P., Fernandes, T., Cohen, L., Morais, J., Kolinsky, R., & Dehaene, S. (2013). Literacy acquisition reduces the influence of automatic holistic processing of faces and houses. *Neuroscience Letters*, *554*, 105–109. https://doi.org/10.1016/j.neulet.2013.08.068

Vizioli, L., Foreman, K., Rousselet, G. A., & Caldara, R. (2011). Inverting faces elicits sensitivity to race on the N170 component: A cross-cultural study. *Journal of Vision*, *10*(1), 15–15. https://doi.org/10.1167/10.1.15

Vizioli, L., Rousselet, G. A., & Caldara, R. (2010). Neural repetition suppression to identity is abolished by other-race faces. *Proceedings of the National Academy of Sciences*, *107*(46), 20081–20086. https://doi.org/10.1073/pnas.1005751107

Walker, P. M., Silvert, L., Hewstone, M., & Nobre, A. C. (2008). Social contact and other-race face processing in the human brain. *Social Cognitive and Affective Neuroscience*, *3*(1), 16–25. https://doi.org/10.1093/scan/nsm035

Walker, P. M., & Tanaka, J. W. (2003). An encoding advantage for own-race versus other-race faces. *Perception*, *32*(9), 1117–1125. https://doi.org/10.1068/p5098

Walle de Ghelcke, A., Rossion, B., Schiltz, C., & Lochy, A. (2021). Developmental changes in neural letter‐selectivity: A 1‐ year follow‐ up of beginning readers. *Developmental Science*, *24*(1). https://doi.org/10.1111/desc.12999

Wang, M.-Y., Kuo, B.-C., & Cheng, S.-K. (2011). Chinese characters elicit face-like N170 inversion effects. *Brain and Cognition*, *77*(3), 419–431. https://doi.org/10.1016/j.bandc.2011.08.016

Wang, Q., Xiao, N. G., Quinn, P. C., Hu, C. S., Qian, M., Fu, G., & Lee, K. (2015). Visual scanning and recognition of Chinese, Caucasian, and racially ambiguous faces: Contributions from bottom-up facial physiognomic information and top-down knowledge of racial categories. *Vision Research*, *107*, 67–75. https://doi.org/10.1016/j.visres.2014.10.032

Wastell, D. G., & Kleinman, D. (1980). Potentiation of the habituation of human brain potentials. *Biological Psychology*, *10*(1), 21–29. https://doi.org/10.1016/0301-0511(80)90004-6

White, D., Phillips, P. J., Hahn, C. A., Hill, M., & O'Toole, A. J. (2015). Perceptual expertise in forensic facial image comparison. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1814), 20151292. https://doi.org/10.1098/rspb.2015.1292

Wiese, H., Kaufmann, J. M., & Schweinberger, S. R. (2014). The neural signature of the own-race bias: Evidence from event-related potentials. *Cerebral Cortex*, *24*(3), 826–835. https://doi.org/10.1093/cercor/bhs369

Wiese, H., Stahl, J., & Schweinberger, S. R. (2009). Configural processing of other-race faces is delayed but not decreased. *Biological Psychology*, *81*(2), 103–109. https://doi.org/10.1016/j.biopsycho.2009.03.002

Willenbockel, V., Fiset, D., Chauvin, A., Blais, C., Arguin, M., Tanaka, J. W., Bub, D. N., & Gosselin, F. (2010). Does face inversion change spatial frequency tuning? *Journal of Experimental Psychology: Human Perception and Performance*, *36*(1), 122–135. https://doi.org/10.1037/a0016465

Wilhelm, O., Schroeders, U., & Schipolowski, S. (2014). Berliner Test zur Erfassung fluider und kristalliner Intelligenz für die 8. bis 10. Jahrgangsstufe (BEFKI 8-10).

Wilmer, J. B., Germine, L., Chabris, C. F., Chatterjee, G., Williams, M., Loken, E., Nakayama, K., & Duchaine, B. (2010). Human face recognition ability is specific and highly heritable. *Proceedings of the National Academy of Sciences*, *107*(11), 5238–5241. https://doi.org/10.1073/pnas.0913053107

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

https://doi.org/10.1167/12.13.7

Wong, A. C.-N., Bukach, C. M., Yuen, C., Yang, L., Leung, S., & Greenspon, E. (2011). Holistic processing of words modulated by reading experience. *PLoS ONE*, *6*(6), e20753. https://doi.org/10.1371/journal.pone.0020753

Wong, A. C.-N., Wong, Y. K., Lui, K. F. H., Ng, T. Y. K., & Ngan, V. S. H. (2019). Sensitivity to configural information and expertise in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(1), 82–99. https://doi.org/10.1037/xhp0000590

Xu, Y., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature*, *440*(7080), 91–95. https://doi.org/10.1038/nature04262

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), 933–947. https://doi.org/10.1037/0096-1523.28.4.933

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. https://doi.org/10.1080/13506280344000077

Yi, D.-J., Olson, I. R., & Chun, M. M. (2006). Shape-specific perceptual learning in a figure-ground segregation task. *Vision Research*, *46*(6–7), 914–924. https://doi.org/10.1016/j.visres.2005.09.009

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, *81*(1), 141–145. https://doi.org/10.1037/h0027474

Young, A. W., & Burton, A. M. (2018). Are we face experts? *Trends in Cognitive Sciences*, *22*(2), 100–110. https://doi.org/10.1016/j.tics.2017.11.007

Young, S. G., Hugenberg, K., Bernstein, M. J., & Sacco, D. F. (2012). Perception and motivation in face recognition: A critical review of theories of the cross-race effect. *Personality and Social Psychology Review*, *16*(2), 116–142. https://doi.org/10.1177/1088868311418987

Yovel, G., & Duchaine, B. (2006). Specialized face perception mechanisms extract both part and spacing information: Evidence from developmental prosopagnosia. *Journal of Cognitive Neuroscience*, *18*(4), 580–593. https://doi.org/10.1162/jocn.2006.18.4.580

Yovel, G., & Kanwisher, N. (2004). Face perception. *Neuron*, *44*(5), 889–898. https://doi.org/10.1016/j.neuron.2004.11.018

Yum, Y. N., & Law, S.-P. (2019). Impact of characteristics of L1 literacy experience on picture processing: ERP data from trilingual non-native Chinese and English readers. *Cognition*, *183*, 213–225. https://doi.org/10.1016/j.cognition.2018.11.009

Zhong, N., Hsiao, J., Zhou, G., & Hayward, W. (2021). Idiosyncratic eye-movement patterns modulate holistic processing of faces: Evidence from the composite face effect and the inverted face effect. *Journal of Vision*, *21*(9), 1851. https://doi.org/10.1167/jov.21.9.1851

Zhou, G., Cheng, Z., Zhang, X., & Wong, A. C.-N. (2012). Smaller holistic processing of faces associated with face drawing experience. *Psychonomic Bulletin & Review*, *19*(2), 157–162. https://doi.org/10.3758/s13423-011-0174-x

Zhou, X., Itz, M. L., Vogt, S., Kaufmann, J. M., Schweinberger, S. R., & Mondloch, C. J. (2021). Similar use of shape and texture cues for own- and other-race faces during face learning and recognition. *Vision Research*, *188*, 32–41. https://doi.org/10.1016/j.visres.2021.06.014

Zhu, Q., Song, Y., Hu, S., Li, X., Tian, M., Zhen, Z., Dong, Q., Kanwisher, N., & Liu, J. (2010). Heritability of the specific cognitive ability of face perception. *Current Biology*, *20*(2), 137–142. https://doi.org/10.1016/j.cub.2009.11.067

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, *131*(1), 3–29. https://doi.org/10.1037/0033-2909.131.1.3

Zimmer, M., & Kovács, G. (2011). Electrophysiological correlates of face distortion after-effects. *Quarterly Journal of Experimental Psychology*, *64*(3), 533–544. https://doi.org/10.1080/17470218.2010.501964

Zion-Golumbic, E., & Bentin, S. (2007). Dissociated neural mechanisms for face detection and configural encoding: evidence from N170 and induced gamma-band oscillation Effects. *Cerebral Cortex*, *17*(8), 1741–1749. https://doi.org/10.1093/cercor/bhl100

Declaration

This work was supported by a doctoral scholarship from the China Scholarship Council (CSC) under Grant [No. 201708330288] to Xiaoli Ma. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt,

1. dass keine Zusammenarbeit mit gewerblichen Promotionsberatern stattfand,
2. dass ich die Dissertation auf der Grundlage der angegebenen Hilfsmittel und Hilfen selbstständig angefertigt habe,
3. dass ich mich nicht anderwärts um einen Doktorgrad beworben habe bzw. einen entsprechenden Doktorgrad besitze,
4. dass mir die dem angestrebten Verfahren zugrunde liegende Promotionsordnung der Lebenswissenschaftlichen Fakultät vom 05. März 2015, veröffentlicht im Amtlichen Mitteilungsblatt der Humboldt-Universität zu Berlin Nr. 12/2015 bekannt ist,
5. dass die Dissertation oder Teile davon nicht bereits bei einer anderen wissenschaftlichen Einrichtung eingereicht, angenommen oder abgelehnt wurden,
6. dass die Grundsätze der Humboldt-Universität zu Berlin zur Sicherung guter wissenschaftlicher Praxis eingehalten wurden.

Frankfurt, den 16. June, 2022

Xiaoli Ma