

UC Irvine

UC Irvine Electronic Theses and Dissertations

Title

A Mnemonic Discrimination Account for the Behavioral and Neural Correlates Underlying the Other-Race Effect

Permalink

<https://escholarship.org/uc/item/5012d5sv>

Author

Yaros, Jessica L

Publication Date

2021

Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NonCommercial-NoDerivatives License, available at <https://creativecommons.org/licenses/by-nc-nd/4.0/>

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

A Mnemonic Discrimination Account for the Behavioral and
Neural Correlates Underlying the Other-Race Effect

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Neurobiology and Behavior

by

Jessica L Yaros

Dissertation Committee:
Professor Michael A. Yassa, Chair
Distinguished Professor Elizabeth F. Loftus
Professor Craig E. Stark
Professor Georg Striedter

2021

DEDICATION

To

my soulmate,
my soulcats,
my family,
biological
and
adopted.

TABLE OF CONTENTS

	Page
LIST OF FIGURES	v
LIST OF TABLES	vi
ACKNOWLEDGEMENTS.....	vii
VITA.....	viii
ABSTRACT	ix
Introduction.....	1
Chapter 1: Cognitive Frameworks for the Other-Race Effect.....	5
Early Explanatory Theories of the Other-Race Effect.....	5
Current Theories Accounting for the Other-Race Effect.....	8
<i>Perceptual Expertise Theories</i>	8
<i>Perceptual Expertise Mechanisms</i>	11
<i>Social-Cognitive Theories</i>	16
<i>Social-Cognitive Mechanisms</i>	18
<i>Integrated Theories</i>	20
Chapter 2: Neural Correlates of the Other-Race Effect	23
Perceptual Expertise and Holistic Processing	23
Attentional and Social Modulation	28
Chapter 3: A Memory Model of the ORE: Differential Mnemonic Discrimination of Same and Other-Race Faces.....	31
Introduction	31
Methods	35
Results.....	39
Discussion	47
Chapter 4: Transitioning from the Medial-Temporal Lobe to the Whole Brain in Investigating Neural Correlates of the Other-Race Effect.....	54
Nonoptimal Behavioral Results for Testing Pattern Separation Hypotheses.....	54
Novelty of Mnemonic Discrimination Paradigm.....	56
Exploratory Data Analysis Informing Subsequent Analysis Methods	57
Chapter 5: Review of Relevant Works in Network Neuroscience	61
Chapter 6: Graph Theoretical Analysis Reveals Reconfigurations in Network Efficiency across Same- and Other-Race Facial Recognition	67

Introduction.....	67
Methods.....	71
Results.....	81
Discussion.....	96
Conclusion.....	102
Abbreviations.....	110
References.....	111

LIST OF FIGURES

	Page
Figure 3. 1 Mnemonic Discrimination and Match-to-Sample task designs	39
Figure 3. 2 Discriminability Results from Mnemonic Discrimination and Match-to-Sample Tasks.	41
Figure 3. 3 Comparison of ORE in Mnemonic Discrimination and Match-to-Sample tasks	43
Figure 3. 4 Reaction Time Differences	45
Figure 4. 1 Contrasts of encoding activity for successful and unsuccessful subsequent trials	59
Figure 6. 1 Mnemonic Discrimination and Perceptual Discrimination Task Designs.....	74
Figure 6. 2 Differences in Recognition Memory for Same and Other-Race faces.....	83
Figure 6. 3 Local and Global Efficiency during Encoding of faces with subsequent distractors	90
Figure 6. 4 Race differences in local efficiency across networks during accurate and inaccurate encoding	94
Figure 6. 5 Accuracy differences in local efficiency across networks during same and other-race face encoding.....	95

LIST OF TABLES

	Page
Table 3. 1 Linear Analysis Results	43
Table 6. 1 Descriptive statistics for global and local efficiency during each task condition	86
Table 6. 2 Estimates for the interaction of race and accuracy across all conditions	88
Table 6. 3 Descriptive Statistics for Local Efficiency Across Intrinsic Connectivity Networks.....	91

ACKNOWLEDGEMENTS

I would like to thank Professor Michael Yassa for his unconditional support throughout the entirety of my PhD career. I could not have asked for a more inspiring or understanding mentor. I am so fortunate to have worked under an advisor who let me follow my passions, despite not being the major focus of the lab's work, and who encouraged my extracurricular interests in science communication and data science.

Thank you as well to my committee members, Professor Beth Loftus, Professor Craig Stark and Professor Georg Striedter for guiding and improving my research, making observations that I had missed and advising me on approaches to better support my claims.

Thank you to Professor Paul Rapp and Professor Gilpin whose class first exposed me to the concept of graph theoretical analysis, inspiring the analysis approach used for the neuroimaging portion of this dissertation.

Thank you to Luis Colon-Perez and Sarah Kark for brainstorming sessions and advice on how to approach analyzing the neuroimaging data.

Thank you to many other members of the Yassa lab family for their role in shaping the work included in this dissertation. In no particular order, thanks to Zach Reagh, Allen Chang, and Craig Stark for advice on experimental design; Derek Delisle and John Janecek for help programming the experiments; Diana Salama and Edge Yu for generation of experimental stimuli; Liv McMillan and Elizabeth Murray for managing the experiments and their protocols, and keeping the lab in one piece; Jessica Noche Lingad, Myra Larson, and Martina Hollearn for participant recruitment and overseeing data entry and research assistants; Bethany Houalla for management, recruitment, and scheduling of participants, as well as analysis of behavioral data from the MRI study; Lea Stith, Rachon Sweiss, and Edge Yu for flyering; Joren Adams, Martina Hollearn, Blake Miranda, and Mithra Sathiskumar for training on the MRI scanner; Robert Jirsarai for developing the neuroimaging protocol; Steve Granger, Mithra Sathiskumar, Sarah Kark, Natalie DiProspero, Caro Chwiesko, Soyun Kim, and Jenna Adams for fielding neuroimaging analysis questions; Megan Smith and Kate Kirby for statistical consulting; Fellow graduate students Miranda Chappel-Farley, Natalie DiProspero, Freddie Marquez, Maria Montchal, Sandra Gattas, Steve Granger, Logan Harriger, and Rebecca Stevenson for lending their advice and friendship over the years.

Thank you to David Beier for editing emails, applications, and publications over the years.

And finally, thank you to all the undergraduate research assistants who collected data, and without whom none of this work would be possible. While I no doubt am missing people and apologize for that, thank you to Will Cabanela, Liam Campbell, Haitham Chehabi, Vishali Kapoor, Vincent Le, Abanoub Mikhail, Saum Naderi, Ali Ozgur, Rachon Sweiss, Hung Tran, and Edge Yu.

VITA

Jessica L Yaros

- 2010 Research Assistant, Cognitive Development Laboratory,
University of California San Diego
- 2011 Research Assistant, Systems Neuroscience Laboratory,
University of California San Diego
- 2011 Teaching Assistant, Cognitive Science Department,
University of California San Diego
- 2011-12 Research Assistant, Laboratory for Sleep and Behavioral Neuroscience,
University of California San Diego
- 2012 B.S. in Cognitive Science with a specialization in Neuroscience,
University of California, San Diego
- 2013-15 Project Assistant, Laboratory of Neuro Imaging,
University of California, Los Angeles, and University of Southern California
- 2016-18 Teaching Assistant, Department of Biological Science and Department of
Neurobiology and Behavior, University of California, Irvine
- 2016-18 Educational Science Podcast Host, Science for The People
- 2018 M.S. in Neurobiology and Behavior, University of California, Irvine
- 2018 – 19 Science Writing Fellowship, The Loh Down on Science, Science
Communications Initiative, University of California Irvine
- 2020-21 Capstone Project: Development of Clinical Literature Database,
Medtronic
- 2021 Ph.D. in Neurobiology and Behavior

PUBLICATIONS

Yaros JL, Salama DA, Delisle D, Larson MS, Miranda BA, Yassa MA. A Memory Computational Basis for the Other-Race Effect. *Scientific Reports* 9, 19399 (2019).

ABSTRACT

A Mnemonic Discrimination Account for the Behavioral and
Neural Correlates Underlying the Other-Race Effect

by

Jessica L. Yaros

Doctor of Philosophy in Neurobiology & Behavior

University of California, Irvine, 2021

Professor Michael A. Yassa, Chair

People often recognize and remember faces of individuals within their own race more easily than those of other races. While behavioral research has long suggested that this so-called Other-Race Effect (ORE) is due to extensive experience with one's own race group, the neural mechanisms underlying the effect remain unclear. Prior neuroimaging research has explored differences in perceptual processing of same and other-race faces, overlooking the important contributions of mnemonic processes in shaping the ORE. The work comprised in this dissertation attempts to fill this gap, characterizing a framework in which asymmetries in memory mechanisms give rise to the ORE. In a series of experiments employing mnemonic discrimination (pattern separation) paradigms based on computational models of episodic memory, we uncovered both behavioral and neural correlates of the ORE. In an initial set of studies, the ORE was demonstrated to be driven by differences in successful memory discrimination across races as a function of degree of interference between face stimuli. In follow-up studies we characterized different functional brain network properties conducive to successful same and other-race face recognition. Together these findings suggest that the ORE may emerge in part due to tuned

memory mechanisms that may enhance same-race, at the expense of other-race face detection. Furthermore, brain network connectivity during memorization that predicted successful same-race recognition was found to be maladaptive for other-race recognition, suggesting employment of different cognitive strategies may be optimal depending on the race of face being processed. Given that cross-race eyewitness identifications disproportionately contribute to wrongful convictions by the criminal justice system, this work should further motivate the development of procedures to mitigate the impact of the ORE on eyewitness testimony.

Introduction

The Other-Race Effect (ORE) is the tendency to be worse at recognizing and remembering faces outside of one's own race group. The effect was first academically documented in 1914, in an early study of environmental influence on visual discrimination¹. Even then, it was anecdotally "well known that, other things being equal, individuals of a given race are distinguishable from each other in proportion to our familiarity, to our contact with the race as a whole." Still, over a half-century passed before the effect was first empirically studied in 1969. Since then, the ORE has become one of the most replicated phenomena in face perception^{2,3}.

Researchers of the effect have never settled on a unified name. While I choose to use 'Other-Race Effect' throughout this dissertation, the phenomenon has also been coined as the 'own-race bias', 'own-race effect', 'cross-race effect', 'cross-race bias', 'own-group-bias', and 'own-ethnicity-bias'. Whichever name is subscribed to, one thing is agreed upon—the effect is dangerous. On an interpersonal and social level, being mistaken for someone else due to one's race affiliation contributes to anxiety and social isolation^{4,5}. But greater than that, the ORE can put individuals at risk when it manifests in the context of person identification. Verification of identity in airports, security-sensitive government or business buildings, and banks all rely on personnel accurately matching the face of a live person to a photograph. However, research continually demonstrates people struggle with this task, even when identifying same-race faces⁶. Identity matching of other-race faces becomes even more challenging^{7,8}. And if matching a live face to a photo is difficult, making an

accurate identification from memory is even more challenging. Therefore eye-witness testimony can be a controversial topic. Research has shown that memory is fallible, and the ORE only serves to complicate matters.

The first academic documentation of the ORE acknowledged this much, and it should be no surprise it was published in what is now called, the Journal of Criminal Law and Criminology¹:

“I admit that the identification of a foreigner in the same environment in which, not he, but a member of his race had been seen before, might result in false recognition¹ (50).”

Despite a seeming public acknowledgement of the ORE, for years the lack of empirical evidence supporting it likely took a great toll on its victims. The Quincy Five was one such case in 1971 involving five Black men wrongfully convicted based solely on eyewitness testimony. Allegedly the presiding judge would not allow a prominent psychologist to serve as an expert witness since only three studies on the ORE had been conducted at the time³. Fortunately, the five were pardoned (several of whom became the US's first ever cases of exoneration from death row) and the actual perpetrators charged. And yet today, with a wealth of research validating the ORE and signaling its risk in the context of eyewitness testimony, the impact of the effect is still quite apparent through the statistics of wrongful convictions.

To date the legal organization, The Innocence Project reports that amongst 350 exonerations by DNA evidence, 42% involved mistaken cross-race identifications⁹. Similarly The National Registry of Exonerations reports that 67% of overturned sexual

assault convictions involved mistaken identification, in half of which the defendant was Black and the victim was white^{10,11}.

There have been several attempts to quantify the damages caused by wrongful convictions. Unsurprisingly, exonerees' lives are severely compromised. A comprehensive survey found that impacts include (but are not limited to) loss of identity, damage to reputation, post-traumatic stress disorder, and physical health problems⁵. But damages are not restricted to exonerees and their families. Data has shown that even original crime victims (i.e. the original eye-witnesses) suffer, reporting feelings of guilt, fear, helplessness, devastation and depression¹². Even the American taxpayer suffers. Numerous states across the US have put together reports of the economic burden associated with these convictions. For example, the cost of incarceration, civil settlements, and compensation over approximately a decade cost Illinois and California taxpayers 214 and 220 million dollars, respectively^{13,14}. And finally, there are the unquantifiable damages caused to future victims of the actual perpetrators who remain at large and unpursued.

Given the widespread and diffuse consequences of mistaken identification of other-race persons, it is imperative to better understand the ORE and its psychological and neural correlates. Fortunately, decades of perceptual and memory research have finally begun to impact certain policies in police institutions across the country. This is in part thanks to a 2014 landmark report by the National Academies chronicling the wealth of research surrounding fallibility in eye-witness reports and providing data driven recommendations to improve testimony¹⁵. However, while the article mentions the ORE in several sections, it does not offer recommendations specific to combatting it.

A whole report itself could be written on the fallibility surrounding testimony where race is a factor, as it is far more complex and less well understood than many of the other variables discussed in the National Academies Report. As opposed to perceptual phenomena (viewing angle, brightness of lighting), race is convoluted by psychological phenomena such as stereotyping, sheer familiarity with faces, social categorization, etc. The report acknowledges this much, concluding:

“Additional research is needed to identify procedures that may help estimate the degree of own-race biases in individual eyewitnesses following an identification procedure. Until the scientific basis for these effects is better understood, great care may be warranted when constructing lineups in instances where the race of the suspect differs from that of the eyewitness (97).”

In keeping with this call for more research, the current dissertation comprises of two additional studies on the ORE. The first study (chapter 3) poses a mechanistic account for the ORE and provides a potential metric to measure its severity. The second (chapter 6) employs neuroimaging analysis to examine differences in brain connectivity when same and other-race faces are processed. The remaining chapters serve to introduce the current state of psychological and neuroscientific research on the ORE (chapter 1 and chapter 2), as well as outline the processes informing our neuroimaging analysis approaches (chapter 4 and chapter 5).

Chapter 1: Cognitive Frameworks for the Other-Race Effect

Early Explanatory Theories of the Other-Race Effect

The Other-Race Effect (ORE) is a highly replicated phenomenon in face perception and memory research. It has been reproduced across a variety of paradigms including those testing standard recognition and working memory, as well as perceptual matching ability and gender and race categorization¹⁶. Across mnemonic variations of the task, the effect is often characterized by a deficit in other-race relative to same-race performance within the following four measures: i) decreased hits (reduction in correctly identifying a face as repeated), ii) increased false alarms (increase in falsely recognizing a new face as seen before), iii) decreased discriminability (reduced ability to discriminate between repeated faces and new faces) and a more liberal response criterion (increased willingness to identify a face as repeated). Put less academically, when a perceiver interacts with a member of another race, they are less likely to accurately remember the individual's face later on, and more likely to falsely recognize a new face as belonging to the original identity. These coinciding factors manifest themselves in well-known phrases describing the experience, such as 'They all look the same to me,' and 'All [Insert Group Here] Look Alike'.

While today there is a general belief that no one race looks more alike than the other, this was not always the case. In the early days of research into this effect (circa 1969), the two dominant theories of its cause were divided into psychophysical and psychosocial camps.

Psychophysical explanations proposed that the reflectance of dark-skin and higher morphological similarity between non-white faces led to the observed reduction (in white individuals) of other-race recognition¹⁷. These theories were likely influenced by opinions in line with Social Darwinism which still to this day promote scientifically racist studies¹⁸. Even an early study finding evidence against the psychophysical account references

Darwin:

"From a Darwinian point of view, it could be argued that feature variability should be closely related to the 'usefulness' of the feature as a discriminant stimulus in everyday face recognition. Stated in another way, the degree of variance of facial features should be correlated with their ecological importance as discriminant stimuli for facial recognition in ordinary interpersonal interactions.¹⁹"

Considering the discriminatory relevance of features from evolutionary point of view is valid. Later in this section, variation in featural importance will be discussed. However, it is interesting to consider the historical context of this statement, suggesting researchers who subscribed to the psychophysical hypothesis may have quietly considered non-white individuals less physically individuated and potentially less evolutionarily fit as result.

In contrast to the psychophysical hypothesis, psychosocial explanations attributed decreased other-race recognition to prejudice and lack of familiarity with other-race faces due to relatively segregated and mono-race societies. Large support for psychosocial accounts came from the well-known anecdotal reciprocal nature of the effect (acknowledged as early as the year 1914¹) as well as experimental evidence beginning around the 1970s^{20,21}. Unfortunately, one of the inaugural experiments found that even Black individuals recognized White faces better, perpetuating the psychophysical theory²². The authors admitted that their stimuli were not controlled for physical homogeneity

across each race, which would be a necessary manipulation to test whether stimulus selection confounded their results. Therefore, their findings may have resulted from unintentional selection of Black faces that were less variable than White faces. The ORE is an interesting phenomenon, where its manifestation in experimenters themselves can unconsciously impact the study design, and in turn bias results. I believe the issue of non-random stimulus selection and even lack of diverse stimulus availability likely accounts for discrepancies across many papers published on the ORE over the last half-century.

Ten years from that initial study, the researcher quoted above as referencing Darwin set out to settle the psychophysical/psychosocial debate by analyzing craniofacial data. The proposed motivation for the study was that, although the evidence “weakly” favored the psychosocial theories, “enough residual doubt remains to warrant a study of the facial topography of various racial groups to explore the extent of within-group feature variability.” Data on the variation in facial feature measurements within specific races had already been collected in disparate facial anthropological studies. These were aggregated and compared across race, though not in a statistically rigorous manner. It was concluded that while the average values for facial features differed across the three races, the variance across means were matched across each race. I.e., faces were equally heterogeneous across races. It was therefore not likely that physical homogeneity of non-white faces was a contributing factor to the ORE. It was likely a combination of this study, replications of ORE’s reciprocal nature, and modernizing opinions that led the psychophysical theory to fade into history.

Current Theories Accounting for the Other-Race Effect

As the psychophysical explanation of the ORE fell out of favor, the disparate explanations united under the psychosocial umbrella splintered into new factions. These competing theories evolved into the two dominant frameworks of today: the Perceptual Expertise and Social-Cognitive Hypotheses.

Perceptual Expertise Theories

The perceptual expertise model of the ORE is based in the observation that asymmetric experience with same-race relative to other-race groups manifests in differential abilities to recognize faces corresponding to these groups. A variety of models fall under the perceptual expertise umbrella – all sharing the core assumption that humans are experts in processing faces, but this expertise does not generalize to unfamiliar categories of faces²³.

A popular line of perceptual expertise research is in the developmental origins of the ORE. Findings indicate the development of facial processing has core similarities with language development in that both are quite egalitarian at first²⁴; Just as newborns enter the world with an ability to process all possible units of sound (i.e. phonemes), they also indiscriminately process all categories of faces. However as early as 3 months of age, infants begin to pay more attention to faces within their race²⁵. Other studies indicate the ORE in attention allocation emerges later, between 6 and 9 months of age²⁶. Regardless of the granularity of timing, these observations support a perceptual narrowing account, in which sensitivity to visual information is broadly tuned at birth and then narrows rapidly

within the first year of life²⁷. This tuning is believed to influence the persistence of the other-race recognition deficit into later childhood and adulthood.

Because environmental interaction is core to the perceptual expertise hypothesis, multiple researchers in developmental psychology have tested how critical the early years of life are in establishing the ORE. A major line of interest being, if experience is so integral, perhaps perceptual narrowing may be mitigated or even reversed by exposure to other-race faces. Studies have resoundingly supported this logic. One case demonstrated that briefly habituating Caucasian three-month-olds to Asian faces resulted in their immediate ability to discriminate between familiar and novel Asian faces—i.e. an elimination of the ORE²⁸. However, this study did not implement any follow-ups to determine the mitigation's duration. A more recent experiment set out to test this longevity by exposing 6-month-old Caucasian infants to Asian faces in a picture book for a total of one-hour spread across a 3-month period²⁹. At 9-months of age, as opposed to the age-matched control group these infants displayed a strong novelty preference for new Asian faces, demonstrating they could discriminate individual exemplars from this race. Therefore, the ORE in infancy may be reduced with relatively simple but intermittent intervention.

Results of such experimental intervention are bolstered by evidence that natural exposure to other-race faces improves recognition for those faces. For instance, in one study Korean children adopted into Caucasian European families between the age of 3 – 7 demonstrated a reversal of the ORE³⁰: In a match to sample working memory paradigm, they remembered Caucasian faces better than Asian faces. This was not observed in a control group of non-adopted Koreans. Given these individuals presumably preferred Asian faces prior to

adoption, these results suggest that the ORE can be tuned across child development, and even reversed/flipped by complete overhauls in environmental experience. More subtle evidence of facial recognition's malleability comes from one study where children's deviations in same-race/other-race recognition was associated with how racially integrated their home-neighborhood was. The more diverse the child's community, the less extreme their ORE was.

There is additional evidence that this experience-dependent ORE extends into adulthood. A meta-analytic study of 39 of the earliest experiment found that exposure accounted for a small but reliable amount of variation in the ORE¹⁶. More recent work has shown that length of stay in a non-native country predicted the magnitude of the ORE³¹. Furthermore, training studies have established that the ORE in adults, just like infants, is reduceable. Following several days of individuation training involving learning of face-name (or face-letter) associations, participants demonstrated reduced OREs in recognition memory of novel exemplars relative to baseline performance³²⁻³⁴. Meanwhile, a control condition found that mere exposure and categorization of faces was insufficient to reduce the ORE. This supports the opinion that in adulthood, the combination of face with identity information may be especially important in maintaining robust face representations in memory. One limitation of this work is the short timescale at which recognition is evaluated-- generally one day after successive days of training. To my knowledge the longevity of such trainings remains to be tested. A recent comprehensive study sheds doubt on the promise of such paradigms; Researchers found that self-reported contact with other-race people over participants' lifetimes predicted the magnitude of the ORE, where

extensive contact with diverse communities prior to the age of 12 was the only factor associated with a reduction in the ORE. Furthermore, adults who had lived in new countries for several years with many other-race friends were not found to have reductions relative to controls. This serves as evidence that the ORE may be somewhat fixed after primary school, and that interventions should likely focus on children, rather than adults.

Perceptual Expertise Mechanisms

These studies lend strong evidence supporting the role of experience and development in modulating the magnitude of the ORE. But the cognitive processes underlying expertise in same relative to other-race faces is still up for debate. Dominant explanations tend to agree that asymmetrical experience leads to differential processing mechanisms or differential mental representations that give rise to the ORE in recognition memory

Configural/Holistic Processing

Much of the perceptual expertise literature posits that extensive experience within a category of objects elicits an enhanced ability to extract meaningful information and make fine discriminations between novel instances of those objects³⁵. A foundational study found that bird and car experts' abilities to discriminate differences in their respective objects of expertise were dramatically reduced when those objects were inverted³⁶. In contrast, novice viewers with no expertise in birds or cars displayed relatively limited detriments in response to inversion. Object inversion is a manipulation known to impact configural processing which is the ability to extract spatial-relational information between features in visual input, allowing for more holistic mental representations³⁷. The results of this

research therefore implicated configural processing as a candidate mechanism underlying visual expertise. Within several years the inversion effect was applied to study the ORE; The researchers believed if the ORE was actually an expertise effect for same-race faces, viewers would have a greater deficit in face discrimination upon inversion of same-race but not other-race faces³⁸. The results supported this hypothesis, and helped solidify the perceptual expertise theory of the ORE. Within the framework, regular interaction with same-race individuals leads to expertise with same-race faces, enhancing the capacity to encode such faces in a configural manner. Meanwhile, other-race faces are encoded more similarly to non-face objects with a greater reliance on the individual components of the input, but not the spatial relations amongst these.

The Whole/Parts paradigm is another experimental manipulation that supports the role of configural encoding in the ORE^{39,40}. The paradigm finds that after encoding same-race faces, participants have an advantage in discriminating subsequent changes in features when they are embedded within a whole face, rather than displayed in isolation. Because this observation is not as strong for other-race faces, this is interpreted as further evidence that same-race faces are processed in a more holistic manner.

The relative difference in recruitment of configural processes for same-race relative to other-race faces is generally believed to give rise to the ORE in recognition memory. However, to date only a few studies have tested the association between configural/holistic encoding and the ORE. One such study⁴⁰ reported the the same-race whole-part advantage in Caucasian participants was not correlated with the same-race memory advantage. Meanwhile, Chinese participants demonstrated no differences in holistic processing for

same-race and other-race faces, despite having a robust ORE in memory. However, a later similar whole/part study found the greater the advantage in holistic processing of same-race relative to other-race faces, the greater the advantage in own-race recognition memory⁴¹. The researchers attributed their sensitivity in this study to a regression method employed, instead of the more traditional difference score used previously. A third study that instead used the inversion paradigm found that participant race and a cross-race difference score in configural processing could significantly predict 38% of variance in the ORE⁴². The results of these studies suggest that configural/holistic processing mechanism is likely tied to the ORE, though cannot account for it entirely.

Face-space model

The face-space model^{43,44}, proposes that mental representations of learned faces are embedded within a multidimensional similarity space. The spatial location of faces is dependent on the distinctiveness of the face at encoding, where faces that are quite typical and similar to one another cluster together while distinct faces are situated more distantly. In the context of a recognition experiment, a participant is more likely to accurately recognize a distinctive than typical face, since the former has carved out its own location in face-space, while the latter is positioned quite closely to other faces. The sparser spatial distribution of distinct faces allows for less confusability (or interference) at retrieval.

When first developed, this model attempted to explain not just the distinctiveness memory advantage, but also the own-race advantage in memory, i.e. the ORE. Face-space fits well into the perceptual-expertise framework, because its dimensions are proposed to be selected and scaled to optimize discrimination within the population of faces that an

individual interacts with. Therefore, if an individual has only interacted with members of their own race, their face-space will be modeled to resolve interference only between faces of that race. The ORE is proposed to arise when other-race faces are encoded within a psychological space that was not tuned to the subtle variations in faces from that race. (The model assumes that different races have different dimensions upon which they vary most). This results in other-race faces being encoded more similarly, and are therefore more densely clustered and less discriminable despite their actual physiognomic variability. In this way, many other-race faces may suffer the similarity disadvantage found for only the most prototypical/average of same-race faces.

Several simulation and behavioral experiments support the face-space model. One study trained two separate autoassociative neural networks on Caucasian and Asian faces respectively to test the perceptual expertise model of the ORE⁴⁵. In such a system the neural network model learns faces by developing connection weights that optimize storage capacity of a matrix representing memory for those faces. Put another way, the model will learn features that are important for orthogonalizing face representations where discriminatory features are given more importance. After training the separate models, principal component analysis was used to reconstruct faces and calculate cosine similarity, showing that minority-trained faces were represented more similarly to one another than majority trained faces. Several years later a similar study that generated projections of reconstructed faces into three-dimensional space (corresponding to the first three eigenvectors), confirmed the spatial density assumptions of the face-space model⁴⁶; Face vectors for the majority-trained race were more elaborated in space relative to densely

clustered vectors for the minority-trained race. Furthermore, both studies showed that models performed more poorly on minority trained races in simulated recognition memory and gender discrimination tasks respectively, supporting the hypothesis that representational spatial density is associated with the ORE.

The face-space model is bolstered by findings in human behavioral studies as well. In one experiment participants rated the similarity between hundreds of same and other-race faces⁴⁷. A multidimensional scaling technique was then used to derive spatial configurations of the faces within feature-space, confirming that--within the first dimension--other-race faces were densely clustered together and regionally distinct from same-race faces. In addition, similarity ratings were found to account for 80% of variance in performance in a subsequent individuation task. Interestingly the model seemed to rely more heavily on several race-agnostic dimensions of face-space-- evidence that individuation task-demands may shift reliance away from categorical race information. A second study taking a different approach used the reaction time from speeded pairwise face discriminations as an index of perceived similarity between faces⁴⁸. The resulting first two dimensions of face-space reproduced greater inter-face distances between same-race faces, as well as a category boundary between races. Of note, there was no memory task employed so this study cannot comment on the ORE directly. Regardless, the similar findings across such different task paradigms suggest that differential distribution of faces in feature space may contribute behavioral differences in same-race and other-race face recognition.

Social-Cognitive Theories

While perceptual expertise theories can account for many observations within the ORE, there are certain limitations. For instance, levels of interracial contact and holistic processing do not always map onto the measured ORE^{40,49}. While one large metanalytic study found that race contact did account for the ORE, only 2% of variation in the effect was described. These results imply that a more complete accounting of the ORE likely involves factors beyond those described under the perceptual expertise model.

The most popular alternative accounts for the ORE fall under the social-cognitive umbrella. Social-cognitive theories draw from psychological research on in-group/out-group membership. These models propose that in-group (or same-race) faces are processed in an individuated manner, while out-group (or other-race) faces are processed in a categorical manner. In-group processing relies upon extracting identity-specific characteristics of input, while out-group processing relies on extracting group membership²³. The ORE is therefore not as much about race in itself, but rather the identification with one group as opposed to another, thus influencing cognitive processing strategies.

Evidence of the power of social categorization on recognition memory comes from experiments where manipulation of group-membership can generate disadvantages in memory like the ORE. In one experiment, an in-group memory advantage was found for faces that were labeled with an institution that participants attended⁵⁰. A subsequent study employed a composite face paradigm and found that same-race faces with in-group university affiliation were processed more holistically than same-race faces with out-group

university affiliation⁵¹. A conceptually similar study found that racially ambiguous faces were processed holistically when labeled as same-race, rather than other-race faces⁵².

Thus, holistic processing, previously only connected with perceptual expertise theories was demonstrated to be recruitable by merely altering category membership. This suggests engagement of configural mechanisms do not necessarily require years of perceptual tuning, and that the holistic processing mechanisms found in encoding same-race faces could reasonably be attributed to labeling them as in-group.

Moreover, several studies have found that orthogonalizing relevant group membership away from race is enough to attenuate the ORE. For instance, participants who studied black and white faces grouped by university affiliation rather than race, not only had an advantage in memory for in-group university faces but displayed no differences in memory for same and other-race faces⁵³. Additionally, memory for in-group other-race faces was greater than out-group same-race faces-- evidence that group membership overcame and even flipped the ORE. A similar study assigning high-power and low-power occupations to faces found that other-race faces with high-status jobs were as memorable as same-race faces⁵⁴. Meanwhile in the low-power condition, the standard ORE was found. Collectively these studies suggest that a pure expertise model cannot account for the ORE entirely since lack of expertise may be overcome by shifting attention away from race-categorization.

Social-Cognitive Mechanisms

While social-cognitive research has more recently found that holistic processing is not just under the purview of perceptual expertise models⁵¹, before elaborating on frameworks that accordingly integrate these disparate models, traditional social-cognitive mechanistic accounts will be reviewed.

Feature-selection model

The feature selection model posits that the out-group membership of other-race individuals leads their faces to be processed with more emphasis on category-relevant information^{55,56}. This leads to an asymmetry in feature-selection when encoding faces, where race is treated as a visual feature exclusively in other-race faces. Furthermore, the focus on race-category comes at the expense of coding individuating details. Therefore, the ORE is caused by a selection of features important for classification, rather than subsequent recognition. Most of the experimental evidence for this theory comes from visual search experiments finding that participants can locate an other-race face in a constellation of same-race faces faster than they can perform the reverse scenario. An overemphasis on race information in out-group faces is believed to facilitate this speed advantage.

This other-race categorization advantage has been associated with the ORE in recognition memory. In one study, subjects who demonstrated a greater ORE also had a larger speed advantage for detecting other-race faces in visual search⁵⁵. Conceptually similar studies have described a speed advantage in merely labeling the race of an out-group face. One study identified a mirror image effect in the speed of categorization versus recognition:

Same-race faces are recognized faster than other-race faces, while other-race faces are categorized faster than same-race ones⁵⁷. There is an inverse correlation between same-race recognition speed, and same-race categorization speed. However, actual recognition memory was not correlated with the speed of categorization, and the authors concluded that the same-race recognition advantage in memory may not be clearly related to the other-race categorization advantage.

Cognitive Disregard

A separate social-cognitive account suggests that individuals deem individuating information of out-group faces as task-irrelevant and may not allocate cognitive resources beyond categorical information extraction. There is evidence that attention (as a proxy for allocation of resources) to faces is modulated by race membership. In one study, participants made fewer (but also longer) eye fixations, with greater pupil dilation for other-race faces⁵⁸. They also focused on different features depending on race of face. Furthermore, as they were given more time to encode faces, participants reduced their relative time spent on other-race faces. However, another study found that when given the chance to self-pace their study, subjects spent as much time encoding other-race as same-race faces⁵⁹. Despite this, subjects still demonstrated robust OREs in memory. One caveat of this latter study was that it did not employ eye-tracking so it is still very possible that differences in eye fixations and saccades could have contributed to the ORE, despite no overall differences in processing time. Taken together, evidence for cognitive disregard in the context of the ORE is sparse, and mixed. Further, the model cannot rule out that perceptual expertise, rather than social categorization, contributes to cognitive disregard.

Integrated Theories

While both perceptual expertise and social cognitive theories of the ORE account for certain observations in the experimental literature, neither alone can sufficiently characterize the effect. For instance, perceptual expertise models cannot easily refute evidence that social context alone can eliminate the ORE. Likewise, social-cognitive models struggle to account for the very early developmental emergence of the ORE, well before social constructs of category membership are believed to emerge. Because of the relative strengths of both models there have been several integrated theories that attempt to reconcile them.

The Dual-Process model (drawing from dual process theory in memory research) suggests that the higher social-value placed on same-race faces results in more effortful encoding and subsequent recollection of episodic information from same-race faces⁶⁰. In contrast, other-race faces are more shallowly encoded, leading to a reliance on feelings of familiarity. The standard behavioral measures of the ORE map onto well-established predictions of dual-process memory models: Same-race recognition is characterized by increased target recognition and decreased false alarms, consistent with recollection-based strategies. Meanwhile, other-race face recognition is characterized by prevalence of false alarms, consistent with familiarity strategies. Since recognition and familiarity are considered differentially taxing, the dual-process approach is in line with social-cognitive theories of differential allocation of cognitive/attentional resources. Meanwhile, perceptual expertise is nodded to with the proposed ease of extracting information from same-race faces, potentially improving the fidelity of encoded representations.

A strength of the next two integrative models is the incorporation of findings that in-group identification can engage holistic processing strategies^{51,52}. The In-group/Out-group model proposes that by default, individuals encode faces in a configural manner⁶¹. However, if during presentation of a face, out-group membership is rapidly established, holistic processing may be bypassed in favor of categorical processing. This cue to out-group categorization could be physical, like race. However it could also be dictated by task demands, such as university affiliation in the case of the aforementioned studies^{50,51,53}. In line with perceptual expertise models, out-group faces will be processed more featurally, resulting in a subsequent memory disadvantage.

The most recent proposal, the Categorization-Individuation model is similar to the in-group/out-group model in the importance placed on holistic processing during encoding of in-group faces⁶². The model outlines three factors that can influence recruitment of these processes in face recognition: At first, there is a tendency –especially for outgroup faces-- to attend to category specifying information, leading to perceptual homogenization of faces within a category. Because context can dictate category activation, same-race faces may also be processed categorically^{50,53}. Secondly, motivation (i.e. task demands or power dynamics⁵⁴) can redirect attention towards extracting individuating (holistic) information, promoting subsequent accurate identity retrieval. Third, expertise within a category will promote extraction of identity diagnostic information. Therefore, several factors interact to direct elective attention towards configural information, ultimately dictating subsequent memory accuracy.

...

To summarize, there are a wide range of theories and mechanisms that have been proposed to account for the ORE. Perceptual expertise theories can greatly describe the early emergence of the ORE in development. Social-cognitive theories can explain general biases related to in-group/out-group processing and have more power to explain the malleability of the ORE. Both frameworks have their strengths, resulting in the integration of each into more all-encompassing hybrid models.

Chapter 2: Neural Correlates of the Other-Race Effect

Neuroscientific investigations of the Other-Race Effect (ORE) are sparse considering the size of the social-psychological corpus. They also have been less focused on settling the debate between the multiple theoretical accounts for the ORE. Instead, they are more focused on identifying anatomical regions that are differentially engaged for same and other-race faces, and their correspondence to behavioral metrics of the ORE. Despite the relative agnosticism on the theoretical accounts, these studies will be discussed and organized within the context of the prevailing cognitive frameworks, where possible.

Also note that this review is focused on functional magnetic resonance imaging (fMRI) as opposed to electroencephalography (EEG) investigations, due to the employment of fMRI methods within my own research. There is an extensive body of research on the temporal-neural components of face processing, and a smaller subset of studies on the deviations between same-race/other-race face processing within event-related potentials that is beyond the scope of this dissertation.

Perceptual Expertise and Holistic Processing

Early neuroscientific research on the ORE came at the heels of a debate on the evolutionary purpose of a highly face-preferential patch in the fusiform gyrus of the brain. A seminal study in 1997 coined this region the 'Fusiform Face Area' (FFA) due to its greater activation (i.e. oxygen consumption) when participants viewed pictures of faces relative to houses and hands⁶³. After this finding, a competing hypothesis emerged influenced by the

observation that objects of expertise (i.e. birds and cars³⁶), just like faces, were subject to inversion effects in experts (chapter 1). With the argument that faces were akin to objects with which people had extensive expertise, researchers demonstrated that the FFA was more engaged when experts as opposed to novices viewed such objects^{64,65}. This triggered an ongoing debate whether the fusiform gyrus is specialized for face processing specifically, or rather serves a more general-purpose role of extracting information crucial to discrimination between instances of objects with which we have great experience⁶⁶. This debate is beyond the scope of this dissertation, but from here-on I will refer to this face-selective patch as the FFA, not because of adherence to the specialized face-processing theory, but due to its continued conventional use.

Given the predominance of perceptual expertise theories in the ORE literature, it was not long before the FFA's recruitment during processing of same-race and other-race faces was tested⁶⁷. The researchers hypothesized that if the FFA was more engaged for objects of expertise, it would be more engaged during encoding of same-race relative to other-race faces. Participants were scanned while they passively viewed faces and were administered a surprise memory test afterwards. As predicted, the results confirmed differential recruitment of the right FFA. They further found that the magnitude of the left FFA signal was correlated with an advantage in memory for same-race relative to other-race faces. This served as strong evidence of the FFA's role in the neurological basis of the ORE, however the differential observations in left and right FFA complicated interpretations. It should also be noted that traditional memory regions (hippocampus and parahippocampal gyrus) were also correlated with the ORE, with the strongest association in the

parahippocampal gyrus. Despite this finding, there has been relatively little attention paid to the role of medial-temporal lobe regions in the ORE.

The discovery of the FFA further prompted hypotheses of its role in holistic processing, given its demonstrated importance in face and expertise tasks. Several studies have employed holistic processing paradigms testing FFA engagement during tasks believed to engage holistic as opposed to parts-based or featural processing. In one such study, participants performed a working memory task where they were shown whole faces or scrambled faces with features re-arranged into non-natural locations to disrupt holistic processing⁶⁸. For whole face presentations, accuracy in performance was associated with similarity of multivariate patterns of activity in the FFA between initial and test faces. Meanwhile, earlier visual and face processing areas (such as the Occipital Face Area (OFA)) demonstrated more similar patterns associated with accuracy independent of whether faces were whole or scrambled. These results are interpreted as an emergent behaviorally relevant role for the FFA specific to processing whole faces – i.e. holistic processing.

Several studies specifically investigated the role of holistic processing in the FFA as a neural correlate of the ORE. One experiment scanned subjects while they performed a simple 1-back recognition task during viewing of same and other-race faces⁶⁹. Later they were administered a surprise memory test. The researchers then calculated same-race advantages in memory as well as neural signal activation by computing difference scores (other-race metrics subtracted from same-race.) The behavioral and neural same-race advantages for whole faces were strongly correlated in the FFA but not the lower-level feature processing OFA. This suggests that differential engagement of the FFA during

viewing of same-race and other-race faces may uniquely explain behavioral variation in subsequent memory. The lack of behavioral relevance in the OFA may also suggest that the ORE emerges based on higher-level holistic processing deficits, as opposed to lower-level facial feature processing. However, it should be noted that experimenters did not find statistically significant overall differences in same-race and other-race recognition memory in this study.

Another study finding neural differences in response to same-race and other-race faces was designed to exploit repetition suppression, a reduction in magnitude of neural signal when stimuli are repeated⁷⁰. The researchers scanned subjects during performance on the composite face paradigm, another classic technique that illuminates same-race relative to other-race reliance on holistic processing⁷¹. This paradigm often finds that participants cannot easily discriminate between two same-race faces that have duplicate top halves, unless the bottom halves are offset from center, disrupting holistic processing. Participants do not have the same discrimination difficulty between other-race composite faces that are aligned, suggesting that we rely more on additive featural than holistic processes for other-race discrimination. In this study, the FFA demonstrated a release of suppression for aligned faces specific to the same-race condition. This suggests the FFA was more sensitive to identity changes of same-race faces. Meanwhile the OFA was sensitive to identity change in other-race but not same-race faces. Together this may suggest that the OFA's role in feature processing may be especially important for identity recognition of other-race faces, while reliance on holistic processing in the FFA is more important for same-race recognition.

While most studies investigating neural correlates of holistic processing have been focused on face preferential regions like the FFA and OFA, one analysis expanded its scope to other regions⁷². The researchers employed a composite face task, defining additional regions known to be involved in processing of scenes and objects, as well as perceptual grouping. In contrasting activity during aligned and misaligned conditions, the authors found significant differences in the retrosplenial cortex and parahippocampal place area, but not the FFA.

While the above⁷² study had no race manipulation, it importantly demonstrates that it is possible for more than one region to be involved in holistic processing, and that regions that are not highly and exclusively active to faces may still play a role in face cognition. For instance, note the observation of holistic processing in the parahippocampal place area. Recall that the first-ever fmri investigation of the ORE found that the parahippocampal gyrus was even more strongly correlated with the behavioral ORE than the FFA⁶⁷. Furthermore, there is a large body of object-recognition research focused on identifying brain areas that integrate feature information into holistic representations. Medial-temporal lobe (MTL) regions are often implicated, including the parahippocampal, perirhinal and entorhinal cortices as well as the hippocampus^{73,74}. Given the role of MTL in memory encoding and feature integration that is often overlooked in the context of face recognition, a potential role for the MTL's involvement in the ORE will be outlined in the third chapter of this dissertation.

Attentional and Social Modulation

Despite evidence that holistic processing can be dictated by task demands and that attentional allocation may differ between same and other-race faces, there has been relatively little focus on brain regions involved in attention. As evidenced in the prior section, fMRI research has greatly focused on studying visual or face-preferential regions within the context of the ORE. However, one novel study tested the relation between attentional and cognitive control regions, and the ORE in memory⁷⁵. While in the scanner participants were tasked with creating elaborate stories for faces during encoding to promote successful and more ecologically realistic recognition than typically studied in the lab setting. The next day subjects performed a recognition test, so that encoding activity could be sorted based upon subsequent accurate or inaccurate performance. The analysis found differential engagement of attentional and control regions for same and other-race faces. Specifically, reduced activity in the cognitive control network was associated with failure to recognize other-race but not same-race faces. Furthermore, functional connectivity between the FFA and Dorsal Attention networks was more predictive of success for same-race recognition. These results demonstrate not only the importance of attentional and executive control regions during encoding, but also a disparity between how same-race and other-race faces are successfully processed. Less engagement of these regions could be consistent with the shallower encoding processes proposed in social-cognitive and integrated accounts. Meanwhile, greater top-down attention and cognitive control could give rise to more elaborated representations of same-race faces subserving subsequent mnemonic discrimination, as predicted by the face-space model (chapter 1).

A quite different study demonstrates the interaction between implicit social attitudes and the neural representation of faces⁷⁶. Because the FFA is associated with holistic processing and identity encoding, the authors wondered whether patterns of activity in the FFA represent race, or if solely lower-level feature processing regions reflected race-categorical information. They were also curious whether implicit attitudes towards Black people would influence how white participants represented Black faces. They reasoned that social bias might change the way faces are encoded, such that people with higher bias might have larger differences in representation of white and Black faces. The authors used a logistic regression classification algorithm to test how discriminable neural face representations were across a series of early visual processing regions as well as a variety of brain regions implicated in face processing (including amygdala, posterior superior temporal sulcus, inferior frontal gyrus, orbital-frontal cortex, etc). They found that the race of the faces was only predictable based on patterns in the visual occipital pole, the OFA, and FFA. However, in the FFA, increasing accuracy of the algorithm was correlated with increasing bias in the subjects. When subjects were split at the median and the logistic classifier was retrained, neural patterns in the FFA were only separable in the high-bias group, while race was still discriminable in the occipital pole and OFA in both groups. These results suggest that preservation of race-diagnostic information in the later visual processing stream may be socially modulated. Within the context of the face-space model (see perceptual expertise mechanisms), it is interesting to consider that along a dimension of race, faces may have clustered into separate neighborhoods and were separable in the occipital pole, OFA, and FFA until implicit bias was considered. Once subjects were separated by bias, the face-space of the FFA in the low-bias group no longer had distinct enough neighborhoods of

clusters to be discriminable by race. While this finding is incredibly interesting, it is unclear whether the separability of race representations impacts memory for faces, because the ORE in recognition was not tested. Furthermore, despite the reported accuracies in the logistic regression model being statistically above chance, they are quite low (with accuracy levels around 60%) and results would be strengthened by selecting a stronger classification algorithm.

...

To summarize, the ORE has not been extensively studied with fMRI. When it has been investigated, there is often only interest in testing the role of a select few preferential face processing regions. Furthermore, studies have not always been able to reproduce an ORE in behavior, or they have been more focused on underscoring differences in neural activity independent of demonstrating behavioral relevance. Importantly, only a few studies have identified the involvement of regions beyond the putative face processing network in holistic and same-race/other-race face processing. There is therefore a great need for exploring the ORE from perspectives not exclusively focused on face preferential and visual processing regions.

Chapter 3: A Memory Model of the ORE: Differential Mnemonic Discrimination of Same and Other-Race Faces

A large portion of the content in this chapter was originally published under the title 'A Memory Computational Basis for the Other-Race Effect' in 2018 in Scientific Reports. The full citation and list of authors is included in the Vita and References ⁷⁷.

Introduction

It may be apparent from the preceding chapters that the theoretical frameworks of the Other-Race Effect (ORE) are highly focused on perceptual and social mechanisms, rather than memory mechanisms. Despite this, the ORE is considered a memory deficit and is indexed by performance on recognition paradigms. Only three of the studies discussed in chapter 2 employed memory tasks, and this is representative of neuroimaging research as a whole^{67,69,75}. Therefore, it is unclear how many of the findings on differential neural representation and processing of same and other-race faces relate to the ORE in memory.

Since the ORE is defined as a recognition deficit, attempts to understand it can benefit from considering the wealth of research in the field of learning and memory. The dual-process framework proposed sixteen years ago (chapter 1), is one of the only attempts at characterizing and reconciling the ORE with the findings in memory literature, proposing differential recruitment of recognition and familiarity strategies for same and other-race faces.⁶⁰ This framework remains obscure to my knowledge; I only discovered it when conducting the extensive literature review for this dissertation. What follows in this chapter is our group's similarly motivated effort to propose and test a memory-informed model of the ORE. We suggest that the potential contributions of not only perceptual and attentional, but also mnemonic (memory-based) processing should be assessed for its

impact on the ORE. We address this gap using models of mnemonic interference reduction that are becoming increasingly popular in memory research due to backing by strong neurobiological evidence⁷⁸.

Our approach to studying the ORE is fundamentally informed by computational models of hippocampal contributions to episodic memory. The hippocampus as well as the surrounding medial-temporal neocortical regions of the brain play a well-established role in the formation of episodic memories⁷⁹. Computational and rodent work suggests that the hippocampus --and more recently the perirhinal cortex-- are involved in pattern separation, a neurocomputational process that allows for detailed encoding of similar experiences by reducing overlapping mnemonic 'interference' across similar inputs^{78,80-91}. Functional MRI studies have also shown distinct patterns of activity in the hippocampus perirhinal, parahippocampal and entorhinal cortices during memory encoding, consistent with pattern separation⁹²⁻⁹⁹.

Behaviorally, pattern separation is thought to underlie the ability to discriminate among similar experiences, or more simply put, to assist in the individual recall of similar items.¹⁰⁰ For example, remembering where you parked your car today versus yesterday requires pattern separation; these two experiences are largely similar and need to be stored independently of each other. Mnemonic discrimination tasks have been used frequently to assess this capacity to remember similar experiences, by testing subjects' memory for various common objects that have been independently rated for relative similarity to one another⁷⁸. Like highly similar objects, faces share a general configuration of features, with no one component ideal for consistent successful differentiation. To efficiently remember

faces despite this baked-in ambiguity, a facial processing system must have mechanisms in place to resolve high interference between distinct experiences (i.e. pattern separation).

A critical facet of the pattern separation computation indexed by mnemonic discrimination, is that it operates as an input-output transfer function that is nonlinear. The input to the system is similar sensory experiences (e.g. similar faces), and the output is the response of the system (e.g. whether the faces are stored as distinct from one another or as instances of the same face). Ideally, an efficient memory system should be able to discriminate among faces that are similar but belong to different individuals (i.e. pattern separate) but also be tolerant of variability in inputs of the same face across encounters despite minor context-dependent differences (i.e. pattern complete). These two conditions demonstrate the need in facial recognition for a nonlinear input-output transfer function that allows for distinct enough stimuli to be separable from one another but is also stable (robust to change) when stimuli belong to the same identity. Several studies have used input-output transfer functions to characterize visual and mnemonic discrimination for both object and facial recognition. Neuroscience literature suggests that rodents' discrimination behavior in response to manipulated environmental contexts is best described by a sigmoidal transformation.⁸⁰ Facial recognition has also been described as sigmoidal in both behavioral and neural computational work^{101,102}. Other research has characterized object mnemonic discrimination using more curvilinear input-output transfer functions^{103,104}.

Traditional face-recognition tasks used to assess the ORE do not manipulate mnemonic interference—or similarity—and therefore cannot produce input-output transfer functions. However, mnemonic discrimination tasks parametrically vary the similarity of

lure stimuli allowing a thorough characterization of the transformation between stimulus similarity (experience) and neurobehavioral responses (representation). Thus, mnemonic discrimination tasks are an ideal tool for characterizing facial recognition amidst mnemonic ambiguity in facial processing. The paradigm further accommodates visualization and comparison of input-output transfer functions for different experimental groups or stimulus types. For instance, if mnemonic discrimination of same-race and other-race faces is not comparable, we would expect the emergence of diverging transfer function trends. This would allow us to pinpoint where recognition fails for other-race relative to same-race faces along the spectrum of mnemonic overlap, and to infer MTL computational differences in processing faces across race.

In addition to developing a mnemonic discrimination face task, we created a match-to-sample task, where subjects held one face in memory briefly before being prompted to make same/different discriminations on repeated or lure faces with the same manipulated parametric interference presented in the mnemonic discrimination task. This allowed us to compare performance on same-race and other-race discriminations as a function of similarity between face-pairs when subjects were required to internally represent and maintain only one face in memory at a time for several seconds. This paradigm therefore reduces the proactive interference found in the mnemonic discrimination task that occurs naturally with generation and storage of increasing information. Furthermore, this task allowed us to establish the extent to which deficits in resolving interference between other-race faces might arise in perception or attention, without placing strong demands on episodic memory mechanisms such as pattern separation.

In the current study, we hypothesized that mnemonic discrimination is altered in facial recognition of one's own relative to another race, and therefore is characterized by distinct input-output transfer functions for same-race and other-race faces. We predicted the same-race input-output transfer function would be significantly higher than the other-race function at high mnemonic interference levels. In other words, when faces are highly similar, subjects should perform significantly better on same-race relative to other-race mnemonic discriminations. However, when interference is low enough, or faces are more distinct from one another, same-race and other-race discriminations should be comparable in accuracy. Meanwhile in a match-to-sample task we expected subjects to demonstrate relatively little or no differences in input-output transfer functions for same-race and other-race faces. That is, the two transfer functions should be relatively similar if we believe that the ORE is dependent on the compounding effects of mnemonic in addition to perceptual and attentional processes. Meeting these predictions would suggest that altered efficiency of computational pattern separation processes for same-race relative to other-race faces may promote the emergence of the ORE.

Methods

Participants

This study protocol was approved by the Institutional Review Board (IRB) at the University of California, Irvine, and complies with IRB guidelines and regulations. Participants provided informed consent in accordance with the board and received course credit or monetary compensation. Ninety-nine healthy volunteers (77 Female; 22 Male; mean age of

20.62, SD 2.83) were recruited from the University of California, Irvine community. All participants were between 18 and 36 years of age and were screened for major neurological and psychiatric conditions (exclusionary criteria). These subjects performed a mnemonic discrimination facial recognition task. Seven participants were excluded for missing over 10% of test trials. An additional three participants were excluded due to very poor performance on the task (two or more standard deviations below the group mean), suggesting a lack of engagement or misunderstanding of the instructions. Performance was measured using the sensitivity index, d' , calculated as $z(\text{target hit rate}) - z(\text{lure false alarm rate})$. These exclusions resulted in a final sample of 89 subjects (68 Female, 21 Male; mean age 20.63, SD 2.92). These subjects were divided into three groups according to their self-identified race: (75 Asian, 12 Caucasian and 2 Black). Due to the small sample size of non-Asian participants, the analysis presented focuses on Asian subjects' performance for a total sample of 75 subjects between 18 and 36 years of age. (57 Female, 18 Male; mean age of 20.47, SD 2.59). Though more females volunteered for this study, there were no gender differences in performance.

To control for perceptual and attentional contributions to the ORE, we also conducted a match-to-sample task in an independent sample that included 34 subjects between the ages of 18 and 31 (26 Female, 8 Male; mean age of 21.62, SD 2.53) who were subjected to the same screening procedures as the experimental subjects above. Again, all subsequent analysis is restricted to the Asian participant data for a total sample of 24 subjects between the ages of 18 to 25 (19 Female, 5 Male, mean age of 21.04, SD 1.85).

Stimuli

A database of face stimuli was created using FaceGen Modeller 3.5. A set of 272 faces were generated, evenly distributed across gender and two races: Asian and Black. (Caucasian faces were also generated but were not included in the version of the task administered to Asian subjects.) For each race, 88 faces were created using the FaceGen Generate function. Of the 88 faces, 48 were randomly selected as 'parent faces' to serve as templates for 48 face lures. Lures were created by running the Genetic Randomness algorithm on parent faces, to apply normally distributed perturbations with means proportional to an inputted value. Equal numbers of lure stimuli were generated to create four lure bins at 20%, 30%, 40%, and 50% perturbations from parent faces. A 20% perturbation results in a face lure that is highly similar, or nearly identical to the parent face, whereas a 50% perturbation generates a more dissimilar- looking face (Fig. 3.1).

Behavioral Tasks

The following procedures were designed to present participants with both faces of their own race and another race. In the remainder of this chapter, Asian faces are referred to as same-race and black faces as other-race.

Mnemonic Discrimination Task

All experiments were programmed in PsychoPy v1.85.2. Participants performed a blocked task, where each of 8 blocks included an encoding, followed by a test phase (Fig. 3.1a). In the encoding phase, subjects were asked to explicitly memorize each of 22 presented faces. Faces were randomized, presented consecutively and evenly divided amongst same-race

and other-race categories. In a following test phase, participants viewed a second series of faces, half of which were identical to the memorized faces. The remaining faces were lures spread across all four bins. Participants were asked to identify which faces were 'same'/'old' (Target Repeats), and which were 'different'/'new' (Lure Distractors), using corresponding button presses on the keyboard. A response of 'same' to a target repeat indicated successful recognition, while 'different' for a lure distractor indicated successful mnemonic discrimination, or a correct rejection. In both encoding and test phases, stimuli were presented for 3.0 seconds with a 1.5s ITI. After completion of 4 blocks, subjects were given a short break.

Match-to-Sample Task

A separate group of participants performed a match-to-sample task, which required minimal long-term memory retention. In each of 8 blocks subjects were shown one face, followed by a 2.5 second dynamic mask, and a second face that received subject input (Fig. 3.1b). The test faces were divided evenly into Target Repeat and Lure Distractor trials. Subjects were asked to make the same 'old'/'new' judgments described above. The exact same stimulus dataset, trials per block, trial durations and ITI were used for the match-to-sample as the mnemonic discrimination task.

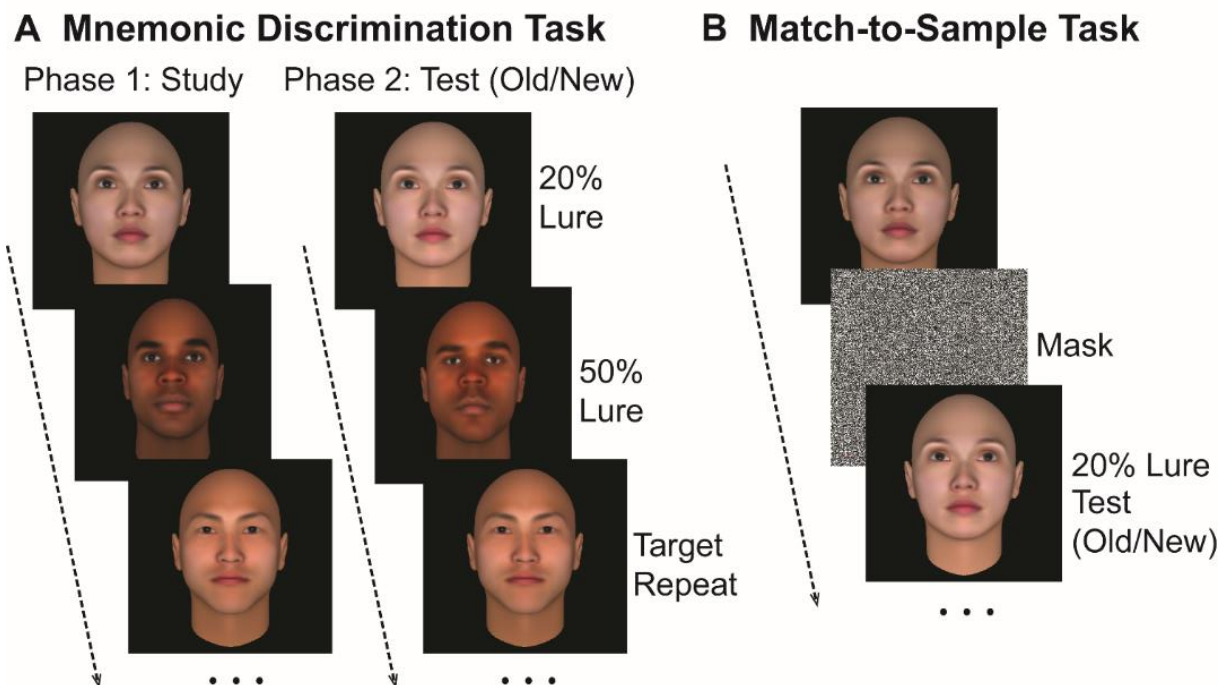


Figure 3.1 Mnemonic Discrimination and Match-to-Sample task designs.

(A) The mnemonic discrimination task comprised of an initial encoding phase followed by a test phase. Stimuli at test were either exact Target Repeats or similar Lure Distractors deviated from faces at encoding by 20%, 30%, 40%, or 50% perturbations. Subjects indicated whether test faces were the same or different from faces presented in the encoding phase. (B) The match-to-sample task comprised of only one phase where subjects saw a face followed by a mask for 2.5 s, followed by either a Target Repeat or Lure Distractor. They indicated whether the second face was the same or different from the face prior to the mask. On both mnemonic discrimination and match-to-sample tasks, stimulus duration = 3.0 s and ITI = 1.5 s.

Results

For both mnemonic discrimination and match-to-sample tasks, performance was calculated using subjects' proportions of target hits ('Same'| Target Repeat) and false alarms ('Same'| Lure Distractor). The sensitivity index (d') was calculated as $z(\text{target hit rate}) - z(\text{lure false alarm rate})$ to evaluate the ability to discriminate between old repeated faces and new distractor faces. First, we confirmed the canonical measures of the ORE: a reduced d' and increased proportions of false alarms for other-race faces. Mnemonic

discrimination accuracy (d') for same-race faces was significantly greater than for other-race faces [$t(74) = 4.755, p < 0.0001, r^2 = .234$] (Fig. 3.2a). In addition, subjects false alarmed more to other-race than same-race faces [$t(74) = 4.166, p < 0.0001, r^2 = .19$]. A 2 X 2 repeated measures ANOVA revealed significant main effects of stimulus race [$F(1,74) = 22.26, p < 0.0001, \eta^2 = .05$] and mnemonic interference [$F(3,222) = 51.38, p < 0.0001, \eta^2 = .14$] as well as an interaction [$F(3,222) = 9.868, p < 0.0001, \eta^2 = .03$] (Fig. 3.2b). Post hoc Sidak multiple comparison tests revealed that same-race performance was better than other-race performance for the first three interference levels [20% $p = .0011$, 30% $p < 0.0001$, 40% $p = .0001$] (Fig. 3.2b). The same analyses were run on the match-to-sample version of the task demonstrating no effect of stimulus race on performance [$t(23) = 0.8563, p = .4007, r^2 = .03$; ($F(1,23) = .7332, p = .4007, \eta^2 = .00$)] and a main effect of interference, as expected. [$F(3,69) = 70.08, p < 0.0001, \eta^2 = .4$] (Fig. 3.2c-d).

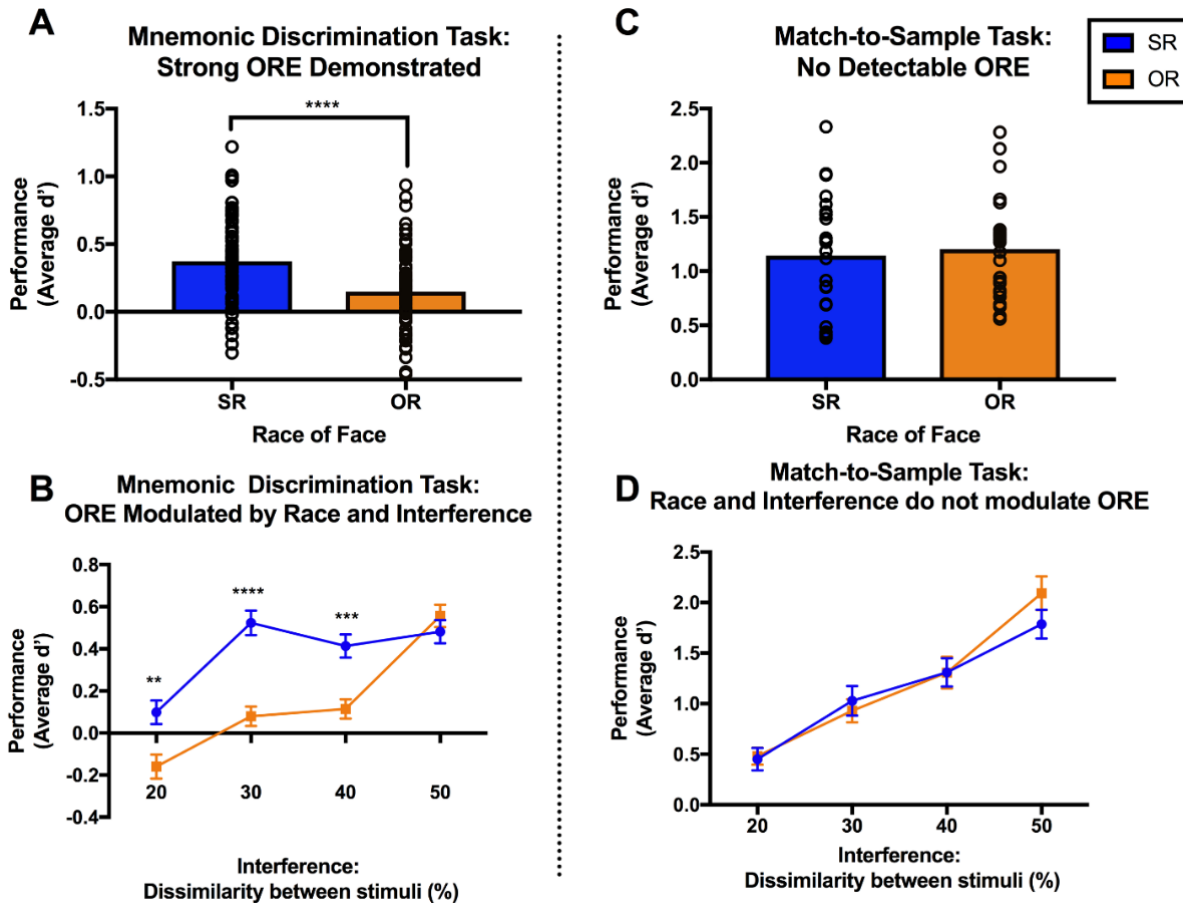


Figure 3.2 Discriminability Results from Mnemonic Discrimination and Match-to-Sample Tasks.

The ORE is present in a mnemonic discrimination but not match-to-sample task, suggesting increasing proactive and mnemonic interference may contribute to the effect. (A) In the mnemonic discrimination task, accuracy for same-race faces was significantly greater than for other-race faces ($p < 0.0001$). (B) In the mnemonic discrimination task subjects performed more accurately on same-race faces for all but the highest interference level ([20% $p = 0.0011$, 30% $p < 0.0001$, 40% $p = 0.0001$]). (C) In the match-to-sample task subjects recognize other-race faces as well as same-race faces. (D) In the match-to-sample task subjects perform equally on same-race and other-race faces, regardless of mnemonic interference level.

Due to increased recruitment of female relative to male subjects, we confirmed that the gender skew did not impact results. There were no significant differences between female and male lure discrimination performance for same-race [$t(74) = 1.241$, $p = .22$, $r^2 = .02$] or other-race faces [$t(74) = .3914$, $p = .70$, $r^2 = .00$]. Additionally, a 2x2 ANOVAs revealed no main effect of gender on performance across interference levels for same-race [$F(1,74) = 1.961$, $p = .17$, $\eta^2 = .07$] or other-race discrimination [$F(1,74) = .33$, $p = .57$, $\eta^2 = .01$].

To further investigate the modulation of the ORE by task type, we collapsed d' across interference levels and ran a 2 X 2 ANOVA with task type as the between-subject factor and stimulus race as the within-subject factor. This analysis revealed significant main effects of task-type, $[F(1,97) = 154.88, p < 0.0001, \eta^2 = 0.64]$, stimulus race $[F(1,97) = 4.60, p = 0.0345, \eta^2 = 0.12]$ as well as a significant interaction of the two $[F(1,97) = 7.14, p = 0.0089, \eta^2 = 0.07]$ (Fig. 3.3a). A post hoc Sidak comparison indicates a significant difference in same-race and other-race performance for the mnemonic discrimination ($p < 0.0001$), but not the match-to-sample task ($p = .9437$). Because the group sizes differ between task types, we ran an additional linear analysis that is robust to sample size and variance differences across groups, to confirm these results. A model was fit using generalized estimating equations, where d' was modeled as a linear combination of race, task, and the interaction of race and task. This produced similar results to the analysis of variance (Table 3.1), including significant differences in the estimated d' means for same-race and other-race faces in the mnemonic discrimination task [Table 3.1a: $\beta' = .22, S.E. = .05, 95\% CI = (0.13, 0.31), p < .0001$] but not match-to-sample task [Table 3.1b: $\beta' = -.02, S.E. = .07, 95\% CI = (-0.16, 0.12), p = .73$], Further, there remains a significant interaction between task and race; In the mnemonic discrimination task, the difference in the estimated d' between participant's recognition of same-race and other-race faces was .25 larger than the difference in the estimated d' between same-race and other-race faces in the match-to-sample task [Table 3.1c: $\beta' = .25, S.E. = .09, 95\% CI = (0.08, 0.42), p < .005$].

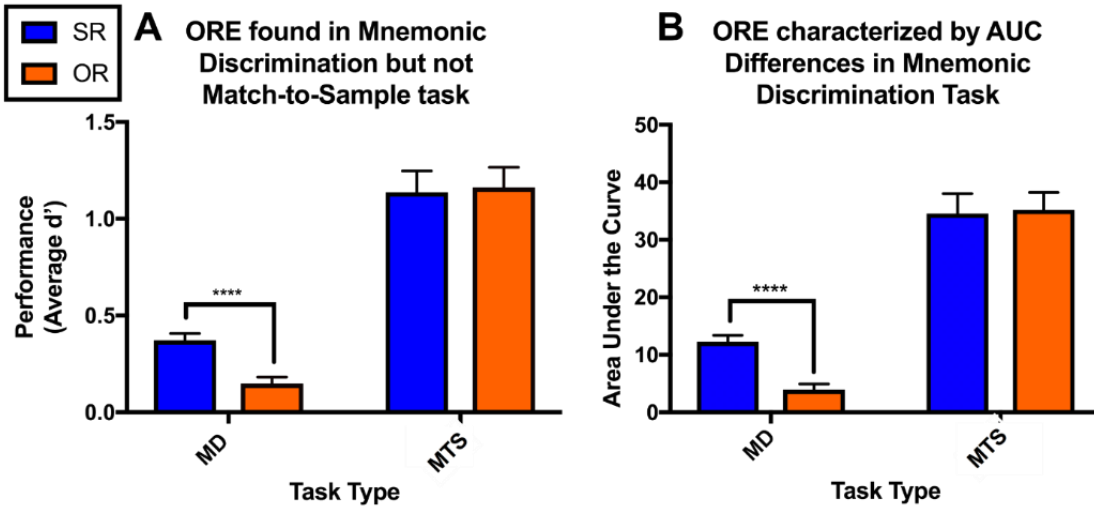


Figure 3.3 Comparison of ORE in Mnemonic Discrimination and Match-to-Sample tasks

Both d' and Area Under the Curve measures find ORE in Mnemonic Discrimination but not Match-to-Sample (MTS) tasks. (A) An Analysis of Variance finds significant main effects of task ($p < 0.0001$), stimulus race, ($p = 0.0345$) and an interaction of the two ($p < 0.0089$) on performance. A post hoc multiple comparisons test finds a significant difference in performance for same-race and other-race faces only during mnemonic discrimination ($p < 0.0001$, labeled on figure A). (B) Comparing area under the curves (AUCs) for same-race and other-race faces in both mnemonic discrimination and match-to-sample tasks reveals a strong ORE only in the mnemonic discrimination task. The larger the significant difference in same-race and other-race AUCs, the greater the ORE. Analysis revealed significant main effects of both task type ($p < 0.0001$) and stimulus race ($p < 0.05$) on performance, as well as a significant interaction of the two ($p < 0.005$). A post hoc test found significant AUC differences for same-race and other-race input-output transfer functions for the mnemonic discrimination but not match-to-sample tasks ($p < 0.0001$, labeled on figure B). Abbreviations: MD – mnemonic discrimination; MTS – match to sample.

Table 3.1 Linear Analysis Results

	Estimate	Standard Error ^{a*}	95% CI	p
a. Mnemonic Discrimination: Same-race vs. Other-race face	0.22	0.05	(0.13, 0.31)	<0.0001
b. Match-to-Sample: Same-race vs. Other-race face	-0.02	0.07	(-0.16, 0.12)	0.7308
c. Interaction between task and race	.25	0.09	(0.08, 0.42)	<0.005

A linear analysis was run to account for differences that may be attributable to unmatched sample sizes. The model was fit using generalized estimating equations, which are robust to sample size differences across groups. d' was modeled as a linear combination of race, task, and the interaction of the two. (A) and (B) The analysis reveals significant differences in estimated population means for same-race and other-race faces in the mnemonic discrimination but not match-to-sample task. (C) There is a significant interaction between task and race with the difference in the estimated d' between participant's recognition of same-race and other-race faces 0.25 larger in the mnemonic discrimination task than the difference in the estimated d' between same-race and other-race faces in the match-to-sample task. ^{a*} Heteroscedasticity-consistent "sandwich" standard errors are used to allow for differences in the variance of model errors across different participant subgroups.

In addition, we tested whether subject-specific input-output transfer functions could be used to calculate a metric of the ORE by calculating the areas under the same-race and other-race curves in both tasks. We did this by using the summed average of d' at each interference level, added to the prior level for both same-race and other-race functions in every subject. The larger the net AUC (area under the curve) value, the more accurate the performance. t-tests comparing AUC for same-race and other-race functions indicated a strong ORE in the mnemonic discrimination [$t(74) = 5.869, p < .0001, r^2 = .31$], but not the match-to-sample task [$t(23) = .3208, p = .75, r^2 = .00$]. In further support, a 2 X 2 repeated measures ANOVA with task type as the between-subject factor and stimulus race as the within-subject factor and the AUC values as the outcome measure revealed significant main effects of task type [$F(1,194) = 211.6, p < .0001, \eta^2 = .50$] and stimulus race, [$F(1,194) = 4.3, p < .05, \eta^2 = .01$] as well as a significant interaction between the two [$F(1,194) = 5.968, p < .0005, \eta^2 = .01$]. A post hoc Sidak comparison indicates a significant difference in same-race and other-race performance for the mnemonic discrimination ($p < .0001$), but not the match-to-sample task ($p = .9718$) (Fig. 3.3b).

The ORE was also apparent in the mnemonic discrimination task when using reaction time (RT) as the outcome measure. In general, subjects required more time to correctly reject other-race than same-race lures. [$t(74) = 2.533, p < .05, r^2 = .08$]. On average same-race faces were correctly rejected after 1.43 seconds, while other-race faces were correctly rejected after 1.48 seconds. Further, RT was associated with better lure discrimination performance for other-race faces but not same-race faces. T-tests show that subjects spent significantly more time on other-race lure correct rejections ($\mu = 1.48$ s) than false alarms

($\mu = 1.40$ s) [$t(74) = 3.435$, $p = .0010$, $r^2 = .14$]. No such RT relationship was found for same-race lure correct rejections and false alarms [$t(74) = 1.677$, $p = .0978$, $r^2 = .04$]. These results were recapitulated using a 2 X 2 repeated measures ANOVA, reporting main effects of race [$F(1,74) = 4.31$, $p = .0414$, $\eta^2 = .0034$], and correctness [$F(1,74) = 9.26$, $p = .0032$, $\eta^2 = .016$] on reaction time means, but no interaction [$F(1,74) = 2.908$, $p = .0923$, $\eta^2 = .0023$] (Fig. 3.4a). Mean reaction times were longer overall for other-race faces [$\mu = 1.44$ s] than same-race faces [$\mu = 1.41$ s], and correct responses were longer on average [$\mu = 1.46$ s] than incorrect ones [$\mu = 1.39$ s]. Post hoc Sidak comparisons revealed that reaction time averages were significantly different between correct rejections and false alarms for other-race [$p = .0002$] but not same-race faces [$p = .2262$].

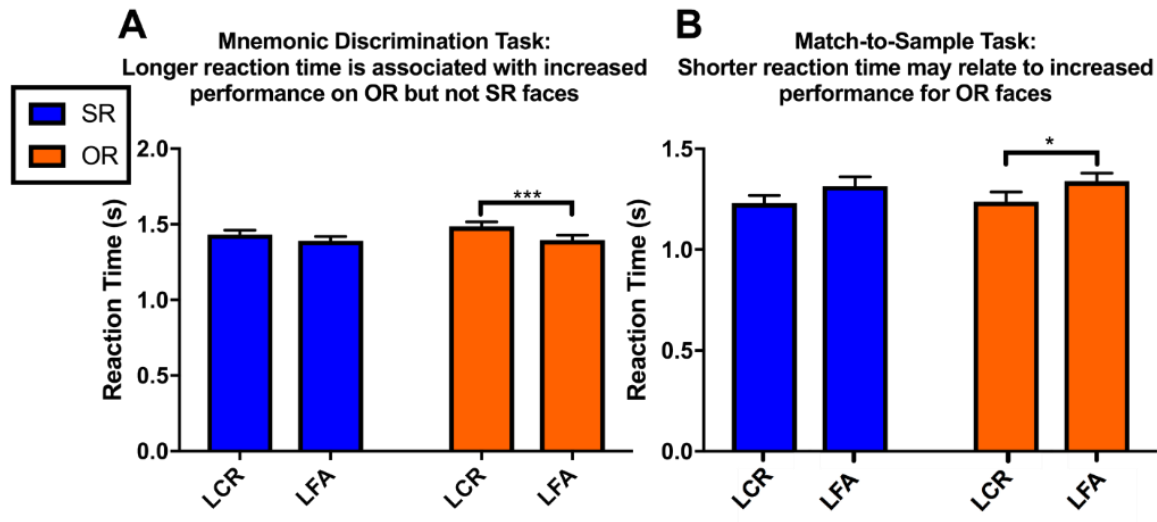


Figure 3.4 Reaction Time Differences

Reaction time differences found for same-race and other-race recognition in Mnemonic Discrimination task and to a lesser extent in the Match-to-Sample task. (A) Race ($p < 0.05$) and correctness ($p < 0.005$) significantly affect reaction time means in the mnemonic discrimination test, however there is no interaction of the two. A post hoc multiple comparisons test finds reaction time differences are associated with accuracy for other-race but not same-race face discriminations, where longer responses are linked to correct discriminations (labeled as ***). (B) An analogous Analysis of Variance of the Match-to-Sample data finds only a main effect of correctness on reaction time ($p < 0.05$) However, a post-hoc test finds reaction times are associated with accuracy for other-race but not same-race faces ($p < 0.05$), where quicker responses are linked to accuracy (labeled as *).

Though the ORE was not detectable in the match-to-sample task using accuracy measures (Fig. 3.3), we tested whether reaction time could detect early differences in same-race and other-race face processing. In a 2 x 2 repeated measures ANOVA (analogous to that run in the mnemonic discrimination task) there was a main effect of correctness [$F(1, 23) = 6.015, p=0.0222, \eta^2 = .05$], where subjects spent less time on correct rejections [$\mu = 1.234$] than false alarms [$\mu = 1.327$] (Fig. 3.4b). Interestingly, this trend was the reverse of the mnemonic discrimination findings, where longer reaction times were associated with correct responses. In addition, there was no main effect of race on response time [$F(1, 23) = 0.849, p=0.3664, \eta^2 = .0014$], nor an interaction of race and correctness [$F(1, 23) = 0.1448, p=0.7070, \eta^2 = .00$]. Despite this, a post hoc Sidak comparison echoed the mnemonic discrimination results though to a lesser extent, finding reaction times differed more greatly between correct and incorrect other-race ($p = 0.0196$) relative to same-race faces ($p = 0.0679$). This was the only deviation found between same-race and other-race behavior in the match-to-sample task. In a final analysis, we tested whether there was a reaction time difference in same-race and other-race faces across the mnemonic discrimination and match-to-sample tasks by running a 2 x 2 repeated measures ANOVA with stimulus race as a within-subject factor and task as between-subject factor. Unlike the d' analysis, there was no main effect of race on reaction time across tasks [$F(1, 97) = 2.812, p=0.0968, \eta^2 = .00$]. There was a main effect of task type [$F(1,97) = 8.912, p = .0036, \eta^2 = .08$] but no interaction of task and race [$F(1,97) = .2619, p = .61, \eta^2 = .00$].

Discussion

We characterized the ORE using a mnemonic discrimination task, which unlike standard recognition tasks introduced face lures of varying similarity from previously presented faces. This task is sensitive to pattern separation, a neural computation that supports discrimination among similar experiences. This afforded us the opportunity to characterize recognition accuracy in terms of the ability to resolve mnemonic interference between prior face memories and new experiences of faces. Specifically, we found that facial recognition is modulated by race and stimulus similarity. Our results supported our prediction that subjects would demonstrate enhanced recognition accuracy for same-race over other-race stimuli at intermediate interference levels, and even at the highest interference levels where distractor faces were maximally similar to the originals. Same-race recognition was significantly better than other-race recognition at all but the lowest interference level; Only when faces were as little as 50% similar to one another, could subjects discern differences in other-race faces as readily as same-race faces.

A major question we sought to answer was whether reducing proactive interference would reduce or abolish the ORE. In contrast to the clear ORE we observed in the mnemonic discrimination task, subjects demonstrated equal accuracy on same-race and other-race face recognition judgments in a match-to-sample task. Performance increased as face pairs became more distinct from one another, however this was independent of stimulus race. Subjects therefore demonstrate no deficit in resolving interference between other-race face representations when faces were internally represented and maintained one at a time for several seconds. The ORE was only observable when proactive interference was increased

by the generation and storage of multiple overlapping face representations in memory. These results support the hypothesis that the ORE is related to deficits in interference resolution during episodic memory processing for other-race relative to same-race faces.

It is possible that this interpretation of our results is limited by non-matched task difficulty across the mnemonic discrimination and match-to-sample paradigms. That is, more taxing demands in the mnemonic discrimination task could be giving rise to the ORE, and perhaps a similar effect could be produced by a match-to-sample paradigm if it were made comparably challenging. However, even when performance (as an index of task difficulty) is matched across both tasks at approximately d' of 0.5 (Fig. 3.2b,d), there is no ORE present in the match-to-sample task, indicating difficulty alone does not elicit an ORE. Additional support comes from a similar study to ours, which did not detect an ORE in a match-to-sample task, even with retention intervals of over 12 seconds and high face-pair similarity.¹⁰⁵ Their results also suggest that long retention intervals alone may not generate enough proactive interference to elicit an ORE. Only when their study disrupted maintenance of internal representations with trivia questions did an advantage for same-race recognition emerge.

We are not suggesting that perceptual and attentional encoding processes do not play a role in the emergence of the effect. For instance, in the match-to-sample data, a post hoc reaction time analysis captured what could be an indicator of the ORE in attention— where reaction time was related to accuracy in other-race but not same-race faces. These results suggest that while perceptual and attentional processes may not always facilitate an ORE in early behavior, they may still give rise to qualitative differences in face processing or

representations. For instance, studies finding no behavioral differences in working memory for same-race and other-race faces still find differences in EEG components during maintenance of those faces^{106,107}. And as discussed in chapter 2, numerous studies have found neural differences in same-race and other-race face processing in perceptual tasks. Such differences may contribute an emergence of the ORE in recognition memory, though this has not been extensively tested. Furthermore, there are several studies where OREs were behaviorally detectable in traditional working memory tasks^{108,109}. In these cases, it is possible there is no involvement of mnemonic mechanisms, in line with the classic view that working memory does not recruit brain regions associated with long-term memory. However, it is also possible that under certain task demands and/or growing proactive interference, the same mnemonic mechanisms implicated in long-term memory tasks could be engaged. There is certainly emerging evidence that selective attention processes during maintenance may act on mnemonic in addition to perceptual representations, recruiting the long-term memory associated medial-temporal lobe (MTL) when the task demands it¹¹⁰⁻¹¹³.

Reconciliation with prior ORE frameworks

Our suggestion of a mnemonic component to the ORE is not incompatible with perceptual expertise and social cognitive theories (see chapter 1). Expertise accounts suggest specialized holistic processing mechanisms are tuned exclusively to same-race faces, while social-cognitive accounts suggest automatic processing of race categories leads to more deeply encoded and higher fidelity same-race representations relative to shallowly encoded and impoverished other-race representations^{16,23}. Our results suggest that the

tuning of same-race relative to other-race representations extend to mnemonic mechanisms, and that this tuning allows for same-race representations to be more discriminable than other-races ones at retrieval. Our study presents a novel approach to study this. We plotted multiple levels of mnemonic interference against accuracy to produce input-output transfer functions for facial recognition. The ORE was operationalized as the disparity between same-race and other-race transfer functions. Higher same-race performance along the input-output transfer function is likely reflective of memory mechanisms that have been tuned via years of predominant interaction with and privileged social individuation of one's own race group to optimally discriminate and generalize between same-race faces. At the same time, experience with other-race individuals may be impoverished and compounded by suboptimal attentional encoding due to implicit labeling as "other", resulting in an other-race input-output transfer function much reduced from the same-race one. This divergence may reflect a system sub-optimally tuned for other-race face recognition.

It is worth noting that the nonlinear tuning of input-output transfer functions we observed here are similar to results in another recent study examining the relationship between physical fitness and mnemonic discrimination¹⁰³. The researchers found a curvilinear input-output transfer function for highly fit relative to more sedentary subjects. The authors interpreted this finding as a possible enhancement of pattern separation processes resulting from long-term physical activity and exercise. By the same logic, our results could highlight an enhancement of pattern separation processes for same-race relative to other-race faces resulting from increased experience with and attention paid to same-race

individuals. If this is the case, the ORE may emerge in part as a result of altered efficiency for neural pattern separation of faces from distinct race groups.

Because our study highly controlled the amount of interference between every face pair, it appears that other-race face pairs that are *more* physiologically distinct than same-race ones, are still often encoded and remembered as more similar, leading to deficits in mnemonic discrimination. These findings, and the backing of computational and experimental mnemonic discrimination and pattern separation literature are interestingly consistent with the face-space model^{43,44} (chapter 1). This model suggests that within latent psychological space, same-race faces are encoded with a widespread spatial distribution while other-race faces are clustered and more confusable. Our proposed mnemonic discrimination separation model similarly suggests that same-race face representations are more effectively separated/disentangled than other-race ones. A greater ability to orthogonalize same-race representations would result in the non-overlapping and distinct same-race face representations proposed by the face-space model. Similarly, a reduced ability to disentangle and discriminate between other-race representations is highly consistent with the proposal that other-race faces are clustered and therefore confusable in face-space. Pattern separation could plausibly serve as one mechanism contributing to the unique organization of representations within face-space. An interesting line of future work could be to test whether spatial distance within experimentally determined face-space is associated with mnemonic discrimination performance as well as neuronal indices of pattern separation.

Suggestions for future neuroimaging research

Characterizing the ORE in terms of mnemonic in addition to perceptual and attentional mechanisms paves the way for a more inclusive neurobiological approach to uncovering the neural basis of the ORE. The majority of neuroimaging studies of the ORE focus on visual processing regions alone – specifically the fusiform face area (FFA, lateral fusiform gyrus) of the inferior temporal cortex. However, while the FFA seems greatly involved in differential representations of race^{67,76,114,115} its activity has not consistently predicted recognition accuracy– the behavioral metric of the ORE (chapter 2). Given the role of assessing the ORE using memory tasks, it is surprising that studies have not explored the involvement of the medial temporal lobe. At the root of this may be a widespread modular perspective on visual and memory processing regions in the brain, where occipito-temporal areas are associated with perception and medial temporal regions with memory. However, there is growing evidence that the functional boundaries of perceptual and mnemonic processes are blurred across anatomical lines, and that regions are recruited based on the complexity of representations and information they contain, which are necessary to complete the task at hand^{73,116-121}. This ‘Representational Hierarchical’⁷³ perspective is supported by work finding that the perirhinal cortex (a region traditionally involved in memory processing) is integral to facial recognition,^{116,120,122-126} and has been shown to be sensitive to facial discrimination accuracy regardless of the perceptual or mnemonic nature of the task^{120,124}. With these results in mind, we propose that regions typically associated with episodic memory, including the hippocampus and rhinal cortex play a role in generating the ORE. Due to the mnemonic discrimination task’s tendency to

engage the medial temporal cortex, our results here suggest the ORE may be in part facilitated by the different extent to which perirhinal or hippocampal pattern separation mechanisms may be recruited for same-race and other-race faces. We suggest that future studies focus on the role of the hippocampus and perirhinal cortex in generating the ORE.

...

In conclusion, we developed a mnemonic discrimination paradigm that evaluates the role of memory processes in the ORE. Our findings suggest that the ORE is not a purely perceptual or attentional phenomenon and is exacerbated when faces must be held in memory amidst temporal and visual interference. Our task additionally improves upon standard ORE recognition paradigms by evaluating accuracy as an incremental function rather than a single measure, which offers a richer means by which to quantify the ORE and how it changes with training. These results pave the path to a more detailed neurobiological investigation of the ORE (addressed in the remaining chapters of this dissertation), as well as interventional studies attempting to reduce or eliminate the impact of the ORE.

Chapter 4: Transitioning from the Medial-Temporal Lobe to the Whole Brain in Investigating Neural Correlates of the Other-Race Effect

Nonoptimal Behavioral Results for Testing Pattern Separation Hypotheses

As outlined in the neuroimaging section of chapter 3's discussion, we were interested in testing whether mnemonic-processing regions involved in pattern separation (including the perirhinal cortex and hippocampus) were differentially engaged with same-race and other-race faces.

To test this, the behavioral study was adapted for use in the MRI scanner, and data was collected for 27 subjects. (Details are reported in chapter 6.) However, in reviewing the data, it became clear that the distribution of behavioral responses across the four interference levels (20%, 30%, 40%, and 50%) were not ideal. In the analysis approach generally used for MRI data, there should be at least 10 instances of a condition to reliably model its impact on the blood-oxygen response from the brain. However, participants performed more poorly in the MRI version of the task, often resulting in too few target hits ('Same' | Target Repeat) and correct rejections ('Different' | 'Lure Distractor') in the higher interference trials (20% and 30%). The reverse was sometimes found in the low interference conditions (40% and 50%) with too few target misses ('Different' | Target Repeat) and false alarms ('Same' | Lure Distractor). Even when trials were binned together into high (20% and 30%) and low (40% and 50%) interference conditions, enough subjects showed a skewed distribution of trial counts that it was not tenable to continue with an analysis specific to mnemonic interference. To improve distribution of conditions in future

studies, I would recommend reducing the number of lure bins and increasing trial counts per bin. Alternatively, one could maintain four lure bins and increase the number of trials in each, but this which would require lengthening the study – a risk in an already difficult and fatiguing task.

Conclusions related to pattern separation in the brain require comparing brain regional activity across different levels of mnemonic interference, since pattern separation is the ability to resolve interference. Generally, pattern separation-consistent activity is reflected when the brain's signal in response to a lure is similar to signal in response to an entirely new item – suggesting that the lure is neurologically being represented as new. In contrast, a relative reduction in signal to a lure is analogous to the reduction in signal found in response to a repeated item, suggesting the lure is being represented as old.

Much of the face recognition literature exploits these novelty signals to index when a new identity is detected^{70,127-129} 127. However, to my knowledge no study has tested whether novelty signals (also referred to as signal facilitations) to lures in the FFA differs between stimuli with lower or higher confusability from originally encoded faces. We anticipated that the perirhinal cortex and hippocampus would demonstrate unique novelty facilitation to high interference lures at retrieval, suggesting an involvement of pattern separation in MTL regions for difficult face recognition contexts that the FFA may not support on its own. Furthermore, we expected a decrement in medial-temporal lobe novelty facilitation for other-race lures – evidence of reduced pattern separation that might predict the other-race recognition deficit. However, without sufficiently distributed data across interference

levels, it was not possible to test these hypotheses. Therefore, the hypothesis that pattern separation is reduced in other-race relative to same-race faces remains an open question.

Novelty of Mnemonic Discrimination Paradigm

While data could not be analyzed across interference levels, once levels were binned together there was more than enough trials to analyze neural activity differences between accurate and inaccurate encoding and retrieval of same and other-race faces. Even with all the interference levels collapsed, our paradigm preserves its novelty relative to many past studies of the ORE because of the inclusion of lures. Traditional ‘old’/‘new’ recognition studies of the ORE test memory via inclusion of directly repeated images (‘old’/‘same’), or entirely new images (‘new’/‘different’) during a test phase. This allows researchers to contrast activity during encoding that leads to successful or unsuccessful responses to the repeated faces⁶⁷. A common finding in recognition memory literature is that there are certain brain regions that display ‘subsequent memory effects’ where greater activity during encoding is associated with better subsequent recognition performance. The novelty of our study, (which already allows us to test the subsequent memory effect) is the inclusion of lures making possible identification of regions integral to “subsequent mnemonic discrimination effects,” i.e. encoding activity associated with successful correct rejections of distractor lures.

Furthermore, independent of the novel lure condition, the mere inclusion of a recognition paradigm in a neuroimaging study of the ORE is surprisingly rare (see chapter 2). Only three fMRI studies of the ORE to my knowledge have employed traditional memory

tasks^{67,69,75}, and exclusively scanned participants during face encoding but not retrieval. The recording of retrieval activity during face recognition is therefore another unique contribution of our study.

Exploratory Data Analysis Informing Subsequent Analysis Methods

In transitioning the question of interest away from neurologically validating the pattern separation hypothesis, I performed a whole-brain exploratory data analysis to visually compare activity in participants brains when they performed same versus other-race task conditions. Observations from this analysis will be briefly discussed given their pivotal role in motivating my ultimate approach. However detailed methods are not reported as they are not central to this dissertation. For details on the task design, subject population and final analysis, refer to the study presented in chapter 6.

Several group-level contrasts of brain activity were modeled in AFNI^{130,131} to visualize brain areas that were more active for accurate than inaccurate conditions across subjects. We contrasted encoding trials within each race based upon subsequent performance at retrieval. This produced four statistical maps (Fig. 4.1) of the difference in encoding activity for: same-race target hit - same-race target miss (4.1a, top); other-race target hit - other-race target miss (4.1a, bottom); same-race lure correct rejection - same-race lure false alarm (4.1b, top); other-race lure correct-rejection - other-race lure false alarm (4.1b, bottom). Put another way, these maps can be conceptualized as tracking the same-race subsequent memory advantage, other-race subsequent memory advantage, same-race subsequent discrimination advantage, and other-race subsequent discrimination

advantage, respectively. Within these maps the values in each voxel indicate the extent to which the stimulus conditions account for the signal variation across subjects. For instance, in the same-race memory advantage map, a larger positive value indicates a voxel is more responsive during the accurate condition across subjects, while a larger negative value indicates a voxel is more responsive during the inaccurate condition. These maps were loaded and visualized in the AFNI Graphical User Interface. For visualization purposes an arbitrary liberal threshold was applied to mask out voxels that responded similarly to both conditions within a contrast. This resulted in visualizations where orange and blue blobs correspond to encoding activity that was later associated with successful or inaccurate performance, respectively. Therefore, regions that may support subsequent memory effects and subsequent mnemonic discrimination effects appear orange. What was most striking was the relatively sparse activity in the other-race contrasts (Fig. 4.1 bottom panel) compared to the same-race ones (Fig. 4.1 top panel). Regardless of the accuracy or the trial type the same-race conditions appeared to maintain more voxels at the applied threshold.

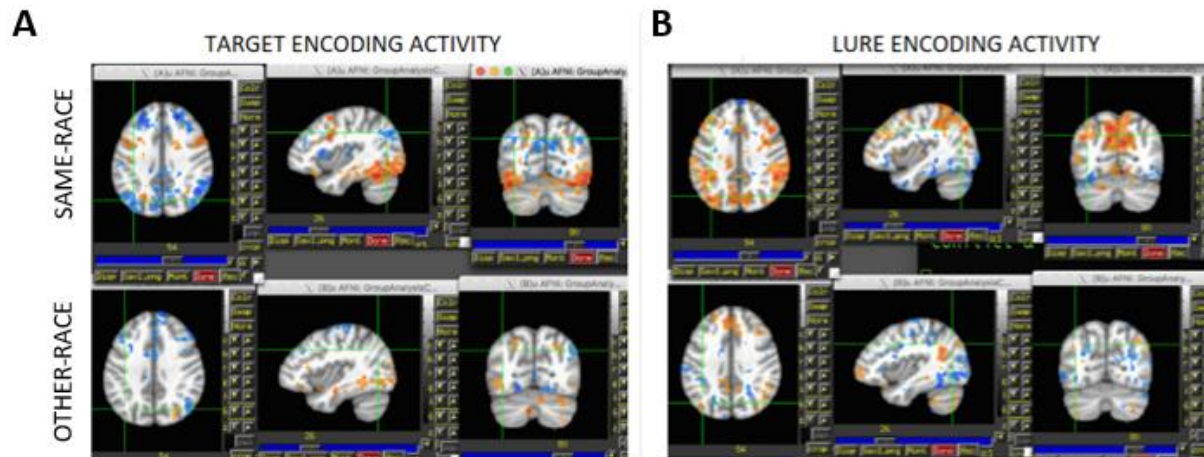


Figure 4. 1 Contrasts of encoding activity for successful and unsuccessful subsequent trials
 (A) Contrasts of target hits and target misses for same-race (top) and other-race (bottom) faces. (B) Contrasts of lure correct rejections and false alarms for same-race (top) and other-race (bottom) faces. Legend: Orange - greater activity during accurate trials; Blue - greater activity during incorrect trials.

While laying out the contrasts and comparing them side-by-side emphasizes differences between them, it is important to highlight that these maps were liberally thresholded. A much higher threshold would be needed to determine the regions that are *significantly* more active in accurate relative to inaccurate conditions. This threshold increase is necessary, given the possibility of falsely determining a voxel is highly active given the vast number of statistical tests that are run in whole-brain analysis. To determine the threshold at which significance can be established, a multiple comparison correction will generally need to be applied. This makes whole-brain analysis a conservative approach that may be underpowered to detect subtle differences in same and other-race face processing.

This prompted an investigation of alternative and more sensitive methods for whole-brain analysis. For the last decade plus, the emerging field of network neuroscience has pioneered methods well suited to analyze the brain as a large complex system¹³². These approaches are informed by graph theory, a mathematical framework that allows for

modeling pairwise relations between objects in a system. Within the context of neuroscience, the pairwise relationships between regions of the brain can be modeled using functional connectivity analysis. Importantly many algorithms have been developed that reduce the dimensionality of complex systems into metrics reflecting various aspects of network organization. In this way it is possible to test how brain regions coordinate activity to facilitate complex behavior. Such metrics can be used to compare whole-brain (as well as region-of-interest) function across task-conditions. By removing the need for statistical tests at the voxel level, it is possible to circumvent the stringent and prohibitive multiple comparison corrections associated with traditional whole-brain analysis. Such an approach may therefore be more sensitive to detect significant differences in how the brain interacts with information from same and other-race faces.

...

In summary, the results of the MRI version of the mnemonic discrimination task were not optimal for testing our original hypotheses of pattern separation's role in producing the ORE. An exploratory whole-brain analysis was performed, motivating an ultimate interest in analyzing the brain from a graph theoretical perspective. The following chapter reviews current applications of graph theory in neuroscience to understand how the brain functionally reorganizes to support diverse behavioral tasks.

Chapter 5: Review of Relevant Works in Network Neuroscience

The wealth of brain imaging methodologies designed to measure functional connectivity reflects a growing acknowledgement that no brain region works alone. Functional connectivity is defined as the temporal coincidence of spatially distant neurophysiological events, where regions demonstrate connectivity if there is a statistical relationship between them¹³³. For instance, a large body of research finds that the at rest, the brain functionally segregates into separate ‘intrinsic connectivity networks,’ where spatially distant regions couple with one another¹³⁴. A focus on magnitude of activity in individual regions (as performed in the whole-brain exploratory analysis in chapter 4), is unable to capture crucial information on how regions of the brain are coordinated with one another. The growing field of network neuroscience is focused on this connectivity between brain regions, exploiting powerful tools provided by the mathematical framework of graph theory.

Network neuroscience draws its methods from the more general field of network science – a discipline exploring how the organization of component parts of a system leads to emergent and complex properties. Network science uses mathematical algorithms from graph theory to understand complexity across systems as diverse as transportation networks, the power grid, social networks, and brain networks.

At the core of graph theory is the modeling of some complex entity using graph representations^{132,135}. A graph is made up of a series of nodes that are connected by edges. In certain domains, the definition of nodes and edges are relatively simple; For instance, in

transportation a metro connection between two neighborhoods is binary – it either exists, or it does not. In contrast, modeling the brain requires more complex decisions about what constitutes a brain region and furthermore, when we can infer that brain regions are connected. Nodes can be defined with varying granularities, from segments as small as one voxel, to multiple anatomical regions grouped together. Edges may be structural and based on fiber tracts, or functional and based on a variety of functional connectivity methods including resting state connectivity, generalized psychophysiological interaction or mutual-information-based analyses, to name a few. There are further decisions to be made regarding the use of binary or weighted graph representations, and a variety of algorithms developed and optimized for only one or the other, often requiring hyperparameter selection. Therefore, the field of network neuroscience is still developing its own best practices, and gold standards do not yet exist^{132,135}.

Regardless of the network neuroscientist's methodology of choice, an area of great interest has been the discovery that the brain demonstrates small-world network topology. A landmark paper found that many types of complex and effective systems are characterized by high clustering of nodes, but also relatively short path lengths between nodes¹³⁶. This topology lies somewhere in between networks that had already been formalized in graph theory – regular lattices with high clustering of nodes and random graphs with short path lengths between unconnected nodes. With an integration of the structural qualities of these former graph configurations, small-world networks provide a cost-efficient configuration that maintains an ability to quickly integrate information between distinct and distant regions. It is perhaps unsurprising that so many networks demonstrate small-world

topology, from the brain structure of *C. Elegans* to the Western US power grid and the collaboration of film actors.

Shortly thereafter a seminal paper introduced algorithms that quantify clustering and path length qualities of small-world networks into metrics approximating how efficiently systems exchange information at both global and local levels¹³⁷. The authors defined efficiency in communication between a pair of nodes as inversely proportional to the shortest path length between those nodes. I.e., the shorter the path (or number of steps) between two nodes, the greater efficiency with which those can exchange information. This calculation can be applied to characterize how well all nodes within a network are integrated with one another (global efficiency), but also how fault tolerant a network's various modules of nodes are to perturbation (local efficiency). (The formal equations for these metrics are introduced in the methods section of chapter 6.) Small-world topology is characterized by both high global and high local efficiency, demonstrating an ability to easily integrate, but also segregate information when necessary.

Graph theoretical metrics like efficiency may help serve as biomarkers of clinical conditions. Numerous studies have found topological metrics can distinguish between subjects with healthy brains and those with conditions known to alter brain structure and function. For instance, Alzheimer's disease is associated with reduced small-world topology, showing reductions of clustering coefficients and increased path lengths and reduced global efficiencies¹³⁸⁻¹⁴⁰. Patients with epilepsy have been shown to have reduced global and local efficiencies relative to healthy controls, indicating compromised network integration and segregation¹⁴¹. In Autism Spectrum disorder, patients have overly

integrated networks, indicated by increased global efficiency at the expense of reduced local efficiency¹⁴². The opposite has been observed in children with ADHD with overly segregated brain networks having high local efficiency but low global efficiency. The deviations in small-world topology found across these conditions underscore the importance of this configuration to healthy cognitive function.

However not all alterations in small-worldness are maladaptive. Healthy brains rely on dynamic reconfigurations of network integration and segregation to meet shifting behavioral demands¹⁴³. One study demonstrating this interplay compared brain connectivity at rest to connectivity during motor sequence tapping and working memory n-back tasks¹⁴⁴. Local efficiency was greater at rest, whereas global efficiency was uniquely high during working memory relative to resting state and sequence tapping tasks.

Furthermore, higher local efficiency predicted poorer performance in both tasks, while higher global efficiency was associated with increased working memory. A separate study using an n-back test similarly found better working memory performance is associated with decreased local efficiency across young and older populations¹⁴⁵. However only young subjects showed a positive relationship between global efficiency and working memory capacity; In fact, such integration was related to slight performance decrements in the aging group. Overall, these studies suggest that more cognitively demanding tasks may rely on greater information integration across the brain via a reduction in local efficiency and facilitation in global efficiency. Furthermore aging, known to alter brain topology may also dictate the topological organizations conducive for certain task demands.

Studies on the topological correlates of longer-term memory storage and retrieval are rarer, though they have demonstrated a similar relationship between network integration and performance. One study had subjects memorize words, and later scanned them during retrieval¹⁴⁶. Reduced brain modularity (a metric of segregation) was associated with improved memory and tracked individual differences in behavior.

These studies paint a picture that global integration is key in complex tasks that require maintenance and memory retrieval. One might instead argue that increased global integration may be associated with high task difficulty in experimental relative to control conditions, rather than the recruitment of memory-specific mechanisms. However, one highly difficult task found that functional connectivity demonstrated increased modularity prior to subjects successfully perceiving hard to hear sounds¹⁴⁷. While this task paradigm is incredibly different than the memory paradigms, it serves to highlight that not all behaviors benefit from more integrated network architectures, and that task difficulty is not always synonymous with increased integration.

In fact, there is evidence that improved performance on difficult tasks is associated with increased modularity over time. A recent study tested whether longitudinal training on a working memory paradigm would shift network topology from more integrated to segregated configurations¹⁴⁸. The results of a six-week training supported this hypothesis, finding a relationship between increasing network segregation and behavioral improvements. The authors suggest that increasing modularity may be associated with greater experience and task automation rather than reduced complexity of tasks. The

authors suggested that experiential tuning towards a more modular network may provide a more cost effective and efficient functional network configuration over time.

...

The application of network neuroscience to task-related connectivity is still quite nascent. While integrated functional network architectures tends to be associated with successful memory performance, there is also evidence that training and task automation is associated with increasing network segregation. To my knowledge no study has investigated the role of topological organization during successful encoding and retrieval of faces. Because the ORE is experientially tuned and involves recruitment of complex mechanisms, it may very well be associated with alterations in global and local efficiency. The following chapter discusses the application of these graph theoretical metrics to characterize network topologies supporting same and other-race face processing.

Chapter 6: Graph Theoretical Analysis Reveals Reconfigurations in Network Efficiency across Same- and Other-Race Facial Recognition

Introduction

As reviewed in chapter 5, the brain functionally reorganizes across resting state, visuomotor, and working memory tasks to meet the demands of shifting behavioral contexts^{143,144,148,149}. However, no study to my knowledge has tested network reconfigurations on two contexts as similar as same and other-race face recognition. One might argue that face recognition, regardless of the race of a face, involves a very similar set of cognitive processes, such that no differences would be evident in whole-brain functional architecture. However as previously discussed, memory performance is associated with brain topological changes, and it is therefore plausible that differences in same-race/other-race memory are associated with altered network topologies. Furthermore, the cognitive frameworks (chapter 1) and neuroscientific evidence (chapter 2) in the other-race effect (ORE) literature find mechanistic and neural differences which would plausibly result in network alterations on the systems level. We therefore set out to test whether differentials in brain-wide network efficiency correspond to the ORE in behavior.

While no study to our knowledge has explored the contributions of efficiency in network topology to the ORE, several experiments have employed graph theoretical methods to study the brain's face network. The face network is comprised of a series of regions that demonstrate preferential activity-- i.e. more oxygen consumption during presentation of faces than other image categories-- across a variety of neuroimaging paradigms¹⁵⁰. These

include a core system containing the most face selective regions-- the Occipital Face Area (OFA), Fusiform Face Area (FFA), and posterior superior temporal sulcus (pSTS). These areas have been implicated in both featural (OFA) and holistic processing (FFA) as well as allocation of attention towards faces (pSTS). However, many other regions across the brain have been identified as important for face processing. For instance, areas crucial to spatial attention including the intraparietal sulcus and frontal eye fields are implicated in extracting identity-relevant information from faces. The amygdala, and insula are involved in tracking emotional content of faces. Temporal lobe regions like the ventral anterior temporal face patch are recruited during processing of semantic and biographical information related to face identities^{151,152}.

Network neuroscientific studies of face-processing have tended to focus on the connectivity between these highly face-selective regions. One study tested how the face network is reconfigured across development by scanning the brains of subjects ranging in age from childhood to early adulthood¹⁵³. They found that face regions grouped into several submodules that significantly reorganized across development, demonstrating both network segregation and integration at different timepoints. In later childhood and adulthood, a stable module formed that included early visual regions (primary visual cortex) and temporal and frontal regions (hippocampus and inferior frontal gyrus). Because these latter regions are involved in domain-general memory and biographical semantic face processing, the authors suggested this submodule might support the development of perceptual expertise known to be tuned into adulthood.

While the prior study observed integration in the face network across the lifespan, the authors acknowledge they focused on the face network in isolation. Therefore, it is plausible that these regions might form a unified network that becomes more segregated from the rest of the brain during development. A separate study, while still focused on face preferential regions, tested the relationship between these regions and the rest of the brain by calculating graph theoretical centrality measures for each face node¹⁵⁴. Centrality indicates a region's relative importance within an entire network, based on how connected it is to other nodes. Nodes with high centrality play a role in integrating information across different modules¹⁵⁵. Between childhood and early adulthood, there was a decrease in centrality of eight out of eleven face-preferential regions. The authors speculated that these nodes played a greater role in information dissemination in childhood, but then become more segregated from the rest of the brain allowing for improved domain-specific processing of face information. However, three nodes did increase their centrality --- evidence that face regions are not uniformly segregated from other brain networks. It is therefore plausible that a reduction in centrality of certain regions reflects cost-efficient developmental pruning, given that the more central nodes in adulthood may be able to propagate information to the remainder of nodes in the face network. This likely reflects a more efficient topological configuration overall.

These two studies demonstrate that across development the face network is topologically reconfigured. As discussed in chapter 1, perceptual expertise theories suggest such tuning is selective to the visual input we are most experienced with. Therefore, we can begin to ask questions relevant to our topic of interest: If the topological configuration of the face

network is tuned for same-race recognition, will it still be effective in relaying information relevant to other-race recognition? Would the network reconfigure for other-race face processing, or just perform less optimally?

Moreover, how would studying the brain as an entire network during face processing differ from analyzing the connectivity of a select few highly face-selective regions? One study reviewed in chapter 2 found that the extent of connectivity between the FFA and cognitive control and attention networks was more predictive of same-race than other-race memory. Yet the areas in these latter networks are not associated with face-preferential activity. Clearly that does not negate their involvement in facial memory and the ORE. This is evidence that regions need not be highly selective to faces to play a mediating role in face memory. Therefore, it may be somewhat limiting to study only the topology of the face-preferential network, at least in the context of the ORE.

This evidence in the literature, as well as our own exploratory observations of group-wide whole-brain differences in regional engagement between same and other-race faces (chapter 4) poses the question of how the brain as a system interacts with information to make complex decisions about different categories of faces. By pairing whole-brain graph theoretical analysis with our mnemonic discrimination study (chapter 3), we tested whether the ORE can be described by alterations in efficiency of information processing in the brain. The exact nature of topological differences across same and other-race faces is hard to predict owing to conflicting results in the literature surrounding when segregation versus integration is behaviorally adaptive (see chapter 5). Therefore, I propose several possible topologies that could be observed.

1. Because small-world architecture has been associated with healthy and improved cognition, processing of same-race faces might be supported by more small-world network configurations. This would be apparent with higher local and global efficiencies during accurate same-race relative to other-race face recognition.
2. Because accuracy on complex tasks has been associated with increased integration and decreased segregation, same-race recognition could be associated with relatively high global efficiency and low local efficiency. Because other-race recognition is not as accurate, we might expect relative deviations from this configuration, such as reduced global efficiency, and increased local efficiency.
3. Because developmental tuning and behavioral training evidence have been associated with increased segregation reflecting task automation, the expertise participants have for same-race faces could be associated with a more locally efficient network overall. Meanwhile other-race recognition which might be more difficult and less automated for subjects, may still require a more integrated and globally efficient network to achieve success.

Methods

Participants

This study protocol was approved by the Institutional Review Board (IRB) at the University of California, Irvine, and complies with IRB guidelines and regulations. Participants were screened for eligibility through a secure online questionnaire using REDCap electronic data capture tools¹⁵⁶. Strict inclusionary criteria required participants to be right-handed with

normal corrected vision, self-identify as East Asian or Southeast Asian, have no MRI contraindications such as metal implants, and no major neurological, psychiatric, or substance-use conditions. 27 recruited subjects (14 Female, 13 Male; mean age 19.93; SD 1.36; age range 18 -22) provided written informed consent and were compensated for their participation. During the study, the participants filled out a series of questionnaires before and after scanning, as well as performed facial recognition tasks in the scanner. Because awareness of the true nature of the study could bias performance^{157,158}, participants were only informed that they would be administered a facial recognition task, with no mention of the race component of the study. After the scan they were debriefed about the full study purpose and consented once again.

Of the initial sample, five participants were excluded from the analysis. Reasons for exclusion included participants not meeting final inclusionary criteria, demonstrating chance performance, missing 20% or more of trials, as well as technical difficulties with the scanner. This yielded a final sample of 22 subjects (11 Female, 11 Male; mean age 19.52, SD 1.29; age range 18 - 22). Of the final participants, 9 identified as East Asian and 13 as Southeast Asian. Two of these subjects only completed three of the four scan runs.

Mnemonic Discrimination Task

This experiment was designed to test participants' retrieval memory for faces of their own and another race. The task was designed for a previous study and adapted for use in the MRI scanner⁷⁷. Research from unpublished pilot studies found that Asian participants displayed greater memory deficits for Black faces than white faces. We therefore limited

our design to include only Asian and Black faces to focus on the larger behavioral differential in our sample. Since all subjects identified as Asian, for the purposes of this manuscript we refer to Asian faces as 'Same-Race' (SR) and Black faces as 'Other-Race' (OR).

Stimulus Set

This experiment used the same stimulus set developed for our prior study⁷⁷. All faces were generated using FaceGen Modeller 3.5. Asian and Black faces were created using the 'Generate' tool within sub-groups for 'Asian' and 'African' racial origins. Half of the faces were randomly selected as 'parent' stimuli to serve as templates for lure distractors. These lures were created by running the Genetic Randomness algorithm on the parent faces, to apply normally distributed perturbations with means proportional to inputted values of 20%, 30%, 40%, and 50%. This introduced variation in how similar lure faces were to 'parent' faces. The current study groups all similarity levels together to increase trial counts for subsequent neuroimaging analysis.

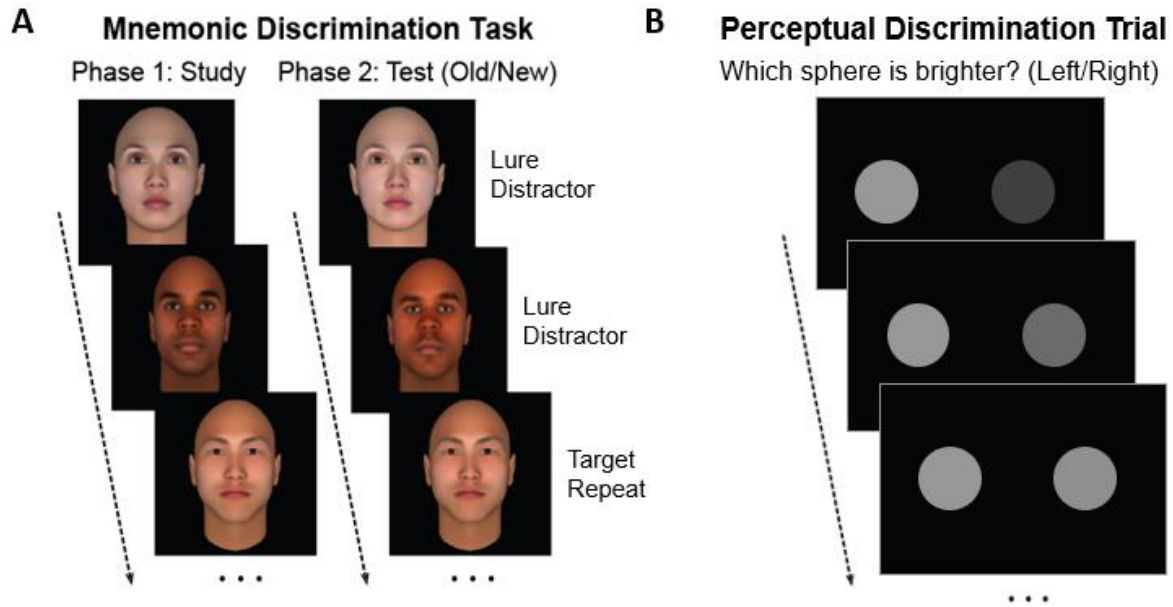


Figure 6.1 Mnemonic Discrimination and Perceptual Discrimination Task Designs

(A) The Mnemonic Discrimination task contained 8 blocks, with each block containing a Study and Test Phase. During the Study Phase participants memorized consecutively presented faces. They simultaneously indicated whether faces were shifted to the Left or the Right of the screen’s center (not displayed). During the immediately following retrieval phase they were asked if each of the presented faces were shown during the Study phase. Half of the faces were previously shown during Study (target repeats), and the other half were new (lure distractors). (B) 3 perceptual discrimination trials were randomly displayed during each phase of the Mnemonic Discrimination task. In these trials participants indicated which of two gaussian blurred spheres were brighter (Left or Right). If correct, the brightness of the spheres became more similar, and thus more difficult to discriminate. If incorrect, the spheres would become more distinct in brightness, and hence more discriminable. Participants performed these discriminations for the entire trial duration. For all trial types, stimulus duration was 3.0 s and ITI was 1.5s.

Task Design

This task, programmed in PsychoPy v1.85.2¹⁵⁹, was structured as an event-related design with 8 interleaved study and test phases (Fig 6.1a). In Study Phases, participants were instructed to memorize each of 22 consecutively presented faces. They were simultaneously expected to indicate via button press whether faces were shifted to the left or right of the screen, to ensure they were attending to the task. In each Test Phase participants were again shown a series of 22 consecutive faces, half of which were directly repeated (target repeats) from the Study Phase just prior. The other half of the faces were

new (lure distractors). For each Test trial participants were asked to identify via button press whether the face was an exact repeat from the preceding Study Phase ('Same'/'Old') or whether they had never seen it before ('New'/'Different'). A response of 'same' to a repeated face indicated successful recognition (i.e., target hit) while a response of 'different' to a new face indicated successful mnemonic discrimination (i.e., lure correct rejection). Across both Study and Test Phases, face stimuli were pseudorandomized and evenly divided amongst race and gender categories. Each stimulus was presented for 3.0s with a 1.5s intertrial interval (ITI).

In addition to trials of interest, the task also employed perceptual discrimination trials to serve as an implicit baseline when subsequently modeling the fMRI BOLD response (Fig 6.1b)¹⁶⁰. The perceptual discrimination trials required subjects to indicate which of two gaussian-blurred circles appeared 'brighter'. The trials were adaptive, adjusting based on participant response during a 3-second trial duration. When answered correctly, the circles would become more similar in luminosity, and hence more difficult to discriminate. When incorrect, the circles would diverge in luminosity. Three perceptual baseline trials were pseudorandomly presented per phase.

MRI Data Acquisition

Neuroimaging data were acquired on a 3.0 Tesla Siemens MAGNETOM Prisma scanner, using a 32-channel head coil at the Facility for Imaging and Brain Research (FIBRE), part of the Campus Center for Neuroimaging (CCNI) at the University of California, Irvine. A high-resolution three-dimensional (3D) magnetization-prepared rapid-gradient echo (MP-

RAGE) structural scan was acquired at the beginning of the session, with the following parameters: 2300 ms repetition time (TR), 2.38 ms echo time (TE), 240 slices, 0.8 mm isotropic resolution, 256 mm field of view (FoV), and 8-degree flip angle. Parallel acquisition was conducted in the GRAPPA mode with reference line phase encoding (PE) of 24 and an acceleration factor of 3. Each of four functional MRI scans were acquired using a multiband echo-planar imaging (EPI) sequence with the following parameters: TR = 1500 ms, TE = 34 ms, 64 slices, 2.1 mm isotropic resolution, 202 mm FoV, 75-degree flip angle, and multiband acceleration factor of 8. Visual stimuli were presented on a BOLDScreen32 LCD monitor mounted onto the back of the bore. Participants viewed the monitor via a single mirror attached to the head coil.

During the session, structural scans were collected first, and used to align all subsequent scans to the anterior commissure - posterior commissure line (AC-PC line). Following that, two blocks of study and test phases were presented during each functional run. Half-way between the four runs, subjects were given a break from the task during collection of a resting state scan.

Preprocessing and Denoising

Structural and functional scans were preprocessed and denoised using the CONN Toolbox¹⁶¹. Functional scans were realigned and unwarped, centered, slice-time corrected, flagged for outliers, segmented, aligned to MNI space, and smoothed using a spatial convolution Gaussian kernel of 4mm full width half maximum (FWHM). Structural scans were skull-stripped, segmented, and aligned to MNI space. A denoising pipeline was

implemented on each scan run to create regressors for potential confounding effects. These nuisance regressors included: CSF and white matter noise components using anatomical component-based noise correction (aCompCor)¹⁶²; Parameters to minimize motion-related variability via 6 realignment regressors and their first order derivatives; Censoring covariates using ART-based flagging of outlier volumes with framewise displacement above 0.9 mm or global signal change above 5 standard deviations from the mean¹⁶³. Additionally, regressors were created for the experimental task effects by convolving stimulus onset and duration with a canonical hemodynamic response function (HRF). These included regressors for each of 16 conditions of interest corresponding to the unique combinations of race (Same-Race or Other-Race), phase (Study or Test), trial type (Target Pair or Lure Pair) and accuracy (Correct or Incorrect), as well as task effects of no interest (Non-response trials, and task instruction reminders). Perceptual discrimination trials were not explicitly modeled, serving as an implicit baseline against which to compare increases or decreases in connectivity during the event types of interest¹⁶⁰. Lastly a linear component was added to model scanner drift. After regression, high-frequency information was preserved using a high-pass temporal filter [0.008 inf].

Quality control checks were run after preprocessing and denoising was complete. We found that across all subjects and runs, motion and global signal change were minimal. Mean framewise displacement was .13 mm with a standard deviation of .04 and mean global signal change was .83 with a standard deviation of .03. Because two subjects were missing one run due to scanner glitches, they were outliers for quantity of included volumes. However, they both maintained enough trials to safely model each condition

(means of 16.19 and 15.89 and minimums of 9 and 12 trials per condition, respectively). Additionally, using the IQR outlier detection method two subjects were established as potential outliers for mean motion (.23 and .24 mm). Given that these are still relatively low mean framewise displacements, we chose to include all data from these subjects as well.

Whole-brain Generalized Psycho-Physiological Interaction Analysis

In order to derive whole-brain task-modulated connectivity data for subsequent graph theoretical analysis, we ran a generalized Psycho-Physiological Interaction Analysis (gPPI)¹⁶⁴ on seed and target regions of interest (ROIs) across the entire brain^{165,166}. The brain was partitioned into 246 ROIs using the Human Brainnetome Atlas parcellation scheme¹⁶⁷. The fully preprocessed and denoised timeseries for each voxel were concatenated across runs, and then averaged within each of the ROIs. Next a general linear model (GLM) was fit for each target ROI including as predictors: a) the HRF-convolved main task effects (i.e., the psychological factors), b) the timeseries for the seed ROI (i.e., the physiological factor), and c) the product of the psychological factors and the physiological factor (i.e., the interaction term). Modeling task-dependent connectivity for all pairs of ROIs required constructing 245 GLMs for each of 246 target ROI, for a total of 60,270 separate regressions. PPI terms for the 16 conditions of interest were then assembled into separate 246x246 connectivity matrices.

Because the estimated effects of a seed ROI on a target ROI are not necessarily equal to the effects of the reverse couplings, gPPI functional connectivity matrices are not symmetric.

This quality distinguishes gPPI models from the traditional bivariate-correlation weighted GLMs used in resting state connectivity analysis. This lack of symmetry cannot be used to infer directionality of information flow between source and target ROIs, or whether coupling is direct or mediated by other ROIs ¹⁶⁸. Because there is no consensus on how to interpret this directedness in gPPI estimates, we followed the example of several other studies and symmetrized all connectivity matrices by averaging the estimates across the matrix diagonal^{169,170}.

Graph Representation

In graph theory, a graph is a mathematical structure that models pairwise relations between objects. Graphs (or networks) are defined as sets of nodes connected by edges. In the context of network neuroscience, we can model brain regions as nodes and associations between regions as edges ¹³².

Graphs are represented by adjacency matrices, which are square matrices with quantities of rows and columns equal to the number of brain regions (here, 246). The values of elements in a matrix corresponds to the associations between each pair of brain ROIs. Here, our associations are the interactions between brain regions during task conditions. Because those interactions are symmetric, these networks are considered undirected, such that only one edge is possible between any pair of ROIs.

Commonly in network analysis, adjacency matrices are binarized to contain only values of 1 and 0, indicating whether an association (or edge) exists. This requires an a priori decision regarding what constitutes a meaningful, or strong enough connection strength to

be represented within a network. Rather than choosing an arbitrary cut-off connection strength, we instead thresholded each individual connectivity matrix at five distinct costs from 5% - 25% in 5% increments^{135,144,171,172}. For each subject, each of the 16 connectivity matrices were described by five binarized adjacency matrices representing the strongest 5%, 10%, 15%, 20%, and 25% of functional interactions in the network. This range of costs was selected because complexity and non-random topology of the brain is most observable at lower connection densities¹³⁵. This allowed us to calculate network topological metrics, while ensuring that the number of edges per graph was consistent and comparable across conditions and subjects¹⁷². Once topological metrics were calculated, they were averaged across costs to derive single measures per each condition¹⁴⁴. Matrix manipulation and analysis was all performed in MATLAB R2020A¹⁷³.

Graph Metrics

Graph theoretical metrics were calculated using the Brain Connectivity Toolbox¹⁷⁴ (brain-connectivity-toolbox.net), a MATLAB toolbox for network analysis of structural and functional brain connectivity data. Our analysis focused on characterizing global and local efficiency of networks during the task conditions¹³⁷.

In graph theory, the efficiency of a pair of nodes is defined as the multiplicative inverse of the shortest path between that pair¹³⁷. The global efficiency of each brain network was calculated by taking the average efficiency of all combinations of node pairs. This computation is formalized in Eq. 1, where N is the number of nodes in graph G , and $1/d_{ij}$ is the inverse shortest path length between a given pair of nodes, i and j . In contrast. the local

efficiency of each network was calculated by computing the average efficiency of each node's subgraph (a node's immediate neighbors) and meaning across all nodal subgraphs for a network wide metric. This is represented in Eq. 2, where G_i is the subgraph of a given node i . While both metrics characterize efficiency of information exchange, global efficiency reflects how well integrated a network is, while local efficiency is a segregation metric that reflects how fault tolerance a network is to removal of individual nodes.

$$E_{global} = \frac{1}{N(N-1)} \sum_{i \neq j \in G} \frac{1}{d_{ij}} \quad (1)$$

$$E_{local} = \frac{1}{N_{G_i}(N_{G_i}-1)} \sum_{i \neq j \in G_i} \frac{1}{d_{ij}} \quad (2)$$

Results

Behavior

There are several common behavioral indices of the ORE. These include greater target hits to same-race faces, greater false alarms to other faces, greater sensitivity (d') to same-race faces, and more liberal response criterion (c) to other-race faces. Because we expected effects to take on these directions, we ran 1-tailed paired t-tests. Participants accurately identified faces as repeated (target hits) for same-race (SR) and other-race (OR) faces with the following mean proportions: [SR: $\bar{x} = .552$, $\sigma = .111$; OR: $\bar{x} = 0.577$, $\sigma = .109$]. Subjects incorrectly recognized lure distractors as repeated (lure false alarms) with the following proportions: [SR: $\bar{x} = .485$, $\sigma = .097$; OR: $\bar{x} = 0.536$, $\sigma = .093$]. While it was not expected that the mean proportions for target hits would be larger for other-race than same-race faces,

lure false alarm ratios followed the anticipated pattern. The proportions were used to calculate sensitivity (eq. 3) and response criterion (eq. 4) indices from signal detection theory:

$$d' = z(\text{target hit rate}) - z(\text{lure false alarm rate}) \quad (3)$$

$$c = -\frac{z(\text{target hit rate}) + z(\text{lure false alarm rate})}{2} \quad (4)$$

The sensitivity, or discriminability index d' corrects for the tendency to respond “old/same” inaccurately, providing a measure of participants’ ability to discriminate between old repeated and new distractor faces. The mean d' for the subjects were [SR: $\bar{x} = 0.1726$, $\sigma = 0.183$; $\bar{x} = 0.1047$, $\sigma = 0.2253$] (Fig. 6.2a). A paired t-test finds that same-race discriminability trends slightly larger than OR discriminability [$t(21) = 1.542$, $p = 0.069$, $r^2 = .102$, 95% CI = -0.1594 to 0.02364]. The response criterion, independent from sensitivity indicates participants’ biases towards responding “old/same”. The mean response criterions were [SR: $\bar{x} = -0.047$, $\sigma = 0.257$; $\bar{x} = -0.148$, $\sigma = 0.239$] (Fig. 6.2b). Overall, subjects were more likely to respond that a face was remembered for both races as evidenced by negative mean criterion values. However this bias was significantly greater for other-race than same-race faces [$t(21) = 1.733$, $p = 0.0489$, $r^2 = 0.1251$, CI = -0.2227 to 0.02023], though the effect was small.

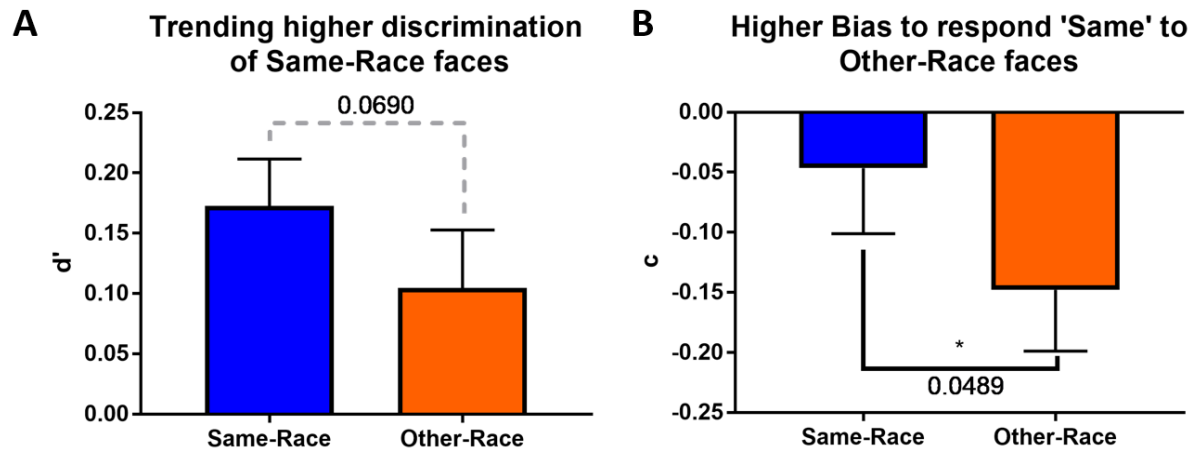


Figure 6.2 Differences in Recognition Memory for Same and Other-Race faces

While this task proved difficult for participants, behavioral patterns consistent with the ORE still emerged. (A) Sensitivity d' trended higher for same-race faces [$t(21) = 1.542$, $p = 0.069$, $r^2 = .102$, 95% CI = -0.1594 to 0.02364]. (B) Response criterion was significantly more liberal for other-race faces [$t(21) = 1.733$, $p = 0.0489$, $r^2 = 0.1251$, CI = -0.2227 to 0.02023].

Taken together these results indicate that this task was difficult – more so than the behavioral version reported in chapter 3⁷⁷, but that differences consistent with the canonical ORE still emerge. Decreased performance on this task may have resulted from the increases in time commitment, discomfort and fatigue associated with studies employed in MRI scanners. Furthermore, there were several amendments to this task that may have increased difficulty including requiring indication of left/right face position during encoding as well as the addition of perceptual discrimination trials. Despite this, participants were significantly more likely to answer that they had seen other-race faces before, and they demonstrated a trending improved ability to discriminate between old and new same-race faces. Furthermore, increased difficulty may have impacted same-race more than other-race recognition, as we may infer based on differences in discriminability performance on the earlier version of this task. In the behavioral task (see chapter 3 results), the mean d' for same and other-race recognition were .3719 and .1484,

respectively. Therefore, in the scanner version of the task there was a 53.59% decrease in same-race discriminability and a smaller 29.45% decrease in other-race discriminability. (Though, it can be argued that this is less meaningful since other-race performance was already closer to floor.) Furthermore, as difficult as the scanner version of the task was, subjects' average same-race discriminability was 14.02% greater than the average other-race discriminability in the easier behavioral task. (Response criterion was not calculated for the behavioral task and is therefore not compared.) This highlights that even in incredibly difficult contexts, same-race recognition may be superior to other-race recognition in less demanding contexts.

Neuroimaging

After running through the pipeline described in methods, each of the 16 unique conditions per participant were characterized by global and local efficiency metrics. Collapsed across conditions, the mean and standard deviation and range of the measures were [Global efficiency: $\bar{x} = 0.5402$, $\sigma = 0.0060$, range = 0.0476; Local Efficiency: $\bar{x} = 0.4977$, $\sigma = 0.0289$, range = 0.2144]. Amongst the dataset there was one subject with an outlier condition that was 6.36 standard deviations below the global efficiency mean, and 4.05 standard deviations above the local efficiency mean. This was the only condition present as an outlier for both metrics and was 2.68 and 1.27 standard deviations removed from the next largest outliers for global, and local efficiency, respectively. Furthermore, it had large influence over the relationship between global and local efficiency, inducing a trending correlation which was not present with its removal. [With outlier, $r = -0.0909$, $p = .0885$; Without outlier, $r = -.0192$, $p = .7201$]. Therefore, the extreme outlier was removed,

resulting in final descriptive statistics of [Global efficiency: $\bar{x} = 0.5403$, $\sigma = 0.0056$, range = .03172; Local Efficiency: $\bar{x} = 0.4974$, $\sigma = 0.0283$, range = 0.1778].

Each of the 16 task conditions corresponds to a distinct phase of the task (study or test), and trial type (target pair or lure pair). Data were sub-grouped based on these conditions to create four distinct categories: A) Study Target, B) Study Lure, C) Test Target, D) Test Lure. Respectively these categories correspond to the conditions under which subjects A) studied faces that would later be repeated, B) studied faces that would later have similar looking lure distractors, C) were tested for memory of repeated faces, and D) were tested for ability to correctly reject similar but never-before-seen faces. Within each of these four conditions, data were divided into factors of stimulus race (same-race (SR) or other-race (OR)) and accuracy (correct or incorrect). Because each face at study had a paired repeat or lure during the Test phase, we were able to evaluate both brain connectivity during memorization as a function of subsequent performance and during retrieval. The mean and standard deviations of global and local efficiency for each of these conditions are reported in Table 6.1.

Table 6. 1 Descriptive statistics for global and local efficiency during each task condition

Task Condition	Accuracy	Global Efficiency		Local Efficiency	
		mean	std	mean	std
Study Target SR	Subsequently correct	0.5392	0.0071	0.5013	0.0293
Study Target OR		0.5404	0.0045	0.4934	0.0216
Study Target SR	Subsequently incorrect	0.5421	0.0048	0.5082	0.0293
Study Target OR		0.5391	0.0071	0.5127	0.0268
Study Lure SR	Subsequently correct	0.5423	0.0052	0.4886	0.0314
Study Lure OR		0.5387	0.0072	0.5079	0.0221
Study Lure SR	Subsequently incorrect	0.5419	0.0031	0.4996	0.0279
Study Lure OR		0.5413	0.0040	0.4906	0.0222
Test Target SR	Correct	0.5382	0.0055	0.4961	0.0247
Test Target OR		0.5404	0.0050	0.4849	0.0300
Test Target SR	Incorrect	0.5398	0.0063	0.5038	0.0224
Test Target OR		0.5392	0.0067	0.5122	0.0289
Test Lure SR	Correct	0.5380	0.0071	0.4885	0.0323
Test Lure OR		0.5422	0.0037	0.5019	0.0225
Test Lure SR	Incorrect	0.5406	0.0053	0.4868	0.0251
Test Lure OR		0.5411	0.0035	0.4816	0.0337

Statistics are organized by task phase (Study or Test), Trial type (Target or Lure), Race (Same-race (SR) or Other-race (OR)), and Accuracy (Correct or Incorrect). For the Study conditions, Target and Lure assignment is based upon whether the face shown was part of a target pair, where the face was later repeated at Test, or a lure pair, where a similar distractor lure was shown at Test. Therefore, accuracy during the study phase is based upon performance on the corresponding item in the pair during the test phase.

Within each of the four major categories, a separate regression was run on global and local efficiency data, resulting in 8 models. Specifically, the effects of stimulus race and accuracy on the topological metrics were estimated using generalized linear models (GLMs) with generalized estimating equations (GEEs)¹⁷⁵. GLMs do not assume normality of dependent variables, and can handle missing values (i.e., the removed outlier), and when implemented with GEEs can handle repeated measures data. We characterized the correlated measures in our data using an exchangeable covariance matrix indicating that all observations per subject were equally correlated to one another. Each regression analysis modeled the linear combination of race, accuracy, and the interaction of race and accuracy on the dependent graph metric (Eq. 5).

$$Y_{efficiency\ metric} = \beta_0 + \beta_1 X_{Accuracy} + \beta_2 X_{Race} + \beta_3 X_{Accuracy*Race} \quad (5)$$

Because the ORE is behaviorally characterized by differences in performance relative to the race of faces, we were specifically interested in identifying whether efficiency of information processing in the brain is modulated by race and accuracy. We therefore directed our focus to the interaction of race and accuracy ($\beta_3 X_{\text{Accuracy} * \text{Race}}$). Given the small sample size (22 observations per factor), we ran permutation analysis for each model to reduce the likelihood that results were due to chance. For each observed dataset, 10,000 randomized datasets were generated where each of the four corresponding condition labels were shuffled within each subject. The 10,000 null datasets were then modeled using the same pipeline described above. To evaluate significance, we calculated the proportion of times that the permutation analysis estimated larger interaction test statistics, z , than those produced when modeling the true data. All 8 uncorrected p -values were then subjected to the Benjamini-Hochberg procedure to control the false discovery rate. Table 6.2 includes a summary of estimates for the interaction terms across conditions, including beta coefficients, z -scores, 95% confidence intervals, $p_{\text{uncorrected}}$ and p_{adjusted} .

Table 6. 2 Estimates for the interaction of race and accuracy across all conditions.

Phase	Trial Type	Graph Metric	β	z	95% CI	$p_{\text{uncorrected}}$	p_{adjusted}
Study	Target, 1 st Presentation	Global efficiency	0.0042	1.674	[-0.001, 0.009]	0.0586	0.09376
		Local efficiency	-0.0123	-1.332	[-0.030, 0.006]	0.102	0.1099
	Lure, 1 st presentation	Global efficiency	-0.0029	-1.468	[-0.007, 0.001]	0.0869	0.1099
		Local efficiency	0.0282	3.088	[0.010, 0.046]	0.0028	0.0224
Test	Target Repeat	Global efficiency	0.0028	1.279	[-0.002, 0.007]	0.1099	0.1099
		Local efficiency	-0.0196	-2.026	[-0.039, -0.001]	0.029	0.07733
	Lure Distractor	Global efficiency	0.0037	1.800	[-0.000, 0.008]	0.0457	0.0914
		Local efficiency	0.0186	2.402	[0.003, 0.034]	0.0154	0.0616

The columns for β , z, and 95% CIs are results from the original models. In glm/gee models, the z-score statistics are calculated by dividing the beta estimates by robust standard errors, which are larger and more conservative than traditional standard errors. $p_{\text{uncorrected}}$ were calculated by permutation analysis, while p_{adjusted} reflect FDR corrections after the Benjamini-Hochberg procedure.

As indicated in Table 6.2, four of the eight models had significant interaction terms prior to FDR correction. After adjustments, only the Study Lure condition maintained a significant interaction effect on network local efficiency [$p_{\text{uncorrected}} = .0028$, $p_{\text{adjusted}} = .0224$, $z = 3.088$, $\beta = .0282$, $SE_{\text{robust}} = .009$, $CI = [0.010, 0.046]$]. While the other three effects remained trending, they corresponded to retrieval conditions, and have much reduced test statistics relative to the Study Lure condition [Test Target_{local}, $z = -2.026$; Test Lure_{global}, $z = 1.800$; Test Lure_{local}, $z = 2.402$]. Given this, the remainder of our analysis is post-hoc, focusing on efficiency of brain networks while subjects studied faces that were part of lure pairs.

Figure 6.3a depicts the modulation of local efficiency during encoding by both race and accuracy factors. For faces that were encoded well enough to avoid false alarms at retrieval, mean local efficiency was greater for Other-race faces [$\bar{x}_{\text{SR}} = .4886$; $\bar{x}_{\text{OR}} = 0.5079$].

Meanwhile the opposite trend is demonstrated for encoding resulting in false alarms, with greater local efficiency for same-race faces [$\bar{x}_{\text{SR}} = .4996$; $\bar{x}_{\text{OR}} = 0.4906$]. (Refer to table 6.1

for descriptive statistics for all conditions.) This cross over interaction suggests that same-race and other-race recognition are supported by different local properties of brain connectivity and information exchange. Encoding of other-race faces was most successful when the brain network was organized to enhance fault tolerance via greater interconnections of nodes in local modules across the brain.

In contrast, encoding of same-race faces was more successful when local efficiency was decreased. This means that local modules contained less redundant pathways than in the other-race context. Given that all networks were kept comparable by maintaining equal edge counts, this poses the question of how edges were distributed during same-race memorization, and whether they might be rearranging into a more globally integrated network. Though permutation analysis established there was no interaction of race and accuracy on global efficiency during lure encoding, we tested a post-hoc hypothesis that same-race encoding might be characterized by a more globally integrated network in this condition (Fig 6.3b). We reran the regression analysis on the study lure global efficiency data and extracted the estimates for the main effect of race, rather than the interaction. The results suggest that race had a slight impact on global efficiency during the study lure condition [$z = 2.067$, $\beta = 0.0036$, $SE_{\text{robust}} = .002$, $p = 0.039$, $CI = [0.000, 0.007]$]. Figure 6.3b displays this main effect, where global efficiency is enhanced during encoding of SR relative to OR faces. [$\bar{x}_{\text{SR_Corr}} = 0.5423$; $\bar{x}_{\text{OR_Corr}} = 0.5387$; $\bar{x}_{\text{SR_Incorr}} = 0.5419$; $\bar{x}_{\text{OR_Incorr}} = 0.5413$] (Table

6.1). Despite this being a main effect, it appears more strongly driven by the correct rejection condition.

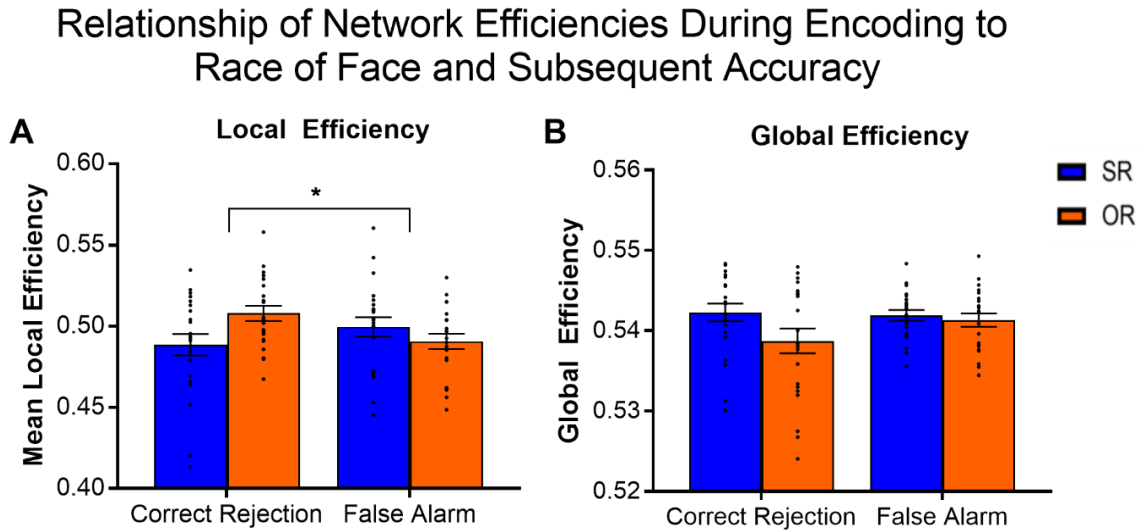


Figure 6.3 Local and Global Efficiency during Encoding of faces with subsequent distractors
 Network efficiency in the brain during encoding of faces that later resulted in correct rejections or false alarms. (A) There was a robust interaction of race and subsequent accuracy on mean local efficiency [$p_{\text{uncorrected}} = .0028$, $p_{\text{adjusted}} = .0224$, $z = 3.088$, $\beta = .0282$, $SE_{\text{robust}} = .009$, $CI = [0.010, 0.046]$]. (B) A post-hoc analysis found a main effect of race on global efficiency [$z = 2.067$, $\beta = 0.0036$, $SE_{\text{robust}} = .002$, $p = 0.039$, $CI = [0.000, 0.007]$].

Thus far, all data reported have been whole-brain network level metrics. Recall, to calculate local efficiency node-level efficiencies were averaged together for the entire network.

Because each node has its own local efficiency, we additionally tested whether the observed pattern of greater network segregation during accurate face encoding of other-race faces was uniform across discrete intrinsic connectivity networks of the brain, or whether any one network particularly drove the overall higher local efficiency. Research has shown that at rest, the brain demonstrates a segregated functional connectivity that results in separate modules¹⁴⁴. The integrity of these modules, often called intrinsic connectivity networks (ICNs), are associated with varying cognitive functions such as

visual perception, bottom up and top down attention, cognitive control, etc¹³⁴. It was therefore of interest to test whether networks associated with certain perceptual and attentional capacities might be differentially engaged for same and other-race encoding.

To test this each node was assigned to an ICN using a preexisting mapping between the Brainnetome atlas and Yeo ICN parcellations^{134,167}. The ICNs included in this parcellation are the Dorsal Attention Network (DAN), Default Mode Network (DMN), Frontoparietal Network (FPN), Limbic Network, Somatomotor (SOM) Network, Ventral Attention Network (VAN), and Visual Network. In the mapping, 207 out of the full sample of 243 nodes corresponded to one of the 7 ICNs. The remaining 36 nodes were removed from the analysis, 33 of which were subcortical and not included in the Yeo Parcellation. The final 3 nodes that were removed corresponded to portions of the cingulate gyrus and insula. The final number of nodes per ICN are described in Table 6.3.

Table 6. 3 Descriptive Statistics for Local Efficiency Across Intrinsic Connectivity Networks

Intrinsic Connectivity Network (ICN)	Number of Nodes	Mean Local Efficiency	Standard deviation local efficiency
Dorsal Attention Network (DAN)	30	0.5019	0.02713
Default Mode Network (DMN)	36	0.504	0.02588
Frontoparietal Network (FPN)	26	0.5085	0.02086
Limbic Network	26	0.4878	0.02047
Somatomotor (SOM)	33	0.5129	0.03139
Ventral Attention Network (VAN)	22	0.5056	0.03261
Visual Network	34	0.4979	0.02405

Of the 246 nodes in the Brainnetome brain networks, 207 mapped onto known intrinsic connectivity networks (ICNs) from the Yeo parcellation. The number of nodes included in each network is shown above, as well as the mean and standard deviation of local efficiency within each network.

Next, ICN-level mean local efficiencies were computed for each subject, by averaging together the local efficiency of nodes within each ICN across the four study lure conditions.

It should be noted that mean local efficiency of an ICN reflects the mean efficiency with

which its nodes' direct neighbors are coupled, independent of the neighbors' ICN affiliation. That is, mean efficiency of an ICN should be interpreted as the overall fault tolerance of modules containing *at least one node* from the given ICN. This calculation does *not* comment on the efficiency with which nodes within a shared ICN are coupled. Mean efficiency was calculated for each separate condition, plotted in Figures 6.4 and 6.5 to emphasize differences across the race and accuracy factors.

Figure 6.4a highlights that the brain-wide increase in local efficiency during accurate encoding of other-race relative to same-race faces (see Fig. 6.3a) was pervasive across intrinsic connectivity networks. Meanwhile, efficiency during inaccurate encoding (Fig. 6.4b) across networks is not as consistently different between same and other-race conditions. The previously observed cross-over interaction between race and accuracy is reproduced in a 3-way repeated measures ANOVA with factors of race, accuracy, and ICN. There are both main effects of ICN [$F(6,588) = 3.33, p = .0031, \eta^2 = .03$] and an interaction between race and accuracy [$F(1,588) = 17.56, p < .0001, \eta^2 = .03$] but no three-way interaction [$F(6,588) = 0.3758, p = .89, \eta^2 = 0$].

We next tested whether specific subnetworks were driving whole-network differences, since certain ICNs display non-overlapping standard errors of means across the race conditions (Fig 6.4a,b). To do this, two-way repeated measures ANOVAS with factors of race and ICN were applied separately on the correct rejection and false alarm conditions. Post-hoc Sidak corrections were then applied to test which subnetworks showed the greatest deviation in local efficiency across race. For subsequent correct rejections there were main effects of race [$F(1,21) = 7.72, p = .011, \eta^2 = .05$] and ICN [$F(6,126) = 2.92, p$

=.011, $\eta^2 = .04$]. After correcting for multiple comparisons across the eight networks, the difference between SR and OR efficiency remained significant in the DAN [SR $\bar{x} = .50$, OR $\bar{x} = .52$, $p = .04$, 95% CI = [-.05 to 0]] and VAN [SR $\bar{x} = .49$, OR $\bar{x} = .53$, $p = 0.005$, 95% CI = [-0.06 to -0.007]]. For the ANOVA applied to the subsequent false alarm condition, there was no effect of race, but the effect of ICN remained [[F(6,126) = 2.43, $p = .03$, $\eta^2 = .03$]. Despite no main effect of race, the FPN demonstrated a significant difference across race after Sidak correction [SR $\bar{x} = .52$, OR $\bar{x} = .49$, $p = 0.02$, 95% CI = [0 to 0.05]]. The overall lack of effect of race in the inaccurate condition suggests that the differences observed in local efficiency across race in the accurate condition may be specifically important to promote successful same and other-race memory.

Local Efficiency of Intrinsic Connectivity Networks During Encoding Leading to Subsequent Correct Rejections and False Alarms

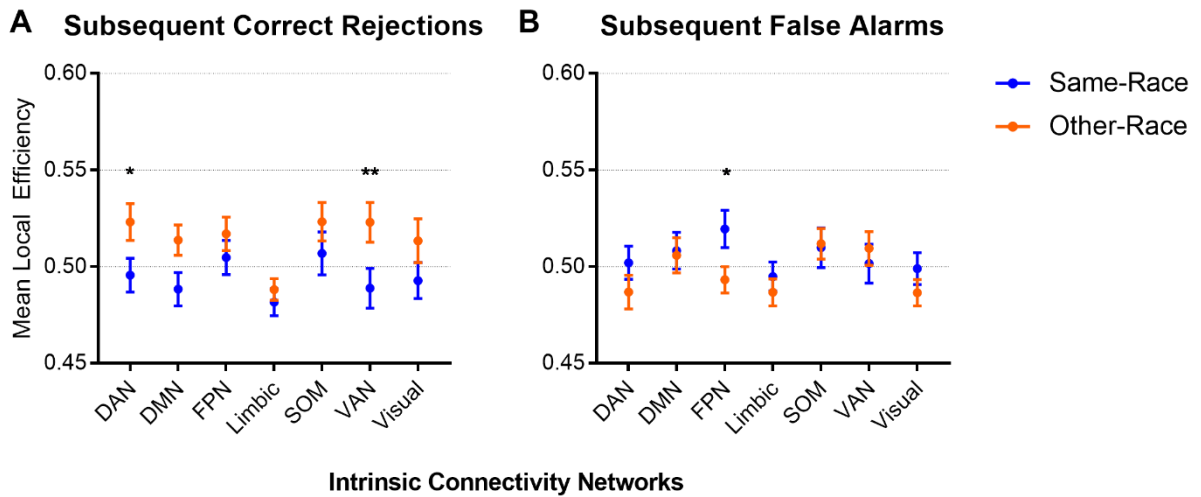


Figure 6.4 Race differences in local efficiency across networks during accurate and inaccurate encoding

A) Across intrinsic networks (ICNs), encoding promoting subsequent correct rejections was associated with greater local efficiency for other-race faces. After an ANOVA and multiple comparison correction, DAN and VAN efficiency remain statistically different across race. B) There is less consistency in encoding promoting subsequent false alarms. Same-race encoding tends to, but does not always demonstrate higher local efficiency. After multiple comparison correction the FPN efficiency remains significantly different across race. Abbreviations: DAN - Dorsal Attention Network; DMN - Default Mode Network; FPN - Frontoparietal Network; SOM - Somatomotor (SOM) Network; VAN - Ventral Attention Network.

We additionally tested whether specific ICNs drove differences within race across accuracy conditions. Figure 6.5 displays local efficiency across accurate and inaccurate encoding conditions for both same -race (Fig. 6.5a), and other-race (Fig. 6.5b) faces. Because the standard error of mean bars overlap for all networks in the same-race condition, no follow-up tests were run. Because network efficiency across accuracy appears more distinct in the other-race condition, a two-way repeated measures ANOVA was applied with factors of accuracy and ICN followed by a post-hoc Sidak correction for multiple comparisons. This resulted in main effects of accuracy [$F(1,21) = 8.21, p = .009, \eta^2 = .04$] and ICN [$F(6,126) = 3.09, p = .007, \eta^2 = .05$]. In addition differences in efficiency across other-race accuracy

conditions survived for the DAN [CR \bar{x} = .52, FA \bar{x} = .49, p = .002, 95 % CI = [0.01 to 0.06]] and the Visual network [CR \bar{x} = .51, FA \bar{x} = .49, p = .045, 95 % CI = [0 to 0.05]].

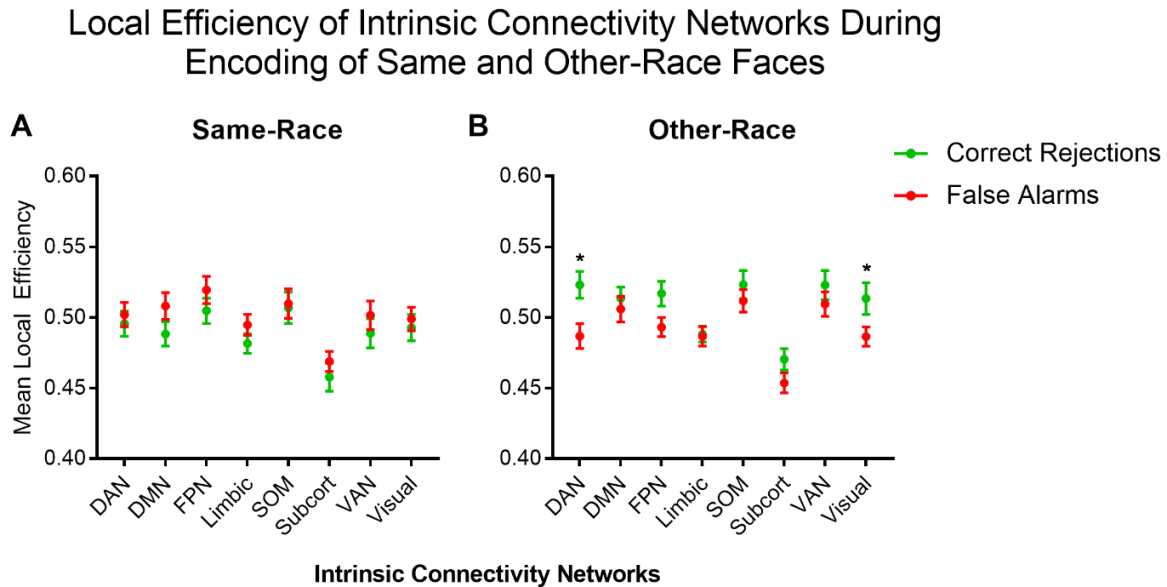


Figure 6.5 Accuracy differences in local efficiency across networks during same and other-race face encoding.

(A) Across intrinsic connectivity networks (ICNs), there appears little difference in local efficiencies during accurate and inaccurate encoding of same-race faces. (B) Meanwhile there is an effect of accuracy on local efficiency in the other-race condition, with higher efficiency for accurate trials. After multiple comparison correction the DAN and Visual networks maintain significantly greater efficiencies for the correct rejection condition. Abbreviations: DAN - Dorsal Attention Network; DMN - Default Mode Network; FPN - Frontoparietal Network; SOM -Somatomotor (SOM) Network; VAN - Ventral Attention Network.

Considering whole-brain and ICN-level efficiencies together, the result of this study suggests that functional reorganization of the brain during encoding influences subsequent capacity to correctly reject face lures. Distinct brain network topologies supported behavior relative to the race of presented faces. Successful same-race encoding was related to a more integrated network infrastructure, while accurate other-race encoding was associated with more network segregation. Greater segregation during successful other-race than same-race encoding was widely distributed across sub-networks of the brain.

Furthermore, local efficiency during encoding across subnetworks provided an accuracy advantage in other-race faces, but not same-race faces.

Discussion

The present study provides the first demonstration that successful same-race and other-race face recognition are supported by different functional brain network topologies. Connectivity in the brain during encoding was greatly modulated by the race of presented faces. This suggests that the context of race alters the behavioral adaptivity of network configurations. Other-race face encoding promoting correct rejections was subserved by a more locally efficient network, while reduced local efficiency was associated with false identifications. In contrast, same-race recognition was associated with the opposite pattern: Reduced local efficiency during encoding supported later correct rejections, while greater local efficiency was related to false alarms. The connection between same-race accuracy and reduced segregation was further underscored by greater global efficiency during encoding of same-race relative to other-race faces. Together these data suggest that a more functionally integrated network is important for same-race correct rejections, while a more segregated network is optimal for other-race correct rejections.

These findings suggest that there is not a one-size fits all network topology for face recognition. Across study and retrieval conditions tested, we found four of eight conditions demonstrated interactions between race and accuracy, though only the study lure condition survived post-hoc correction. This pattern suggests that accuracy in a complex memory task is not unilaterally associated with a more integrated network infrastructure,

in contrast to some past graph theoretical findings¹⁴⁴⁻¹⁴⁶. Another study supporting this conclusion found that a working memory task was initially associated with integrated brain networks but with longitudinal training improvements in performance were related to more modular network configurations¹⁴⁸. These findings suggest that integrated networks may be important for novel tasks, but that segregation may support learning over time.

Reconciling our findings with perceptual expertise accounts of the ORE, we might expect a more specialized and segregated network during same-race face processing afforded by years of tuned expertise with this category of faces. But instead, we saw a more globally integrated network. While this seems counterintuitive, it is still likely that the face network itself is more segregated from the brain as shown in developmental work¹⁵⁴, but that as a larger system overall the brain functionally organizes in a way to promote more 'bridges' between functionally distinct regions during same-race encoding. This could reflect more ease with which information about same-race faces can be propagated across the brain.

In addition to analyzing whole-brain efficiency metrics, we found that specific intrinsic connectivity networks drove network-wide local efficiency differences during same and other-race face encoding. The ventral attention network (VAN) showed the largest increase in local efficiency for successful other-race relative to same-race face encoding. However, all subnetworks demonstrated higher local efficiency to other-race faces, suggesting that widespread reorganization promoting network redundancy may be especially important for other-race encoding. Furthermore, visual and dorsal attention networks (DAN) were more locally efficient in accurate than inaccurate other-race encoding, while local efficiency

across networks during same-race encoding seemed to have less bearing on overall performance.

It is validating that the subnetwork-level analysis implies that network redundancy in visual, top-down (DAN) and bottom up (VAN) attentional networks are specifically important during encoding of other-race faces. Visual regions have been most greatly explored in the context of the ORE (Fusiform Face Area, Occipital Face Area), however nodes within the VAN and DAN are also included in the core and extended face processing networks. For instance, the posterior superior temporal sulcus in the VAN is considered a core face region involved in directing attention during face encoding, and has demonstrated differential activity to same and other-race faces¹⁷⁶. Other VAN regions like the insula and cingulate gyrus are involved in salient emotional face processing^{177,178}. Furthermore, the VAN is globally involved in stimulus-driven attentional control and is likely recruited to orient attention divergently across the same-race and other-race conditions. Meanwhile DAN nodes like the intraparietal sulcus and the middle temporal gyrus are implicated in goal-directed eye gaze and familiar face processing respectively¹⁷⁶. Furthermore, the DAN is involved in biasing responses from lower-level visual regions. These roles are vital for face recognition, so it is possible that specialized and redundantly organized modules may form between nodes of these networks, enabling more successful encoding of other-race faces.

It is interesting that there were no major differences in local efficiency across networks associated with accurate and inaccurate same-race encoding. It is possible that perceptual expertise reduces the need to encode information as redundantly, since a more integrated

network could compensate for reduced local fault tolerance. This integration could involve highly central nodes/ hubs which could computationally lead to reduced local efficiency due to an inability for networks to be resilient to their removal. However, the extent to which several hubs could influence mean efficiencies in entire sub-networks is unclear. A separate analysis would be needed to test whether certain hubs exist during same-race but not other-race recognition, and whether this results in reduced local efficiency during same-race face processing.

The intrinsic connectivity level results found here are bolstered by findings from a study testing the importance of top-down attentional and control networks in shaping the ORE⁷⁵. Differences in activity in the intraparietal sulcus (associated with DAN and Fronto-parietal cognitive control networks) were particularly related to failures to remember other-race faces. Furthermore, functional connectivity between the right fusiform cortex and intraparietal sulcus regions in the DAN were significantly greater for same-race than other-race faces. It is important to note the study design, analyses and results diverge from ours in that success was defined as target recognition (as opposed to lure correct rejection), and analyses were not explored from a graph theoretical perspective. Still, both studies find that there is altered behaviorally relevant recruitment of the DAN during encoding of same and other-race faces. Our more exploratory study contributes the novel observation that altered topology in the VAN and visual intrinsic connectivity networks may also play a role in the ORE. Future studies may be designed to focus on whether the interactions (or lack-there-of) of these networks might relate to the emergence of same-race and other-race recognition disparities.

It is important to note that the main findings of this study are specific to memory encoding that supported successful correct rejections relative to false alarms at retrieval. While our analysis produced three other trending interactions, we did not explore their directionality and can therefore not comment on whether they took on the same pattern of results. It is possible that same-race recognition is always supported by integrated networks and other-race by segregated networks during encoding. It is also possible that task demands at retrieval modulate the most optimal network configuration during encoding, such that successful memorization of targets could be altered by race quite differently than found in the lure condition. Teasing apart the relative contributions of functional connectivity to the ORE, given the competing demands of recognition and mnemonic discrimination will require additional examination.

In hindsight, future network neuroscientific investigations of the ORE will benefit from the following considerations. While the simultaneous inclusion of encoding, retrieval, target, and lure conditions was novel it greatly reduced our power to detect significant effects given the need for stringent post-hoc corrections. Subsequent studies should consider more targeted hypotheses and experimental designs. In addition, the large number of conditions increased the length of this study, likely contributing to it being overly difficult for participants. Despite low performance, a highly conservative analysis found that differences in brain network topologies across encoding lure contexts were likely not due to chance. Still, future studies can be improved by shortening the task design and increasing physiognomic differences between faces in lure pairs. Furthermore, while there are a variety of ways to analyze task-based connectivity in the brain, no large systematic

study has looked at differences in resulting graph topologies based on choice of functional connectivity method. Because graph theoretical analysis may be sensitive to selection of brain atlas, preprocessing methods, and functional connectivity analysis, it is important to study the network-based correlates of the ORE across a variety of paradigms and methods to ensure that the findings reported here are reliable and reproducible.

In conclusion, participants performed a mnemonic discrimination paradigm in the MRI scanner, allowing us to evaluate the role of face recognition memory processes on the functional network organization of the brain. Our findings suggest that same-race and other-race face encoding are supported by distinct network topologies. The ability to not mistakenly recognize faces at retrieval was supported by redundantly organized modules during encoding of other-race faces, and a more integrated network during encoding of same-race faces. The relatively greater importance of network segregation for other-race faces was supported by greater efficiency in visual, bottom-up, and top-down attentional networks during accurate compared to inefficient encoding. As the first graph theoretical analysis specific to the ORE, these results demonstrate that same and other-race face recognition differentially tap into multiple networks, instigating largescale whole-brain differences in functional architecture. These findings should serve to motivate research beyond face-preferential regions to understand how system-wide connectivity differences promote the ORE.

Conclusion Summary

Within the present dissertation, factors in both the behavioral Other-Race Effect (ORE) and neural correlates underlying same and other-race face processing were examined.

Chapters 1 and 2 present a comprehensive literature review of theoretical frameworks and neural research on the ORE. Both historical as well as current explanatory models were discussed, highlighting the importance of developmental experience and social-attentional modulation in the emergence of the effect. While the ORE is defined as a recognition memory deficit for other-race faces that may be modulated by attentional and motivational contexts, studies into the neural correlates have focused on a select few visual processing regions that cannot account for its complexity. Furthermore, many of these studies do not employ recognition paradigms, and can therefore only stipulate how perceptual correlates of same and other-race face processing impact subsequent memory. For the few studies that have evaluated memory, results indicate involvement of traditional mnemonic⁶⁷ and attentional processing⁷⁵ regions in addition to face-preferential regions.

Motivated by the lack of a focus on memory in the literature, we developed a theoretical framework that positions the ORE as a deficit in mnemonic interference resolution for other-race faces (chapter 3). This was influenced by a large body of computational and neurobiological work studying the conditions under which similar objects are successfully discriminated from one another in memory⁷⁸⁻⁹¹. This so called 'mnemonic discrimination' is believed to rely upon a pattern separation mechanism in the hippocampus and some surrounding medial-temporal cortices⁹²⁻⁹⁹. A failure to mnemonically discriminate

between similar objects results in false alarms. Because one of the hallmark indices of the ORE is an increased ratio of false alarms for other-race (relative to same-race) faces, it is plausible that a reduction in successful pattern separation is partially to blame. Such a failure to disentangle other-race neural representations from one another could underly the phenomenological experience of ‘all other-race faces looking alike.’

To test this, we adapted a traditional mnemonic discrimination paradigm for use with face instead of object stimuli. Participants were asked to study a series of faces, followed by a test phase where they responded whether a separate set of faces had been seen before. Lure distractor faces in the Test Phase were varied in similarity to originally presented ‘parent’ faces, allowing us to characterize discrimination ability as a function of mnemonic interference. The results demonstrated a mnemonic discrimination deficit for other-race faces that persisted for all but the lowest interference trials. To counter claims that this deficit could be accounted for by deficits in initial perceptual encoding we employed a control procedure requiring participants to maintain faces in working memory for several seconds before identifying whether a second presentation was the same or a new face. The findings confirmed no performance differences for same and other-race faces, suggesting that the ORE is not accounted for simply by differences in visual encoding. It is possible that behaviorally irrelevant differences at perceptual encoding may be temporally exacerbated during consolidation and retrieval processes. Taken together these findings suggest that other-race faces do not all *look* alike, but rather are *remembered* alike.

To confirm whether pattern separation-consistent mechanisms are involved in the mnemonic homogenization of other-race faces, we ran this task on a separate group of

subjects in the MRI scanner. For methodological reasons discussed in chapter 4, the behavioral results were not optimal to test the pattern-separation hypothesis. We therefore provided recommendations for improving future experimental paradigms to examine this question more successfully in the future. Fortunately, the inclusion of lure pairs in our study is still new to the ORE field, allowing us to make novel contributions even with an analysis where the interference levels were collapsed together.

Thereafter our approach to this data evolved greatly, shifting focus from very specific regions in the medial temporal lobe to the entirety of the brain. Based on theories that the ORE results from greater optimization of processing for same-race faces we tested the hypothesis that the brain operates less efficiently when processing other-race information. This hypothesis was motivated by literature in the burgeoning field of network neuroscience, reviewed in chapter 5.

To test this, we employed graph theoretical analysis to construct graph representations of the entire brain during performance of each face condition, where regions demonstrating strong functional connectivity were connected by edges (see chapter 6). For each graph, we calculated metrics characterizing both global and local efficiency in how these edges were distributed during each context. Our major findings were specific to encoding of faces that were paired with similar looking distractors in the test phase. Same-race and other-race face memorization were supported by entirely inverted network configurations; Successful same-race encoding was characterized by lower local efficiency and higher global efficiency. Successful other-race encoding in contrast was supported by higher local efficiency and lower global efficiency. This implies that the network topology optimal for

encoding of one race, was detrimental to encoding of the other. Successful mnemonic discrimination seemingly required rapid network reconfigurations, with greater integration for same-race faces and greater segregation for other-race faces.

We additionally found that the segregated functional architecture supporting other-race mnemonic discrimination, while distributed across subnetworks, was driven most by visual, top-down, and bottom-up attentional processing networks. This suggests that redundant connectivity in modules involving visual and attentional processing regions might be integral for other-race face recognition. Perhaps these regions were engaged during more effortful encoding of other race faces by participants.

Future Neuroimaging Research

The research included in this dissertation poses a variety of further questions worth exploring. As mentioned, the pattern-separation account of the ORE remains to be tested. But perhaps even more interesting would be an integration of pattern separation with the face-space model (further explored in chapter 3's discussion). While computational and experimental evidence supports more elaborated encoding of same-race faces and more clustered encoding of other-race faces⁴³⁻⁴⁸, there have been no mechanisms proposed to account for these diverging spatial distributions. Because pattern separation^{78,80-99} is involved in disentangling representations, it is therefore plausible that it could promote the more orthogonalized and spatially distinct same-race representations and more overlapping other-race representations predicted by face-space. Furthermore, face-space models explain excessive false alarms for other-race faces based on the co-activation of

overlapping representations. Certainly, this may be reconciled with computational work on pattern completion and its role in generalization. Representational similarity analysis and multi-dimensional scaling would be powerful tools in testing these predictions neurologically.

Our graph theoretical approach is the second study to relate differential recruitment of attentional networks to the ORE⁷⁵. This occurred despite quite different methods employed in the two studies and helps to validate our use of graph theoretical methods. The interaction between visual, memory, bottom-up and top-down attentional networks should be further explored to flesh out the neural correlates underlying this effect.

Additionally, the cross-over interaction (i.e. mirror image) results we found on the behavioral adaptiveness of oppositional topologies suggests that an ability to rapidly reconfigure between integrated and segregated networks is associated with a reduction in the ORE. There is evidence that dynamic reconfiguration of networks – i.e., network flexibility-- supports learning¹⁷⁹. To test this dynamic hypothesis, single trial analysis could be performed to examine whether getting ‘stuck’ in a configuration optimal for recognition of one race is maladaptive for another. Then, relative flexibility across subjects could be correlated with behavioral metrics of the ORE. Because face recognition performance tends to be quite variable, it is possible that individuals who recognize both same and other-race faces well can rapidly adjust network topology, reflecting an ability to shift strategies to meet the demands of different race contexts.

Indexing Improved Outcomes in Training Paradigms

In addition to future imaging studies, the current study can inform research focused on developing training paradigms to mitigate the ORE. Numerous studies have found that training can reduce the ORE—at least temporarily^{180–182}. First, longitudinal paradigms are necessary to determine whether training results in lasting change. There is further room for improvement in outcomes assessments. The preexisting paradigms have used traditional performance measures that are less able to capture patterns of improvement over time. The mnemonic discrimination task offers a suitable alternative paradigm by evaluating performance across a range of stimulus difficulties, allowing observation of a more complex underlying structure of the ORE. In practice, a person's baseline ORE may be established using the other-race transfer function. Across training, a shift in the other-race function towards the same-race function would indicate improved recognition. This could be summarized by taking the difference in same-race and other-race areas under the curve (AUC). If training is robust, participants should demonstrate reduced and stable AUC differences across periods without practice.

There are myriad applications for facial recognition training paradigms and appropriate evaluation metrics. The ORE arises as young as infancy and just like language, is subject to sensitive learning periods^{26,183–185}. Exposing children to diverse faces as early as possible in their environment or schooling could lead to reduction or elimination of the ORE. In the absence of such experience, it is possible that training paradigms may be regularly employed to reduce the impact of the ORE, especially in situations where the ORE can result in severely negative consequences, e.g. law enforcement.

Mnemonic Discrimination Relevance to the Criminal Justice System

I was initially motivated to study the ORE because of its negative societal impacts. Those of us who study it hope that understanding and publicizing its mechanisms will ultimately influence eyewitness policies in the criminal justice system. Up until now most studies of the ORE have focused on disparities between correctly *remembering* same-race and other-race faces^{16,30-34,40,41,53,54,60}. Within eyewitness procedures a failure to remember a face may lead to a perpetrator being overlooked. However, an arguably greater consequence is a wrongful conviction, which instead occurs when witnesses mistakenly recognize a face as belonging to a perpetrator. In other words, they fail to *mnemonically discriminate* (or correctly reject) the face. Mnemonic discrimination studies therefore are more focused than traditional recognition studies on the very behavior that the justice system should be trying to mitigate. That is, if we are to believe that the system still upholds the Blackstone Formulation¹⁸⁶, a cornerstone of the American Legal system that “it is better that ten guilty persons escape than that one innocent suffers.”

A 2014 report from the National Academies on the fallibility of eye-witness testimony called for the establishment of procedures to estimate the severity of the ORE in individual eyewitnesses following identification of a cross-race person. A mnemonic discrimination task could be well-suited to the task, specifically given the inclusion of multiple interference levels allowing for a more nuanced picture of the recognition deficit as a function of facial similarity. Still, no test alone should be relied upon. Instead, a battery of procedures could be administered drawing from the corpus of existing experimental paradigms in the ORE literature. This would be quite an undertaking, requiring designing

and optimizing tasks for multiple races, validating test-retest reliabilities, a creation of a global scoring system, etc. Furthermore, there are ethical considerations of administering such tasks to victims, and it is possible that the stress of proving one's own validity could call into question testimony that is otherwise accurate, and negatively impact the mental health of witnesses. While great care and further research would be necessary in implementing such policies, they could be vital in preventing wrongful incarcerations in the future.

...

In conclusion, the research included in this dissertation evaluates the role of memory processes in the behavioral and neural ORE. Other-race recognition was shown to be disproportionately affected by demands to resolve mnemonic interference, suggesting a potential role of altered pattern-separation in shaping the ORE. Furthermore, brain network topology during encoding of faces promoting successful mnemonic discrimination was modulated by race, suggesting differential functional connectivity profiles are adaptive for same and other-race face processing. Together these findings should motivate more investigations of the crucial role of mnemonic discrimination in giving rise to the ORE.

Abbreviations

ANOVA	analysis of variance
CR	correct rejection
DAN	Dorsal Attention Network
DMN	Default Mode Network
FA	False alarm
FFA	Fusiform face area
fMRI	Functional magnetic resonance imaging
FPN	Fronto-Parietal Network
gPPI	Generalized psychophysiological interaction analysis
ICN	Intrinsic Connectivity Network
LCR	Lure correct rejection
LFA	Lure false alarm
MTL	Medial temporal lobe
OFA	Occipital face area
OR	Other-race
ORE	Other-race effect
ROI	Region of interest
SOM	Somatomotor Network
SR	Same-race
TH	Target hit
TM	Target miss
VAN	Ventral Attention Network

References

1. Feingold GA. The Influence of Environment on Identification of Persons and Things. *J Am Inst Crim Law Criminol.* 1914;5(1):39-51. doi:10.2307/1133283.
2. Hugenberg K, Young SG, Bernstein MJ, Sacco DF. The categorization-individuation model: An integrative account of the other-race recognition deficit. *Psychol Rev.* 2010;117(4):1168-1187. doi:http://dx.doi.org/10.1037/a0020463.
3. Meissner CA, Brigham JC. Thirty years of investigating the own-race bias in memory for faces: a meta-analytic review. *Psychol Public Policy, Law.* 2001;7(1):3-35. doi:10.1037/1076-8971.7.1.3.
4. McKone E, Dawel A, Robbins RA, Shou Y, Chen N, Crookes K. Why the other-race effect matters: Poor recognition of other-race faces impacts everyday social interactions. *Br J Psychol.* 2021:1-23. doi:10.1111/bjop.12508.
5. Brooks SK, Greenberg N. Psychological impact of being wrongfully accused of criminal offences: A systematic literature review. *Med Sci Law.* 2021;61(1):44-54. doi:10.1177/0025802420949069.
6. White D, Kemp RI, Jenkins R, Matheson M, Burton AM. Passport officers' errors in face matching. *PLoS One.* 2014;9(8). doi:10.1371/journal.pone.0103510.
7. Sporer SL, Gehrke J, McClure KA. Patrolling the Borders: Experience with Multiple Out-Group Members is Associated with the Own-Ethnicity Effect. *J Police Crim Psychol.* 2020. doi:10.1007/s11896-020-09387-x.
8. Robertson DJ, Black J, Chamberlain B, Megreya AM, Davis JP. Super-Recognisers show an advantage for other race face identification. *Appl Cogn Psychol.* 2020;34(1):205-216. doi:10.1002/acp.3608.
9. DNA Exonerations in the United States. The Innocence Project. <https://innocenceproject.org/dna-exonerations-in-the-united-states/>. Published 2021. Accessed June 8, 2021.
10. Exonerations by contributing Factor and Crime. The National Registry of Exonerations. <https://www.law.umich.edu/special/exoneration/Pages/ExonerationsContribFactorsByCrime.aspx>. Published 2021. Accessed June 8, 2021.
11. Race and Wrongful Convictions. The National Registry of Exonerations. <https://www.law.umich.edu/special/exoneration/Pages/Race-and-Wrongful-Convictions.aspx>. Published 2021. Accessed June 8, 2021.
12. Irazola S, Williamson E, Stricker J, Niedzwiecki E. Final Report: Study of Victim Experiences of Wrongful Conviction. *Rep Prep Natl Inst Justice, US Dep Justice.* 2013:1-71.
13. Conroy J, Warden R. Special Investigation: The High Costs of Wrongful Convictions.

- Cent Wrongful Convictions, Better Gov Assoc.* 2011.
<https://www.bettergov.org/news/special-investigation-the-high-costs-of-wrongful-convictions>.
14. Silbert R, Hollway J, Larizadeh D. Criminal (In)justice: A Cost Analysis of Wrongful Convictions, Errors, and Failed Prosecutions in California's Criminal Justice System. *Chief Justice Earl Warren Inst Law Soc Policy, Berkeley Sch Law*. 2015.
 15. Identifying the Culprit: Assessing Eyewitness Identification. *Comm Sci Approaches to Underst Maximizing Validity Reliab Eyewitness Identif Law Enforc Court Natl Res Counc Natl Acad*. 2014. <http://www.nap.edu/catalog/18891/identifying-the-culprit-assessing-eyewitness-identification>.
 16. Meissner CA, Brigham JC. Thirty Years of Investigating the Own-Race Bias in Memory for Faces: A Meta-Analytic Review. *Psychol Public Policy, Law*. 2001;7(1):3-35. doi:10.1037//1076-8971.7.1.3.
 17. Goldstein AG. Race-related variation of facial features: Anthropometric data I. *Bull Psychon Soc*. 1979;13(3):187-190. doi:10.3758/BF03335055.
 18. Saini A. *Superior: The Return of Race Science.*; 2020.
 19. Goldstein AG. Facial feature variation: Anthropometric data II. *Bull Psychon Soc*. 1979;13(3):191-193. doi:10.3758/BF03335056.
 20. Chance J, Goldstein AG, McBride L, McBride L. Differential experience and recognition memory for faces. *J Soc Psychol*. 1975;97(2):243-253. doi:10.1080/00224545.1975.9923344.
 21. Shepherd JW, Deregowski JB, Ellis HD. a Cross-Cultural Study of Recognition Memory for Faces. *Int J Psychol*. 1974;9(3):205-212. doi:10.1080/00207597408247104.
 22. Malpass, R. S. & Kravitz J. Recognition of Faces for Own and Other Race.
 23. Young SG, Hugenberg K, Bernstein MJ, Sacco DF. Perception and Motivation in Face Recognition A Critical Review of Theories of the Cross-Race Effect. *Personal Soc Psychol Rev*. 2012;16(2):116-142. doi:10.1177/1088868311418987.
 24. Watson TL, Robbins RA, Best CT. Infant perceptual development for faces and spoken words: An integrated approach. *Dev Psychobiol*. 2014;56(7):1454-1481. doi:10.1002/dev.21243.
 25. Kelly D, Quinn P, Slater A. Three-month-olds, but not newborns, prefer own-race faces. *Dev Sci*. 2005;6(8):31-36. <http://onlinelibrary.wiley.com/doi/10.1111/j.1467-7687.2005.0434a.x/full>.
 26. Kelly DJ, Quinn PC, Slater AM, Lee K, Ge L, Pascalis O. The Other-Race Effect Develops During Infancy Evidence of Perceptual Narrowing. *Psychol Sci*. 2007;18(12):1084-1089. doi:10.1111/j.1467-9280.2007.02029.x.
 27. Scott LS, Pascalis O, Nelson CA. A domain-general theory of the development of perceptual discrimination. *Curr Dir Psychol Sci*. 2007;16(4):197-201.

- doi:10.1111/j.1467-8721.2007.00503.x.
28. Sangrigoli S, De Schonen S. Recognition of own-race and other-race faces by three-month-old infants. *J Child Psychol Psychiatry*. 2004;45(7):1219-1227. doi:10.1111/j.1469-7610.2004.00319.x.
 29. Heron-Delaney M, Anzures G, Herbert JS, et al. Perceptual Training Prevents the Emergence of the Other Race Effect during Infancy. *PLoS One*. 2011;6(5). doi:10.1371/journal.pone.0019858.
 30. Sangrigoli S, Pallier C, Argenti A-M, Ventureyra VAG, De Schonen ' S. Reversibility of the Other-Race Effect in Face Recognition During Childhood. *Psychol Sci*. 2005;16(6):440-444. doi:10.1111/j.0956-7976.2005.01554.x.
 31. Rhodes G, Ewing L, Hayward WG, Maurer D, Mondloch CJ, Tanaka JW. Contact and other-race effects in configural and component processing of faces. *Br J Psychol*. 2009;100(4):717-728. doi:10.1348/000712608X396503.
 32. Tanaka JW, Pierce LJ. The neural plasticity of other-race face recognition. *Cogn Affect Behav Neurosci*. 9(1):122-131. doi:10.3758/CABN.9.1.122.
 33. Lebrecht S, Pierce LJ, Tarr MJ, Tanaka JW. Perceptual Other-Race Training Reduces Implicit Racial Bias. *PLoS One*. 2009;4(1):e4215. doi:10.1371/journal.pone.0004215.
 34. McGugin RW, Tanaka JW, Lebrecht S, Tarr MJ, Gauthier I. Race-Specific Perceptual Discrimination Improvement Following Short Individuation Training With Faces. *Cogn Sci*. 2011;35(2):330-347. doi:10.1111/j.1551-6709.2010.01148.x.
 35. Gauthier I, Nelson CA. The development of face expertise. *Curr Opin Neurobiol*. 2001;11(2):219-224. doi:10.1016/S0959-4388(00)00200-2.
 36. Diamond R, Carey S. Why Faces Are and Are Not Special: An Effect of Expertise. *J or Exp Psychol Gen*. 1986;115(2):107-117.
 37. Yin RK. Looking at upside-down faces. *J Exp Psychol*. 1969;81(1):141-145. doi:10.1037/h0027474.
 38. Rhodes G, Brake S, Taylor K, Tan S. Expertise and configural coding in face recognition. *Br J Psychol*. 1989;80(3):313-331. doi:10.1111/j.2044-8295.1989.tb02323.x.
 39. Tanaka JW, Kiefer M, Bukach CM. A holistic account of the own-race effect in face recognition: evidence from a cross-cultural study. *Cognition*. 2004;93(1):B1-B9. doi:10.1016/j.cognition.2003.09.011.
 40. Michel C, Caldara R, Rossion B. Same-race faces are perceived more holistically than other-race faces. *Vis cogn*. 2006;14(1):55-73. doi:10.1080/13506280500158761.
 41. DeGutis J, Mercado RJ, Wilmer J, Rosenblatt A. Individual Differences in Holistic Processing Predict the Own-Race Advantage in Recognition Memory. *PLoS One*. 2013;8(4):e58253.

42. Hancock KJ, Rhodes G. Contact, configural coding and the other-race effect in face recognition. *Br J Psychol.* 2008;99(1):45-56. doi:10.1348/000712607X199981.
43. Valentine T. A Unified Account of the Effects of Distinctiveness, Inversion, and Race in Face Recognition. *Q J Exp Psychol Sect A.* 1991;43(2):161-204. doi:10.1080/14640749108400966.
44. Valentine T, B. LM, Hills PJ. Face-space: A unifying concept in face recognition research. *Respir Physiol Neurobiol.* 2004;143(1):21-35. doi:10.1080/17470218.2014.990392.
45. O'Toole AJ, Deffenbacher K, Abdi H, Bartlett JC. Simulating the "Other-race Effect" as a Problem in Perceptual Learning. *Conn Sci.* 1991;3(2):163-178. doi:10.1080/09540099108946583.
46. Caldara R, Abdi H. Simulating the "other-race" effect with autoassociative neural networks: Further evidence in favor of the face-space model. *Perception.* 2006;35(5):659-670. doi:10.1068/p5360.
47. Byatt G, Rhodes G. Identification of own-race and other-race faces: implications for the representation of race in face space. *Psychon Bull Rev.* 2004;11(4):735-741. doi:10.3758/BF03196628.
48. Papesh, Megan and Goldinger S. A multidimensional scaling analysis of own- and cross-race face spaces. 2011;116(2):283-288. doi:10.1016/j.cognition.2010.05.001.A.
49. Malpass, R. S. & Kravitz J. Recognition of Faces for Own and Other Race. 1969.
50. Bernstein MJ, Young SG, Hugenberg K. The Cross-Category Effect Mere Social Categorization Is Sufficient to Elicit an Own-Group Bias in Face Recognition. *Psychol Sci.* 2007;18(8):706-712. doi:10.1111/j.1467-9280.2007.01964.x.
51. Hugenberg K, Corneille O. Holistic processing is tuned for in-group faces. *Cogn Sci.* 2009;33(6):1173-1181. doi:10.1111/j.1551-6709.2009.01048.x.
52. Michel C, Corneille O, Rossion B. Race categorization modulates holistic face encoding. *Cogn Sci.* 2007;31(5):911-924. doi:10.1167/6.6.435.
53. Hehman E, Mania EW, Gaertner SL. Where the division lies: Common ingroup identity moderates the cross-race facial-recognition effect. *J Exp Soc Psychol.* 2010;46(2):445-448. doi:10.1016/j.jesp.2009.11.008.
54. Shriver ER, Hugenberg K. Power, individuation, and the cross-race recognition deficit. *J Exp Soc Psychol.* 2010;46(5):767-774. doi:10.1016/j.jesp.2010.03.014.
55. Levin DT. Race as a visual feature: Using visual search and perceptual discrimination tasks to understand face categories and the cross-race recognition deficit. *J Exp Psychol Gen.* 2000;129(4):559-574. doi:http://dx.doi.org/10.1037/0096-3445.129.4.559.
56. Levin DT, Angelone BL. Categorical Perception of Race. *Perception.* 2002;31(5):567-

578. doi:10.1068/p3315.
57. Ge L, Zhang H, Wang Z, et al. Two faces of the other-race effect: Recognition and categorisation of Caucasian and Chinese faces. *Perception*. 2009;38(8):1199-1210. doi:10.1068/p6136.
 58. Goldinger SD, He Y, Papesh MH. Deficits in Cross-Race Face Learning: Insights From Eye Movements and Pupillometry. *J Exp Psychol Learn Mem Cogn*. 2009;35(5):1105-1122. doi:10.1037/a0016548.
 59. Tullis JG, Benjamin AS, Liu X. Self-pacing study of faces of different races: metacognitive control over study does not eliminate the cross-race recognition effect. *Mem Cognit*. 2014;42(6):863-875. doi:10.3758/s13421-014-0409-y.
 60. Meissner CA, Brigham JC, Butz DA. Memory for own- and other-race faces: A dual-process approach. *Appl Cogn Psychol*. 2005;19(5):545-567. doi:10.1002/acp.1097.
 61. Sporer SL. Recognizing Faces of Other Ethnic Groups: An Integration of Theories. *Psychol Public Policy, Law*. 2001;7(1):36-97. doi:10.1037/1076-8971.7.1.36.
 62. Hugenberg K, Young SG, Bernstein MJ, Sacco DF. The Categorization-Individuation Model: An Integrative Account of the Other-Race Recognition Deficit. *Psychol Rev*. 2010;117(4):1168-1187. doi:10.1037/a0020463.
 63. Kanwisher N, McDermott J, Chun MM. The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for Face Perception. *J Neurosci*. 1997;17(11):4302-4311. <http://www.jneurosci.org/content/17/11/4302>. Accessed February 14, 2017.
 64. Gauthier I, Tarr MJ, Anderson AW, Skudlarski P, Gore JC. Activation of the middle fusiform "face area" increases with expertise in recognizing novel objects. *Nat Neurosci*. 1999;2(6):568-573. doi:10.1038/9224.
 65. Gauthier I, Skudlarski P, Gore JC, Anderson AW. Expertise for cars and birds recruits brain areas involved in face recognition. *Nat Neurosci*. 2000;3(2):191-197. doi:10.1038/72140.
 66. Tovée MJ. Is face processing special? *Neuron*. 1998;21(6):1239-1242. doi:10.1016/S0896-6273(00)80644-3.
 67. Golby AJ, Gabrieli JD, Chiao JY, Eberhardt JL. Differential responses in the fusiform region to same-race and other-race faces_annotated2. *Nat Neurosci*. 2001;4(8):845-850. doi:10.1038/90565.
 68. Zhang J, Li X, Song Y, Liu J. The fusiform face area is engaged in holistic, not parts-based, representation of faces. *PLoS One*. 2012;7(7). doi:10.1371/journal.pone.0040390.
 69. Zhao M, Cheung S hang, Wong ACN, et al. Processing of configural and componential information in face-selective cortical areas. *Cogn Neurosci*. 2014;5(3-4):160-167. doi:10.1080/17588928.2014.912207.
 70. Zhou G, Liu J, Xiao NG, Wu SJ, Li H, Lee K. The Fusiform Face Area Plays a Greater Role

- in Holistic Processing for Own-Race Faces Than Other-Race Faces. *Front Hum Neurosci.* 2018;12(June):1-16. doi:10.3389/fnhum.2018.00220.
71. Young AW, Hellawell D, Hay DC. Configurational information in face perception. *Perception.* 1987;16(6):747-759. doi:10.1068/p160747.
 72. Foster C, Bühlhoff I, Bartels A, Zhao M. Investigating holistic face processing within and outside of face-responsive brain regions. *Neuroimage.* 2021;226(November 2020). doi:10.1016/j.neuroimage.2020.117565.
 73. Kent BA, Hvoslef-Eide M, Saksida LM, Bussey TJ. The representational-hierarchical view of pattern separation: Not just hippocampus, not just space, not just memory? *Neurobiol Learn Mem.* 2015;129:99-106. doi:10.1016/j.nlm.2016.01.006.
 74. van den Honert RN, McCarthy G, Johnson MK. Holistic versus feature-based binding in the medial temporal lobe. *Cortex.* 2017:1-11. doi:10.1016/j.cortex.2017.01.011.
 75. Brown TI, Uncapher MR, Chow TE, Eberhardt JL, Wagner AD. Cognitive control, attention, and the other race effect in memory. *PLoS One.* 2017;12(3):e0173579. doi:10.1371/journal.pone.0173579.
 76. Brosch T, Bar-David E, Phelps E a. Implicit Race Bias Decreases the Similarity of Neural Representations of Black and White Faces. *Psychol Sci.* 2013;24(2):160-166. doi:10.1177/0956797612451465.
 77. Yaros JL, Salama DA, Delisle D, Larson MS, Miranda BA, Yassa MA. A Memory Computational Basis for the Other-Race Effect. *Sci Rep.* 2019;9(19399). doi:10.1038/s41598-019-55350-0.
 78. Leal SL, Yassa MA. Integrating new findings and examining clinical applications of pattern separation_AN032718. *Nat Neurosci.* 2018;21(February):163-173. doi:10.1038/s41593-017-0065-1.
 79. Squire LR. Declarative and nondeclarative memory: Multiple brain systems supporting learning and memory. *J Cogn Neurosci.* 1992;4(3):232-243. doi:10.1162/jocn.1992.4.3.232.
 80. Guzowski JF, Knierim JJ, Moser EI. Ensemble dynamics of hippocampal regions CA3 and CA1. *Neuron.* 2004;44(4):581-584. doi:10.1016/j.neuron.2004.11.003.
 81. Leutgeb JK, Leutgeb S, Moser M, Moser EI, Moser I. Pattern Separation in the Dentate Gyrus and CA3 of the Hippocampus. *Science (80-).* 2007;315(5814):961-966. doi:10.1126/science.1135801.
 82. Leutgeb S. Distinct Ensemble Codes in Hippocampal Areas CA3 and CA1. *Science (80-).* 2004;305(5688):1295-1298. doi:10.1126/science.1100265.
 83. McClelland JL, McNaughton BL, Reilly RCO. Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol Rev.* 1995;102(3):419-457. doi:10.1037/0033-295X.102.3.419.

84. Neunuebel JP, Knierim JJ. CA3 retrieves coherent representations from degraded input: Direct evidence for CA3 pattern completion and dentate gyrus pattern separation. *Neuron*. 2014;81(2):416-427. doi:10.1016/j.neuron.2013.11.017.
85. Treves A, Rolls ET. Computational analysis of the role of the hippocampus in memory. *Hippocampus*. 1994;4(3):374-391. doi:10.1002/hipo.450040319.
86. Vazdarjanova A. Differences in Hippocampal Neuronal Population Responses to Modifications of an Environmental Context: Evidence for Distinct, Yet Complementary, Functions of CA3 and CA1 Ensembles. *J Neurosci*. 2004;24(29):6489-6496. doi:10.1523/JNEUROSCI.0350-04.2004.
87. Yassa MA, Stark CEL. Pattern separation in the hippocampus. *Trends Neurosci*. 2011;34(10):515-525. doi:10.1016/j.tins.2011.06.006.
88. Marr D. Simple Memory: A Theory for Archicortex. *Philos Trans R Soc B Biol Sci*. 1971;55(July):23-81.
89. Burke SN, Wallace JL, Hartzell AL, Nematollahi S, Plange K. BCA deficits in pattern separation functions of the perirhinal cortex: a cross-species consensus. *BN* 2011;125(6):836-847. doi:10.1037/a002623. Age-Associated Deficits in Pattern Separation Functions of the Perirhinal Cortex: A Cross-species Consensus. *Behav Neurosci*. 2011;125(6):836-847. doi:10.1037/a0026238.Age-Associated.
90. Miranda M, Kent BA, Morici JF, et al. Molecular Mechanisms in Perirhinal Cortex Selectively Necessary for Discrimination of Overlapping Memories, but Independent of Memory Persistence Magdalena. *eNeuro*. 2017;4(5). doi:10.1523/ENEURO.0293-17.2017.
91. Miranda M, Bekinschtein P. Plasticity Mechanisms of Memory Consolidation and Reconsolidation in the Perirhinal Cortex. *Neuroscience*. 2018;370:46-61. doi:10.1016/j.neuroscience.2017.06.002.
92. Bakker A, Kirwan CB, Miller M, Stark CEL. Pattern Separation in the Human Hippocampal CA3 and Dentate Gyrus. *Science (80-)*. 2008;319(March):1640-1642. doi:10.1126/science.1152882.
93. Yassa MA, Stark SM, Bakker A, Albert MS, Gallagher M, Stark CEL. High-resolution structural and functional MRI of hippocampal CA3 and dentate gyrus in patients with amnesic Mild Cognitive Impairment. *Neuroimage*. 2010;51(3):1242-1252. doi:10.1016/j.neuroimage.2010.03.040.
94. Reagh ZM, Watabe J, Ly M, Murray E, Yassa M a. Dissociated signals in human dentate gyrus and CA3 predict different facets of recognition memory. *J Neurosci*. 2014;34(40):13301-13313. doi:10.1523/JNEUROSCI.2779-14.2014.
95. LaRocque KF, Smith ME, Carr VA, Witthoft N, Grill-Spector K, Wagner AD. Global Similarity and Pattern Separation in the Human Medial Temporal Lobe Predict Subsequent Memory. *J Neurosci*. 2013;33(13):5466-5474. doi:10.1523/JNEUROSCI.4293-12.2013.

96. Berron D, Schutze H, Maass A, et al. Strong Evidence for Pattern Separation in Human Dentate Gyrus. *J Neurosci*. 2016;36(29):7569-7579. doi:10.1523/JNEUROSCI.0518-16.2016.
97. Kyle CT, Stokes JD, Lieberman JS, Hassan AS, Ekstrom AD. Successful retrieval of competing spatial environments in humans involves hippocampal pattern separation mechanisms. *Elife*. 2015;4(NOVEMBER2015):1-19. doi:10.7554/eLife.10499.
98. Doxey CR, Kirwan BC. Structural and functional correlates of behavioral pattern separation in the hippocampus and medial temporal lobe. *Hippocampus*. 2015;25(4):524-533. doi:10.1002/hipo.22389.
99. Reagh ZM, Yassa MA. Object and spatial mnemonic interference differentially engage lateral and medial entorhinal cortex in humans. *Proc Natl Acad Sci*. 2014;111(40):E4264-E4273. doi:10.1073/pnas.1411250111.
100. Schurgin MW. Visual memory, the long and the short of it: A review of visual working memory and long-term memory. *Attention, Perception, Psychophys*. 2018;80(5):1035-1056. doi:10.3758/s13414-018-1522-y.
101. Chang A, Murray E, Yassa MA. Mnemonic discrimination of similar face stimuli and a potential mechanism for the “other race” effect_annotated2. *Behav Neurosci*. 2015;129(5):666-672. doi:10.1037/bne0000090.
102. Carlin JD, Kriegeskorte N. Adjudicating between face-coding models with individual-face fMRI responses. *PLoS Comput Biol*. 2017;13(7):1-28. doi:10.1371/journal.pcbi.1005604.
103. Suwabe K, Hyodo K, Byun K, et al. Aerobic fitness associates with mnemonic discrimination as a mediator of physical activity effects: Evidence for memory flexibility in young adults /631/378/1595/1554 /631/378/1595/1554 /631/378/1595/1554 /631/378/1595/1554 /631/477/2811 /141 article. *Sci Rep*. 2017;7(1):1-10. doi:10.1038/s41598-017-04850-y.
104. Reagh ZM, Ho HD, Leal SL, et al. Greater loss of object than spatial mnemonic discrimination in aged adults. *Hippocampus*. 2016;26(4):417-422. doi:10.1002/hipo.22562.
105. Papesh, Megan H and Goldinger SD. Deficits in Other-Race Face Recognition: No Evidence for Encoding-Based Effects_an_090717. 2010;63(4):253-262. doi:10.1037/a0015802.Deficits.
106. Sessa P, Tomelleri S, Luria R, Castelli L, Reynolds M, Dell’Acqua R. Look out for strangers! Sustained neural activity during visual working memory maintenance of other-race faces is modulated by implicit racial prejudice. *Soc Cogn Affect Neurosci*. 2012;7(3):314-321. doi:10.1093/scan/nsr011.
107. Sessa P, Dalmaso M. Race perception and gaze direction differently impair visual working memory for faces: An event-related potential study. *Soc Neurosci*. 2016;11(1):97-107. doi:10.1080/17470919.2015.1040556.

108. Walker PM, Tanaka JW. An encoding advantage for own-race versus other-race faces. *Perception*. 2003;32(9):1117-1125. doi:10.1068/p5098.
109. Stelter M, Degner J. Investigating the other-race effect in working memory. 2018:1-22. doi:10.1111/bjop.12304.
110. Eriksson J, Vogel EK, Lansner A, Bergström F, Nyberg L. Neurocognitive Architecture of Working Memory. *Neuron*. 2015;88(1):33-46. doi:10.1016/j.neuron.2015.09.020.
111. Newmark RE, Schon K, Ross RS, Stern CE. Contributions of the hippocampal subfields and entorhinal cortex to disambiguation during working memory. *Hippocampus*. 2013;23(6):467-475. doi:10.1002/hipo.22106.
112. Axmacher N, Mormann F, Fernandez G, Cohen MX, Elger CE, Fell J. Sustained Neural Activity Patterns during Working Memory in the Human Medial Temporal Lobe. *J Neurosci*. 2007;27(29):7807-7816. doi:10.1523/jneurosci.0962-07.2007.
113. Olson IR. Working Memory for Conjunctions Relies on the Medial Temporal Lobe. *J Neurosci*. 2006;26(17):4596-4601. doi:10.1523/jneurosci.1923-05.2006.
114. Natu V, Raboy D, O'Toole AJ. Neural correlates of own- and other-race face perception: Spatial and temporal response differences_AN. *Neuroimage*. 2011;54(3):2547-2555. doi:10.1016/j.neuroimage.2010.10.006.
115. Kim JS, Yoon HW, Kim BS, Jeun SS, Jung SL, Choe BY. Racial distinction of the unknown facial identity recognition mechanism by event-related fMRI. *Neurosci Lett*. 2006;397(3):279-284. doi:10.1016/j.neulet.2005.12.061.
116. Mundy ME, Downing PE, Graham KS. Extrastriate cortex and medial temporal lobe regions respond differentially to visual feature overlap within preferred stimulus category. *Neuropsychologia*. 2012;50(13):3053-3061. doi:10.1016/j.neuropsychologia.2012.07.006.
117. Bussey TJ, Saksida LM. The organization of visual object representations: a connectionist model of effects of lesions in perirhinal cortex. *Eur J Neurosci*. 2002;15(2):355-364. doi:10.1046/j.0953-816x.2001.01850.x.
118. Bussey TJ, Saksida LM, Murray EA. Perirhinal cortex resolves feature ambiguity in complex visual discriminations. *Eur J Neurosci*. 2002;15(2):365-374. doi:10.1046/j.0953-816x.2001.01851.x.
119. Barense MD, Henson RNA, Lee ACH, Graham KS. Medial temporal lobe activity during complex discrimination of faces, objects, and scenes: Effects of viewpoint. *Hippocampus*. 2010;20(3):389-401. doi:10.1002/hipo.20641.
120. O'Neil EB, Cate AD, Kohler S. Perirhinal Cortex Contributes to Accuracy in Recognition Memory and Perceptual Discriminations. *J Neurosci*. 2009;29(26):8329-8334. doi:10.1523/JNEUROSCI.0374-09.2009.
121. Bussey TJ, Saksida LM, Murray EA. The perceptual-mnemonic / feature conjunction model of perirhinal cortex function. 2005:269-282.

doi:10.1080/02724990544000004.

122. Lee ACH, Bussey TJ, Murray EA, et al. Perceptual deficits in amnesia: Challenging the medial temporal lobe “mnemonic” view. *Neuropsychologia*. 2005;43(1):1-11. doi:10.1016/j.neuropsychologia.2004.07.017.
123. Smith CN, Jeneson A, Frascino JC, Kirwan CB, Hopkins RO, Squire LR. When recognition memory is independent of hippocampal function. *Proc Natl Acad Sci U S A*. 2014;111(27):9935-9940. doi:10.1073/pnas.1409878111.
124. Mundy ME, Downing PE, Dwyer DM, Honey RC, Graham KS. A Critical Role for the Hippocampus and Perirhinal Cortex in Perceptual Learning of Scenes and Faces: Complementary Findings from Amnesia and fMRI. *J Neurosci*. 2013;33(25):10490-10502. doi:10.1523/JNEUROSCI.2958-12.2013.
125. O’Neil EB, Barkley VA, Köhler S. Representational demands modulate involvement of perirhinal cortex in face processing. *Hippocampus*. 2013;23(7):592-605. doi:10.1002/hipo.22117.
126. Lee ACH, Buckley MJ, Pegman SJ, et al. Specialization in the medial temporal lobe for processing of objects and scenes. *Hippocampus*. 2005;15(6):782-797. doi:10.1002/hipo.20101.
127. Webster MA, Kaping D, Mizokami Y, Duhamel P. Adaptation to natural facial categories. *Nature*. 2004;428(April):557-561. doi:10.1038/nature02361.1.
128. Natu V, Raboy D, O’Toole AJ. Neural correlates of own- and other-race face perception: Spatial and temporal response differences. *Neuroimage*. 2011;54(3):2547-2555. doi:10.1016/j.neuroimage.2010.10.006.
129. Reggev N, Brodie K, Cikara M, Mitchell JP. Human face-selective cortex does not distinguish between members of a racial outgroup. *eNeuro*. 2020;7(3). doi:10.1523/ENEURO.0431-19.2020.
130. Cox RW. AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res*. 1996;29(3):162-173. doi:10.1006/cbmr.1996.0014.
131. Cox RW. AFNI: What a long strange trip it’s been. *Neuroimage*. 2012;62(2):743-747. doi:10.1016/j.neuroimage.2011.08.056.
132. Bullmore E, Sporns O. Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci*. 2009;10(3):186-198. doi:10.1038/nrn2575.
133. Friston KJ. Functional and Effective Connectivity: A Review. *Brain Connect*. 2011;1(1):13-36. doi:10.1089/brain.2011.0008.
134. Thomas Yeo BT, Krienen FM, Sepulcre J, et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol*. 2011;106(3):1125-1165. doi:10.1152/jn.00338.2011.

135. Bullmore ET, Bassett DS. Brain graphs: Graphical models of the human brain connectome. *Annu Rev Clin Psychol*. 2011;7:113-140. doi:10.1146/annurev-clinpsy-040510-143934.
136. Watts DJ, Strogatz SH. Collective dynamics of “small-world” networks. *Nature*. 1998;393(June):440-442. <https://www.ncbi.nlm.nih.gov/pubmed/9623998>.
137. Latora V, Marchiori M. Efficient behavior of small-world networks. *Phys Rev Lett*. 2001;87(19):198701-1-198701-198704. doi:10.1103/PhysRevLett.87.198701.
138. delEtoile J, Adeli H. Graph Theory and Brain Connectivity in Alzheimer’s Disease. *Neuroscientist*. 2017;23(6):616-626. doi:10.1177/1073858417702621.
139. Stam CJ, Jones BF, Nolte G, Breakspear M, Scheltens P. Small-world networks and functional connectivity in Alzheimer’s disease. *Cereb Cortex*. 2007;17(1):92-99. doi:10.1093/cercor/bhj127.
140. Rauchmann B, Ersoezlue E, Stoecklein S, et al. Resting-State Network Alterations Differ between Alzheimer’s Disease Atrophy Subtypes. *Cereb Cortex*. 2021;(bhab130):2021. doi:10.1093/cercor/bhab130.
141. Vlooswijk MCG, Vaessen MJ, Jansen JFA, et al. Loss of network efficiency associated with cognitive decline in chronic epilepsy. *Neurology*. 2011;77(10):938-944. doi:10.1212/WNL.0b013e31822cfc2f.
142. Rudie JD, Brown JA, Beck-Pancer D, et al. Altered functional and structural brain network organization in autism. *NeuroImage Clin*. 2013;2(1):79-94. doi:10.1016/j.nicl.2012.11.006.
143. Shine JM, Poldrack RA. Principles of dynamic network reconfiguration across diverse brain states. *Neuroimage*. 2018;180(March):396-405. doi:10.1016/j.neuroimage.2017.08.010.
144. Cohen JR, D’Esposito M. The segregation and integration of distinct brain networks and their relationship to cognition. *J Neurosci*. 2016;36(48):12083-12094. doi:10.1523/JNEUROSCI.2965-15.2016.
145. Stanley ML, Simpson SL, Dagenbach D, Lyday RG, Burdette JH, Laurienti PJ. Changes in brain network efficiency and working memory performance in aging. *PLoS One*. 2015;10(4):1-17. doi:10.1371/journal.pone.0123950.
146. Westphal AJ, Wang S, Rissman J. Episodic memory retrieval benefits from a less modular brain network organization. *J Neurosci*. 2017;37(13):3523-3531. doi:10.1523/JNEUROSCI.2509-16.2017.
147. Sadaghiani S, Poline JB, Kleinschmidt A, D’Esposito M. Ongoing dynamics in large-scale functional connectivity predict perception. *Proc Natl Acad Sci U S A*. 2015;112(27):8463-8468. doi:10.1073/pnas.1420687112.
148. Finc K, Bonna K, He X, et al. Dynamic reconfiguration of functional brain networks during working memory training. *Nat Commun*. 2020;11(2435).

doi:doi.org/10.1038/s41467-020-15631-z.

149. Farahani F V., Karwowski W, Lighthall NR. Application of graph theory for identifying connectivity patterns in human brain networks: A systematic review. *Front Neurosci.* 2019;13(JUN):1-27. doi:10.3389/fnins.2019.00585.
150. Ishai A, Ungerleider LG, Haxby J V. Distributed neural systems for the generation of visual images. *Neuron.* 2000;28(3):979-990. doi:10.1016/S0896-6273(00)00168-9.
151. Haxby J V., Hoffman EA, Gobbini MI. The distributed human neural system for face perception. *Trends Cogn Sci.* 2000;4(6):223-233. doi:10.1016/S1364-6613(00)01482-0.
152. Collins JA, Olson IR. Beyond the FFA: The role of the ventral anterior temporal lobes in face processing_an1. *Neuropsychologia.* 2014;61(1):65-79. doi:10.1016/j.neuropsychologia.2014.06.005.
153. Joseph JE, Swearingen JE, Clark JD, et al. The changing landscape of functional brain networks for face processing in typical development. *Neuroimage.* 2012;63(3):1223-1236. doi:10.1016/j.neuroimage.2012.08.021.
154. Joseph JE, Vanderweyen D, Swearingen J, et al. Tracking the development of functional connectomes for face processing. *Brain Connect.* 2019;9(2):231-239. doi:10.1089/brain.2018.0607.
155. Lord LD, Stevner AB, Deco G, Kringelbach ML. Understanding principles of integration and segregation using whole-brain computational connectomics: Implications for neuropsychiatric disorders. *Philos Trans R Soc A Math Phys Eng Sci.* 2017;375(2096). doi:10.1098/rsta.2016.0283.
156. Harris P, Taylor R, Thielke R, Payne J, Gonzales N, Conde J. Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inf.* 2009;42(2):377-381.
157. Hugenberg K, Miller J, Claypool HM. Categorization and individuation in the cross-race recognition deficit: Toward a solution to an insidious problem. *J Exp Soc Psychol.* 2007;43(2):334-340. doi:10.1016/j.jesp.2006.02.010.
158. Pica E, Warren AR, Ross DF, Kehn A. Choosing Your Words and Pictures Wisely: When Do Individuation Instructions Reduce the Cross-Race Effect? *Appl Cogn Psychol.* 2015;29(3):360-368. doi:10.1002/acp.3112.
159. Peirce J, Gray JR, Simpson S, et al. PsychoPy2: Experiments in behavior made easy. *Behav Res Methods.* 2019;51(1):195-203. doi:10.3758/s13428-018-01193-y.
160. Stark CEL, Squire LR. When zero is not zero: The problem of ambiguous baseline conditions in fMRI. *Proc Natl Acad Sci U S A.* 2001;98(22):12760-12765. doi:10.1073/pnas.221462998.
161. Whitfield-Gabrieli S, Nieto-Castanon A. Conn: A Functional Connectivity Toolbox for

- Correlated and Anticorrelated Brain Networks. *Brain Connect.* 2012;2(3):125-141. doi:10.1089/brain.2012.0073.
162. Behzadi Y, Restom K, Liau J, Liu TT. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *Neuroimage.* 2007;37(1):90-101. doi:10.1016/j.neuroimage.2007.04.042.
 163. Power JD, Mitra A, Laumann TO, Snyder AZ, Schlaggar BL, Petersen SE. Methods to detect, characterize, and remove motion artifact in resting state fMRI Jonathan. *Neuroimage.* 2014;84:320-341. doi:doi:10.1016/j.neuroimage.2013.08.048. Methods.
 164. McLaren DG, Ries ML, Xu G, Johnson SC. A generalized form of context-dependent psychophysiological interactions (gPPI): A comparison to standard approaches. *Neuroimage.* 2012;61(4):1277-1286. doi:10.1016/j.neuroimage.2012.03.068.
 165. Gerchen MF, Bernal-Casas D, Kirsch P. Analyzing task-dependent brain network changes by whole-brain psychophysiological interactions: A comparison to conventional analysis. *Hum Brain Mapp.* 2014;35(10):5071-5082. doi:10.1002/hbm.22532.
 166. Markett S, Jawinski P, Kirsch P, Gerchen MF. Specific and segregated changes to the functional connectome evoked by the processing of emotional faces: A task-based connectome study. *Sci Rep.* 2020;10:4822. doi:10.1038/s41598-020-61522-0.
 167. Fan L, Li H, Zhuo J, et al. The Human Brainnetome Atlas: A New Brain Atlas Based on Connectional Architecture. *Cereb Cortex.* 2016;26(8):3508-3526. doi:10.1093/cercor/bhw157.
 168. O'Reilly JX, Woolrich MW, Behrens TEJ, Smith SM, Johansen-Berg H. Tools of the trade: Psychophysiological interactions and functional connectivity. *Soc Cogn Affect Neurosci.* 2012;7(5):604-609. doi:10.1093/scan/nss055.
 169. Di X, Biswal BB. Psychophysiological interactions in a visual checkerboard task: Reproducibility, reliability, and the effects of deconvolution. *Front Neurosci.* 2017;11(OCT):1-16. doi:10.3389/fnins.2017.00573.
 170. Di X, Zhang Z, Biswal BB. Understanding psychophysiological interaction and its relations to beta series correlation. *Brain Imaging Behav.* 2020. doi:10.1007/s11682-020-00304-8.
 171. Achard S, Bullmore E. Efficiency and cost of economical brain functional networks. *PLoS Comput Biol.* 2007;3(2):0174-0183. doi:10.1371/journal.pcbi.0030017.
 172. Bassett DS, Bullmore E, Verchinski BA, Mattay VS, Weinberger DR, Meyer-Lindenberg A. Hierarchical organization of human cortical networks in health and Schizophrenia. *J Neurosci.* 2008;28(37):9239-9248. doi:10.1523/JNEUROSCI.1929-08.2008.
 173. MATLAB. version 9.8.0.1451342 (R2020a). Natick, Massachusetts: The MathWorks Inc. 2020.
 174. Rubinov M, Sporns O. Complex network measures of brain connectivity: Uses and

- interpretations. *Neuroimage*. 2010;52(3):1059-1069.
doi:10.1016/j.neuroimage.2009.10.003.
175. Pekár S, Brabec M. Generalized estimating equations : A pragmatic and flexible approach to the marginal GLM modelling of correlated data in the behavioural sciences. *Ethology*. 2017;124(4):86-93. doi:10.1111/eth.12713.
 176. Liu J, Wang Z, Feng L, Li J, Tian J, Lee K. Neural Trade-Offs between Recognizing and Categorizing Own- and Other-Race Faces. *Cereb Cortex*. 2015;25(8):2191-2203. doi:10.1093/cercor/bhu025.
 177. Zhen Z, Fang H, Liu J. The Hierarchical Brain Network for Face Recognition. *PLoS One*. 2013;8(3). doi:10.1371/journal.pone.0059886.
 178. Fan J, Gu X, Liu X, et al. Involvement of the anterior cingulate and frontoinsula cortices in rapid processing of salient facial emotional information. *Neuroimage*. 2011;54(3):2539-2546. doi:10.1016/j.neuroimage.2010.10.007.
 179. Bassett DS, Wymbs NF, Porter MA, Mucha PJ, Carlson JM, Grafton ST. Dynamic reconfiguration of human brain networks during learning. *Proc Natl Acad Sci U S A*. 2011;108(18):7641-7646. doi:10.1073/pnas.1018985108.
 180. DeGutis J, DeNicola C, Zink T, McGlinchey R, Milberg W. Training with own-race faces can improve processing of other-race faces: Evidence from developmental prosopagnosia. *Neuropsychologia*. 2011;49(9):2505-2513. doi:10.1016/j.neuropsychologia.2011.04.031.
 181. McGugin RW, Ryan KF, Tamber-Rosenau BJ, Gauthier I. The Role of Experience in the Face-Selective Response in Right FFA_an032718. *Cereb Cortex*. 2017:1-14. doi:10.1093/cercor/bhx113.
 182. McGugin RW, Tanaka JW, Lebrecht S, Tarr MJ, Gauthier I. Race-Specific Perceptual Discrimination Improvement Following Short Individuation Training With Faces_annotated. *Cogn Sci*. 2011;35(2):330-347. doi:10.1111/j.1551-6709.2010.01148.x.
 183. Montoya L, Westerlund A, Troller-Renfree S, Righi G, Nelson CA. The effect of heterogeneous race exposure during infancy. *Cogn Dev*. 2017:1-10. doi:10.1016/j.cogdev.2017.02.014.
 184. Kelly DJ, Liu S, Lee K, et al. Development of the other-race effect during infancy: Evidence toward universality? *J Exp Child Psychol*. 2009;104(1):105-114. doi:10.1016/j.jecp.2009.01.006.
 185. Liu S, Xiao WS, Xiao NG, et al. Development of visual preference for own- versus other-race faces in infancy. *Dev Psychol*. 2015;51(4):500-511. doi:http://dx.doi.org/10.1037/a0038835.
 186. Blackstone W. *Commentaries on the Laws of England*. doi:10.4324/9781315053837-7.