Title
The Relation Between Race-based Stressors and Cognitive Control Processes in Latinx College Students

Author
Vazquez, Salvador Roberto

Publication Date
2023

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

The Relation Between Race-based Stressors and Cognitive Control Processes in Latinx College Students

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Education

by

Salvador Roberto Vazquez

2023

ABSTRACT OF THE DISSERTATION

The Relation Between Race-based Stressors and Cognitive Control Processes

in Latinx College Students

by

Salvador Roberto Vazquez

Doctor of Philosophy in Education

University of California, Los Angeles, 2023

Professor Jennie Katherine Grammer, Chair

Latinx is the fastest growing ethnic group in the United States. As a result, the demographic landscape of colleges across the nation is also changing to reflect this new reality. However, many Latinx students are still not finishing college at the same rate as other ethnic groups. One area of research that is lacking regarding Latinx students is their experiences around learning, studying, and testing. While many studies illustrate the unique circumstances that Latinx students face when going to college (e.g., likely to be first-generation student, strong sense of family obligation, experiencing racial discrimination) few have explored how these experiences relate directly to cognitive processes associated with learning.

The testing effect paradigm has been well studied by cognitive psychologists, who have demonstrated self-testing is a very effective way to learn. Although the testing effect has been well researched in predominately Caucasian college students, little is known about how well it translates to other groups. The present study aims to look at how the heterogenous educational experiences of Latinx students at a Southern California university relates to cognitive processes of learning. Using a well-studied cognitive phenomenon known as the Testing Effect, this study used survey measures, experimental tasks (a Flanker task and a Spatial Working Memory Task), and electroencephalography (EEG) to assess which sociocultural experiences relate to measures of cognitive processes involved in learning on a testing effect task. Results indicate that race-based stress may account for overall lower accuracy on the testing effect task, but Latinx participants benefitted from testing as evidenced by better accuracy and reactions times on test condition items. The Error Related Negativity and Feedback Related Negativity were not present in the testing effect task, but both the P300 and alpha power were present and associated with race-based stress. Findings from this study will contribute to a better understanding of the importance of studying the unique experiences of Latinx college students as they relate to cognitive processes involved in learning and the considerations students and educators may need to take before using testing as a learning tool.

The dissertation of Salvador Roberto Vazquez is approved.

Cecilia Rios Aguilar

Sandra H. Graham

James W. Stigler

Jennie Katherine Grammer, Committee Chair

University of California, Los Angeles

2023

DEDICATION

This work is dedicated to my parents, Salvador Vazquez Sr. and Laura Elena Bustillos. Thank you for your unwavering support and encouragement to pursue a life of happiness.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

This study would not have been possible without the help of some great individuals. Jennie Grammer, thank you for the opportunity to achieve my dream of conducting neuroscience research. I have learned so much about how to be a better scientist and how to navigate academia through your guidance and mentorship. I especially want to thank you for your support and advocacy over the past few years. Graduate school is not easy, but working with you made it fun. I would also like to thank members of my dissertation committee—Jim Stigler, Cecilia Rios-Aguilar, and Sandra Graham—who all provided invaluable feedback to improve the quality of my dissertation. Thank you to the many other faculty and staff who played a part in the development and execution of this project, including Carola Suarez-Orozco, Daniel Solorzano, Agatha Lenartowicz, Joel P. Diaz, Taylor Hazelbaker, Amy Gershon, and Connie Kurtew. Thank you to the wonderful undergraduate research assistants who helped collect, clean, and process data: Pauline Vartany, Jaquelin Aquino, Susy Cisneros, Daniel Huan, Daniel Vega, Maddie Ward, Briana Camacho, Rukhsar Sayani, and Luis Ramirez. And thank you to the participants. Financial support for this project was provided by the UCLA Dissertation Year Fellowship and the California State University Northridge College of Social and Behavioral Science Doctoral Scholar Program.

I feel fortunate to have had the opportunity to learn from some great educators throughout my time as a student. Two in particular made it possible for me to be where I am today. Christia Spears Brown, thank you for being the first to give me the opportunity to learn about psychology research and for providing the best learning experience an undergraduate student could ever ask for. Your mentorship and my time in your lab were the reasons why I felt confident applying to a PhD program. To Patricia Greenfield, you recognized when I was lost and provided the support I

needed to get back on track by creating a writing group that has continued for almost 4 years. Thank you for always being there for me.

I would be very remiss if I did not thank all of my teachers from Micheltorena and Solano Elementary, Thomas Starr King Middle, and John Marshall High School. I wish I had the space to name them all because they all played a part in my development towards becoming an educational researcher. In particular though I would like to thank my 8th grade science teacher Ms. Howard for showing me the value of being a good observer, to my AP European History teacher Mr. Patrick Jeffries for modeling how to be a fun and approachable educator, and to the late Mr. Dan Lucci, my AP Biology teacher, for showing me how fun it is to learn science.

It is very easy to feel lonely during graduate school, but I had the support of a community that helped me get through it all. Thank you to Gerardo Ramirez, Rashmita Mistry, Mitsue Yokota, Stephen Kwok, Yolanda Vasquez-Salgado, Negin Ghavami, Laura Rhinehart, David Taylor, and Jennifer Hites-Thomas. Special thanks to all the members of the UCLA B.E.A.R. Lab for their friendship, mentorship, and emotional support. To Sarah Jo Torgrimson and Keye Xu, I could not have asked for two better people to be with on this journey.

To my closest friends Christopher Lavarias, Jarrad Skinner, Hector Valadez, Helen Leung, and Russell Horning, thank you for all the love and support you have shown me and my family. To my partner Estefania Orelo, thank you for your optimism and humor. I'm looking forward to continuing this journey with you.

Thank you to the Vazquez and Bustillos clan. Thank you to my sister Monique Vazquez and brother Juan Vazquez for their unconditional love and support. And thank you to my father Salvador Vazquez Sr. for showing me the value of persistence and hard work and to my mother Laura Elena Bustillos for showing me how to speak up for myself. I love you all.

VITA

<u>EDUCATION</u>

M.A., Human Development and Psychology, UCLA School of Education	June 2016
B.A., Psychology, University of California, Los Angeles	June 2007

<u>AWARDS AND HONORS</u>

UCLA Department of Education Distinguished Teaching Assistant Award [$1000]	2022
California State University Northridge Doctoral Scholar Program (DSP) [$10,000]	2022-2023
IMBES Travel Grant [$2000]	2022
UCLA SE&IS Mary Elizabeth Jenson Fellowship [$8,750]	2021-2022
UCLA Dissertation Year Fellowship (DYF) [$20,000 + Tuition]	2020-2021
UCLA Collegium of University Teaching Fellow (CUTF) [Salary + Tuition]	2020-2021
Ford Foundation Dissertation Fellowship – Honorable Mention	2020-2021
Build Poder External Mentor Award [$2,000]	2019
UCLA Graduate Research Mentorship Fellowship [$20,000 + Tuition]	2017-2018
UCLA Graduate Summer Research Mentorship Award [$6,000]	2017
UCLA SE&IS Louise Tyler Scholarship [$4,000]	2016

<u>PUBLICATIONS</u>

Rhinehart, L., Vazquez, S. R., & Greenfield, P. M. (2021). The impact of screen-free zones in an undergraduate psychology classroom: Assessing exam performance and instructor evaluations in two quasi-experiments. *Teaching of Psychology*. https://doi.org/10.1177/00986283211017443

Vazquez, S. R., & Greenfield, P. M. (2021). The influence of social class on family participation in children's education: A case study. *Revista Colombiana de Psicología*, *30*(1), Article 1. https://doi.org/10.15446/rcp.v30n1.89185

Vazquez, S. R., Ermeling, B. A., & Ramirez, G. (2020). Parental beliefs on the efficacy of productive struggle and their relation to homework-helping behavior. *Journal for Research in Mathematics Education*, *51*(2), 179–203. JSTOR. https://doi.org/10.5951/jresematheduc-2020-0019

<u>**SELECTED PRESENTATIONS** (peer reviewed) **denotes undergraduate student*</u>

Vazquez, S. R., *Aquino, J., *Ward, M., Diaz, J. P., & Grammer, J. K. (2022, July 21). *Relations between minority status stress, the testing effect, the late positive potential, and the feedback related negativity* [poster presentation]. International Mind Brain and Education Society Conference, Montreal, QC, Canada.

*Huang, D., Vazquez, S. R., & Grammer, J. K. (2022, May 24). *The relation of inter and intra-group discrimination to college students' academic and cognitive performance* [video presentation]. UCLA Undergraduate Research Week Showcase, Virtual conference.

*Vega, D., Vazquez, S. R., & Grammer, J. K. (2022, May24). *Latent class analysis of college students' experience with microaggressions* [video presentation]. UCLA Undergraduate Research Week Showcase, Virtual conference.

*Vartany, P., Vazquez, S. R., & Grammer, J. K. (2022, May 6). *Race-based stress predicts college student health outcomes* [poster presentation]. UCLA Psychology Undergraduate Research Conference, Los Angeles, CA, United States.

Vazquez, S. R. & Grammer, J. K. (2022, April 21-26). *Validating the presence of the ERN, FRN, and P300 during a testing effect task* [poster presentation]. American Educational Research Association, San Diego, CA, United States.

Vazquez, S. R., Rapp, A. M., Tan, P., & Grammer, J. K. (2021, September 17-21). *Failure mindset predicts error-related negativity during a go/no-go task in young adults* [poster presentation]. Flux, Virtual conference.

Vazquez, S. R., Rapp, A. M., Tan, P., & Grammer, J. K. (2018, May 25-28). *Neural correlates of behavioral measures of motivation* [poster presentation]. Association for Psychological Science, San Francisco, CA, United States.

Vazquez, S.R., & Grammer, J.K. (2018, May 3). *A comparative case study of parental math homework-helping strategies and the use of productive struggle* [poster presentation]. UCLA Research and Inquiry Conference, Los Angeles, CA, United States.

Vazquez, S.R., Ramirez, G., & Ermeling, B. (2017, May 25-28). *Parental beliefs on the efficacy of productive struggle and their relation to homework-helping behavior* [poster presentation]. Association for Psychological Science, Boston, MA, United States.

INVITED TALKS

CSUN College of Social and Behavioral Science Dean's Colloquium (11/17/2022)
Guest author roundtable at UCLA Department of Psychology (5/19/2022)
Lecture for Educational Psychology course at UCLA School of Education (5/11/2022)
Collegium University Teaching Fellows Seminar - Alumni Panel (11/10/2021)
Guest author roundtable at UCLA Department of Psychology (2/11/2021)
Lecture for Cognitive Development and Schools at UCLA School of Education (3/4/2020)

LEAD INSTRUCTOR

Teaching Preparation in Education	2021-2023
How the Stress of Racism Affects Learning: Examining Academic and Psychological Outcomes of Racial-Ethnic Disparities in Education	W21

TEACHING ASSISTANT

Research in Developmental Psychology	S23
Research Methods in Psychology	Summer C 2021, F22
Early Childhood Development	Summer A 2021
Applied Positive Neuroscience	F19
Digital Media and Human Development	W19
Cognitive Development and Schools	F18, W20
Developmental Psychology	Summer A 2018
Culture and Human Development	S16, S17, F17, S18, S19, S20, S21, W23

PROFESSIONAL EXPERIENCE AND SERVICE

EEG Consultant; Culture, Health, and Development Lab CSUN Department of Psychology	2022-Present
Graduate Student Researcher, UCLA College of Humanities, E.P.I.C.	Fall 2020
UCLA Graduate Division - Graduate Student and Postdoctoral Academic Success Retention Initiatives Officer Search Committee	Feb. 2022

The Relation Between Race-based Stressors and Cognitive Control Processes in Latinx College Students

Latinx is the fastest growing ethnic minority group in the United States, accounting for 52% of total population growth as of 2018 (Flores et al., 2019). In colleges and universities, Latinx enrollment has increased by 180% from 1999 to 2016 (Gromlich, 2017). Yet, while college completion rates for Latinx students have increased in recent years, only 20% of Latinx high school sophomores will eventually earn a bachelor's degree (Carnevale & Fasules, 2017). When it comes to better understanding the social experiences of Latinx students in higher education that may contribute to their rates of success in college, studies have made great strides looking at complex social issues such as immigration status (Guyll et al., 2010), racism and microaggressions (Minikel-Lacocque, 2013), and funds of knowledge (Gonzalez et al., 1995). While these studies and others like them have made a concerted effort to disaggregate the Latinx experience and investigate the nuance of the group, research in cognitive psychology tends to treat Latinx people as a monolithic and homogenous group. As a result, Latinx students have rarely been the focus of studies on factors influencing their cognition and learning in educational settings and even less so in studies using electrophysiological instruments. Studies that can establish whether there is a link between the social experiences of Latinx students and cognitive processes can simultaneously begin to fill the large gap between contextualized research in Education and the lab-based research common in Cognitive Psychology, and better understand the nuance of how social experiences among Latinx college students relate to their abilities to study and learn in school settings.

Towards this aim, this study will employ a White comparison sample group. There are valid criticisms drawing attention to why more research needs to focus on the heterogeneity within non-White ethnic and racial groups (Aceves et al., 2022; Ayala & Chalupa, 2016), including the ideas that between group comparisons can highlight differences that further stigmatize one group, they can encourage deficit-oriented interpretations, and they often help maintain the “White-centric” approach that is dominant in the social sciences (Aceves et al., 2022). One of the goals of this study is to contextualize the Latinx experience in higher education by highlighting that Latinx students are more likely to face race-based stressors than White students. The cumulative effects of race-based stressors and common stressors experienced by most students regardless of race may have detrimental effects on cognitive processes that facilitate effective learning strategies. By contextualizing these experiences with a White comparison group, it is the intent of the author to bring to light potential avenues for future research on the mechanisms involved in complex interactions between the social environment and cognition in educational settings.

Literature Review

Race-based Stressors Reported by Latinx College Students

Studies on the experience of Latinx students when entering college have reported that Latinx students often face stressors that can make it difficult to succeed. Three forms of stressors that are commonly reported in the literature by Latinx students include racism in the form of microaggressions, acculturative stress, and minority stress.

Microaggressions

Students of color in college settings are less likely to report overt forms of racism because such forms of racism are less likely to be socially condoned. Instead students of color are more likely to report experiencing covert forms of racism which are subtle insults that can be automatically or unconsciously conveyed verbally, nonverbally, and/or visually (Solorzano et al., 2000). As a result of microaggressions being able to hide in everyday interactions and their seemingly innocuous nature, the cumulative effects of microaggressions have the tendency to reduce academic performance among students of color due to a depletion of emotional and psychological resources (Sue, 2005; Sue & Constantine, 2007).

Studies focusing specifically on Latinx students have found that in college settings they are more likely to encounter microaggressions in the form of verbal and nonverbal racial affronts from faculty, teaching assistants, and students; racial jokes; and from the institution as characterized by the university structures, practices, and discourses that support a racial climate that is hostile to people of color (Yosso et al., 2009). Having to deal with microaggressions when entering college adds an additional step towards adjusting to the new environment that requires Latinx students to resist and reject the racism they encounter by drawing on cultural resources. Yosso and colleagues (2009) emphasize that in academic settings, microaggressions are likely to function much like stereotype threat (Steele & Aronson, 1995) by causing underperformance on tests and other academic assignments. However, considering the pervasiveness of microaggressions, its effects may be more enduring.

Acculturative Stress

Acculturative stress generally occurs when the member of a minority group must contend with the prevailing cultural values of the dominant group (Born, 1970; Mena et al., 1987). As such, it is most often experienced by recent immigrants but has also been reported among later generations (Mena et al., 1987; Padilla et al., 1986). Acculturative stress can result from problems with adapting to the cultural values, practices, and language of the dominant culture, in addition to discrimination, thus potentially resulting in anxiety and depression (Gil et al., 1994; Williams & Berry, 1991).

Studies have demonstrated that Latinx college students do contend with acculturative stress and it often leads to those same students experiencing some form of psychological distress (Crockett et al., 2007; Rodriguez et al., 2000). The link between acculturative stress and academic performance among Latinx college students has not been extensively researched. Mena and colleagues (1987) did find that a higher level of acculturative stress was linked to academic underperformance among college students, but the sample was only 4% Latinx. Among adolescents, negative correlations between acculturative stress and academic performance has been found in Latinx middle and high school students (Albeg & Castro-Olivo, 2014; Roche & Kuperminc, 2012; Schwartz et al., 2007).

Minority Stress

Minority stress is the result of negative experiences in the campus environment that minority students perceive to be linked to social, physical, or cultural attributes of said minority students (Arbona & Jimenez, 2014; Smedley et al., 1993). Studies have linked minority stress with mental health outcomes such as psychological distress, depression, and anxiety (Jones et al., 2007; Kessler et al., 1999; Neville et al., 2004). However, many of these studies have focused on samples of individuals from different ethnic backgrounds or on Black students at predominantly White institutions.

While it has been found that at colleges where Latinx students are not a minority that minority stress does not negatively affect psychological adjustments, considering the current political climate, Latinx students' experiences with minority stress may be particularly salient at the moment, even at institutions that are not predominantly white (Rodriguez et al., 2000). A more recent study indicates that minority stress can contribute to diminished mental health at universities where Latinx are not the minority (Arbona & Jimenez, 2014). More importantly, it has been indicated that minority stress is distinct from other forms of stress in college and that Latinx college students are more likely to experience it during their first year in college (Smedley et al., 1993; Wei et al., 2011).

The literature clearly indicates that Latinx college students are more likely to contend with stressors above and beyond academic stress experienced by many college students. These stressors in turn are believed to lead to psychological distress and other serious conditions such as anxiety and depression that may be interfering with their ability to study, learn, or perform well on tests. The next section examines how these stressors may be impacting the learning process for Latinx students.

Linking the Latinx College Experience to Learning and Cognition

While studies in education have made a concerted effort to explore the experiences of Latinx as a group, research in cognitive psychology tends to treat Latinx people as monolithic and homogenous when examining cultural factors. As a result, Latinx students have rarely been the focus of studies on factors influencing their cognition and learning in educational settings. This is problematic because it can prevent researchers from focusing on the unique circumstances of Latinx students that need to be addressed. As mentioned earlier, the ever-growing prevalence of the Latinx group within our schools and society deem it important to have a clear understanding of the Latinx student experience and its relations to learning and cognition. It has already been highlighted that Latinx college students are more likely to report instances of racism in the form of microaggressions, acculturative stress, and minority stress. However, less is known about how these stressors might affect Latinx students' ability to learn in academic settings. While the link between stressors and psychological distress has been established in some studies, there are fewer studies that have examined the link between the stressors and cognition.

Relation between Stereotype Threat and Executive Functions as a Case Study

The purpose of this section is to examine the relation between stereotype threat and executive functions as a case study to illustrate gaps in the methodology and what we stand to gain from research linking the social experiences of students of color and their cognition in academic situations.

Previously, Yosso and colleagues (2009) have suggested that examining the links established between stereotype threat, academic performance, and cognition would be a helpful step toward better understanding how race-based stressors such as microaggressions, acculturative stress, and minority stress might affect learning and cognition. Stereotype threat occurs when a member of a group fears they might confirm a negative stereotype of poor performance in a given domain, often leading to poorer performance in that domain (Steele, 1997). The very first study on stereotype threat demonstrated that simply having Black college students write in their race on a form before a test was enough to prime any negative stereotypes they may have had about their ability to perform well and this led to underperformance on a difficult verbal test (Steele & Aronson, 1995).

In addition to replicating these findings with different ethnic, racial, and gender groups, studies have since also focused on underlying mechanisms that lead to underperformance in a stereotype threat condition (Wheeler & Petty, 2001). Particular attention has been paid to cognitive processes known as executive functions. Executive function often refers to three cognitive processes necessary for goal directed behavior: updating/monitoring of working memory, attention shifting, and response inhibition (Miyake et al., 2000). Some studies have suggested that stereotype threat interferes with an individual's working memory. Schmader and

Johns (2003) found that for both a group of women and Latinx college students, priming negative stereotypes correlated with reduced working memory capacity and the reduction in working memory mediated the stereotype threat effect and reduced performance on a math test for women. Beilock and colleagues (2007) also demonstrated that participants working on math problems with high working memory demands were more likely to fail only in the stereotype threat condition.

With regards to attention shifting, most stereotype threat studies suggest that underperformance as a result of not wanting to confirm a negative stereotype about one's group can be interpreted as a strong motivation to avoid failure (Schmader et al., 2008). This results in individuals in a stereotype threat condition to focus more on their own behavior and actively attend to signs of failure. For instance, well learned and automatic processes are susceptible to stereotype threat except when participants are distracted with another task (Beilock et al., 2006). In a study that employed a task similar in design to the one that was used in this study, participants in the stereotype threat condition were distracted by negative feedback to math problems they solved incorrectly and consequentially were less likely to attend to learning feedback that provided them with potential solutions to the problems for a later re-test (Mangels et al., 2012).

Although there seem to be no studies directly linking response inhibition to stereotype threat, studies do suggest that stereotype threat activates negative thoughts. For example, it has been speculated that to perform well on a task, people who experience stereotype threat must actively suppress those thoughts, thus engaging cognitive resources that are used to inhibit prepotent responses (Schmader et al., 2008). However, studies that attempt to measure underlying processes during stereotype threat conditions often had trouble doing so with self-report measures. For example, participants in a study by Bosson et al. (2004) demonstrated significantly more anxiety during stereotype threat when anxiety was measured using an implicit indicator of underlying anxiety, in this case body language, than using a self-report measure of anxiety when compared to a non-stereotype threat control group (Bosson et al., 2004). Therefore, it is safe to reason that individuals in stereotype threat situations may find it more difficult to inhibit certain prepotent responses and that one reason why this has not been well studied is because it may be difficult to observe with existing measures.

Since stereotype threat is so closely tied to an individual's racial and ethnic identity (among many other identities), it stands to reason that experiencing microaggressions, acculturative stress, and minority stress may similarly result in a depletion of cognitive processes which in turn can lead to diminished abilities to study and learn in academic settings. However, a feature of stereotype threat studies is that they examine the effects of race-based stress by priming individuals at the moment or just before engaging with an academic task. It is more likely that Latinx college students are primed by race-based stressors throughout their college going experience and hence may present a cumulative effect on their cognition. Therefore, to understand this general effect rather than the effect as it relates to a specific priming event, this study will use surveys to assess experience with race-based stress.

Another thing that is clear from the existing stereotype threat literature focusing on cognitive outcomes is that researchers cannot rely solely on self-report measures. I now turn to the potential merits of using electroencephalography (EEG) to measure covert cognitive processes.

Measuring Cognitive Processes with EEG

EEG studies have helped establish well-studied neural correlates of cognition that can be used to measure covert cognitive processes (Deouell & Knight, 2009). For instance, a study on children in kindergarten and first grade found changes due to schooling in event related potentials (ERPs) that reflect attention and response monitoring but no analogous behavioral differences in accuracy or response time (Morrison et al., 2019). This suggests that there are certain cognitive processes (e.g., attention, working memory, and response inhibition) that are related to learning that can be measured with neural methods but not with existing behavioral tasks and self-report scales. In other words, EEG allows researchers to see in real time neural responses that give us insight into processes that we cannot observe and that are difficult to measure behaviorally. Considering these limitations in self-report and behavioral measures, this study will use EEG to allow for a different level of measurement that is distinct from self-report or observer ratings. The aim is that using EEG to measure covert cognitive processes may expand our understanding of how experiences like racism and race-based stressors might affect learning and the brain, especially in situations where participants are unable to articulate such effects. To do this, this study will also employ the use of a well vetted learning phenomenon known as the Testing Effect.

The Testing Effect

One area where these considerations might be particularly important involves the study of the Testing Effect (TE). The TE is an area of study in cognitive psychology that has the potential to help students to learn more efficiently but has yet to be widely adopted by students and educators (Vaughn & Kornell, 2019). The TE occurs when testing oneself repeatedly results in better long-term recall than having studied the material by less elaborative means. In other words, the TE is an act of successfully retrieving information stored in memory that makes that information more recallable in the future (R. A. Bjork, 1975). Although the term used in the literature is testing effect, the use of the word testing is a misnomer because the task serves as a form of studying in addition to serving as a form of assessment. When engaging with a Testing Effect task, the person is practicing the act of retrieving the information from memory. By using this task, the goal is to examine the process of learning and not any behavioral phenomena associated with taking a test, such as test anxiety. As a result, this task really elicits moments in which researchers can confidently observe and measure underlying learning processes.

In empirical studies, with the testing effect task often involves two conditions: testing and studying. In the testing condition, participants are exposed to a study block followed by at least one testing block in which they attempt to recall the previously learned stimuli. The study condition involves no testing at all, only re-study of the stimuli in each additional presentation block. After all the study and testing blocks have been administered, participants are then given an immediate retention test. This is also followed by a longer delay retention test (usually up to one week later) to test long-term retention. In all these studies participants in the test condition outperform those in the study condition on delayed retention tests. Although tests are typically used in school contexts to assess a student's knowledge about a specific topic or area, the testing effect suggests that testing is also a useful learning tool because of its ability to modify memory (R. A. Bjork, 1975).

One of the distinguishing features of testing as a study aid when compared to other more commonly used methods of studying (e.g., re-reading, highlighting, and taking notes) is the level of discomfort or unease that students often report. This feature of the testing effect is actually what some researchers consider to be one of the benefitting mechanisms that leads to increased gains in memory retention (Pyc & Rawson, 2009; Vaughn & Kornell, 2019). As a result, TE studies are situated in a broader body of work that looks at learning processes that require effort and may not lead to immediate gains, such as desirable difficulties (E. L. Bjork & Bjork, 2011), productive struggle (Hiebert & Grouws, 2007; Warshauer, 2011), productive failure (Kapur, 2008), and exploratory/discovery learning (Alfieri et al., 2011; DeCaro & Rittle-Johnson, 2012).

All these phenomena involve putting students in uncomfortable situations where they are asked to work with material that is just above their level of comprehension with the goal of inducing them to struggle with the material. It is believed that in this act of struggle, students are exploring multiple potential solutions and connections, many of which lead to dead ends, but in the end allow them to learn from the many fruitless paths that were taken. It is often this struggle that leads students to report feeling uncomfortable or as if they are not learning. Studies looking at student's judgements of learning indicates that students actually have a better sense of whether they know something or not after engaging with a testing effect task (Kornell & Son, 2009). This may be in part due to the elevated difficulty causing one to reevaluate their own knowledge, whereas those who re-study by re-reading a passage before taking a test on it, for example, are more likely to report feeling confident in part because re-reading gives them a false sense of understanding the material. While there are benefits to the TE that suggest all students should be adopting it as a method to study and learn, its ability to make students feel as though they are not learning may be problematic for students who report above normal levels of stress at school due to experiences with racial discrimination.

Testing Effect and Latinx Students

Although the TE has been consistently demonstrated in several studies, most research on the TE has been conducted with homogenous samples of convenience. This over-reliance on convenience and W.E.I.R.D. (western, educated, industrialized, rich, and democratic) samples (Henrich et al., 2010; Rad et al., 2018) has resulted in studies on student learning and cognition that attempt to generalize findings to an increasingly diverse U.S. population while based mostly on middle-class and European-American individuals (Sears, 1986). While this approach seems to assume that cognitive processes are universal, research in education and neuroscience has demonstrated that the social context affects the brain and cognitive development (May, 2011; Morrison et al., 2019).

Considering the TE in particular, although repeated replications of the TE (see Carpenter, 2012; Roediger & Butler, 2011; Roediger & Karpicke, 2006b for reviews) have led to the recommendation that testing can be a powerful method for learning, these studies have mostly ignored the role that social and individual factors may have on the learning process which may make testing more beneficial for some students but not all. This is particularly evident for Latinx students when in addition to the aforementioned studies on microaggressions, acculturative stress, and minority stress, studies have demonstrated that they face particular issues around stereotype threat (Guyll et al., 2010) and standardized testing (Contreras, 2005). Therefore, suggesting that testing is a suitable study method for all students can be problematic without studies that first assess how different groups of students perform on a TE task.

Testing Effect and EEG

Most of the TE research has involved the use of behavioral measures. However, while these methods provide measurable outcomes that indicate learning has occurred (e.g., accuracy, forgetting), it is not possible to measure cognitive processes associated with this TE by observing behavior alone. Using EEG methods, which yield measures of brain activity during task performance, make it possible to examine covert cognitive processes – including how the test items are processed and how participants respond when they are making a mistake – in real time.

ERP studies using similar testing paradigms have demonstrated the merit and feasibility of combining a TE task with EEG to measure cognitive processes of attention and working memory (Bai et al., 2015; Ernst & Steinhauser, 2012; Mangels et al., 2012; Pastötter & Bäuml, 2016; Rosburg et al., 2015). For example, attention during studying can be indexed by examining the P300, a well-researched ERP component associated with attentional processes and updating working memory (Polich, 2011). In addition, it is also possible to explore what happens when students make errors. These errors are particularly interesting when examined using ERP methods, as there have been well documented individual differences in response to errors – as indexed by the Error Related Negativity (ERN) – that are linked to anxiety (Moser et al., 2013). Stereotype threat has been found to affect how people attend to feedback after having made an error by measuring the Feedback Related Negativity (FRN; Mangels et al., 2012). Participants in stereotype threat conditions are more likely to have larger FRN amplitudes when attending to negative feedback after having made an error, therefore making the FRN a component of interest when examining how race-based stressors might affect performance on the testing effect task.

Some cognitive processes are more appropriately measured continuously, allowing for an examination of how processes change across intervals of time. In this investigation, alpha band oscillations, which occur in the range of 8 to 12Hz, were used to measure fluctuations in attention. Studies have indicated that lower alpha power is indicative of increased attentional processes (Klimesch, 2012; Lenartowicz et al., 2019; Ray & Cole, 1985). Examining alpha power during learning and working memory tasks can provide insight into covert attentional processes that may be associated with learning and performance on the task.

Thus, by examining EEG components when individuals are engaged in a TE task and related cognitive tasks, it is possible to explore questions specific to the role that factors like anxiety and race-based stress might have in undermining student performance.

The Current Study

The overarching goal of this study is to determine if and how experiences with microaggressions, acculturative stress, and minority stress as reported by Latinx college students, relates to student's ability to learn using a testing effect task and to neural correlates of covert cognitive processes such as attention and working memory. This study assessed stress due to racial discrimination using three different surveys on acculturative stress, minority stress, and microaggressions and related those findings to behavioral/neural measures of cognition and learning as measured by the testing effect task, a flanker task, and a spatial working memory task.

Study aims included:

Study Aim 1: Replicate the Testing Effect with Latinx and White College Students

The first study aim is to replicate the testing effect with a sample of Latinx college students. The testing effect has been replicated many times over several decades by cognitive psychologists, however reported sample demographics are largely homogenous. Here the testing effect is examined in ethnic minority college students who are not well represented in previous research.

Hypothesis 1: The testing effect's long history of replicability suggests that there will be evidence of a testing effect with fewer test condition items being forgotten than study condition items during a 1-week delay recall test.

Study Aim 2: Relate Race-based Stressors to Testing Effect Performance

The second study aim is to assess a relation between the level of stressors Latinx college students report (i.e., microaggressions, acculturative stress, and minority stress) and their performance on the testing effect task. Much like stereotype threat, there is evidence suggesting that an individual, or combination of, stressor(s) can lead to underperformance on a demanding task.

Hypothesis 2: It is hypothesized that students who report higher levels of microaggression, acculturative stress, and/or minority stress will perform worse on the TE task than students who rate lower on the stressors.

Study Aim 3: Relate Race-based Stressors to Correlates of Cognitive Processes

The third study aim is to assess whether there are differences on cognitive control tasks associated with students' ratings of microaggressions, acculturative stress, and/or minority status stress.

Study Aim 3.1: First, associations between behavioral measures of attentional shifting/error monitoring and working memory (as indexed by performance on a flanker and spatial working memory task; SWM) and race-based stressors will be examined.

Hypothesis 3.1: It is hypothesized that students with higher ratings of microaggressions, acculturative stress, and minority stress will exhibit diminished attentional shifting, and working memory as measured by behavioral correlates.

Study Aim 3.2: The second sub aim is to assess if neural correlates of cognitive control are associated with students' ratings of race-based stress.

Hypothesis 3.2: It is hypothesized that participants with higher ratings of microaggressions, acculturative stress, and/or minority stress will have larger FRN amplitudes after incorrect responses in the TE task and smaller P300 amplitudes when presented with learning feedback. Similarly, we would expect to see corresponding small P300 amplitudes in the SWM high load trials for individuals with high race-based stress and high alpha power during feedback trials of the TE task and high load encoding and maintenance of the SWM task.

Study Aim 4: Relate Neural Correlates of Cognition to Testing Effect Task Performance

The fourth study aim is to assess if neural correlates of cognition measured during the TE task are associated with behavioral measures of performance in the same task.

Hypothesis 4: Participants with larger FRN, smaller P300 amplitudes and higher alpha activity as elicited by feedback to an incorrect response on the TE task will have lower accuracy and slower reaction times on the testing effect task.

Methods

Overview

To examine the relation between self-reported demographic, social, and cultural factors and Latinx and White students' cognition and learning, survey measures and experimental tasks were used. During the experimental tasks, EEG data were also collected. Participation involved in-person visits, and online surveys/testing.

Participants

Conducting an exact power analysis using a predetermined specific effect size requires a strong empirical basis (Miller & Yee, 2015). As the questions posed in this study are novel, this basis is not available for all aspects of the present study. However, sample size was determined from power analyses using G*Power (Faul et al., 2007) software, following the traditional recommendation of 0.80 for power at alpha = 0.05 and analyses were computed for a medium and a large effect size. Considering a regression-based data analytic plan, it is estimated that a sample size of 74 is adequately powered to detect medium effects under each research aim (as is routinely seen in similar EEG and cultural research).

The final sample consisted of 59 Latinx and 49 White undergraduate college students recruited from UCLA. Participants were recruited via flyers, email, and word of mouth. Participants recruited via the Psychology department's subject pool received 2 course credits (n = 65) and participants recruited via flyers, emails, and word of mouth were compensated with $20 cash ($n$ = 43). Potential participants were screened based on the following criteria: they had to be an undergraduate student at UCLA, they self-identified as being White or Latinx, and they had to have no knowledge of the Japanese language. These criteria were assessed using a questionnaire.

One hundred and eight undergraduate students (M_{age} = 19.8 years, SD = 2.3) from UCLA who met the screening criteria gave signed consent to participate in the study (See Table 1 for participant demographics). All 108 participants included in the final sample completed the survey, however sample size varied across individual survey measures and tasks. Participant data were retained for survey analyses if they answered all the questions for the survey of interest and were retained for EEG analysis if they met all inclusion criteria for the task of interest (see Electrophysiological Tasks and Measures and Appendix G for further details on exclusion criteria for experimental tasks and see Table 4 for sample sizes per survey scale).

General Procedure

Participants who met the screening criteria initially scheduled a 2-hour appointment to attend a laboratory located on campus. The day before the appointment participants were sent either a reminder email or text message with details about when and where to attend their appointment and how to prepare for the EEG recording. Upon arrival, participants were given a verbal summary of the consent document (IRB #20-000374) and given as much time as they needed to read the document and ask any questions. The assessment began once participants provided signed consent.

Before beginning assessment tasks, participants were outfitted with 32 Ag/AgCl scalp electrodes using the Brain Products actiCAP attached to an actiCHamp amplifier (actiCHamp, Brain Products GmbH, Gilching, Germany). EEG setup took 15 to 20 minutes. Participants completed 3 experimental tasks – including a Flanker task, the Testing Effect task, and the Sternberg Spatial Working Memory task– which were counterbalanced and opportunities to take short breaks were given in-between tasks.

Following the completion of the experimental tasks, participants were asked to complete the survey portion of the study on Qualtrics using the same computer they used for the previous tasks. While participants were completing the surveys, research assistants began removing the EEG cap and electrodes for cleaning and disinfection. Participants who were compensated with cash were then asked to sign up for a brief 10-minute follow-up appointment 5 to 7 days later and were given the opportunity to ask any questions that they had about the study but were told that they would receive a full debrief during the follow-up appointment. Participants who were compensated with course credit were told they would receive an email with a link 5 days later and they would have 2 days to complete the study at home from their own computers. Approximately 5 to 7 days later, participants compensated with cash returned to the lab to take a surprise retention test of the Testing Effect task and a brief survey. Upon finishing they were fully debriefed, given the opportunity to ask questions, and were compensated for their participation. Participants who were compensated with course credit completed the surprise retention test and a brief survey on their home computers. They were debriefed with a written statement and told they could email the researcher with any additional questions they might have.

Table 1

Participant Demographics (N = 108)

Variable	Latinx (n = 59)	White (n = 49)	Total		Group Difference
	Freq.	Freq.	Freq.	%	
Gender					$X^2(3, 108) = 0.70, p = .873$
Male	14	10	24	22.2	
Female	43	36	79	73.1	
Questioning	1	1	2	1.9	
Non-conforming/Queer	1	2	3	2.8	
Family Income (US$)					$\mathbf{X^2(7, 107) = 65.88, p < .001}$
Less than 15,000	5	1	6	5.6	
15,000 to 34,999	13	0	13	12.1	
35,000 to 49,999	14	0	14	13.1	
50,000 to 74,999	12	1	13	12.1	
75,000 to 99,999	6	4	10	9.4	
100,000 to 199,999	8	14	22	20.6	
200,000 to 299,999	1	12	13	12.1	
300,000 or more	0	16	16	15	
Parents' Education (n = 214)					$\mathbf{X^2(9, 108) = 62.36, p < .001}$
Less than elementary	13	0	13	6.1	$\mathbf{X^2(10, 106) = 65.24, p < .001}$
Elementary	14	0	14	6.5	
Middle/Junior HS	25	1	26	12.1	
HS, no diploma	15	1	16	7.5	
HS or GED	19	3	22	10.3	
Some college, no degree	6	5	11	5.1	
Trade/technical/vocational training	0	3	3	1.4	
Associate degree	3	4	7	3.3	
Bachelor's degree	13	31	44	20.6	
Advanced degree (e.g., Master's, MD, Ph.D., etc.)	5	50	55	25.7	
Other/Unknown	3	0	3	1.4	
Home language other than English					$\mathbf{X^2(1, 108) = 51.66, p < .001}$
Yes	56	14	70	64.8	
No	3	35	38	35.2	
Immigration Generation Status					$\mathbf{X^2(5, 108) = 58.38, p < .001}$
1st	14	0	14	12.9	
2nd	38	12	50	46.3	
3rd	5	2	7	6.5	
4th	2	14	16	14.8	
5th	0	18	18	16.7	
International	0	3	3	2.8	
Year in School					$X^2(3, 108) = 6.10, p = .107$
Freshman	18	21	39	36.1	
Sophomore	13	16	29	26.9	
Junior	15	6	21	19.4	
Senior	13	6	19	17.6	
Transfer Student					$X^2(1, 108) = 3.38, p = .066$
Yes	14	5	19	17.6	
No	45	44	89	82.4	
First Generation Student					$\mathbf{X^2(1, 108) = 61.69, p < .001}$
Yes	47	2	49	45.4	
No	12	47	59	54.6	

Note: Most participants reported education level of two parents. Frequencies and percentages for Parents' Education reflect total number of parents. Group comparison for parents' education level includes comparisons for each parent.

Survey Measures

The present section will provide an overview of all the survey measures used in the study. Each measure will be described with a brief review of its origin, existing statistics of internal consistency for populations like this study's sample, and when appropriate analyses examining internal consistency and constructs within the current sample.

Demographic and Background Questions

Demographic information regarding participants' age, gender, race/ethnicity, and measures to assess socioeconomic status were collected. In addition, background information on participants' family/home (e.g., language spoken, immigration generation, and parent occupation/income) and education (e.g., year in school, transfer status, and first-generation status) were also collected (see Table 1).

Race-based Stress Measures

Prior to the investigation, decisions were made to use the race-based stress measures for all participants, knowing that some of the survey items may not apply to the experiences of White college students. This was an intentional choice informed by Causadias et al. (2018) that demonstrated that psychological research tends to underemphasize the role of culture in the behavior of White participants and overemphasize culture among racial/ethnic minorities. Therefore, this choice was made in the attempt to avoid assuming that White college students do not experience race-based stress. This decision led to some challenges that are discussed in the current section and in the Limitations section.

Three surveys were used to assess three aspects of race-based stress: acculturative stress, minority status stress, and microaggressions, including:

The Multidimensional Acculturative Stress Inventory (MASI). The MASI is a 36-

item scale that was developed to measure acculturative stress in individuals of Mexican origin (Rodriguez et al., 2002). Originally constructed and validated with a community sample of adults, the MASI has since also been validated with a group of adolescent children of Mexican descent between the ages of 14 and 20 (Rodriguez et al., 2015). Use with different samples has yielded different constructs through exploratory factor analysis, including four constructs from Rodriguez et al. (2002): Spanish competency pressures, English competency pressures, pressure to acculturate, and pressure against acculturation. Factors had high internal consistency (Cronbach α between .77 and .93) and the entire scale had an overall Cronbach α of .90 (Rodriguez et al., 2002).

This study administered the MASI scale to both Latinx and White participants. To my knowledge, this is the first study to attempt measuring how White participants respond to acculturation stress measures that are designed for a Latinx population. This resulted in some challenges when aggregating scores and identifying constructs. The first obvious challenge was that not all White participants responded to all the items. This was in part because some items did not apply to them (e.g., Spanish comprehension pressure items) and some constructs were difficult to adapt for White participants. For instance, items representing the pressure to acculturate construct would be defined as feeling stress due to pressure to adopt the majority culture, which in the case for most White participants would be adopting their own culture. Similarly, the pressure against acculturation items would imply that White participants feel stress from pressure to adapt to cultures other than the dominant culture. The author believes that these concepts are not irrelevant to the experience of being White in the United States. Similar constructs such as white fragility and replacement conspiracy theories among white supremacists all highlight fears and concerns over interactions with members from different racial groups and the threat of a majority-minority population in the United States (DiAngelo, 2016; Krosch et al., 2022; Obaidi et al., 2022). However, items in the MASI are more deeply rooted in the day-to-day experiences of Latinx individuals in the United States while the analogous concepts for White participants seem to be rooted in their racial identity and how that affects their perception of their environment. Therefore, when comparing Latinx and White participants on the MASI scale, only the mean score of all 18-items that both groups answered (α = .91; see Appendix A) are used for analyses.

The Minority Status Stress Scale (MSSS). The MSSS is a 33-item measure of ethnic minority stress that was originally validated with a sample of African American, Latinx, and Filipino college students attending a predominantly White university (Smedley et al., 1993). Two studies assessing minority stress with a diverse sample of Latinx college students have demonstrated that not all 33-items in the original scale pertain to Latinx students (Arbona & Jimenez, 2014; Rodriguez et al., 2000). Rodriguez et al. (2000) used a short version consisting of 19 items (see Appendix B) which loaded onto 3 factors: college climate (α = .88), ethnic discrimination (α = .90), and intra-ethnic group pressures (α = .80).

Using statistical procedures suggested by Brown (2015), confirmatory factor analysis using the same model from Rodriguez et al. (2000) examined model fit with three index categories: exact fit, parsimony correct index, and comparative fit index. Overall test of exact fit was significant, $X^2(149) = 316.30$, $p < .001$, with an SMSR = 0.04. Parsimony correct index RMSEA = 0.10, 90% CI [0.08, 0.12], and comparative fit index CFI = 0.92. Due to the relatively small sample size and evidence suggesting that model fit statistics can vary according to many different parameters (Beauducel & Wittmann, 2005) the following cutoffs were used to determine model fit: SMSR < .08, CFI > .90, and RMSEA < .10 (Bentler, 1990; Brown, 2015;

MacCallum et al., 1996). These results suggested a mediocre fit with the same three factor model from Rodriguez et al. (2000). Further analysis revealed very high significant correlations between the three factors (.85 to .90, $p < .001$), therefore means scores of the entire scale were used for analyses.

Racial Microaggression Scale. A new microaggression measure was designed for this study (see Appendix C). Participants were first presented with definitions of racial microaggressions from Sue & Constantine (2007) with examples of racial microaggressions sourced from items on the Racial and Ethnic Microaggression Scale (REMS; Nadal, 2011). The first of two questions asked participants to answer yes or no if they had ever experienced microaggressions during their time at UCLA. Participants who responded in the affirmative were then asked how often they experienced microaggressions using a 6-point Likert type scale (see Appendix C). For the purposes of analysis, responses from both items were collapsed into a single item such that participants who responded "no" to the first item were coded as a zero on the scale from the second item.

Covariates

To demonstrate that race-based stressors are associated with performance and neural activity during the experimental tasks, other forms of stress and related symptomology were statistically controlled during regression analyses. The following scales were used to measure stress related to college, anxiety, and depression.

College Stress Scale (CSS). As a control for other forms of stress undergraduate students are likely to experience, participants responded to the 18-item *College Stress Scale* with three subscales measuring academic stress, social stress, and financial stress. The CSS was originally validated with a sample of Latinx college students and had strong internal consistency in all three subscales (α = .80 to .85; Rodriguez et al., 2000).

Confirmatory factor analysis with the existing sample indicated that the previous three factor model was not a good fit for this sample, $X^2(132) = 246$, $p < .001$, with an SMSR = 0.07, RMSEA = 0.09, 90% CI [0.07, 0.11], CFI = 0.86. Mean scores of the entire scale were used for analyses (α = .89).

Brief Symptom Inventory (BSI) for Anxiety and Depression. The BSI-18 is a shortened version of the 53-item BSI (Derogatis, 1993) that focuses on screening for the most common mental health disorders (i.e., anxiety, depression, and somatization; see Asner-Self et al., 2006 for more detailed history).

Anxiety and Depression subscales from the BSI-18 were used in this study. Previous studies have suggested using a composite score from the entire scale, however confirmatory factor analysis indicated statistically good model fit for this sample with a two-factor model of anxiety and depression, X^2 (53)= 65.7, p = 0.11, SRMR = 0.04, CFI = 0.98, RMSEA = 0.05, 90% CI [0.00, 0.08], so both scales were used (Asner-Self et al., 2006). Cronbach α was .86 for each subscale.

Electrophysiological Procedures

EEG Recording

Continuous EEG was recorded with 32 Ag/AgCl scalp electrodes using the Brain Products actiCAP attached to an actiCHamp amplifier (actiCHamp, Brain Products GmbH, Gilching, Germany). The electrooculogram generated from blinks was recorded using two Ag/AgCl electrodes placed below each eye. Impedances were maintained below 50kOhms. During recording, data was referenced to electrode site Cz, digitized at 500Hz, and offline re-referenced to the average of all electrode sites (exclusive of face electrodes) at 500 Hz.

EEG Preprocessing

All preprocessing procedures were conducted using BrainVision Analyzer software (*BrainVision Analyzer*, 2019). Recorded EEG data from all tasks and participants were first visually inspected for bad channels and artifacts. Any between task data was then marked for exclusion during analysis (e.g., data recorded before and after the task, practice trials, etc.). For ERP analyses, an infinite impulse response (IIR) band pass filter with a high cutoff of 0.1Hz and low cutoff of 30Hz was applied to waveforms. For spectral power data, an IIR band pass filter with a high cutoff of 1Hz and a low cutoff of 48Hz was applied. All data was re-referenced from Cz to an average of all scalp electrodes. Extreme artifacts were then automatically rejected according to the following criteria: a) the absolute voltage range for any individual channel was greater than 500µV and b) activity was less than 0.5µV. Independent component analysis was used to identify and correct for ocular movement artifacts.

ERP Preprocessing. After ocular correction, data was segmented into response and stimulus segments as described per individual tasks below. All segments with artifacts were automatically rejected according to the following criteria: a) gradient larger than 50µV/ms, b) an absolute voltage range greater than 300µV, and c) activity less than 0.5µV. Response and stimulus segments were baseline corrected in the range of -200ms to -100ms and -100ms to 0ms, respectively. Segments were then averaged, and low pass filtered at 30Hz.

Spectral Power Preprocessing. After ocular correction, data was segmented. All segments with artifacts were automatically rejected according to the following criteria: a) gradient larger than 50µV/ms, b) an absolute voltage range greater than 300µV, and c) activity less than 0.5µV. Total spectral power was calculated using a continuous wavelet transform with the following criteria: a) morlet complex, b) minimum and maximum frequencies of 3Hz and

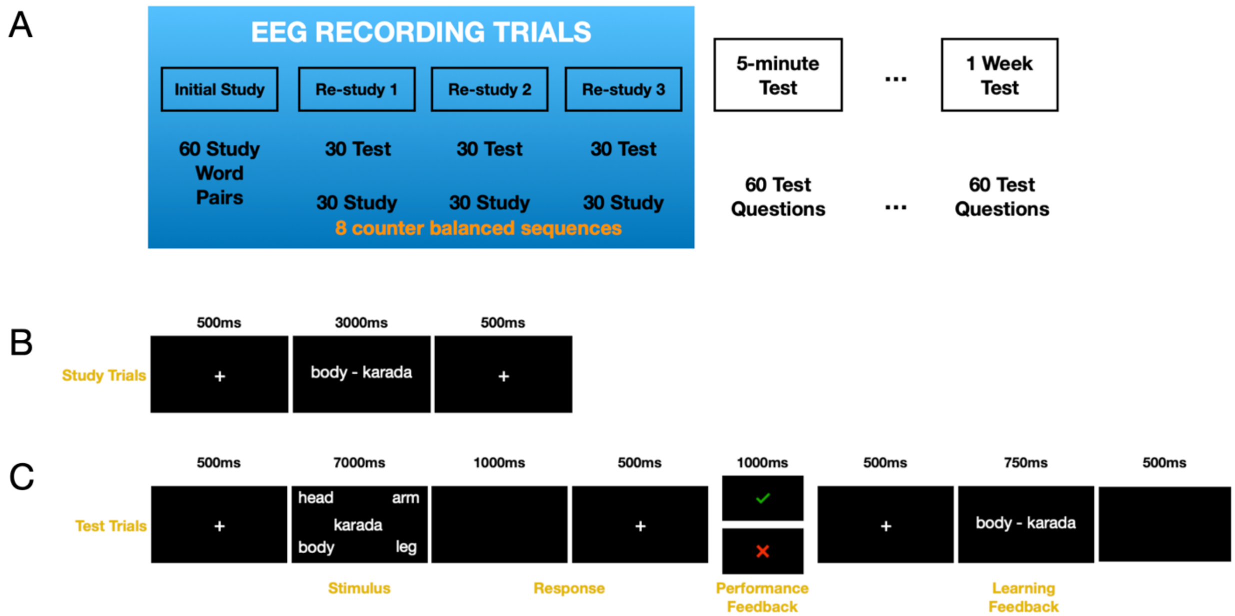
48Hz, respectively, c) 45 logarithmic frequency steps, d) instantaneous amplitude wavelet normalization, and e) a morlet parameter of 5. Spectral power was logarithmically normalized relative to a reference interval (SWM: -1200ms to -600ms before target stimulus, TE Task: -1000 to -500ms before performance feedback stimulus) and reported in decibels (dB).

Experimental Tasks

Testing Effect Task. The testing effect task that will be used in this study is based on similar experimental tasks that have been successfully used in EEG studies to elicit ERPs (Ernst & Steinhauser, 2012; Mangels et al., 2012; Pastötter & Bäuml, 2016). Implementing a within-subjects design, all participants completed both the testing condition (experimental) and study condition (control). Both the testing condition and study condition were comprised of a sequence of slides presenting either 30 English-Japanese word translation multiple-choice questions (testing condition) or 30 English-Japanese word pairs (study condition) from a list of 60 English-Japanese translation pairs. The task included four blocks in which every participant received all 60 English-Japanese word pairs during the first block to study and were asked to remember them. During blocks 2, 3, and 4, participants received 30 English-Japanese word pairs to study and 30 English-Japanese word translation questions to test themselves on. Study and test condition blocks were counterbalanced and word stimuli presentation within each block was randomized. Test condition stimulus slides had a Japanese word in the center and four English translations words at each corner (1 correct response and 3 distractors). Participants were asked to select the correct English translation word by pressing a corresponding keyboard key. Study condition stimuli slides had the correct English-Japanese translation pairs and participants were asked to memorize them. Word pairs and corresponding word questions were presented an equal number of times in each condition (see Figure 1 for task design and parameters).

Figure 1

Testing Effect Task Design and Parameters



Note. A) Testing Effect task design illustrating which blocks were counter balanced and when EEG was recorded. B) Study trial parameters. C) Test trial parameters.

Word Selection. The 60 English-Japanese word pairs were selected from a list of the most frequently used Japanese nouns (Tono et al., 2013). All selected nouns are among the top 1000 most frequently used nouns in the Japanese language. Only nouns representing concrete objects (e.g., car) that could be easily imagined by the participants were selected (see Appendix D for list of selected nouns). Any Japanese nouns considered to be a part of the American English lexicon (e.g., sushi, samurai, manga, etc.) were not included in the final list of 60-word pairs. One hundred and eighty unique distractor nouns were randomly selected from this same list of frequently used nouns (see Appendix E). Only nouns representing concrete objects were kept in the final set of distractors.

Parameters and Procedure. A testing condition trial was comprised of 1) a pre-stimulus fixation point that lasted 500ms, 2) a stimulus with the Japanese word in the center and four

English answer choices (3 distractors and 1 correct response) at each corner of the slide for 7000ms or until the participant made a response, 3) a 1000ms blank screen, 4) a 500ms fixation point, 5) performance feedback with a green check mark for correct responses and a red cross for incorrect responses in the center of the slide for 1000ms, 6) a 500ms fixation point, 7) learning feedback with the correct English-Japanese word pair for 750ms, and 8) and a 500ms intertrial blank screen. Using the EEG recording software, time markers were placed when the participant made a response to the question, when the performance feedback was presented, and when the learning feedback was presented to assess and analyze the ERN, FRN, and P300 at those time points.

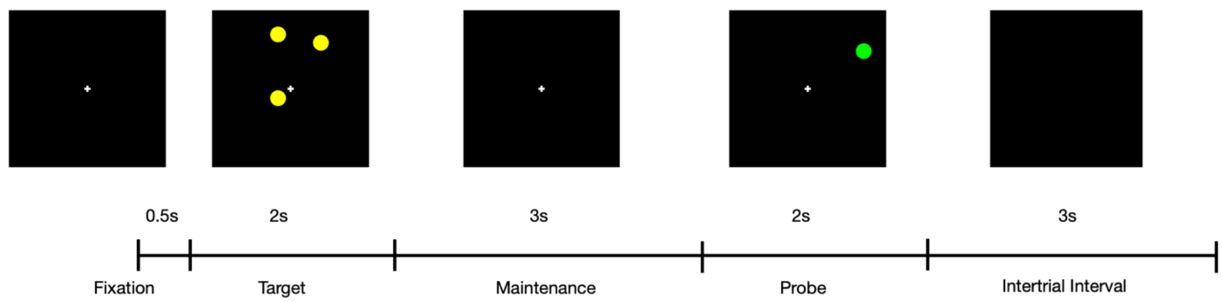
A study condition trial was comprised of 1) a pre-stimulus fixation point that lasted for 500ms and 2) a stimulus with an English-Japanese word pair for 3000ms. Time markers were placed when the stimulus was presented to assess and analyze the P300 and alpha power.

After a 5-minute distraction period, participants took an immediate retention test in which they were presented with test questions for all 60 word pairs to assess how many they remembered. One week later, participants were asked to come in again for a brief follow up in which they were given a surprise re-test or were sent a link to take a follow-up survey in which they were presented with a surprise re-test at home.

Flanker Task. Version 0.9 of the ERP Core Eriksen Flanker task was used to measure the ERN as a correlate of error monitoring and attentional shifting (Kappenman et al., 2020). The task consisted of 10 blocks of 40 trials each for a total of 400 trials. Participants were presented with stimuli consisting of 5 arrows and asked to fixate on the center arrow (see Kappenman et al., 2020 for task illustration). During each trial participants pressed a left button if the center arrow was pointed left and a right button if the center arrow was pointed right. During congruent trials all arrows were facing the same direction and during incongruent trials the 4 surrounding arrows were facing opposite of the center arrow. Each trial was comprised of 1) a 1200 to 1400ms fixation point/response window and 2) a stimulus slide with the arrows for 200ms. The ERN was measured when the participant committed a response.

Figure 2

Spatial Working Memory Task Parameters



Spatial Working Memory Task. A computerized variation of the "Sternberg" spatial working memory task was used to measure participants' working memory (Lenartowicz et al., 2021). One advantage of using this task is that it is already designed to capture both behavioral and neural correlates of working memory. Each trial contained 1) a fixation point for 500ms, 2) an encoding display of 1, 3, 5, or 7 yellow dots for 2000ms, 3) a fixation point for a duration of either 3000ms, and 4) a probe display of a single dot for 2000ms in either a location that matched one of the previous dots or in a location that did not match any of the previous dots (see Figure 2 for illustration of task parameters). Participants were instructed to press a left or right button on a keyboard if the probe location matched or did not match a location from the target/encoding display, respectively (Lenartowicz et al., 2016). There was a total of 40 trials, 10 for each load, and trials were presented in a random order to each participant. In the analyses, data from trials with 1 and 3 dots were combined to create a low load variable and trials with 5 and 7 dots were combined to create a high load variable.

Electrophysiological Measures

This study's approach is to use well studied EEG measures to assess the covert cognitive processes underlying the TE. Since no studies to the author's knowledge have looked at these specific EEG components in relation to a TE task, the study is employing the flanker and SWM tasks to provide convergent evidence that the neural activity occurring in the TE task is similar to neural activity from well-studied experimental tasks.

Table 2

Summary of Neural Correlates of Cognitive Processes

EEG Measure	Cognitive Process	Experimental Task
Error Related Negativity (ERN)	error monitoring, attention shifting	Flanker, Testing Effect
Feedback Related Negativity (FRN)	feedback appraisal	Testing Effect
P300	attention, working memory updating	Testing Effect, SWM
Alpha Spectral Power	attention, working memory	Testing Effect, SWM

Error Related Negativity (ERN)

The ERN will be measured in both the flanker task and the TE task and serves as a correlate for error monitoring and selective attention. Although the ERN is typically measured in speeded response tasks like the flanker, examining it in a learning task may provide some insight into processes related to error detection and adjusting one's performance when attempting to learn something new. The mean amplitude for the ERN was measured from -50ms to 50ms following a committed error and relative to a pre-response baseline of -200ms to -100ms at the midline electrode sites for both the Flanker and Testing Effect task.

Feedback Related Negativity (FRN)

The FRN is similar to the ERN in that it is a negative waveform that occurs when making an error, but it is locked to presentation of a feedback stimulus. A larger FRN amplitude after receiving negative feedback on the TE task may be an indication of feedback appraisal as part of a reinforcement learning system that is used to alert a person when they have made a correct or incorrect response (Gehring et al., 2011). The FRN mean amplitude was measured from 150ms to 250ms following performance feedback at midline electrode sites during the Testing Effect task.

P300

The P300 has been implicated with several cognitive processes including attention, working memory, and memory updating (Polich, 2007). For the TE task, measuring the P300 when receiving learning feedback (i.e., given the correct answer after having made an incorrect response) may provide insight into knowing if participants are attending to the learning feedback and if that attention results in improved performance on the task. Evidence from the SWM task indicates that with increased working memory load the P300 amplitude increases during encoding (Lenartowicz et al., 2021). Larger P300 amplitudes on the TE task may suggest similar cognitive processes are occurring as in the SWM task. The P300 mean amplitude was measured from 350ms to 500ms after learning feedback for the Testing Effect task and 350ms to 500ms after target stimuli in the Spatial Working Memory task, both at midline electrode sites.

Alpha

Alpha oscillations are a continuous waveform that can be measured across an extended period. Unlike ERPs which are time-locked to a specific event, alpha can be measured over several seconds or minutes. This makes it ideal for measuring processes that do not have a definitive beginning or ending such as attention and working memory. A decrease in alpha power in the SWM task is associated with increased attentional processes that tend to occur when attending to tasks that require a lot of working memory capacity (Lenartowicz et al., 2016, 2021). Similarly, lower alpha oscillations during negative performance feedback and learning feedback in the TE task may indicate an increase in attentional processes.

Alpha spectral power was quantified as the mean normalized power values of EEG oscillations in the 8 to 12Hz frequency band occurring over parietal/occipital electrodes. Alpha was recorded during negative performance feedback and learning feedback for incorrect response trials in the Testing Effect task and during high load encoding and maintenance trials in the Spatial Working Memory task. EEG time series recordings were divided and averaged into appropriate time windows.

EEG Exclusion Criteria

Appendix F provides a summary of the number of participants that were excluded in the analyses of ERPs and alpha power for each task. In tasks that involved quantifying an EEG measure from incorrect responses, a minimum of 6 incorrect responses were required to maintain statistical power (Boudewyn et al., 2018). Similarly, any participant that was missing more than 25% of trials either due to participant behavior or presence of artifacts in the EEG data were also excluded from analyses (Luck, 2014).

Confirming Presence of ERP Components

Before conducting analyses to address the proposed research questions, the presence of individual ERP components was confirmed. The identification of ERPs was done using parameters defined in the existing literature, visually inspecting ERP waveforms derived from this investigation, and repeated measures analyses of variance (RMANOVA). Specifically, for each component of interest a 7 (midline electrode sites: Fz, FCz, Cz, CPz, Pz, POz, Oz) x 2 (task

Table 3
Confirming Presence of ERP's of Interest

Measures	Sum of Squares	df	Mean Square	F	p	η^2_p
Flanker Task						
ERN and CRN						
Correct vs Incorrect Response	440.57	1	440.57	135.07	**< .001**	0.57
Residual	332.70	102	3.26			
Electrode Site Mean Activity	3084.15	1.86	1655.33	103.49	**< .001**	0.50
Residual	3039.86	190.04	16.00			
Response x Electrode Site	1619.49	1.75	926.99	131.10	**< .001**	0.56
Residual	1260.06	178.20	7.07			
Testing Effect Task						
ERN and CRN						
Correct vs Incorrect Response	0.10	1	0.10	0.10	0.753	0.00
Residual	98.05	101	0.97			
Electrode Site Mean Activity	339.21	1.88	180.37	29.99	**< .001**	0.23
Residual	1142.54	189.94	6.02			
Response x Electrode Site	0.93	1.73	0.54	0.20	0.788	0.00
Residual	469.83	174.91	2.69			
FRN - Performance Feedback						
Positive vs Negative Feedback	1.65	1	1.65	1.58	0.211	0.02
Residual	105.17	101	1.04			
Electrode Site Mean Activity	750.56	1.55	483.87	34.43	**< .001**	0.25
Residual	2202.06	156.67	14.06			
Feedback x Electrode Site	7.98	1.82	4.38	1.70	0.188	0.02
Residual	473.92	183.96	2.58			
P300 - Learning Feedback						
Correct vs Incorrect Response	12.82	1	12.82	7.27	**0.008**	0.07
Residual	177.99	101	1.76			
Electrode Site Mean Activity	4515.41	1.99	2269.43	119.48	**< .001**	0.54
Residual	3816.88	200.96	18.99			
Response x Electrode Site	96.84	2.09	46.35	10.73	**< .001**	0.10
Residual	911.24	211.04	4.32			
Spatial Working Memory Task						
P300 - Target Stimulus						
Low Load vs High Load	0.01	1	0.01	0.01	0.907	0.00
Residual	105.26	100	1.05			
Electrode Site Mean Activity	1384.16	1.62	853.42	31.34	**< .001**	0.24
Residual	4416.31	162.19	27.23			
Target Load x Electrode Site	69.70	1.95	35.77	10.63	**< .001**	0.10
Residual	655.41	194.84	3.36			

conditions) RMANOVA was conducted (see Table 3 for a summary of results). All RMANOVAs violated the assumption of sphericity; therefore, all degrees of freedom are reported using Greenhouse-Geisser estimates. Any ERP components that cannot be confirmed will be used in the analyses, but findings will be interpreted with caution.

Flanker Task

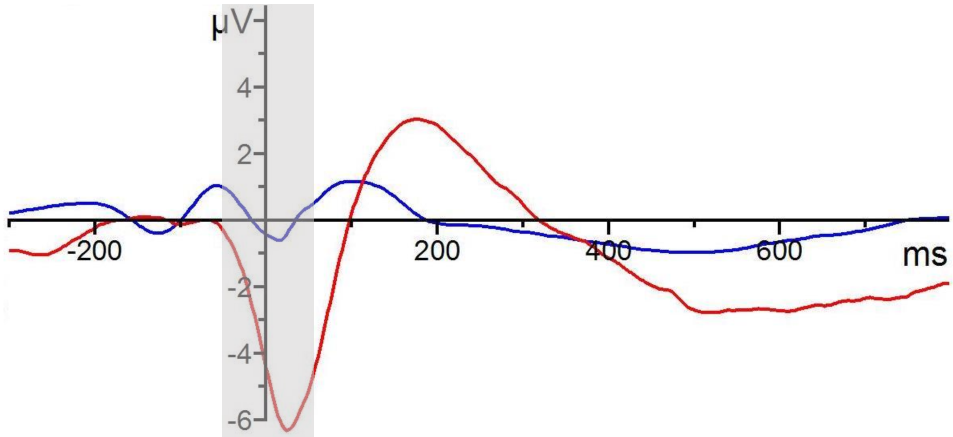
ERN. As seen in Figure 3, visual inspection of the grand-average ERP waveforms for correct and incorrect response trials reveals an enhanced negative deflection around time of error commission relative to correct responses at FCz. To assess significant differences in mean amplitude at electrode sites and between response types, a 7 (electrode sites) x 2 (correct and incorrect responses) RMANOVA was conducted. There was a significant main effect of response type, indicating that mean amplitude was more negative during incorrect responses. A significant main effect of electrode site indicated that mean amplitude was significantly more negative at frontal sites Fz and FCz than at more posterior sites. Both main effects are consistent with the literature on ERNs (Gehring et al., 2011). The interaction between response type and electrode site (see Figure 4) is significant, indicating that amplitude at frontal sites was on average more negative during incorrect responses on average relative to amplitude at posterior sites. Mean amplitude during incorrect responses was maximally negative at FCz, M = -3.87uV, SE = 0.26, 95% CI [-4.39, -3.37], and the difference between incorrect and correct responses was also maximal at FCz as confirmed by post hoc analysis, $t(102) = 15.54$, $p_{tukey} < .001$. These findings confirm the presence of the ERN at FCz and all further analyses examining the ERN during the Flanker task will use mean amplitude at FCz.

Testing Effect Task

ERN. Visual inspection of the grand-average ERP waveforms for correct and incorrect response trials revealed that there was no difference in mean amplitude (see Figure 5). Results from the 7 (electrode sites) x 2 (correct and incorrect response) RMANOVA indicated a significant main effect of electrode site, suggesting that mean amplitude was more negative at frontal sites and more positive at more posterior sites (see Figure 6), but there was no significant

Figure 3

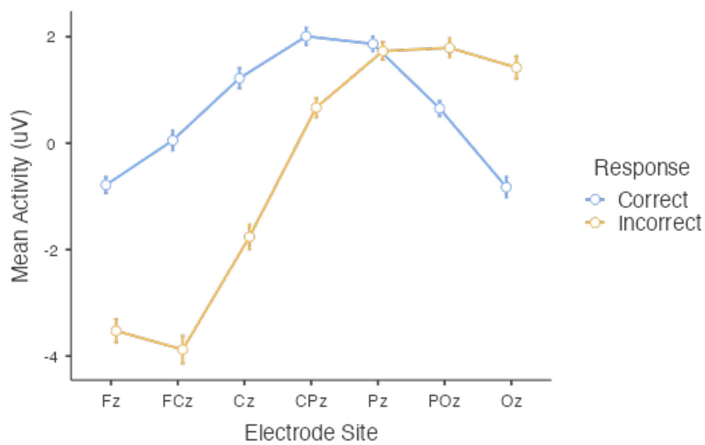
Flanker Response Trials Grand Average ERP at FCz



Note. Correct trials are in blue and incorrect trials are in red.

Figure 4

Flanker Response Trials Mean Amplitude Across Midline Electrodes

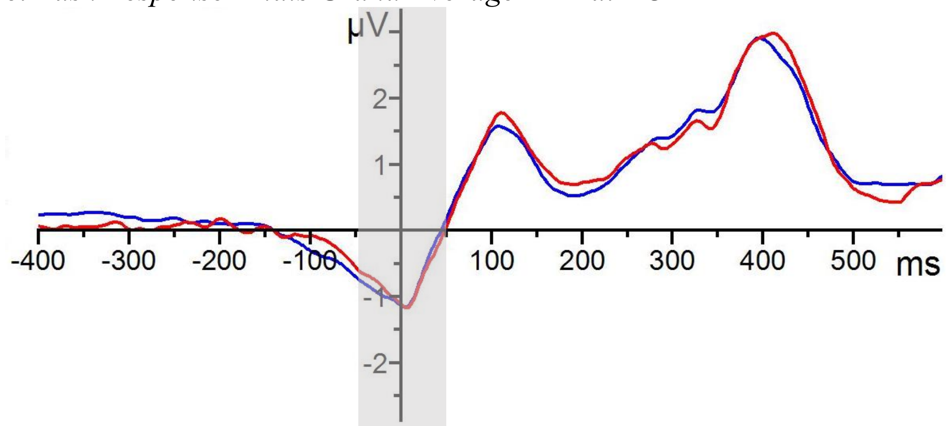


Note. Electrode cites along the x-axis listed in order from front to posterior beginning with Fz. Bars are standard error.

main effect of response type and no significant interaction. Negative deflection shortly after error commission during incorrect trials is maximally negative at FCz, (*M* = -0.73, *SE* = 0.13, 95% CI [-0.98, -0.48], indicating the potential presence of an ERN. However, because there is no significant difference between amplitudes on error relative to correct trials, it is difficult to conclude whether the ERN is present. Accordingly, any significant findings pertaining to the ERN during the TE task will be interpreted with caution.

Figure 5

Testing Effect Task Response Trials Grand Average ERP at FCz

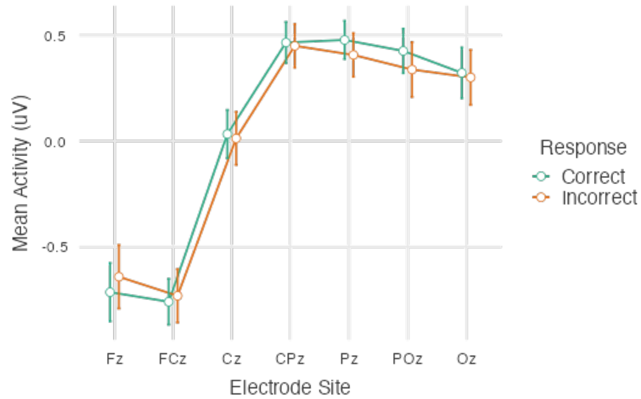


Note. Correct trials are in blue and incorrect trials are in red.

Performance Feedback FRN. The Feedback Related Negativity is a negative deflection observed approximately 200ms after receiving negative feedback (see Figure 7). However, as can be seen in Figure 5 a more prominent negative deflection occurs after positive feedback, which is inconsistent with the literature (Glazer et al., 2018). The 7 (electrode site) x 2 (positive and negative performance feedback) RMANOVA yielded a significant main effect for electrode sites but no main effect for performance feedback and no significant interaction (see Table 3). Electrode site main effect indicates that mean amplitude was on average significantly more positive at frontal sites than posterior sites (see Figure 6). Similar to the ERN findings above, there was not adequate evidence to suggest that the FRN is present so any significant findings will be interpreted with caution.

Figure 6

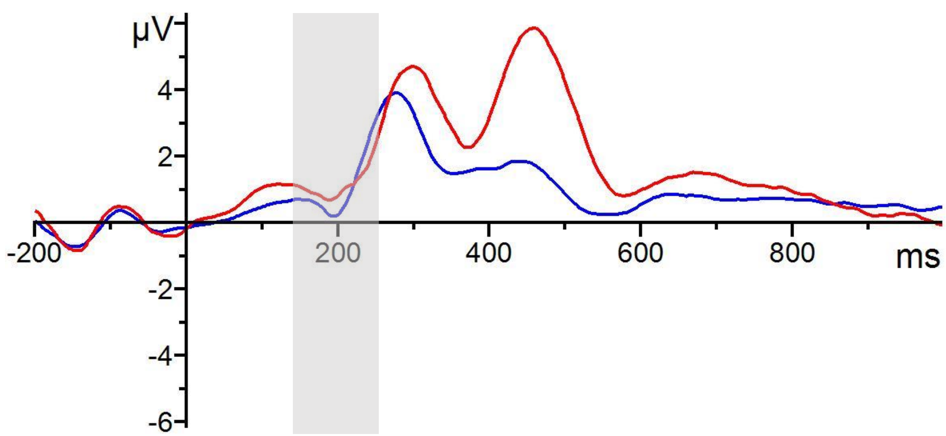
Testing Effect Task Response Trials Mean Amplitude Across Midline Electrodes



Note. Electrode cites along the x-axis listed in order from front to posterior beginning with Fz. Bars are standard error.

Figure 7

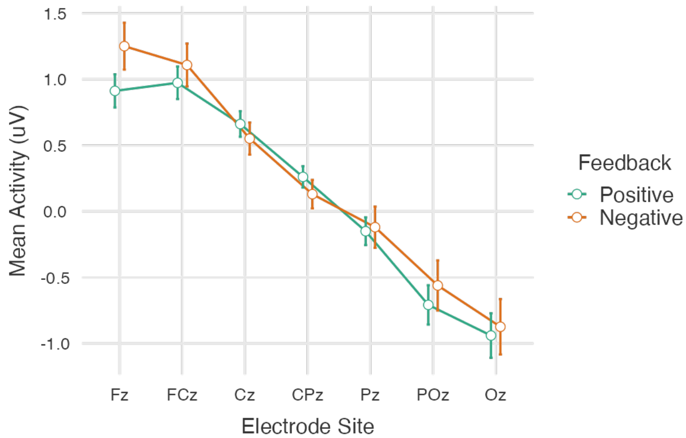
Testing Effect Task Performance Feedback Trials Grand Average ERP at FCz



Note. Positive feedback trials are in blue and negative feedback trials are in red.

Figure 8

Testing Effect Task Performance Feedback Trials Mean Amplitude Across Midline Electrodes

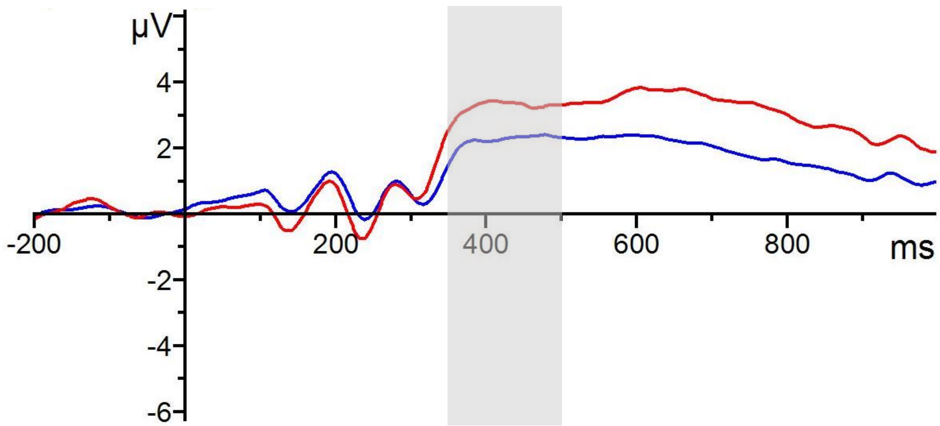


Note. Electrode cites along the x-axis listed in order from front to posterior beginning with Fz. Bars are standard error.

Learning Feedback P300. Visual inspection of the mean amplitude after receiving learning feedback indicates a prominent positive deflection occurring at parietal sites peaking at approximately 400ms (see Figure 9). The 7 (electrode sites) x 2 (learning feedback after correct and incorrect responses) RMANOVA yielded significant main effects and a significant interaction (see Table 3). Main effect of electrode site indicated that mean amplitude was more positive at central/parietal sites and the main effect of learning feedback indicated that mean amplitude was more positive for learning feedback received after an incorrect response. The interaction indicates that mean amplitude was more positive for learning feedback received after an incorrect response at parietal to occipital sites and more negative at frontal sites when compared to learning feedback received after correct responses (see Figure 10). Mean amplitude was maximal at Pz for learning feedback received after an incorrect response, $M = 3.26$, $SE = 0.20$, 95% CI [2.86, 3.66], and the difference between feedback after correct and incorrect responses was also maximal at Pz, $t(101) = -6.86$, $p_{tukey} < .001$. Further analyses on the P300 for learning feedback in the testing effect task will use mean amplitude from Pz.

Figure 9

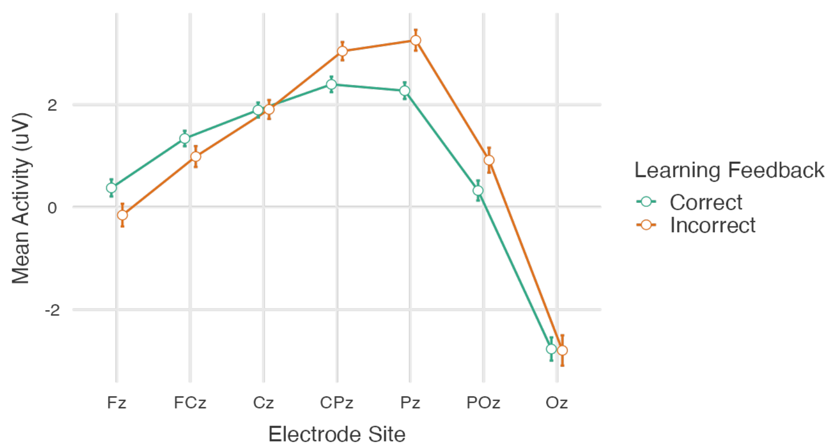
Testing Effect Task Learning Feedback Trials Grand Average ERP at Pz



Note. Feedback during correct response trials is in blue and during incorrect response trials is in red.

Figure 10

Testing Effect Task Learning Feedback Trials Mean Amplitude Across Midline Electrodes



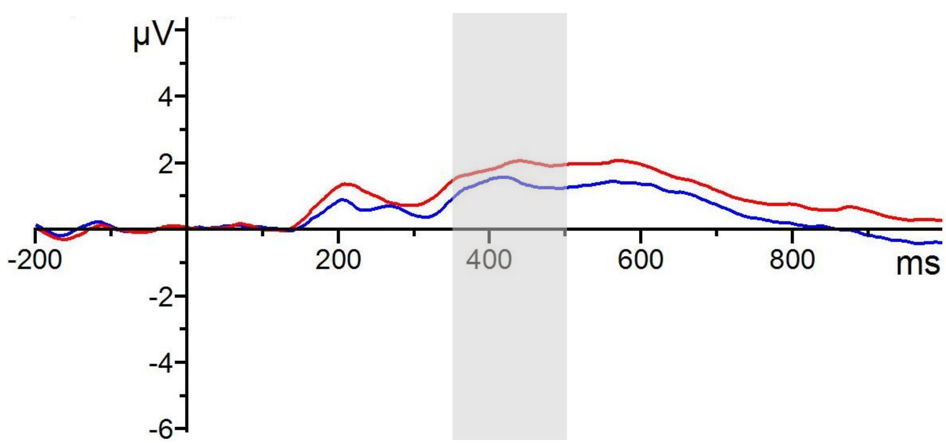
Note. Electrode cites along the x-axis listed in order from front to posterior beginning with Fz. Bars are standard error.

Spatial Working Memory Task

Target Stimulus P300. Visual inspection of mean amplitude at central and parietal sites after presentation of the target stimulus shows a positive deflection occurring at approximately 300ms (see Figure 11). The 7 (electrode sites) x 2 (low and high load target stimulus) RMANOVA yielded a significant main effect of electrode site and a significant interaction, but there was no significant main effect of target stimulus load (see Table 4). Electrode site main effect indicated that mean amplitude was more positive over central sites and the significant interaction indicated that mean amplitude for high load targets was higher at central sites but lower at frontal and posterior sites when compared to low load targets (see Figure 12). Maximal mean amplitude was at CPz, $M = 1.85$, $SE = 0.16$, 95% CI [1.53, 2.17], and difference between low and high load target mean amplitude was only significant at CPz, $t(100) = -4.38$, $p_{tukey} = .002$. Further analyses on the P300 for target stimuli in the SWM task will use mean amplitude from CPz.

Figure 11

SWM Task Target Stimulus Trials Grand Average ERP at CPz



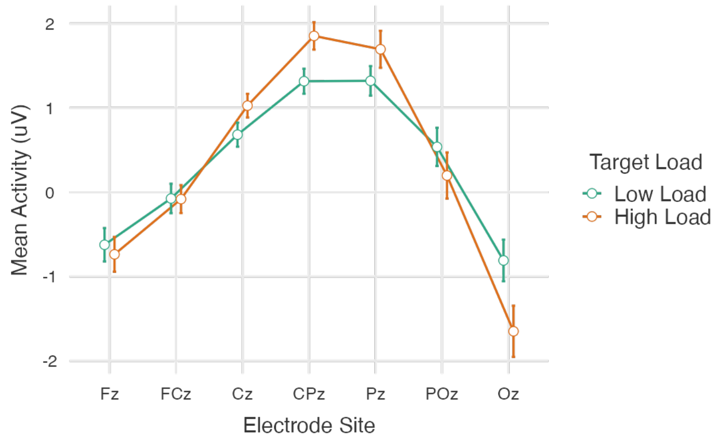
Note. Low load trials are in blue and high load trials are in red.

Summary

Using a multi-step process that included visual inspection of ERP data, examining existing literature, and statistical analysis, the following components were included in the final

Figure 12

SWM Task Target Stimulus Trials Mean Amplitude Across Midline Electrodes



Note. Electrode cites along the x-axis listed in order from front to posterior beginning with Fz. Bars are standard error.

analyses. For the Flanker task, mean amplitude at FCz from -50ms to 50ms relative to participants' incorrect response was used to quantify ERN. For the Testing Effect task, there was insufficient evidence for the ERN but mean amplitude relative to incorrect responses from -50ms to 50ms will be used and interpreted with caution. Similarly, it is not very clear that an FRN can be quantified for performance feedback during the Testing Effect task, but mean amplitude will be examined at FCz from 150ms to 250ms relative to negative performance feedback and interpreted with caution. The P300 for the Testing Effect task was quantified using mean amplitude at Pz from 350ms to 500ms relative to learning feedback presentation. For the Spatial Working Memory task, the P300 was quantified using mean amplitude at CPz from 350ms to 500ms relative to presentation of target stimuli.

Results

An overview of descriptive statistics and group comparisons for demographics, survey scales, task behavior, and relevant EEG components will be followed by results organized by study aims. All statistical analyses were conducted using R packages via the Jamovi graphical user interface (The Jamovi Project, 2022).

Descriptive Results

Demographic Comparisons

Gender. The size of gender groups was skewed towards female, which is in line with demographic characteristics of the participant pool from which they were recruited. Specifically, participants consisted of Psychology students, and more than half were Latinx students. University of California statistics for the Fall 2022 quarter, for example, shows that social sciences are comprised of approximately 68% female undergraduate students and among Latinx students in the social sciences, approximately 72% identify as female (University of California, 2023). In addition, among Latinx college students, researchers have recently highlighted the trend of fewer Latinx males in higher education (Saenz & Ponjuan, 2009). There was no significant difference in gender distribution between the Latinx and White participant sample groups for this study (see Table 1).

Family Income. Family income was significantly different between the Latinx and White participant groups (see Table 1). More Latinx participants reported a family income below $75,000 per year and White participants reported more family incomes of $100,000 or greater per year.

Parents' Education. Participants were asked to report education level on up to two parents. There were group differences for both parent one and two (see Table 1). Overall, more Latinx participants reported that both parents had an education level of HS Diploma/GED or below while White participants reported that their parents had a bachelor's degree or an advanced degree.

Home Language and Generation Status. More Latinx students reported that they spoke a language at home other than English and that they had immigrated to the United States or have immigrant parents (1st and 2nd generation status) when compared to White participants (see Table 1).

College Status. There were no group differences in terms of year in school and transfer student status. More Latinx participants reported that they were first-generation college students than White participants (see Table 1).

Table 4

Survey Descriptive Statistics and Group Comparisons

	Latinx		White		Total			
Scales	n	Mean (SD)	n	Mean (SD)	N	Mean (SD)	Group Difference	Cohen's d
MASI (18-items)	59	1.09 (0.67)	48	0.43 (0.30)	107	0.79 (0.63)	**$t(105) = 6.28, p < .001$**	**1.22**
Microaggressions	59	1.71 (1.47)	49	0.18 (0.63)	108	1.02 (1.39)	**$t(106) = 6.50, p < .001$**	**1.26**
Minority Status Stress Scale (MSSS)	57	2.52 (1.02)	49	0.63 (0.46)	106	1.64 (1.25)	**$t(104) = 11.90, p < .001$**	**2.31**
College Stress Scale (CSS)	58	3.12 (0.72)	49	2.54 (0.63)	107	2.85 (0.74)	**$t(105) = 4.43, p < .001$**	**0.86**
BSI Anxiety	58	2.75 (0.95)	49	2.26 (0.91)	107	2.53 (0.96)	**$t(105) = 2.72, p = .008$**	**0.53**
BSI Depression	55	2.41 (1.00)	48	2.10 (0.87)	103	2.27 (0.95)	$t(101) = 1.69, p = .093$	0.33

Note. Cohen's d small ~ 0.2, medium ~ 0.5, large > 0.8

Survey Descriptive Statistics and Group Comparisons

Descriptive statistics for the survey scales are reported in Table 4, including the full sample, disaggregated by race/ethnicity, and group comparison statistics. Differences between Latinx and White participants on the survey scales were statistically significant on all except the BSI Depression scale. On average Latinx participants in this sample reported experiencing more acculturative stress, more microaggressions, more minority stress, experience more college related stress, and experience more anxiety when compared to White participants.

Experimental Tasks Descriptive Statistics and Group Comparisons

Descriptive statistics for the experimental tasks are reported in Appendix G, including the full sample, disaggregated by race/ethnicity, and group comparison statistics.

Flanker Task. Descriptive statistics from performance on the Flanker task indicate that the task functioned as expected. On average, reaction times were faster for incorrect responses than correct responses, $t(107) = -26.28$, $p < .001$, Cohen $d = -2.53$, and slower for incongruent trials in comparison to congruent trials, $t(107) = -26.07$, $p < .001$, Cohen $d = -2.51$. Accuracy was also as expected with congruent trials garnering more accurate responses than incongruent trials, $t(107) = 16.09$, $p < .001$, Cohen $d = 1.55$. Overall, very few trials were missed ($M = 2.56$, $SD = 6.15$) indicating that participants were engaged and responding within the appropriate time window.

Latinx and White participants performed significantly different on reaction time on correct responses and incorrect responses, accuracy on congruent trials, reaction time on both congruent and incongruent trials, reaction time on correct congruent and correct incongruent trials, reaction time on incorrect incongruent trials, and on the number of missed trials. In summary, White participants were faster, more accurate, and missed fewer trials than Latinx participants.

Testing Effect Task. Participants performed very well on the retention tests for the TE task. As expected, performance was better on the 5-min delay retention test than on the 1-week delay retention test, $t(98) = 13.87$, $p < .001$, Cohen $d = 1.39$. Somewhat unexpected was that accuracy on test condition items at the 5-min delay test was better than accuracy on the study condition items, $t(98) = 4.57$, $p < .001$, Cohen $d = 0.46$. Participants responded more slowly to study condition items than test condition items on both retention tests ($t_{5\text{-}min}(98) = -5.89$, $p < .001$, Cohen $d = -0.59$; $t_{1\text{-}wk}(98) = -7.58$, $p < .001$, Cohen $d = -0.76$) suggesting a potential link to ease of retrieval for test condition items.

Significant differences were observed between Latinx and White participants on all measures of the TE task, with the exception of reaction time on study condition items during the 5-min delay retention test. Overall, White participants were more accurate and had faster reaction times on all measures of the TE task compared to Latinx participants.

Spatial Working Memory Task. In line with previous research, performance on the low load trials was more accurate, $t(104) = 17.77$, $p < .001$, Cohen $d = 1.73$, and faster, $t(104) = -24.12$, $p < .001$, Cohen $d = -2.35$, than on high load trials. Significant differences between groups were observed in the accuracy of high load trials and reaction time for incorrect low load trials with White participants being more accurate and faster compared to Latinx participants.

EEG Measures Descriptive Statistics and Group Comparison

ERPs. ERP descriptive and group comparison statistics are outlined in Table 7. Average ERP amplitudes for each group were significantly different for the TE P300 when receiving learning feedback after an incorrect response and high target P300 on the SWM task. All significantly and marginally different ERPs were more positive for the White participants, which in the case of the P300 corresponds with increased amplitudes and more neural activity.

Alpha Power. Alpha power descriptive and group comparison statistics are outlined in Table 8. No significant group differences in alpha spectral power for either the TE task or the SWM task were observed.

Summary

Taken together, descriptive analyses and group comparisons indicate that there are significant group differences on survey scales, experimental task performance, and neural correlates of cognition. Next, study aims were addressed to assess associations between race-based stress and both testing effect performance and correlates of cognition.

Table 5

EEG Components' Descriptive Statistics and Group Comparisons

	Latinx					White					Total						
				Range					Range					Range			
	n	Mean	SD	Min.	Max.	n	Mean	SD	Min.	Max.	N	Mean	SD	Min.	Max.	Group Difference	Cohen's d
Flanker ERN	55	-3.94	2.71	-11.40	1.53	48	-3.81	2.53	-10.55	0.81	103	-3.88	2.61	-11.40	1.53	$t(101) = -0.23, p = .815$	-0.05
TE ERN	54	-0.83	1.13	-3.50	2.31	48	-0.63	1.43	-3.61	2.77	102	-0.73	1.28	-3.61	2.77	$t(100) = -0.79, p = .433$	-0.16
TE FRN Neg Perf FB	54	0.94	1.41	-2.36	3.86	48	1.30	1.86	-2.17	5.98	102	1.11	1.64	-2.36	5.98	$t(100) = -1.12, p = .264$	-0.22
TE P300 Inc Learn FB	54	2.32	1.45	-0.57	5.14	48	4.32	2.11	0.39	8.52	102	3.26	2.05	-0.57	8.52	$\mathbf{t(100) = -5.62, p < .001}$	**-1.12**
SWM P300 High Load Target	55	1.36	1.34	-1.83	4.21	47	2.41	1.75	-1.01	8.40	102	1.84	1.62	-1.83	8.40	$\mathbf{t(100) = -3.41, p < .001}$	**-0.68**
TE Neg Perf FB Alpha	53	-0.63	1.91	-4.35	4.88	47	-1.18	1.52	-4.75	2.70	100	-0.89	1.76	-4.75	4.88	$t(98) = 1.59, p = .116$	0.32
TE Inc Learn FB Alpha	53	-0.63	1.86	-4.28	5.07	47	-1.00	1.61	-5.56	2.76	100	-0.8	1.74	-5.56	5.07	$t(98) = 1.07, p = .287$	0.21
SWM High Load Target Alpha	53	-4.23	1.84	-10.46	-1.50	42	-3.91	2.01	-8.87	-0.81	95	-4.09	1.91	-10.46	-0.81	$t(93) = -0.81, p = .418$	-0.17
SWM High Load Maintenance Alpha	53	-3.14	1.56	-7.78	0.56	42	-3.01	1.87	-6.50	1.43	95	-3.08	1.70	-7.78	1.43	$t(93) = -0.37, p = .710$	-0.08

Note. ERPs are measured in microvolts (μV) and alpha spectral power in decibels (dB).

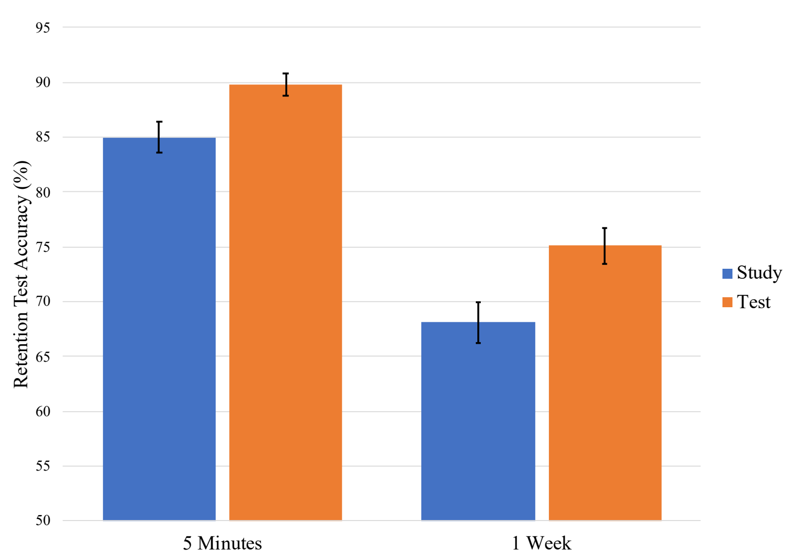
Aim 1: Replicate the Testing Effect with Latinx College Students

Testing Effect Accuracy

Whole Sample. A 2 (test and study conditions) x 2 (retention test delay: 5 minutes and 1-week) RMANOVA was first conducted using the entire sample to assess if there was an overall testing effect. There was a significant main effect for task condition, $F(1, 98) = 48.05$, $p < .001$, $\eta^2_p = 0.33$, indicating that participants on average were able to recall more words in the test condition than words in the study condition across both delay retention tests (see Appendix G for means and Figure 13). There was also a significant main effect for retention test delay, $F(1,98) = 192.49$, $p < .001$, $\eta^2_p = 0.66$, indicating that participants on average forgot more words during the 1-week delay retention test than during the 5-minute delay retention test (see Figure 13). These findings were not supported by a significant interaction of task condition by retention test delay, $F(1, 98) = 3.04$, $p = .084$, $\eta^2_p = 0.03$, indicating that the typical pattern of results observed in testing effect tasks in which recall of the study condition words is higher during more immediate recall and lower during further delayed recall was not observed in this study.

Figure 13

Testing Effect Task Accuracy on Retention Tests



Post hoc analysis confirmed these findings. On the 5-minute retention test words in the study condition were recalled more than words in the test condition, $t(98) = -4.57$, $p_{tukey} < .001$. Findings met expected patterns during the 1-week retention test where words in the test condition were recalled more than words in the study condition, $t(98) = -6.53$, $p_{tukey} < .001$. These findings suggest that participants on average forgot more words in the study condition than in the test condition over the period of one week. Despite the lack of a significant interaction, patterns suggest that testing is still beneficial for long term retention.

Latinx Participants. Findings on the testing effect task were similar for the Latinx sample. The 2 x 2 RMANOVA yielded a significant main effect of task condition, $F(1, 53) = 26.36$, $p < .001$, $\eta^2_p = 0.33$, a main effect of retention test delay, $F(1, 53) = 127.12$, $p < .001$, $\eta^2_p = 0.70$, but no significant interaction, $F(1, 53) = 2.60$, $p = .112$, $\eta^2_p = 0.05$. Latinx participants recalled more words in the test condition than in the study condition during the 5-minute retention test, $t(53) = -2.86$, $p_{tukey} = 0.03$, and during the 1-week retention test, $t(53) = -5.43$, $p_{tukey} < .001$.

White Participants. The testing effect task results were similar for the White sample. The 2 x 2 RMANOVA yielded a significant main effect of task condition, $F(1, 44) = 21.48$, $p < .001$, $\eta^2_p = 0.33$, a main effect of retention test delay, $F(1, 44) = 74.04$, $p < .001$, $\eta^2_p = 0.63$, but no significant interaction, $F(1, 44) = 0.54$, $p = .468$, $\eta^2_p = 0.01$. White participants recalled more words in the test condition than in the study condition during the 5-minute retention test, $t(44) = -4.04$, $p_{tukey} = 0.001$, and during the 1-week retention test, $t(44) = -3.74$, $p_{tukey} = .003$.

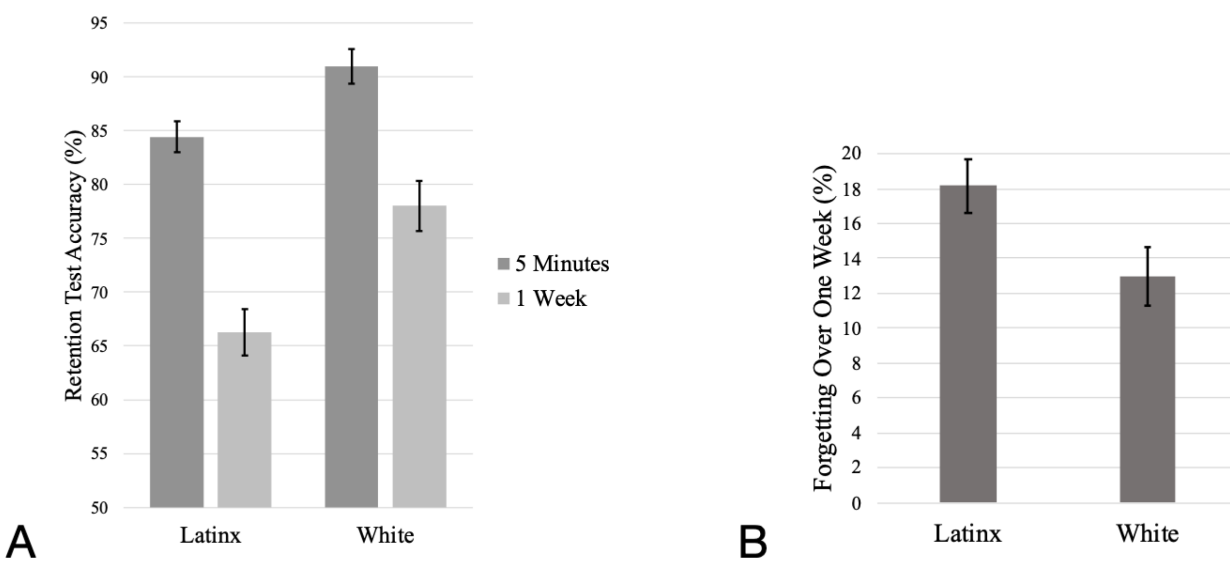
Between Group Effects. To assess if there were any differences between Latinx and White participants on the TE task, a between subjects 2 (task condition) x 2 (retention test delay) RMANOVA was conducted. Findings indicated a main effect of race, $F(1, 97) = 13.67$, $p < .001$,

$\eta^2_p = 0.12$, suggesting there are differences in performance on the TE task between Latinx and White participants. This was supported by a significant interaction between race and retention test delay, $F(1, 97) = 5.38$, $p = .023$, $\eta^2_p = 0.05$ (see Figure 14A), which indicated that on average Latinx participants forgot significantly more words 1-week later when compared to White participants (See Figure 14B).

While the results do not mirror typical findings in TE tasks studies where study condition words are remembered better than test condition words during the immediate retention test and test condition words are remembered better one week later, the results do confirm that test words were remembered better than study words one week later, confirming the benefit of testing as a learning strategy. Therefore, the first hypothesis is confirmed as there is evidence that both the Latinx and White participants benefitted from testing because on average they forgot fewer test condition words during the 1-week retention test.

Figure 14

Testing Effect Performance by Race



Note. A) Accuracy on 5-min delay and 1-week delay retention tests by race. B) Percent of words forgotten after 1 week.

Testing Effect Reaction Time

Reaction time on retention tests may also be an indicator of how well something is learned and can provide more nuance on the potential effects of testing on learning outcomes. A between subjects 2 (test and study condition) x 2 (5-min and 1-week delay) RMANOVA was conducted to assess how reaction time was different for each group and changed over time. Overall, there was a main effect of delay retention test, a main effect of condition, and a three-way interaction between delay, condition, and race.

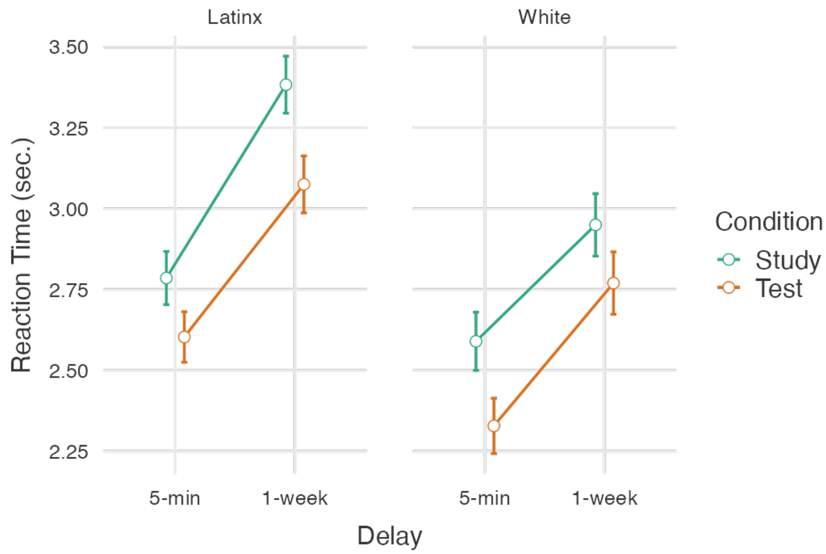
The delay main effect, $F(1, 97) = 71.59$, $p < .001$, $\eta^2_p = 0.42$, indicates that reaction time was significantly faster for both groups during the 5-min delay test than during the 1-week delay test, $t(97) = -8.46$, $p_{tukey} < .001$.

The condition main effect, $F(1, 97) = 71.31$, $p < .001$, $\eta^2_p = 0.42$, indicates that reaction time was significantly faster for both groups when answering questions that were learned in the test condition than words learned in the study condition, $t(97) = 8.44$, $p_{tukey} < .001$.

The three-way interaction between delay, condition, and race, $F(1, 97) = 5.67$, $p = .019$, $\eta^2_p = 0.06$, suggests that the change in reaction time from the 5-min to the 1-week delay retention test was moderated by participants' race (see Figure 15). White participants had faster reaction times overall, $t(97) = 2.84$, $p_{tukey} = .005$, but reaction times for both conditions became slower from the 5-min delay to the 1-week delay retention test, 0.36 seconds slower for study condition and 0.44 seconds slower for test condition. Latinx participants had slower reaction times overall, but reaction times became slower at a faster rate for study condition items (0.59 seconds slower) than test condition items (0.47 seconds) as can be observed in Figure 15. This is supported by post hoc analyses that show that for White participants, the difference between study and test reaction times at the 5-min delay test, $t(97) = 4.73$, $p_{tukey} < .001$, $MD = 0.26$, became smaller at the 1-week delay test, $t(97) = 3.73$, $p_{tukey} = .007$, $MD = 0.18$. The opposite was true for Latinx participants who saw an increase in the difference between study and test reaction times from the 5-min delay test, $t(97) = 3.61$, $p_{tukey} = .011$, $MD = 0.18$, to the 1-week delay test, $t(97) = 7.00$, $p_{tukey} < .001$, $MD = 0.31$.

Figure 15

Testing Effect Reaction Time Three-way Interaction



Summary

Results from accuracy and reaction times on the TE task suggests that both groups benefitted from testing when learning new Japanese vocabulary, but Latinx participants may have benefitted more with regards to reaction time as it suggests that their rate of forgetting for test condition words was slower than their rate of forgetting for study condition words.

Aim 2: Relate Race-based Stressors to Testing Effect Performance

Correlation and Multiple regression analyses were used to assess the predictive power self-reported microaggression, acculturative stress, and minority stress have on performance on a testing effect task. All stress variables will be included in a single model to assess if and how much more race-based stress predicts performance on the TE task.

Bidirectional correlation analyses were conducted to assess the relation between the self-reported race-based stress scales, covariates, and performance TE task (see Table 6 for correlation matrix). Acculturative stress was significantly correlated with test condition item accuracy on the 5-min delay test and reaction time on study condition items on the 1-week delay test. Microaggressions was significantly correlated with accuracy on both condition items at both retention tests and it was not correlated with reaction time on any condition items and on either retention tests. Minority stress was correlated with accuracy on both condition items at both retention tests and with reaction times on test condition items on the 5-min delay test and study condition items on the 1-week delay test. The CSS and anxiety were the only two covariates that correlated with any measures of performance on the testing effect task. All significant correlations were in the expected direction: more stress correlated with less accuracy and slower reaction times on the TE task.

Multiple regression analyses were conducted to assess if the three race-based stress constructs predicted performance on the TE task while controlling for other stressors (CSS: academic, social, and financial), anxiety, and depression (see Table 7 for multiple regression statistics). All multiple regression models were conducted using the same predictor variables and only the outcome variable changed for each model. Results indicate that race-based stressors did predict performance on the TE task above and beyond stressors measured by the CSS and both BSI Anxiety and Depression scales. Accuracy on both the study and test conditions during the 5-minute delay retention test were significantly predicted by minority stress and test condition accuracy during the 5-minute delay retention test was marginally predicted by acculturative stress. All standardized coefficients were negative, indicating that participants who reported higher amounts of minority and acculturative stress were less accurate on the 5-minute delay

Table 6

Correlation Matrix for Testing Effect, Race-based Stressors, and Covariates

	1	2	3	4	5	6	7	8	9	10	11	13	15
1 5-min Test Condition Accuracy	—												
2 5-min Study Condition Accuracy	0.67 ***	—											
3 1-week Test Condition Accuracy	0.70 ***	0.62 ***	—										
4 1-week Study Condition Accuracy	0.62 ***	0.67 ***	0.82 ***	—									
5 5-min Test Reaction Time	-0.48 ***	-0.25 *	-0.46 ***	-0.37 ***	—								
6 5-min Study Reaction Time	-0.39 ***	-0.43 ***	-0.46 ***	-0.48 ***	0.81 ***	—							
7 1-week Test Reaction Time	-0.25 *	0.00	-0.20 *	-0.15	0.64 ***	0.48 ***	—						
8 1-week Study Reaction Time	-0.18	0.02	-0.13	-0.18	0.60 ***	0.51 ***	0.88 ***	—					
9 MASI (18 items)	-0.30 **	-0.11	-0.19	-0.16	0.15	0.10	0.18	0.23 *	—				
10 Microaggressions	-0.31 **	-0.29 **	-0.29 **	-0.31 **	0.14	0.10	0.07	0.17	0.57 ***	—			
11 Minority Status Stress Scale (MSSS)	-0.38 ***	-0.31 **	-0.34 ***	-0.37 ***	0.25 *	0.19	0.17	0.27 **	0.62 ***	0.78 ***	—		
13 College Stress Scale (CSS)	-0.20 *	-0.14	-0.17	-0.18	0.16	0.15	0.16	0.19	0.40 ***	0.45 ***	0.55 ***	—	
15 Anxiety	-0.22 *	-0.13	-0.12	-0.15	0.14	0.06	0.13	0.10	0.35 ***	0.35 ***	0.36 ***	0.55 ***	—
16 Depression	-0.18	-0.03	-0.09	-0.11	0.16	0.06	0.06	0.09	0.33 ***	0.34 ***	0.24 *	0.48 ***	0.68 ***

Note. * p < .05, ** p < .01, *** p < .001

Table 7

Multiple Regression Analyses for Testing Effect Task

DV's	Common Stress						Race-based Stress					
	CSS		Anxiety		Depression		Acculturative Stress		Minority Stress		Microaggressions	
	β	t	β	t	β	t	β	t	β	t	β	t
5-min Test Accuracy	**0.26**	**1.77**†	-0.15	-1.03	-0.13	-0.88	**-0.26**	**-1.97**†	**-0.43**	**-2.25***	0.16	0.95
	$R^2_{adj.}$ = 0.19, F(6, 79) = 4.34***											
5-min Study Accuracy	0.26	1.65	-0.18	-1.2	0.02	0.16	0.04	0.26	**-0.49**	**-2.41***	0.04	0.23
	$R^2_{adj.}$ = 0.09, F(6, 79) = 2.44*											
1-wk Test Accuracy	-0.16	0.99	-0.06	-0.37	-0.07	-0.48	-0.07	-0.47	**-0.37**	**-1.77**†	0.01	0.06
	$R^2_{adj.}$ = 0.07, F(6, 79) = 2.04†											
1-wk Study Accuracy	0.19	1.21	-0.08	-0.56	-0.09	-0.60	0.04	0.26	**-0.41**	**-2.00***	-0.05	-0.30
	$R^2_{adj.}$ = 0.10, F(6, 79) = 2.52*											
5-min Test Reaction Time	0.00	0.01	-0.00	-0.03	0.17	1.07	0.05	0.37	0.30	1.43	-0.15	-0.84
	$R^2_{adj.}$ = 0.03, F(6, 79) = 1.40											
5-min Study Reaction Time	0.01	0.08	-0.09	-0.57	0.17	1.07	0.06	0.38	0.33	1.54	-0.23	-1.24
	$R^2_{adj.}$ = 0.00, F(6, 79) = 1.00											
1-week Test Reaction Time	0.04	0.23	0.09	0.26	-0.02	0.19	0.19	1.32	0.09	0.41	-0.16	-0.83
	$R^2_{adj.}$ = -0.01, F(6, 79) = 0.82											
1-week Study Reaction Time	0.01	0.07	-0.09	-0.56	0.11	0.73	0.19	1.31	0.26	1.23	-0.15	-0.81
	$R^2_{adj.}$ = 0.03, F(6, 79) = 1.48											

Note. † p < .1, * p < .05, ** p < .01, *** p < .001

retention test.

Results were similar for the 1-week delay test but only minority stress significantly predicted performance on study condition items and marginally predicted performance on the test condition items. Standardized coefficients were also negative indicating that more minority status stress was associated with lower accuracy on the 1-week delay test.

For reaction times, neither common stressors nor race-based stressors significantly predicted reaction times during the TE task. Despite significant group differences in reaction time, measures of stress used in this study do not account for any of those differences. Previous group comparisons (see Appendix G) indicate that Latinx participants were overall less accurate on both condition items and during both retention tests. They also reported significantly higher levels of minority and acculturative stress. In summary, evidence gathered from these analyses support the hypothesis that lower performance on the TE task for Latinx participants may be in part a result of high levels of acculturative and minority stress.

Aim 3: Relate Race-based Stressors to Correlates of Cognitive Processes

As in Study Aim 2, correlation and multiple regression analysis were used to assess if microaggressions, acculturative stress, and minority stress relate to behavioral and neural correlates of cognition as measured with the experimental tasks. Mean amplitudes of ERPs of interests and behavioral measures will serve as the outcome variables for the regression models.

Behavioral Correlates of Cognitive Processes

Behavioral measures from both the flanker task and the SWM task were examined as correlates for error monitoring/attentional shifting and working memory, respectively. For the flanker task higher accuracy suggests that participants were able to switch back and forth from congruent and incongruent trials with few errors, post-error slowing is a correlate for a

Table 8

Correlation Matrix for Race-based Stress and Behavioral Correlates of Cognition

	1	2	3	4	5	6	7	8
1 MASI (18 items)	—							
2 Microaggressions	0.57 ***	—						
3 Minority Status Stress	0.62 ***	0.78 ***	—					
4 College Student Stress	0.40 ***	0.45 ***	0.55 ***	—				
5 BSI Anxiety	0.35 ***	0.35 ***	0.36 ***	0.55 ***	—			
6 BSI Depression	0.33 ***	0.34 ***	0.24 *	0.48 ***	0.68 ***	—		
7 Flanker Accuracy	-0.25 *	-0.06	-0.15	-0.20 *	-0.18 †	-0.19 †	—	
8 Flanker Post-error Slowing	-0.06	0.07	-0.03	-0.09	-0.02	-0.09	0.19 †	—
9 SWM High Load Accuracy	-0.09	-0.22 *	-0.18 †	-0.16	-0.22 *	-0.15	0.00	-0.08

Note. † $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$

participants' ability to adjust performance after having committed an error. For the SWM task, accuracy for high load trials was used to assess a participant's working memory processes. Bidirectional correlations were conducted to assess the relation between race-based stressors, control variables, and behavioral measures of cognitive processes from the flanker task and the SWM task (see Table 8). Flanker accuracy was negatively associated with acculturative stress, college student stress, anxiety, and depression. Results indicate that as stressors increased, flanker accuracy decreased. Flanker post-error slowing was not significantly correlated with any of the stress scales. High load accuracy in the SWM task was negatively associated with microaggressions, minority stress, and anxiety. Like the flanker accuracy, increased stress was associated with a decrease in high load accuracy.

Further analysis reveals that after controlling for college stress, anxiety, and depression, only flanker accuracy is significantly predicted by acculturative stress (see Table 9). Both flanker post-error slowing and SWM high load accuracy were not significantly predicted by any of the stress variables when controlling for common sources of stress.

In summary, these results partially support the hypothesis that higher levels of race-based stress are associated with poorer cognitive performance as demonstrated by the significant

Table 9

Multiple Regression Analyses for Behavioral Correlates of Cognition

DV's	Common Stress						Race-based Stress					
	CSS		Anxiety		Depression		Acculturative Stress		Minority Stress		Microaggressions	
	β	t	β	t	β	t	β	t	β	t	β	t
Flanker Accuracy	0.00	0.03	-0.18	-1.23	-0.04	-0.26	**-0.36**	**-2.66****	-0.01	-0.06	0.24	1.40
	$R^2_{adj.}$ = 0.09, *F* (6, 87) = 2.46*											
Flanker Post-erro Slowing	-0.12	-0.80	0.11	0.68	-0.12	-0.79	-0.05	-0.35	-0.11	-0.54	0.23	1.26
	$R^2_{adj.}$ = -0.03, *F* (6, 87) = 0.54											
SWM High Load Accuracy	-0.07	-0.47	-0.12	-0.77	-0.04	-0.24	0.05	0.38	0.05	0.25	-0.21	-1.14
	$R^2_{adj.}$ = 0.01, *F* (6, 84) = 1.20											

Note. † p < .1, * p < .05, ** p < .01, *** p < .001

Table 10

Correlation Matrix for Race-based Stress and EEG Components

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1 MASI (18 items)	—													
2 Microaggressions	0.57 ***	—												
3 Minority Status Stress	0.62 ***	0.78 ***	—											
4 College Student Stress	0.40 ***	0.45 ***	0.55 ***	—										
5 BSI Anxiety	0.35 ***	0.35 ***	0.36 ***	0.55 ***	—									
6 BSI Depression	0.33 ***	0.34 ***	0.24 *	0.48 ***	0.68 ***	—								
7 Flanker ERN	-0.03	-0.04	-0.03	-0.10	-0.07	-0.02	—							
8 TE ERN	-0.04	-0.03	-0.11	-0.19 †	-0.14	0.01	-0.06	—						
9 TE FRN Neg Perf FB	-0.03	-0.03	-0.15	-0.12	-0.04	0.01	-0.17	0.10	—					
10 TE P300 Inc Learn FB	-0.24 *	-0.28 **	-0.39 ***	-0.36 ***	-0.27 **	-0.22 *	0.13	-0.02	0.21 *	—				
11 SWM P300 High Load Target	-0.22 *	-0.35 ***	-0.32 **	-0.28 **	-0.14	-0.27 **	-0.03	-0.12	0.18 †	0.38 ***	—			
12 TE Neg Perf FB Alpha	0.02	0.05	0.16	0.11	0.18 †	0.10	0.03	-0.04	-0.19 †	0.06	-0.04	—		
13 TE Inc Learn FB Alpha	0.11	0.09	0.22 *	0.21 *	0.12	0.14	-0.06	-0.10	-0.17 †	0.02	-0.17 †	0.76 ***	—	
14 SWM High Load Target Alpha	-0.07	0.06	-0.12	0.13	0.00	0.16	-0.07	0.17	0.04	-0.14	-0.30 **	-0.04	0.09	—
15 SWM High Load Maintenance Alpha	0.07	0.07	-0.08	0.13	0.08	0.17	0.05	0.20 †	0.10	-0.06	-0.21 *	0.01	0.09	0.74 ***

Note. † p < .1, * p < .05, ** p < .01, *** p < .001

negative association between flanker accuracy and acculturative stress.

Neural Correlates of Cognitive Processes

To understand the relations between race-based stress and neural correlates of cognition, first bidirectional correlation analyses were conducted to assess the relation between race-based stressors, control variables, and neural correlates of cognitive control during the experimental tasks (See Table 10 for ERP and alpha power correlations).

There were significant associations between all three race-based stressor scales and the TE task P300 when receiving learning feedback after incorrect trials. The same was observed for the SWM task P300 when encoding a high load target. All significant ERP correlations were negative and indicate that with increased acculturative stress, microaggressions, and minority stress, P300 amplitudes are more likely to be smaller or less positive. For alpha spectral power correlations, only minority stress was positively associated with alpha power during the TE task when receiving learning feedback after an incorrect trial. This result suggests that as minority stress increases so does alpha activity.

Multiple regression analyses indicate several marginal and significant relationships between race-based stress and neural correlates of cognition (see Table 11 for statistical results). Specifically, the P300 for learning feedback during incorrect response trials was significantly predicted by a model with all six measures of stress including the three race-based stress scales. Of the three race-based stress scales, only minority status stress was marginally associated with the TE P300 for learning feedback on incorrect trials. The SWM P300 for encoding high load targets was marginally associated with depression, but none of the race-based stressors. As expected, the P300 associations are all in the negative direction, suggesting that as stress increases—both common indicators of stress and race-based stressors—P300 amplitudes decrease or become more negative.

Alpha power for the TE task when receiving negative performance feedback was marginally associated in a positive direction with anxiety. Alpha power when receiving learning feedback after an incorrect response on the TE task was marginally predicted by minority stress in a positive direction. Both models, however, were not significant. The positive associations indicate that alpha power was higher for individuals who experienced more anxiety in the case of negative performance feedback and for those who experienced more minority stress in the case of learning feedback during incorrect trials.

For the SWM task, alpha power during high load target encoding was marginally associated with CSS and microaggressions in a positive direction and significantly associated with minority stress in a negative direction. Alpha during high load maintenance was marginally associated with CSS and microaggression in a positive direction and with minority stress in a negative direction.

In summary, although not all the EEG correlates of cognitive control were significantly associated with race-based stress or other indicators of stress, significant and marginal associations with the P300 and alpha suggest that some race-based stressors may interfere with working memory and attention processes.

Table 11

Multiple Regression Analyses for EEG Components

DV's	Common Stress						Race-based Stress					
	CSS		Anxiety		Depression		Acculturative Stress		Minority Stress		Microaggressions	
	β	t	β	t	β	t	β	t	β	t	β	t
Flanker ERN	-0.10	-0.69	0.05	0.31	-0.02	-0.11	-0.03	-0.20	0.03	0.14	-0.01	-0.03
	$R^2_{adj.} = -0.06$, $F(6, 84) = 0.12$											
TE ERN	-0.19	-1.30	-0.22	-1.39	0.23	1.50	0.08	0.53	-0.08	-0.42	0.06	0.31
	$R^2_{adj.} = 0.01$, $F(6, 82) = 1.12$											
TE Negative Perf. FB FRN	-0.09	-0.57	-0.05	-0.31	0.08	0.67	0.12	0.83	-0.29	-1.44	0.13	0.72
	$R^2_{adj.} = -0.02$, $F(6, 82) = 0.74$											
TE Inc. Learning FB P300	-0.10	-0.69	-0.09	-0.61	-0.09	-0.62	0.04	0.31	**-0.33**	**-1.76†**	0.07	0.38
	$\boldsymbol{R^2_{adj.} = 0.11}$**,** $\boldsymbol{F(6, 82) = 2.76}$*****											
SWM High Load Target P300	-0.13	-0.92	0.19	1.27	**-0.28**	**-1.85†**	0.05	0.34	-0.11	-0.57	-0.18	-1.05
	$\boldsymbol{R^2_{adj.} = 0.10}$**,** $\boldsymbol{F(6, 81) = 2.59}$**†**											
TE Negative Perf. FB Alpha	-0.05	-0.31	**0.27**	**1.70†**	0.00	0.00	-0.18	-1.23	0.32	1.63	-0.17	-0.93
	$R^2_{adj.} = 0.02$, $F(6, 80) = 1.32$											
TE Inc. Learning FB Alpha	0.11	0.75	-0.02	-0.10	0.12	0.77	-0.03	-0.23	**0.38**	**1.91†**	-0.28	-1.52
	$R^2_{adj.} = 0.03$, $F(6, 80) = 1.43$											
SWM High Load Target Alpha	**0.28**	**1.94†**	-0.21	-1.27	0.19	1.22	-0.06	-0.37	**-0.49**	**-2.44***	**0.31**	**1.69†**
	$\boldsymbol{R^2_{adj.} = 0.08}$**,** $\boldsymbol{F(6, 75) = 2.14}$**†**											
SWM High Load Maintenance Alpha	**0.26**	**1.76†**	-0.23	-1.42	0.21	1.33	-0.11	-0.69	**-0.35**	**-1.69†**	**0.33**	**1.77†**
	$R^2_{adj.} = 0.05$, $F(6, 75) = 1.72$											

Note. † p < .1, * p < .05, ** p < .01, *** p < .001

Aim 4: Relate Neural Correlates of Cognition to Testing Effect Task Performance

Correlation and multiple regression analyses were used to assess the association between the neural correlates of cognition and performance measures on the TE task. Mean amplitudes for the P300 and FRN and alpha spectral power during feedback trials served as outcome variables with accuracy and reaction time data from the TE task serving as predictor variables.

Bidirectional correlations were conducted between the neural correlates of cognition and behavioral performance measures from the Testing Effect task. There was only one marginal association between the P300 for learning feedback during incorrect trials and test condition accuracy during the 5-min delay retention test (see Table 12).

Multiple regression analysis revealed a marginal association between alpha power when receiving negative performance feedback and reaction times for study condition items in the 5-min delay retention test, but the regression model was not statistically significant (see Tables 13 and 14 for regression statistics).

In summary, there were no statistically significant associations between neural correlates of cognition and task performance in the TE task.

Table 12

Correlation Matrix for Neural Correlates of Cognition and Testing Effect Task

	1	2	3	4	5	6	7	8	9	10	11	12
1 5-min Test Condition Accuracy	—											
2 5-min Study Condition Accuracy	0.67 ***	—										
3 1-week Test Condition Accuracy	0.70 ***	0.62 ***	—									
4 1-week Study Condition Accuracy	0.62 ***	0.67 ***	0.82 ***	—								
5 5-min Test Reaction Time	-0.48 ***	-0.25 *	-0.46 ***	-0.37 ***	—							
6 5-min Study Reaction Time	-0.39 ***	-0.43 ***	-0.46 ***	-0.48 ***	0.81 ***	—						
7 1-week Test Reaction Time	-0.25 *	0.00	-0.20 *	-0.15	0.64 ***	0.48 ***	—					
8 1-week Study Reaction Time	-0.18 †	0.02	-0.13	-0.18 †	0.60 ***	0.51 ***	0.88 ***	—				
9 TE ERN	0.05	0.04	0.00	0.02	-0.02	0.05	0.01	0.04	—			
10 TE FRN Neg Perf FB	0.01	-0.01	0.00	-0.03	-0.03	0.02	-0.04	0.00	0.10	—		
11 TE P300 Inc Learn FB	0.19 †	0.04	0.13	0.12	-0.11	-0.07	-0.08	-0.15	-0.02	0.21 *	—	
12 TE Neg Perf FB Alpha	-0.04	0.04	0.01	0.02	0.05	-0.08	0.12	0.08	-0.04	-0.19 †	0.06	—
13 TE Inc Learn FB Alpha	0.03	0.08	0.09	0.15	0.04	-0.03	0.01	0.04	-0.10	-0.17 †	0.02	0.76 ***

Note. † $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table 13

Multiple Regression Analyses for Testing Effect EEG Measures and Task Accuracy

DV's	5-min Test Acc.		5-min Study Acc.		1-week Test Acc.		1-week Study Acc.	
	β	t	β	t	β	t	β	t
TE ERN	0.08	0.51	0.02	0.14	-0.09	-0.47	0.03	0.17
	$R^2_{adj.}$ = -0.04, F(4, 90) = 0.13							
TE Negative Perf. FB FRN	0.02	0.13	-0.00	-0.00	0.06	0.32	-0.09	-0.46
	$R^2_{adj.}$ = -0.04, F(4, 90) = 0.06							
TE Inc. Learning FB P300	0.25	1.65	-0.17	-1.17	-0.02	-0.13	0.10	0.55
	$R^2_{adj.}$ = 0.01, F(4, 90) = 1.17							
TE Negative Perf. FB Alpha	-0.14	-0.90	0.11	0.72	0.03	0.14	0.01	0.06
	$R^2_{adj.}$ = -0.03, F(4, 88) = 0.25							
TE Inc. Learning FB Alpha	-0.09	-0.55	0.03	0.17	-0.01	-0.05	0.18	1.00
	$R^2_{adj.}$ = -0.02, F(4, 88) = 0.58							

Note. † p < .1, * p < .05, ** p < .01, *** p < .001

Table 14

Multiple Regression Analyses for Testing Effect EEG Measures and Task Reaction Times

DV's	5-min Test RT		5-min Study RT		1-week Test RT		1-week Study RT	
	β	t	β	t	β	t	β	t
TE ERN	-0.20	-0.96	0.18	1.00	-0.03	-0.14	0.09	0.43
	$R^2_{adj.}$ = -0.03, F(4, 90) = 0.36							
TE Negative Perf. FB FRN	-0.09	-0.46	0.10	0.53	-0.12	-0.53	0.11	0.41
	$R^2_{adj.}$ = -0.04, F(4, 90) = 0.21							
TE Inc. Learning FB P300	-0.18	-0.87	0.11	0.64	0.28	1.26	-0.35	-1.63
	$R^2_{adj.}$ = 0.00, F(4, 90) = 1.00							
TE Negative Perf. FB Alpha	0.25	1.22	**-0.34**	**-1.89†**	0.14	0.60	-0.02	-0.11
	$R^2_{adj.}$ = 0.01, F(4, 88) = 1.30							
TE Inc. Learning FB Alpha	0.23	1.10	-0.21	-1.16	-0.21	-0.92	0.20	0.89
	$R^2_{adj.}$ = -0.02, F(4, 88) = 0.50							

Note. † p < .1, * p < .05, ** p < .01, *** p < .001

Discussion

Cognitive psychologists have demonstrated the effectiveness of testing as a learning strategy, but few of these studies have acknowledged the social context that, in other domains, is often associated with students' ability to study and learn. The goal of this study was to contextualize the types of stress experienced by Latinx students and relate those experiences to cognitive outcomes as measured during a testing effect task.

Results from the current project indicate that both Latinx and White participants performed better on test condition items 1-week after initially learning the Japanese vocabulary, but White participants overall performed more accurately and had faster reaction times than Latinx participants. A main goal of the investigation was to understand the associations between

TE performance and race-based stressors. Latinx participants reported significantly more experience with race-based stressors, however there was a negative association between performance on the TE task and race-based stressors for both Latinx and White participants. When considering the neural correlates of cognitive processes, the P300 and alpha spectral power when receiving learning feedback after incorrect response trials were significantly associated with minority status stress. Lastly, while we see these parallel associations, we did not see these associations within the TE task.

The Social Context for Latinx College Students

Experiencing stress and its detrimental effects is not unique to the Latinx population, however results from this study reinforce findings from previous research indicating that Latinx students contend with stressors related to their ethnic/racial background in addition to stressors related to everyday college experiences (Crockett et al., 2007; Guyll et al., 2010; Minikel-Lacocque, 2013; Rodriguez et al., 2000; Yosso et al., 2009). Results from this study revealed the stark difference in the social context that surrounds Latinx students and White students attending the same university. Latinx students reported experiencing more acculturative stress, minority stress, and microaggressions. These troubling results are compounded by evidence from this study indicating that Latinx students also experience higher levels of common college stressors (e.g., academic, social, and financial) and anxiety.

Other factors that were not examined in this study, such as first-generation status and being a transfer student, may also affect how Latinx college students experience college at a 4-year institution. Future studies will need to examine if the stress related to being a transfer or first-generation student is distinct from race-based stress and relates to cognitive processes.

As UCLA nears its goal of achieving Hispanic-Serving Institution (HSI) status in the next few years, it may seem hopeful that this will lead to contextual changes that support Latinx students, but studies have found that even at institutions where Latinx students are not the minority, such as HSIs, experience with race-based stress still persists (Arbona & Jimenez, 2014; Sanchez, 2019; UCLA HSI Task Force, 2022). Concerted efforts will need to be made to provide effective support to help Latinx students persist towards earning a bachelor's degree.

The Testing Effect and Latinx Students

One of the goals of this study was to examine how Latinx students would perform on a Testing Effect task, considering that no known study has focused on the effectiveness of the testing effect on a group of Latinx students.

As hypothesized, Latinx students did remember more test condition items than study condition items during the 1-week delay retention test, suggesting that Latinx students benefit from the testing effect despite reporting higher levels of stress. In addition, while reaction times for White participants on study and test condition items seem to have converged at the 1-week delay retention test, for Latinx students, reaction times on study condition items slowed down at a quicker rate after 1 week than reaction times on test condition items. This increased efficiency in task performance, as indexed by the reaction time results, might even suggest that Latinx students benefitted more from the testing effect. Reaction time in semantic retrieval tasks like the one used in this study is often equated to ease of retrieval and automaticity; faster reaction times equating to easier and more automatic retrieval processes (Gimbel & Brewer, 2011). This is further reinforced by R. A. Bjork & Bjork's (1992) New Theory of Disuse which posits that the act of retrieving an item from memory both enhances retrieval strength and storage strength. Therefore, repeated retrieval attempts in the testing effect task makes remembering the items easier and hence faster to recall. For Latinx participants, results indicated that the testing effect enhanced both storage strength as evidenced by remembering more test condition items and retrieval strength as evidenced by faster reaction times for test condition items.

While there were similarities in performance across groups, this study also provided evidence that the social context still plays a role in Latinx students' performance. Specifically, in comparison to White participants, Latinx participants remembered fewer Japanese words and had slower reaction times at both time points and forgot more words on the 1-week retention test. These differences in accuracy were predicted by race-based stress, in particular minority status stress, but differences in reaction time were not associated with any stress measure, including college stress, anxiety, or depression. Without direct evidence to address this distinction, the literature on stress and memory allows for some speculation. For instance, studies have demonstrated that memory processes are affected by the temporal proximity of the stressful event (Vogel & Schwabe, 2016). When a stressful event occurs long before memory encoding, the stress is more likely to impair memory formation. When the stressful event occurs shortly before or after encoding, it is likely to enhance memory formation. In addition to temporal proximity, stress that is related to the material that is being learned will enhance memory processes while unrelated stressors are less likely to promote long term retention (Schwabe & Wolf, 2010; Smeets et al., 2007).

Participants in the study were asked to answer questions on minority status stress with reference to their time spent at UCLA, hence these are stressors that likely occurred before the testing effect task. In addition, there is no reason to expect that minority stress is related to the content of the testing effect task. It is possible that minority stress and perhaps other race-based stressors are interfering with encoding processes which would likely lead to lower accuracy on the task. On the other hand, if the testing effect task was stressful for the students, that form of stress was more immediate and related to the task itself, hence it would enhance memory processes taking place during the task and this might be evident in faster reaction times on test condition items. Both Latinx and White participants rate of change on test condition item reaction times across one week were only 0.03 seconds apart, hence it can be said that the testing effect had equal effects on both groups. However, the differences observed between both groups may be attributed to differences in the type of stressors they were experiencing. Overall higher accuracy and faster reactions times for White participants may be a result of lower instances of race-based stress and higher stress due to the testing task itself, respectively. Whereas for Latinx participants higher instances in race-based stress may have resulted in lower accuracy score but stress due to the task resulted in reaction times indicating enhanced retrieval strength.

Researchers have suggested that one way to compensate for the detrimental effects of stress on memory is to use effective learning strategies that can be recalled automatically (Vogel & Schwabe, 2016). The premise here is that even under stressful moments, if people default to learning strategies that are effective, such as testing, then they can counteract those detrimental effects. Whereas testing might be effective at improving retrieval strength, it is often still a very rote process. Using other methods of learning that promote making associations with prior knowledge in conjunction with testing may provide more optimal results. It must also be said that instilling these learning strategies at an early age is necessary to make them more automatic, especially when considering that studies suggest that college students can sometimes be resistant towards using effective learning strategies such as testing (Kornell & Son, 2009). Evidence from this study demonstrates that even if using effective learning strategies may compensate for detrimental effects on learning from stress, it may not be enough to overcome the effects of race-based stress.

EEG and Covert Cognitive Processes

Using neural measures in addition to behavioral performance measures provides another dimension that can capture covert cognitive processes. Studies have demonstrated that behavioral and survey measures can sometimes fail to capture covert processes, especially those that may be difficult for participants to articulate. EEG in particular is well suited for measuring cognitive processes in a TE task as it can provide high temporal resolution and allows for real time measurements.

ERN in the Testing Effect Task

The ERN is an ERP component that is typically found in speeded-response tasks such as the flanker task that was used in this study. An understanding of its functional significance is still debated but major theories suggest that it may signify error detection/comparison, a response conflict, or serve as a signal to improve task performance (Gehring et al., 2011; Weinberg et al., 2012). With these theories in mind, it was expected that errors on the TE task would also elicit an ERN, but this was not the case. Both correct and incorrect responses on the TE task elicited nearly identical negative deflection amplitudes indicating that, at least in terms of brain activity captured when making a response, there was no difference between correct and incorrect responses. In other words, errors were likely not detected.

Looking towards the design of the flanker task and the TE task may provide an explanation for these results. The flanker task requires very simple procedures to make a correct response. Participants must look at the central arrow and press the left or right key on the keyboard when the arrow is pointing left or right, respectively. Because of the simple nature of the task, one might say that participants develop expertise rather quickly, making it possible for them to know when they have committed an error (i.e., press the wrong key). The TE task was not designed to be a speeded-response task, instead it was designed to facilitate the learning of Japanese vocabulary words. Since all the participants in the study were novices in the Japanese language, it was very likely that participants did not know when they made mistakes responding to test condition items until they received feedback telling them that they were incorrect. One would expect that if the study included experts in the Japanese language, then there would be indications of an ERN when they incorrectly answer a test condition item.

Only a few studies have examined expertise using the ERN, but one may provide some insight into this finding. Krigolson et al. (2009) investigated the underlying neural mechanisms of developing perceptual expertise and found that individuals who were classified as low learners because their mean accuracy on the task was below 70% had lower ERN amplitudes than participants classified as high learners (> 70%). In addition, ERN amplitudes for high learners gradually increased, meaning errors they made as they gained more expertise in the task elicited larger ERN amplitudes than errors they made early in the task when they had less expertise. This was not true for low learners. This study supports this speculation about how and when the ERN would appear in a non-speeded learning task and how that relates to learning, but future investigations will be needed to truly assess these associations.

Another interesting parallel between the ERN and the TE task has to do with a theory on the function of the ERN which suggests that it is part of a reinforcement learning system that involves dopamine signals sent to the anterior cingulate cortex to modify performance on the task (Holroyd & Coles, 2002). Students often report feeling like they are not learning when they use testing as a learning strategy (Vaughn & Kornell, 2019). This may be in part due to the hypothesis on expertise. If participants do not know when they have committed an error, then there may be a lack of dopaminergic response that would reinforce the person's ability to learn from their mistake. This lack of reinforcement for making a mistake may lead an individual to feel like they are not learning or making any gains while studying using testing. There is evidence supporting the notion that reduced dopamine response is associated with reduced ERN signals and impaired performance in situations where learning is important (Nieuwenhuis et al., 2002). The biggest take away from assessing the ERN in the TE task is that there remains much to investigate, and future studies should focus on how our ability to monitor performance functions in tasks that are relevant in educational settings.

Minority Stress Interferes with the P300 and Memory Updating

Results indicated a marginal negative association between the P300 when receiving learning feedback after an incorrect response and minority stress, suggesting that because Latinx participants reported significantly higher rates of minority stress, they were also more likely to have smaller P300 amplitudes. The P300 is most commonly associated with attentional processes and subsequent context or memory updating (Hajcak & Foti, 2020; Polich, 2007). When placed into the context of this study's design, the P300 was examined as an indicator that when participants incorrectly answered one of the testing condition items they would attend to the correct answer when it was given as learning feedback and update or re-encode this information so that they could use it on subsequent trials when answering the question again. Evidence from this study suggests that the association between the P300 and minority stress may provide insight into these processes.

Other studies have also investigated the role of the P300 when faced with race-based stressors. Mangels et al. (2012) studied the effects of stereotype threat on women while engaged with a math test. Results from that study also indicated marginal associations between the P300 and learning outcomes for women in the stereotype threat condition but those findings were overshadowed by stronger associations between the Late Positive Potential (LPP) and error correction and the FRN and engagement with learning feedback. Although this study did not investigate the LPP—a sustained positive deflection that is increased in response to emotionally arousing stimuli—the LPP is often associated with the P300 and some even suggest that they are the same waveform (Hajcak & Foti, 2020). This study also did not provide evidence that an FRN was elicited from the TE task. The FRN is strongly associated with the ERN and is often also referred to as the Feedback ERN because it is a negative deflection that occurs in response to negative feedback. The absence of the FRN in the TE task may be associated with reasons why the ERN was also not present.

Although the literature linking EEG components, learning, and race-based stress is small, evidence thus far points in the direction of race-based stress, such as minority stress, may interfere with attention and memory updating processes used to learn from errors. Future studies should expand the scope of the types of race-based stressors being assessed as well as further examine if the P300 alone signifies attention and memory updating or if it should be examined in tandem with the LPP.

Alpha Power and Working Memory

Alpha oscillations during the TE task were found to be positively associated with minority stress only during learning feedback on incorrect response trials. The direction of this association was as expected considering that less alpha activity is associated with increased attentional and working memory processes. As minority stress increased, alpha power also increased, suggesting that minority stress may be interfering with working memory processes during the testing effect task. Results from the SWM task were to provide convergent evidence for processes occurring in the TE task, however findings for this were mixed. During high load encoding and goal maintenance on the SWM task, minority stress was negatively associated with alpha power. This is the opposite of what was expected, and it suggests that high levels of minority stress were associated with less alpha power or better attentional and working memory processes. SWM task alpha power, however, was positively associated with college stress and microaggressions, and the overall regression model for SWM high load encoding suggested a positive relationship between all stressors and alpha power. Studies examining different working memory tasks have found similar inconsistencies in which the same neural correlate relates to working memory differently in separate tasks that are measuring the same constructs (Lenartowicz et al., 2021).

One possible explanation may have to do with the type of task being used to measure working memory processes. For instance, associations between alpha power and working memory processes may function differently in a vocabulary learning task. Pi and colleagues (2023) used frontal and parietal alpha power to assess working memory processing during a foreign language learning task and concluded that alpha power may not be a consistent indicator of cognitive processes during foreign language learning tasks. More studies are necessary to assess if alpha power is an interpretable correlate of working memory processes during a language learning task and if race-based stressors are associated with those processes.

A Note on Neuroscience as a Recruitment Tool

While collecting data for this study, there was a keen interest from participants about the purpose of the study and what it would all mean once the results were analyzed. When debriefing Latinx participants, many asked a lot of follow up questions and some even asked if they could participate as a research assistant for the remainder of the project. It was apparent that the confluence of social issues that were very personal for Latinx participants with neuroscientific methods was a point of interest for many of the Latinx students. When placing this anecdote in the context of limited representation of Latinx individuals in the STEM (Science, Technology, Engineering, and Math) fields, it seems like those who are conducting interdisciplinary research that address the social issues that affect individuals of color have a powerful recruitment tool for getting more students of color involved in the STEM fields. The motivation from students is there, and it is up to researchers to recognize this and create these opportunities for those students interested in participating.

Limitations

A large part of the sample in this study came from the psychology research subject pool. This makes the sample one of convenience and limits generalizability. However, this study specifically focuses on the experiences of college students and therefore findings should not be generalized beyond that population. In particular, the focus on the experiences of Latinx college students, still makes this study one of the few to address if the TE is beneficial for Latinx students and one of the few that does this with electrophysiological instruments. It is the author's hope that this study will inspire more like it in the future.

Some concessions had to be made considering the Covid-19 pandemic and the subsequent remote learning policies at UCLA. Initially, to follow social distancing recommendations, participants were asked to conduct the 1-week retention test at home where researchers had very little control over the research environment. Although there did not appear to be any differences between those who took the test at home and those who conducted both parts of the study in the lab, participants at home may have been exposed to more potential distractions that could have affected their performance. The pandemic also made access to participants difficult because EEG data collection requires in-person participation. As a result, there may have been selection bias in that the study only attracted individuals who were willing to take any potential risks involved with being near other people for an extended period.

Although the decision to administer the race-based stress scales to a sample of White participants was made with the intention of reducing culture (mis)attribution bias, this did result in some constraints in terms of measuring race-based stress for the analyses (Causadias et al., 2018). Many of the items on the MASI and the MSSS were rated as not applicable by White participants. This limited the number of items that were included in the final analyses for the MASI therefore potentially affecting the internal validity of the scale. However, there was slight variation in responses to the race-based stress scales from White participants. Chances are that in some cases this was due to intersectional identities. For example, some White participants who also identified as Jewish reported having experienced microaggressions more frequently than their White peers who did not report any other cultural identity. In line with evidence supporting the notion that culture is rarely attributed to explaining the behavior of White individuals, more research is needed to examine how White individuals in the United States experience race-based stress and to develop appropriate measurement scales.

In addition to this, the study employed both behavioral and neural correlates for cognition but relied on self-report measures for assessing race-based stress. Stress can be measured in many ways and relying on a single method to measure stress may provide a partial understanding of stress. Future iterations of this work should explore using non-invasive biological measures of stress such as cortisol and heart rate variability, in addition to self-report measures.

Conclusion

Latinx and White students are contending with different social context while attending college. Latinx must contend with more stress in the form of race-based stressors and some of these stressors may interfere with their ability to perform well on learning tasks. Testing as a study strategy, though, is beneficial for students who face increased levels of stress and should continue to be promoted as an effective learning strategy. But if we want to improve the situation for Latinx students, the social context needs to be changed. Instead of improving cognitive processes via training and study strategies, it may be more fruitful and rewarding to improve cognitive processes by reducing barriers in the social environment.

Appendix A

The Multidimension Acculturative Stress Inventory (MASI) – 18-items

Response Scale

0 Does not apply to me

1 Not at all stressful

2 A little stressful

3 Somewhat stressful

4 Very stressful

5 Extremely stressful

Items

1. I have a hard time understanding others when they speak English.
2. It bothers me that I speak English with an accent.
3. Since I don't speak English well, people have treated me rudely or unfairly.
4. I have been discriminated against because I have difficulty speaking English.
5. I don't speak English or don't speak it well.
6. I feel pressure to learning English.
7. I feel uncomfortable being around people who only speak English.
8. It bothers me when people don't respect my (ethnicity) values.
9. People look down upon me if I practice (ethnicity) customs.
10. I feel uncomfortable when I have to choose between (ethnicity) and American ways of doing things.
11. It bothers me when people pressure me to assimilate to the American way of doing things.

12. Because of my cultural background, I have a hard time fitting in with Americans.
13. I don't feel accepted by Americans.
14. I feel uncomfortable when others expect me to know American ways of doing things.
15. I have had conflicts with others because I prefer American customs (e.g., celebrating Halloween, Thanksgiving) over (ethnicity) ones.
16. I feel uncomfortable because my family does not know (ethnicity) ways of doing things.
17. I feel uncomfortable when others expect me to know (ethnicity) ways of doing things.
18. People look down upon me if I practice American cultures.

Appendix B

Minority Status Stress Scale (MSSS) – 19-items

Instructions

The following set of questions include items and events that may be sources of stress for college students. Rate the stressfulness of each item based on your experiences since you started at UCLA. If you are a transfer student, think about your most recent experiences at UCLA only.

Response Scale

0 Does not apply

1 Not at all stressful

2

3

4

5 Extremely stressful

Items

1. The university does not have enough professors of my race.
2. Racist policies and practices of the university
3. The university lacks concern and support for the needs of students of my race
4. Seeing members of my race doing low status jobs and Whites in high status jobs on campus
5. Few courses involve issues relevant to my ethnic group.
6. White students and faculty expect poor academic performance from students of my race.
7. The university is an unfriendly place.
8. Having to always be aware of what White people might do.

9. Negative attitudes/treatment of students of my race by faculty
10. Being treated rudely or unfairly because of my race
11. Being discriminated against
12. Others lacking respect for people of my race.
13. Having to "prove" my abilities to others (i.e., work twice as hard)
14. Pressure that what "I" do is representative of my ethnic group's abilities, behavior, and so on.
15. The lack of unity/supportiveness among members of my race at the university
16. Trying to maintain my ethnic identity while attending the university.
17. White people expecting me to be a certain way because of my race (i.e., stereotyping)
18. Relationships between males and females of my race (e.g., lack of available dating partners)
19. Pressures from people of my same race (e.g., how to act, what to believe)

Appendix C

Racial Microaggression Scale

Instructions

The following set of questions will be about your experience with racial microaggressions.

Racial microaggressions are "subtle, stunning, often automatic, and non-verbal exchanges which are 'put downs'" and have also been described as subtle insults delivered through dismissive looks, gestures, and tones (verbal, nonverbal, and/or visual) toward people of color; often automatic or unconscious.

Simply stated, racial microaggressions are brief, everyday exchanges that send denigrating messages to people of color because they belong to a racial minority group.

Here are some examples of microaggressions that are commonly reported by people of color:

- Assumptions of being less educated, poor, not intelligent, having a lower paying job, being from a particular neighborhood.
- Instances of surprise at your accomplishments and being told you are "articulate."
- People avoiding you on the street, holding onto their belongings tighter in your presence, not sitting next to you in public spaces, avoiding eye contact with you, receiving substandard service in stores.
- People claiming that they don't see color/race, they are color-blind, racism doesn't exist anymore, that you complain or think about race too much, or that all racial groups experience the same obstacles.

- Assumptions that you speak a language other than English, that you eat food associated with your race/culture every day, being asked that you teach someone words in your "native language," that all people in your racial group look alike or are all the same, someone wanting to date you only because of your race, having your physical features objectified because of your race, and someone not believing that you were born in the U.S. after telling them.
- Lack of representation of members of your race in media, in prestigious professions, in government, at work, and at school.
- Being ignored or your opinion being overlooked at work or school, being treated in an unfriendly way by people at work or school. Being treated differently than your White counterparts.

Items and Response Scales

1. Have you ever experienced any racial microaggression during your time at UCLA?

Yes or No

2. How frequently would you estimate that you have experienced racial microaggressions during your time at UCLA? Choose an answer that best approximates your experience.

1 I have only experienced racial microaggressions once since I started college

2 Hardly at all or a few times a year

3 A few times a quarter

4 Monthly

5 Weekly

6 Daily

Appendix D

Testing Effect Task Word List

Japanese	English	Japanese	English
kodomo	child	yama	mountain
shigoto	job	byouin	hospital
hito	person	hana	flower
me	eye	ongaku	music
uchi	house	eki	station
hi	sun	eiga	movie
kotoba	word	musume	daughter
te	hand	zu	drawing
kokoro	heart	yume	dream
kane	money	iro	color
kao	face	natsu	summer
machi	city	koukou	high school
koe	voice	kaji	fire
kuruma	car	neko	cat
karada	body	otto	husband
otoko	man	uta	song
onna	woman	sakana	fish
gakkou	school	sensou	war
haha	mother	densha	train
mise	store	doubutsu	animal
atama	head	musuko	son
denwa	telephone	ame	rain
hon	book	mimi	ear
yoru	night	fuyu	winter
kazoku	family	ki	tree
chichi	father	oto	sound
heya	room	umi	ocean
daigaku	university	ashi	foot
asa	morning	shashin	photo
inu	dog	kuchi	mouth

Appendix E

Testing Effect Task Distractor Word List

spring	hip	vegetable	machine	desk
fall	hill	window	fan	plate
nose	valley	door	factory	button
apartment	clinic	forest	pot	castle
moon	cartoon	piano	stairs	bridge
star	portrait	document	coffee	tent
sentence	bus	bird	leaf	nurse
arm	plane	toilet	floor	seed
leg	storm	baseball	bone	hotel
coin	bicycle	team	table	grass
chin	letter	salt	teacher	document
town	boat	bread	brain	farmer
truck	class	meat	restaurant	cow
van	tea	finger	camera	mirror
neck	building	earth	wave	smoke
girl	chest	garden	insect	butter
boy	horse	tool	muscle	soup
office	student	rock	shoe	needle
brother	wind	egg	library	thread
sister	rice	clothes	milk	twig
hair	food	baby	fruit	cabin
magazine	doctor	paper	box	barn
pamphlet	athlete	novel	guitar	pool
evening	parent	knife	taxi	diary
noon	sky	supermarket	lunch	tape
grandfather	park	bath	dinner	ticket
grandmother	medicine	snack	breakfast	notebook
church	stomach	chair	market	shadow
rabbit	ball	gold	menu	garlic
hamster	wall	skin	key	gate
plant	air	preschool	poem	nest
noise	shoulder	seat	refrigerator	race
lake	bank	tooth	airport	firework
river	snow	bed	apple	sand
pond	uncle	sugar	lip	weapon
knee	aunt	wife	weather	claw

Appendix F

Table F1

EEG Data Exclusion Criteria

Measures & Data Type	Exclusion Criteria	Trial Type	Number Excluded
Flanker Task			
ERP	more than 25% missing trials		0
ERP	more than 25% incorrect trials		3
ERP	less than 6 incorrect trials		1
Testing Effect Task			
ERP	less than 6 incorrect trials		3
Spectral Power	more than 25% missing trials		
		Feedback Stimulus	2
Spectral Power	less than 6 incorrect trials		3
Spatial Working Memory Task			
ERP	more than 25% missing trials		
		Low Load Target Stimulus	3
		High Load Target Stimulus	2
Spectral Power	more than 25% missing trials		
		High Load	9

Appendix G

Table G1

Experimental Tasks Descriptive Statistics

Measures	Latinx		White		Total	
	n	Mean (SD)	n	Mean (SD)	N	Mean (SD)
Flanker Task						
Overall Accuracy	59	88.3% (7.0)	49	88.9% (5.3)	108	88.6% (6.2)
Reaction Time - Correct Responses	59	444.53 (54.37)	49	416.41 (38.18)	108	431.77 (49.54)
Reaction Time - Incorrect Responses	59	366.48 (59.05)	49	337.74 (32.00)	108	353.44 (50.56)
Accuracy - Congruent Trials	59	94.4% (6.4)	49	97% (3.1)	108	95.6% (5.4)
Accuracy - Incongruent Trials	59	82.2% (9.8)	49	80.7% (9.3)	108	81.5% (9.5)
Reaction Time - Congruent Trials	59	409.25 (54.77)	49	379.44 (36.05)	108	395.72 (49.31)
Reaction Time - Incongruent Trials	59	464.03 (61.16)	49	436.84 (42.23)	108	451.70 (54.89)
Reaction Time - Correct Congruent Trials	59	411.62 (54.63)	49	380.84 (36.48)	108	397.66 (49.52)
Reaction Time - Correct Incongruent Trials	59	482.93 (55.56)	49	459.51 (39.96)	108	472.30 (50.26)
Reaciton Time - Incorrect Congruent Trials	53	366.79 (91.3)	45	350.20 (105.5)	98	359.17 (97.91)
Reaction Time - Incorrect Incongruent Trials	59	366.94 (59.02)	49	335.86 (27.23)	108	352.84 (49.62)
Missed Trials	59	3.85 (9.03)	49	1 (1.32)	108	2.56 (6.15)
Post-error Slowing	59	50.22 (39.6)	49	43.83 (29.52)	108	47.32 (35.37)
Testing Effect Task						
5-minute Delay Test						
Overall Accuracy	54	84.4% (12.0)	45	90.1% (8.6)	99	87.4% (11.1)
Overall Reaction Time	54	2.69 (0.56)	45	2.46 (0.56)	99	2.59 (0.57)
Accuracy - Test Trials	54	86.8% (10.9)	45	93.4% (7.2)	99	89.8% (9.9)
Accuracy - Study Trials	54	82.0% (15.7)	45	88.5% (11.3)	99	85% (14.2)
Reaction Time - Test Trials	54	2.60 (0.60)	45	2.33 (0.54)	99	2.48 (0.59)
Reaction Time - Study Trials	54	2.78 (0.57)	45	2.59 (0.64)	99	2.7 (0.61)
1-week Delay Test						
Overall Accuracy	54	66.3% (16.3)	45	78% (15.2)	99	71.6% (16.8)
Overall Reaction Time	54	3.23 (0.67)	45	2.86 (0.56)	99	3.06 (0.65)
Accuracy - Test Trials	54	70.2% (16.6)	45	81.0% (14.2)	99	75.1% (16.4)
Accuracy - Study Trials	54	62.4% (17.6)	45	75% (17.8)	99	68.1% (18.7)
Reaction Time - Test Trials	54	3.07 (0.69)	45	2.77 (0.59)	99	2.94 (0.66)
Reaction Time - Study Trials	54	3.38 (0.69)	45	2.95 (0.59)	99	3.19 (0.68)
Spatial Working Memory Task						
Overall Accuracy	58	84.1% (5.9)	47	86.6% (7.3)	105	85.2% (6.7)
Reaction Time - Correct Responses	58	1048.19 (147.53)	47	1009.69 (119.27)	105	1030.95 (136.35)
Reaction Time - Incorrect Responses	58	1290.13 (193.73)	47	1250.54 (147.7)	105	1272.41 (174.96)
Accuracy - Low Load Trials	58	90.3% (6.0)	47	92% (8.0)	105	91.0% (7.0)
Accuracy - High Load Trials	58	77.9% (7.4)	47	81.2% (8.2)	105	79.4% (7.9)
Reaction Time - Low Load Trials	58	995.08 (148.98)	47	944.36 (123.77)	105	972.38 (139.94)
Reaction Time - High Load Trials	58	1157.59 (157.02)	47	1120.16 (119.01)	105	1140.83 (141.87)
Reaction Time - Correct Low Load Trials	58	977.68 (146.74)	47	934.55 (126.56)	105	958.38 (139.11)
Reaction Time - Correct High Load Trials	58	1130.15 (157.5)	47	1094.2 (124.61)	105	1114.05 (144.17)
Reaction Time - Incorrect Low Load Trials	54	1258.37 (296.33)	35	1121.26 (275.13)	89	1204.45 (294.4)
Reaction Time - Incorrect High Load Trials	58	1313.68 (203.24)	47	1297.33 (160.24)	105	1306.36 (184.56)
Missed Trials	58	3.60 (2.51)	47	3.34 (2.73)	105	3.48 (2.6)

Note. Reaction times for Flanker and Spatial working memory task are in milliseconds and in seconds for Testing Effect task.

Table G2
Experimental Tasks Group Comparisons

Measures	Latinx & White Group Difference	Cohen *d*
Flanker Task		
Overall Accuracy	$t(106) = -0.45, p = .653$	-0.09
Reaction Time - Correct Responses	**$t(106) = 3.05, p = .003$**	**0.59**
Reaction Time - Incorrect Responses	**$t(106) = 3.05, p = .003$**	**0.59**
Accuracy - Congruent Trials	**$t(106) = -2.59, p = .011$**	**-0.50**
Accuracy - Incongruent Trials	$t(106) = 0.82, p = .415$	0.16
Reaction Time - Congruent Trials	**$t(106) = 3.27, p = .001$**	**0.63**
Reaction Time - Incongruent Trials	**$t(106) = 2.63, p = .010$**	**0.51**
Reaction Time - Correct Congruent Trials	**$t(106) = 3.37, p = .001$**	**0.65**
Reaction Time - Correct Incongruent Trials	**$t(106) = 2.47, p = .015$**	**0.48**
Reaciton Time - Incorrect Congruent Trials	$t(96) = 0.83, p = .406$	0.17
Reaction Time - Incorrect Incongruent Trials	**$t(106) = 3.40, p < .001$**	**0.66**
Missed Trials	**$t(106) = 2.45, p = .016$**	**0.47**
Post-error Slowing	$t(106) = 0.93, p = .352$	0.18
Testing Effect Task		
5-minute Delay Test		
Overall Accuracy	**$t(97) = -3.06, p = .003$**	**-0.62**
Overall Reaction Time	**$t(97) = 2.08, p = .040$**	**0.42**
Accuracy - Test Trials	**$t(97) = -3.48, p < .001$**	**-0.70**
Accuracy - Study Trials	**$t(97) = -2.13, p = .023$**	**-0.47**
Reaction Time - Test Trials	**$t(97) = 2.37, p = .020$**	**0.48**
Reaction Time - Study Trials	$t(97) = 1.61, p = .111$	0.32
1-week Delay Test		
Overall Accuracy	**$t(97) = -3.69, p < .001$**	**-0.74**
Overall Reaction Time	**$t(97) = 2.91, p = .005$**	**0.59**
Accuracy - Test Trials	**$t(97) = -3.45, p < .001$**	**-0.70**
Accuracy - Study Trials	**$t(97) = 3.53, p < .001$**	**-0.71**
Reaction Time - Test Trials	**$t(97) = 2.33, p = .022$**	**0.47**
Reaction Time - Study Trials	**$t(97) = 3.32, p = .001$**	**0.67**
Spatial Working Memory Task		
Overall Accuracy	$t(103) = -1.94, p = .056$	-0.38
Reaction Time - Correct Responses	$t(103) = 1.45, p = .151$	0.28
Reaction Time - Incorrect Responses	$t(103) = 1.15, p = .251$	0.23
Accuracy - Low Load Trials	$t(103) = -1.22, p = .225$	-0.24
Accuracy - High Load Trials	**$t(103) = -2.18, p = .032$**	**-0.43**
Reaction Time - Low Load Trials	$t(103) = 1.87, p = .065$	0.37
Reaction Time - High Load Trials	$t(103) = 1.35, p = .180$	0.26
Reaction Time - Correct Low Load Trials	$t(103) = 1.59, p = .115$	0.31
Reaction Time - Correct High Load Trials	$t(103) = 1.27, p = .205$	0.25
Reaction Time - Incorrect Low Load Trials	**$t(87) = 2.19, p = .031$**	**0.48**
Reaction Time - Incorrect High Load Trials	$t(103) = 0.45, p = .654$	0.09
Missed Trials	$t(103) = 0.51, p = .609$	0.10

Note. Cohen's d small ~ 0.2, medium ~ 0.5, large > 0.8

References

Aceves, L., Ramos-Olazagasti, M. A., & Guzman, L. (2022, September 8). *Hispanic Center webinar compares research design approaches to transform the narrative on Latino populations* [Fact Sheets and Guides, Webinar]. National Research Center on Hispanic Children & Families. https://www.hispanicresearchcenter.org/research-resources/hispanic-center-webinar-compares-research-design-approaches-to-transform-the-narrative-on-latino-populations

Albeg, L. J., & Castro-Olivo, S. M. (2014). The Relationship between Mental Health, Acculturative Stress, and Academic Performance in a Latino Middle School Sample. *Contemporary School Psychology*, *18*(3), 178–186. https://doi.org/10.1007/s40688-014-0010-1

Alfieri, L., Brooks, P. J., Aldrich, N. J., & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology*, *103*(1), 1–18. http://dx.doi.org/10.1037/a0021017

Arbona, C., & Jimenez, C. (2014). Minority stress, ethnic identity, and depression among Latino/a college students. *Journal of Counseling Psychology*, *61*(1), 162–168. http://dx.doi.org/10.1037/a0034914

Asner-Self, K. K., Schreiber, J. B., & Marotta, S. A. (2006). A cross-cultural analysis of the Brief Symptom Inventory-18. *Cultural Diversity and Ethnic Minority Psychology*, *12*(2), 367–375. https://doi.org/10.1037/1099-9809.12.2.367

Ayala, M. I., & Chalupa, D. (2016). Beyond the Latino Essentialist Experience: Racial and Ethnic Self-Identifications, Gender, and College Attainment. *Hispanic Journal of Behavioral Sciences*, *38*(3), 378–394. https://doi.org/10.1177/0739986316642841

Bai, C.-H., Bridger, E. K., Zimmer, H. D., & Mecklinger, A. (2015). The beneficial effect of testing: An event-related potential study. *Frontiers in Behavioral Neuroscience*, *9*. https://doi.org/10.3389/fnbeh.2015.00248

Beauducel, A., & Wittmann, W. W. (2005). Simulation Study on Fit Indexes in CFA Based on Data With Slightly Distorted Simple Structure. *Structural Equation Modeling: A Multidisciplinary Journal*, *12*(1), 41–75. https://doi.org/10.1207/s15328007sem1201_3

Beilock, S. L., Jellison, W. A., Rydell, R. J., McConnell, A. R., & Carr, T. H. (2006). On the Causal Mechanisms of Stereotype Threat: Can Skills That Don't Rely Heavily on Working Memory Still Be Threatened? *Personality and Social Psychology Bulletin*, *32*(8), 1059–1071. https://doi.org/10.1177/0146167206288489

Beilock, S. L., Rydell, R. J., & McConnell, A. R. (2007). Stereotype threat and working memory: Mechanisms, alleviation, and spillover. *Journal of Experimental Psychology: General*, *136*(2), 256–276. http://dx.doi.org/10.1037/0096-3445.136.2.256

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238–246. https://doi.org/10.1037/0033-2909.107.2.238

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher (Ed.), *Psychology and the real world: Essays illustrating fundamental contributions to society* (pp. 56–64). Worth Publishers.

Bjork, R. A. (1975). Retrieval as a Memory Modifier: An Interpretation of Negative Recency and Related Phenomena. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium* (pp. 123–144). Erlbaum.

Bjork, R. A., & Bjork, E. L. (1992). A new theory of disuse and an old theory of stimulus fluctuation. In A. Healy (Ed.), *From learning processes to cognitive processes: Essays in honor of William K. Estes* (Vol. 2, pp. 35–67). https://books.google.com/books?hl=en&lr=&id=NaXm2AKl2w8C&oi=fnd&pg=PA35&dq=A+new+theory+of+disuse&ots=fyQSTCBpuw&sig=BHUXTG0StujHB0YNaTqEYa4YHTU

Born, D. O. (1970). Psychological adaptation and development under acculturative stress: Toward a general model. *Social Science & Medicine (1967)*, *3*(4), 529–547. https://doi.org/10.1016/0037-7856(70)90025-9

Bosson, J. K., Haymovitz, E. L., & Pinel, E. C. (2004). When saying and doing diverge: The effects of stereotype threat on self-reported versus non-verbal anxiety. *Journal of Experimental Social Psychology*, *40*(2), 247–255. https://doi.org/10.1016/S0022-1031(03)00099-4

Boudewyn, M. A., Luck, S. J., Farrens, J. L., & Kappenman, E. S. (2018). How many trials does it take to get a significant ERP effect? It depends. *Psychophysiology*, *55*(6), e13049. https://doi.org/10.1111/psyp.13049

BrainVision Analyzer (2.2.0). (2019). Brain Products, GmbH.

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (Second edition). The Guilford Press.

Carnevale, A. P., & Fasules, M. L. (2017). *Latino Education and Economic Progress: Running Faster but Still Behind*. https://cew.georgetown.edu/cew-reports/latinosworkforce/

Carpenter, S. K. (2012). Testing Enhances the Transfer of Learning. *Current Directions in Psychological Science*, *21*(5), 279–283. https://doi.org/10.1177/0963721412452728

Causadias, J. M., Vitriol, J. A., & Atkin, A. L. (2018). Do we overemphasize the role of culture in the behavior of racial/ethnic minorities? Evidence of a cultural (mis)attribution bias in American psychology. *American Psychologist*, *73*(3), 243–255. http://dx.doi.org/10.1037/amp0000099

Contreras, F. E. (2005). Access, Achievement, and Social Capital: Standardized Exams and the Latino College-Bound Population. *Journal of Hispanic Higher Education*, *4*(3), 197–214. https://doi.org/10.1177/1538192705276546

Crockett, L. J., Iturbide, M. I., Torres Stone, R. A., McGinley, M., Raffaelli, M., & Carlo, G. (2007). Acculturative stress, social support, and coping: Relations to psychological adjustment among Mexican American college students. *Cultural Diversity and Ethnic Minority Psychology*, *13*(4), 347–355. http://dx.doi.org/10.1037/1099-9809.13.4.347

DeCaro, M. S., & Rittle-Johnson, B. (2012). Exploring mathematics problems prepares children to learn from instruction. *Journal of Experimental Child Psychology*, *113*(4), 552–568. https://doi.org/10.1016/j.jecp.2012.06.009

Deouell, L. Y., & Knight, L. T. (2009). Executive Function and Higher-Order Cognition: EEG Studies. In L. R. Squire (Ed.), *Encyclopedia of Neuroscience* (Vol. 4, pp. 105–109). Oxford: Academic Press.

DiAngelo, R. (2016). White fragility. In *What does it mean to be white? Developing white racial literacy: Revised edition* (Vol. 497). Peter Lang AG. http://www.jstor.org/stable/45157307

Ernst, B., & Steinhauser, M. (2012). Feedback-related brain activity predicts learning from feedback in multiple-choice testing. *Cognitive, Affective and Behavioral Neuroscience; New York*, *12*(2), 323–336. https://doi.org/10.3758/s13415-012-0087-9

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Flores, A., Lopez, M. H., & Krogstad, J. M. (2019, July 8). *U.S. Hispanic population reached new high in 2018, but growth has slowed*. Pew Research Center. https://www.pewresearch.org/fact-tank/2019/07/08/u-s-hispanic-population-reached-new-high-in-2018-but-growth-has-slowed/

Gehring, W. J., Liu, Y., Orr, J. M., & Carp, J. (2011). The Error-Related Negativity (ERN/Ne). In S. J. Luck & E. S. Kappenman (Eds.), *Oxford handbook of event-related potential components*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195374148.013.0120

Gil, A. G., Vega, W. A., & Dimas, J. M. (1994). Acculturative stress and personal adjustment among hispanic adolescent boys. *Journal of Community Psychology*, *22*(1), 43–54. https://doi.org/10.1002/1520-6629(199401)22:1<43::AID-JCOP2290220106>3.0.CO;2-T

Gimbel, S. I., & Brewer, J. B. (2011). Reaction time, memory strength, and fMRI activity during memory retrieval: Hippocampus and default network are differentially responsive during recollection and familiarity judgments. *Cognitive Neuroscience*, *2*(1), 19–26. https://doi.org/10.1080/17588928.2010.513770

Glazer, J. E., Kelley, N. J., Pornpattananangkul, N., Mittal, V. A., & Nusslock, R. (2018). Beyond the FRN: Broadening the time-course of EEG and ERP components implicated in reward processing. *International Journal of Psychophysiology*, *132*, 184–202. https://doi.org/10.1016/j.ijpsycho.2018.02.002

Gonzalez, N., Moll, L. C., Tenery, M. F., Rivera, A., Rendon, P., Gonzales, R., & Amanti, C. (1995). Funds of Knowledge for Teaching in Latino Households. *Urban Education*, *29*(4), 443–470. https://doi.org/10.1177/0042085995029004005

Gromlich, J. (2017, September 29). *Hispanic dropout rate hits new low, college enrollment at new high*. Pew Research Center. https://www.pewresearch.org/fact-tank/2017/09/29/hispanic-dropout-rate-hits-new-low-college-enrollment-at-new-high/

Guyll, M., Madon, S., Prieto, L., & Scherr, K. C. (2010). The Potential Roles of Self-Fulfilling Prophecies, Stigma Consciousness, and Stereotype Threat in Linking Latino/a Ethnicity and Educational Outcomes. *Journal of Social Issues*, *66*(1), 113–130. https://doi.org/10.1111/j.1540-4560.2009.01636.x

Hajcak, G., & Foti, D. (2020). Significance?... Significance! Empirical, methodological, and theoretical connections between the late positive potential and P300 as neural responses to stimulus significance: An integrative review. *Psychophysiology*, *57*(7). https://doi.org/10.1111/psyp.13570

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (Vol. 1, pp. 371–404). Information Age Pub.

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*(4), 679–709. https://doi.org/10.1037/0033-295X.109.4.679

Jones, H. L., Cross, W. E., & DeFour, D. C. (2007). Race-Related Stress, Racial Identity Attitudes, and Mental Health Among Black Women. *Journal of Black Psychology*, *33*(2), 208–231. https://doi.org/10.1177/0095798407299517

Kappenman, E. S., Farrens, J., Zhang, W., Stewart, A. X., & Luck, S. J. (2020). *ERP CORE: An Open Resource for Human Event-Related Potential Research* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/4azqm

Kapur, M. (2008). Productive failure. *Cognition & Instruction*, *26*(3), 379–424. https://doi.org/10.1080/07370000802212669

Kessler, R. C., Mickelson, K. D., & Williams, D. R. (1999). The Prevalence, Distribution, and Mental Health Correlates of Perceived Discrimination in the United States. *Journal of Health and Social Behavior*, *40*(3), 208–230. JSTOR. https://doi.org/10.2307/2676349

Klimesch, W. (2012). Alpha-band oscillations, attention, and controlled access to stored information. *Trends in Cognitive Sciences*, *16*(12), 606–617. https://doi.org/10.1016/j.tics.2012.10.007

Kornell, N., & Son, L. K. (2009). Learners' choices and beliefs about self-testing. *Memory*, *17*(5), 493–501. https://doi.org/10.1080/09658210902832915

Krigolson, O. E., Pierce, L. J., Holroyd, C. B., & Tanaka, J. W. (2009). Learning to Become an Expert: Reinforcement Learning and the Acquisition of Perceptual Expertise. *Journal of Cognitive Neuroscience*, *21*(9), 1833–1840. https://doi.org/10.1162/jocn.2009.21128

Krosch, A. R., Park, S. J., Walker, J., & Lisner, A. R. (2022). The threat of a majority-minority U.S. alters white Americans' perception of race. *Journal of Experimental Social Psychology*, *99*, 104266. https://doi.org/10.1016/j.jesp.2021.104266

Lenartowicz, A., Lu, S., Rodriguez, C., Lau, E. P., Walshaw, P. D., McCracken, J. T., Cohen, M. S., & Loo, S. K. (2016). Alpha desynchronization and fronto-parietal connectivity during spatial working memory encoding deficits in ADHD: A simultaneous EEG-fMRI study. *NeuroImage: Clinical*, *11*, 210–223. https://doi.org/10.1016/j.nicl.2016.01.023

Lenartowicz, A., Truong, H., Enriquez, K. D., Webster, J., Pochon, J.-B., Rissman, J., Bearden, C. E., Loo, S. K., & Bilder, R. M. (2021). Neurocognitive subprocesses of working memory performance. *Cognitive, Affective, & Behavioral Neuroscience*, *21*(6), 1130–1152. https://doi.org/10.3758/s13415-021-00924-7

Lenartowicz, A., Truong, H., Salgari, G. C., Bilder, R. M., MeGough, J., McCracken, J. T., & Loo, S. K. (2019). Alpha modulation during working memory encoding predicts neurocognitive impairment in ADHD. *Journal of Child Psychology and Psychiatry*, *60*(8), 917–926. https://doi.org/10.1111/jcpp.13042

Luck, S. J. (2014). *An Introduction to the Event-Related Potential Technique* (2nd ed.). The MIT Press.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, *1*(2), 130–149. https://doi.org/10.1037/1082-989X.1.2.130

Mangels, J. A., Good, C., Whiteman, R. C., Maniscalco, B., & Dweck, C. S. (2012). Emotion blocks the path to learning under stereotype threat. *Social Cognitive and Affective Neuroscience*, *7*(2), 230–241. https://doi.org/10.1093/scan/nsq100

May, A. (2011). Experience-dependent structural plasticity in the adult human brain. *Trends in Cognitive Sciences*, *15*(10), 475–482. https://doi.org/10.1016/j.tics.2011.08.002

Mena, F. J., Padilla, A. M., & Maldonado, M. (1987). Acculturative Stress and Specific Coping Strategies among Immigrant and Later Generation College Students. *Hispanic Journal of Behavioral Sciences*, *9*(2), 207–225. https://doi.org/10.1177/07399863870092006

Miller, G. A., & Yee, C. M. (2015). Moving Psychopathology Forward. *Psychological Inquiry*, *26*(3), 263–267. https://doi.org/10.1080/1047840X.2015.1037818

Minikel-Lacocque, J. (2013). Racism, College, and the Power of Words: Racial Microaggressions Reconsidered. *American Educational Research Journal*, *50*(3), 432–465. https://doi.org/10.3102/0002831212468048

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The Unity and Diversity of Executive Functions and Their Contributions to Complex "Frontal Lobe" Tasks: A Latent Variable Analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Morrison, F. J., Kim, M. H., Connor, C. M., & Grammer, J. K. (2019). The Causal Impact of Schooling on Children's Development: Lessons for Developmental Science. *Current Directions in Psychological Science*, *28*(5), 441–449. https://doi.org/10.1177/0963721419855661

Moser, J., Moran, T., Schroder, H., Donnellan, B., & Yeung, N. (2013). On the relationship between anxiety and error monitoring: A meta-analysis and conceptual framework. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00466

Nadal, K. L. (2011). The Racial and Ethnic Microaggressions Scale (REMS): Construction, reliability, and validity. *Journal of Counseling Psychology*, *58*(4), 470–480. http://dx.doi.org/10.1037/a0025193

Neville, H. A., Heppner, P. P., Ji, P., & Thye, R. (2004). The Relations among General and Race-Related Stressors and Psychoeducational Adjustment in Black Students Attending Predominantly White Institutions. *Journal of Black Studies*, *34*(4), 599–618. https://doi.org/10.1177/0021934703259168

Nieuwenhuis, S., Ridderinkhof, K. R., Talsma, D., Coles, M. G. H., Holroyd, C. B., Kok, A., & Van Der Molen, M. W. (2002). A computational account of altered error processing in older age: Dopamine and the error-related negativity. *Cognitive, Affective, & Behavioral Neuroscience*, *2*(1), 19–36. https://doi.org/10.3758/CABN.2.1.19

Obaidi, M., Kunst, J., Ozer, S., & Kimel, S. Y. (2022). The "Great Replacement" conspiracy: How the perceived ousting of Whites can evoke violent extremism and Islamophobia. *Group Processes & Intergroup Relations*, *25*(7), 1675–1695. https://doi.org/10.1177/13684302211028293

Padilla, A. M., Alvarez, M., & Lindholm, K. J. (1986). Generational Status and Personality Factors as Predictors of Stress in Students. *Hispanic Journal of Behavioral Sciences*, *8*(3), 275–288. https://doi.org/10.1177/07399863860083006

Pastötter, B., & Bäuml, K.-H. T. (2016). Reversing the testing effect by feedback: Behavioral and electrophysiological evidence. *Cognitive, Affective, & Behavioral Neuroscience*, *16*(3), 473–488. https://doi.org/10.3758/s13415-016-0407-6

Pi, Z., Liu, C., Wang, L., Yang, J., & Li, X. (2023). Cues facilitate foreign language vocabulary learning from instructional videos: Behavioral and neural evidence. *Language Teaching Research*, 136216882311647. https://doi.org/10.1177/13621688231164724

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

Polich, J. (2011). Neuropsychology of P300. In S. J. Luck & E. S. Kappenman (Eds.), *Oxford Handbook of Event-Related Potential Components*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195374148.013.0089

Pyc, M. A., & Rawson, K. A. (2009). Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory? *Journal of Memory & Language*, *60*(4), 437–447. https://doi.org/10.1016/j.jml.2009.01.004

Rad, M. S., Martingano, A. J., & Ginges, J. (2018). Toward a psychology of Homo sapiens: Making psychological science more representative of the human population. *Proceedings of the National Academy of Sciences*, *115*(45), 11401–11405. https://doi.org/10.1073/pnas.1721165115

Ray, W. J., & Cole, H. W. (1985). EEG alpha activity reflects attentional demands, and beta activity reflects emotional and cognitive processes. *Science*, *228*(4700), 750–752. https://doi.org/10.1126/science.3992243

Roche, C., & Kuperminc, G. P. (2012). Acculturative Stress and School Belonging Among Latino Youth. *Hispanic Journal of Behavioral Sciences*, *34*(1), 61–76. https://doi.org/10.1177/0739986311430084

Rodriguez, N., Flores, T., Flores, R. T., Myers, H. F., & Vriesema, C. C. (2015). Validation of the Multidimensional Acculturative Stress Inventory on adolescents of Mexican origin. *Psychological Assessment*, *27*(4), 1438–1451. http://dx.doi.org/10.1037/pas0000125

Rodriguez, N., Myers, H. F., Mira, C. B., Flores, T., & Garcia-Hernandez, L. (2002). Development of the Multidimensional Acculturative Stress Inventory for adults of Mexican origin. *Psychological Assessment*, *14*(4), 451–461. http://dx.doi.org/10.1037/1040-3590.14.4.451

Rodriguez, N., Myers, H. F., Morris, J. K., & Cardoza, D. (2000). Latino College Student Adjustment: Does an Increased Presence Offset Minority-Status and Acculturative Stresses?1. *Journal of Applied Social Psychology*, *30*(7), 1523–1550. https://doi.org/10.1111/j.1559-1816.2000.tb02534.x

Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, *15*(1), 20–27. https://doi.org/10.1016/j.tics.2010.09.003

Roediger, H. L., & Karpicke, J. D. (2006). The Power of Testing Memory: Basic Research and Implications for Educational Practice. *Perspectives on Psychological Science*, *1*(3), 181–210.

Rosburg, T., Johansson, M., Weigl, M., & Mecklinger, A. (2015). How does testing affect retrieval-related processes? An event-related potential (ERP) study on the short-term effects of repeated retrieval. *Cognitive, Affective, & Behavioral Neuroscience*, *15*(1), 195–210. https://doi.org/10.3758/s13415-014-0310-y

Saenz, V. B., & Ponjuan, L. (2009). The Vanishing Latino Male in Higher Education. *Journal of Hispanic Higher Education*, *8*(1), 54–89. https://doi.org/10.1177/1538192708326995

Sanchez, M. E. (2019). Perceptions of Campus Climate and Experiences of Racial Microaggressions for Latinos at Hispanic-Serving Institutions. *Journal of Hispanic Higher Education*, *18*(3), 240–253. https://doi.org/10.1177/1538192717739351

Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology*, *85*(3), 440–452. http://dx.doi.org/10.1037/0022-3514.85.3.440

Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review*, *115*(2), 336–356. http://dx.doi.org/10.1037/0033-295X.115.2.336

Schwabe, L., & Wolf, O. T. (2010). Learning under stress impairs memory formation. *Neurobiology of Learning and Memory*, *93*(2), 183–188. https://doi.org/10.1016/j.nlm.2009.09.009

Schwartz, S. J., Zamboanga, B. L., & Jarvis, L. H. (2007). Ethnic identity and acculturation in Hispanic early adolescents: Mediated relationships to academic grades, prosocial behaviors, and externalizing symptoms. *Cultural Diversity and Ethnic Minority Psychology*, *13*(4), 364–373. http://dx.doi.org/10.1037/1099-9809.13.4.364

Sears, D. O. (1986). College sophomores in the laboratory: Influences of a narrow data base on social psychology's view of human nature. *Journal of Personality and Social Psychology*, *51*(3), 515–530. http://dx.doi.org/10.1037/0022-3514.51.3.515

Smedley, B. D., Myers, H. F., & Harrell, S. P. (1993). Minority-Status Stresses and the College Adjustment of Ethnic Minority Freshmen. *The Journal of Higher Education*, *64*(4), 434. https://doi.org/10.2307/2960051

Smeets, T., Giesbrecht, T., Jelicic, M., & Merckelbach, H. (2007). Context-dependent enhancement of declarative memory performance following acute psychosocial stress. *Biological Psychology*, *76*(1–2), 116–123. https://doi.org/10.1016/j.biopsycho.2007.07.001

Solorzano, D., Ceja, M., & Yosso, T. (2000). Critical Race Theory, Racial Microaggressions, and Campus Racial Climate: The Experiences of African American College Students. *The Journal of Negro Education*, *69*(1/2), 60–73. JSTOR.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, *52*(6), 613–629. http://dx.doi.org/10.1037/0003-066X.52.6.613

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*(5), 797–811. http://dx.doi.org/10.1037/0022-3514.69.5.797

Sue, D. W. (2005). Racism and the Conspiracy of Silence: Presidential Address. *The Counseling Psychologist*, *33*(1), 100–114. https://doi.org/10.1177/0011000004270686

Sue, D. W., & Constantine, M. G. (2007). Racial Microaggressions as Instigators of Difficult Dialogues on Race: Implications for Student Affairs Educators and Students. *College Student Affairs Journal*, *26*(2), 136–143.

The Jamovi Project. (2022). *Jamovi* (2.3). https://www.jamovi.org

Tono, Y., Makawa, K., & Yamazaki, M. (2013). *A frequency dictionary of Japanese: Core vocabulary for learners*. Routledge.

UCLA HSI Task Force. (2022). *Cultivating the seeds of change: Becoming a Hispanic-serving institution*. https://ucla.app.box.com/v/HispanicServingInstitution

University of California. (2023, March 3). *Information center: Fall enrollment at a glance* [Statistics Dashboard]. https://www.universityofcalifornia.edu/about-us/information-center/fall-enrollment-glance

Vaughn, K. E., & Kornell, N. (2019). How to activate students' natural desire to test themselves. *Cognitive Research: Principles and Implications*, *4*(1), 35. https://doi.org/10.1186/s41235-019-0187-y

Vogel, S., & Schwabe, L. (2016). Learning and memory under stress: Implications for the classroom. *Npj Science of Learning*, *1*(1), 16011. https://doi.org/10.1038/npjscilearn.2016.11

Warshauer, H. K. (2011). *The role of productive struggle in teaching and learning middle school mathematics*. http://repositories.lib.utexas.edu/handle/2152/ETD-UT-2011-12-4527

Wei, M., Ku, T.-Y., & Liao, K. Y.-H. (2011). Minority stress and college persistence attitudes among African American, Asian American, and Latino students: Perception of university environment as a mediator. *Cultural Diversity and Ethnic Minority Psychology*, *17*(2), 195–203. http://dx.doi.org/10.1037/a0023359

Weinberg, A., Klein, D. N., & Hajcak, G. (2012). Increased Error-Related Brain Activity Distinguishes Generalized Anxiety Disorder With and Without Comorbid Major Depressive Disorder. *Journal of Abnormal Psychology*, *121*(4), 885–896. https://doi.org/10.1037/a0028270

Wheeler, S. C., & Petty, R. E. (2001). The effects of stereotype activation on behavior: A review of possible mechanisms. *Psychological Bulletin*, *127*(6), 797–826. http://dx.doi.org/10.1037/0033-2909.127.6.797

Williams, C. L., & Berry, J. W. (1991). Primary prevention of acculturative stress among refugees: Application of psychological theory and practice. *American Psychologist*, *46*(6), 632–641. http://dx.doi.org/10.1037/0003-066X.46.6.632

Yosso, T., Smith, W., Ceja, M., & Solórzano, D. (2009). Critical Race Theory, Racial Microaggressions, and Campus Racial Climate for Latina/o Undergraduates. *Harvard Educational Review*, *79*(4), 659–691. https://doi.org/10.17763/haer.79.4.m6867014157m707l