BODY MASS INDEX AND EMOTION RECOGNITION IN YOUNG ADULTHOOD AND ITS ASSOCIATION WITH EXECUTIVE FUNCTIONING

A Dissertation
Submitted to
the Temple University Graduate Board

In Partial Fulfillment
of the Requirements for the Degree
DOCTOR OF PHILOSOPHY

---

by
Elan N. French
August 2023

Examining Committee Members:

Eunice Y. Chen, Ph.D., Advisory Chair, Psychology and Neuroscience
Thomas Olino, Ph.D., Psychology and Neuroscience
Michael McCloskey, Ph.D., Examining Chair, Psychology and Neuroscience
Chelsea Helion, Ph.D., Psychology and Neuroscience
Tania Giovanetti, Ph.D., Psychology and Neuroscience
David Smith, Ph.D., Psychology and Neuroscience

ABSTRACT

BACKGROUND: Obesity is a serious health condition that also a risk factor for socio-emotional challenges and medical problems. Preliminary evidence suggests obesity may also be associated with difficulty in accurately identifying emotions, particularly negative emotions. In addition, poor emotion recognition has been linked to weaker executive functioning skills, which is a common challenge in obesity. The direct relationship between body mass index (BMI) and emotion recognition is poorly understood in young adults and warrants further exploration.

HYPOTHESES: We predicted that **1**) after controlling for sociodemographic, clinical, and executive functioning variables and that **2)** BMI would be negatively associated with emotion recognition accuracy for negative emotions (i.e., anger, sadness, and fear) but not positive emotions.

METHODS: Using a subset of the Human Connectome Project dataset (*N*=799), we conducted a hierarchal linear regression (HLR) to test the relationship between overall emotion recognition and the following predictors, adding in steps: 1) sociodemographic and clinical variables, 2) executive functioning variables, and 3) BMI.

RESULTS: Contrary to our hypotheses, BMI was not significantly associated with overall emotion recognition accuracy. Instead, Hispanic ethnicity, greater cognitive flexibility (Dimensional Change Card Sort task), and larger working memory (List Sorting Working Memory Test) was associated with better overall emotion recognition accuracy. Similarly, these same dimensions, as well as being female, was associated with better negative emotion recognition accuracy.

DISCUSSION: In a large BMI-diverse sample of healthy young adults, greater cognitive flexibility and larger working memory, rather than BMI, enhanced the sample's overall and negative emotion recognition. As this current study is cross-sectional, the field needs longitudinal studies that examine whether challenges in executive functioning earlier in life predict later problems with emotion processing, controlling for BMI.

ACKNOWLEDGMENTS

Thank you to my doctoral mentor, Eunice Y. Chen, Ph.D., for continual support and encouragement throughout my education, who always embraced the dialectal approach of validating my progress, strengths, and challenges, while impressing on me the need for continual growth.

Thank you to each of my committee members for the time and feedback they contributed toward this dissertation. Given the multidisciplinary nature of this project, I have been grateful to benefit from the expertise in eating and weight concerns from Eunice Chen, Ph.D., complex statistical approaches from Thomas Olino, Ph.D., emotion processing from Michael McCloskey, Ph.D., executive functioning from Tania Giovanetti, Ph.D., emotional processing from a developmental and neuroscience perspective from Chelsea Helion, Ph.D., and neuroscience from David Smith, Ph.D.

Thank you to my lab members (Brooke Schleyer and Stephanie Millot) for their review and helpful feedback as well as my family, friends, and cohort for their continual support.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Obesity and Overweight Prevalence and Significance in Young Adulthood.

Obesity is defined of having a Body Mass Index, a measure for a weight adjusted for height, of 30 or above, and is associated with a higher percentage of body fat (Centers for Disease Control and Prevention, CDC, 2022). It is a risk factor for several negative health and psychological outcomes including poorer interpersonal functioning, depression, diabetes, cardiovascular disease, certain cancers, and mortality (Albano et al., 2019; Hruby et al., 2016; Luppino et al., 2010; Pratt & Brody, 2014). While there is a significant increase in risk of mortality for individuals in the overweight BMI range, this risk is greater for individuals in the obese BMI range (Global BMI Mortality Collaboration et al., 2016; Xu et al., 2018). Young adults from ages 19 to 40 (Erikson, 1982) are more likely to gain weight than any other age group and are at risk for excessive weight gain and obesity (Cheng et al., 2016; Dietz, 2017; Munt et al., 2017). In fact, obesity rates increase two-fold between the 6- to 17-year-old age bracket and the 20- to 39-year-old age bracket (Dietz, 2017). Within the 25- to 34-year-old age bracket in the United States, obesity rates are almost 30 percent, and an additional 35 percent are overweight (CDC, 2021). Therefore, elucidating the impact of higher BMI in young adulthood is important to understand possible risks for onset and maintenance factors of obesity and its associated negative outcomes later in life.

Obesity, Socio-emotional Processing, and Emotion Recognition.

Individuals with obesity experience significant discrimination in our society (Puhl & Suh, 2015; Rubino et al., 2020). For example, individuals who are obese are more likely to report bullying and teasing during childhood and adolescence, which may lead to greater loneliness and social isolation (Albano et al., 2019; Puhl & Suh, 2015). This may limit opportunities for social experiences that facilitate the development of emotion processing skills, such as the recognition of emotions in the facial expressions of others.

Emotion recognition is the process of identifying human emotions in others, which is critical to social interaction, playing a key role in empathizing with others, and overall interpersonal communication (Niedenthal & Brauer, 2011). Theories of emotions, such as Ekman's Discrete Emotion Theory (Ekman, 1992; Ekman et al., 1969), have highlighted the universality of emotion recognition, including the similarities in the basic processes of emotion recognition of happiness, fear, anger, sadness, anxiety, and disgust across different cultures and races. Dimensional approaches to emotion recognition, such as that of Barrett (1998), allow emotions to be measured on continuous dimensions rather than chosen based on single emotion categories (Barrett, 1998; Xu et al., 2021). For instance, dimensional approaches to emotion recognition have highlighted the importance of the positive and negative emotion valence of recognizing facial expressions. These theories have noted that specific areas in the brain may be related to processing emotional valence (Lindquist et al., 2016). For instance, greater activity in the ventral lateral prefrontal cortex (vlPFC), the amygdala, and the right middle occipital gyrus is associated with recognizing facial expressions of negative emotions relative to positive emotion, with connectivity of the left fusiform face area and left vlPFC being important in the recognition of both negative and positive emotions (Xu et al., 2021).

Due to fewer opportunities for social interaction resulting from stigma and discrimination, recognition of emotions from facial expressions may be compromised in individuals with obesity. An important line of evidence suggesting that individuals with obesity may have differences in recognition of facial emotions comes from a behavioral meta-analysis. We conducted a meta-analysis of studies on recognition of facial emotions in youth and adults (French & Chen, 2021). We included studies that had an overweight and/or obese BMI group and a weight comparison group. Studies were only included if the primary outcome used visual stimuli (i.e., pictures) of facial and eye expressions and quantified emotion recognition skills. We found 12 publications that met our criteria (N=1,259), including 7 studies using adult participants and 5 using children and adolescents. Our main analysis showed that there was a medium effect size for poorer emotion recognition accuracy (across facial expressions of negative and positive emotions) in the obese or overweight BMI range relative to individuals in the healthy BMI range (d=-0.50), although this effect was not significant (p=0.065; CI: -1.04 to 0.03). Most of the studies included only participants in the obesity range. Exploratory analyses showed that only in tasks requiring recognition of negative emotion facial expression did overweight and obese participants performed significantly more poorly than healthy weight control participants, yielding a small but significant effect (d=-0.28; p =0.002; CI: -0.45 to -0.10). Exploratory analyses showed that when only the studies done with youth samples were examined, overweight and obese participants performed significantly more poorly than healthy weight control participants in tasks of facial emotion recognition (d=-0.88; p =0.019; CI: -1.62 to -0.14). We found only one study that examined BMI-status and facial emotion recognition in young adults. This showed that young adults who were overweight/obese had poorer emotion recognition skills compared to young adults who were in the healthy weight range, whereas the opposite pattern was found in older adults, even when controlling for medical conditions (Balter et al., 2021). Since we determined from our meta-analysis that weight status may be associated with differences in emotion recognition of facial expressions and that there is a paucity of studies examining this in young adults, we thought it important to examine this in a large young adult sample.

We also observed that there were mixed findings as to whether individuals in the overweight/obese BMI range relative to healthy weight differ in their emotion recognition skills of negative or positive emotion faces. Therefore, it is important to examine whether there are differences in recognizing positive or negative emotions. Caldú and colleagues (2019) found that participants aged 16 to 40 years old with BMIs in the obesity range were worse at identifying negative emotions than weight controls, but no differences occurred for positive or neutral emotions. Cserjési et al. (2011) found that participants (mean age 48.8 years) who were obese had more difficulty attending to negative facial expressions than healthy weight controls, although the authors did not report mean emotion accuracy scores. Finally, another study showed that children who were obese made more errors in recognizing emotions from facial expressions than healthy controls, although there was no difference in errors made whether the facial expressions were happy, angry, sad, or neutral (Koch & Pollatos, 2015). Of note, although the difference failed to reach statistical significance, participants in the obese BMI range had poorer emotion recognition accuracy for anger and neutral emotional expressions (Koch & Pollatos, 2015). To our knowledge, no studies have examined the role of valence in facial emotion recognition in young adults who are overweight or obese compared to young adults who are healthy weight.

Obesity/Overweight Status and Executive Functioning.

In addition to emotion recognition challenges, studies show that high BMI is also related to poorer executive functioning skills. Executive functioning represents a set of skills needed to initiate goal-oriented behavior and adapt to our environment (Rabinovici et al., 2015). A large meta-analysis demonstrated that individuals with obesity had broad executive functioning problems that were significant across all measured subdomains, including inhibition, cognitive flexibility, working memory, decision-making, verbal fluency, and planning (Yang et al., 2018). Effect sizes were similar across domains (*g's* =−0.308 to −0.441), demonstrating that participants with BMIs in the obesity range had worse performance on executive functioning tasks than participants with healthy range BMIs. Similarly, studies of temporal discounting, which incorporates multiple executive functioning domains including inhibitory control and decision making, generally (but not always) show an inverse association with BMI in young adults (Tang et al.2019).

Another higher-order problem-solving skill that overlaps with several executive functioning domains, including inhibitory control and decision making, and may be associated with obesity is delayed discounting (da Matta et al., 2012; Green et al., 2007; Lavagnino et al., 2016). Delayed discounting measures how much one favors smaller immediate rewards to larger delayed rewards (Mazur, 1987; Myerson et al., 2001). A recent systematic review by Tang and colleagues (2019) showed that in the 7 studies with young adult samples (N=7,619 total participants), 4 studies (n=7,337) found a significant association between BMI and delayed discounting, 1 (n=60) found an association between delayed discounting and body fat percentage (but not BMI) and 2 studies (n=222) did not find a relationship. The broader findings of the relationship between obesity and executive functioning and the relationship between obesity and delayed discounting suggest that higher level executive functioning is important to consider when assessing other cognitive associations with obesity.

Emotion Recognition and Executive Functioning.

Executive function plays a central role in the face recognition system, as it involves monitoring facial features and retrieving information from memory about facial cues (Rapcsak, 2019). There is also some suggestion that executive functioning and emotional processing are both associated with similar brain regions such as the prefrontal cortex (Damasio, 1994; Rolls, 2013; Xu et al., 2021; Zhang et al., 2021). Recognizing facial emotional expressions relies on some networks specific to emotional processing but also relies on some general executive functioning skills, like selectively attending to salient facial regions for emotion encoding (Leppänen & Nelson, 2006). Especially when cognitive resources are limited or when facial cues are ambiguous, the executive system is needed to guide facial processing (Rapcsak, 2019). For instance, during a timed emotion recognition task, executive dysfunction may impair performance, as evidenced in research with healthy individuals (Kostandov et al., 2012; Liu et al., 2020). Given this, it is important to control for executive functioning when examining the relationship between BMI and emotion recognition.

A range of executive functions may be important in emotion recognition. For example, ***working memory*** is the amount of information that can be held temporarily in the mind and used to complete cognitive tasks (Cowan, 2014). Working memory capacity has been found to be associated with encoding and processing faces in emotion and facial recognition tasks (Lee & Cho, 2019; Liu et al., 2020) and accurately recognizing emotions (Kouklari et al., 2017). When attentional resources are limited, there is a trade-off between working memory and attention (Lee & Cho, 2019).

Inhibition is the ability to curtail a behavior or response. It requires effortful attention, as refraining from a behavior often requires more effort than exhibiting the behavior (Wong, 2013). Inhibition and effortful attention have also been shown to be related to the broad construct of affective theory of mind as well as performance on emotion recognition tasks. In a study of healthy adults, completing an inhibitory cognitive task (i.e., the Go/No-go task) improved subsequent performance on an emotion recognition task (Kostandov et al., 2016), suggesting the importance of inhibitory control in emotion recognition. Furthermore, ***delayed discounting***, as reviewed earlier, involves inhibitory control over actions in favor of acting impulsively in the moment (Green et al., 2007; Kouklari et al., 2017).

Finally, ***cognitive flexibility,*** also known as shifting or task shifting, is the ability to switch between two different concepts, sets, or tasks (Archambeau & Gevers, 2018). It has been shown to be correlated with affect theory of mind in healthy adults (Laillier et al., 2019). Cognitive inflexibility and poor emotion recognition have been found in patient samples, such as individuals with methamphetamine use disorder (Kim et al., 2011) or autism spectrum disorder (Fabio et al., 2020). However, there has been little research looking at the direct association between cognitive flexibility and emotion recognition accuracy in healthy adult samples.

Limited Research Examining the Contribution of Executive Functioning in the Relationship between Obesity and Emotion Recognition Accuracy.

Few studies have examined the contribution of executive functions to explain the relationship between BMI and emotion recognition accuracy. Manderino and colleagues (2015) found that performance on an executive functioning task (i.e., the Austin Maze) predicted emotion recognition skills in bariatric patients with severe obesity. Caldú and colleagues (2019) found participants (aged 16 to 40 years) with BMIs in the obese range performed more poorly on a facial emotion recognition test than participants with healthy range BMIs. Although cognitive flexibility, as measured by performance on the Wisconsin Card Sort task, was reported to be similar between the group with obesity and the group in the healthy weight range, no formal mediation analysis was conducted. This highlights the importance of controlling for executive functioning when examining the relationship between BMI and emotion recognition.

Sociodemographic and Clinical Co-variates Moderating Emotion Recognition.

In addition to executive functioning, several demographic and clinical covariates need to be considered when looking at emotion recognition. Age may be related to emotion recognition skills. In the general population, emotion recognition skills peak between ages 15 and 30, with some complex emotion recognition skills developing throughout young adulthood (Meinhardt-Injac et al., 2020) ,before emotion recognition starts to progressively decline after 30 (Olderbak et al., 2019). Gender may also be related to emotion recognition accuracy, with females on average scoring higher than males in most studies (Abbruzzese et al., 2019; Miller, 2015; Olderbak et al., 2019; Quintana et al., 2012). Psychiatric conditions, including depression, Attentiondeficit/hyperactivity disorder (ADHD), and anxiety, negatively impact emotion recognition accuracy (Bora & Berk, 2016; Dalili et al., 2015; Demenescu et al., 2010). It is unclear what role race and ethnicity may play in the relationship between obesity and emotion recognition skills, as few studies report these effects. One study included in our meta-analysis showed that Black men relative to non-Black men were less likely to recognize fear and disgust on others' faces (Neel, 2014), suggesting the need to control for race. Additionally, some studies found that educational attainment or higher socioeconomic status was positively correlated with emotion recognition skills, suggesting the need to control for these variables (Elfenbein et al., 2002; McCubbin et al., 2011). Finally, as genetics and heritability play a role in Body Mass Index (Elks et al., 2012). we controlled for twin-status.

CHAPTER 2

SPECIFIC AIMS

Our study extended the extant emotion recognition literature by examining the relationship between BMI and emotion recognition accuracy of facial expressions in a large young adult sample, while controlling for executive functioning, as well as relevant sociodemographic and clinical variables, via a series of hierarchal linear regressions. We did this using the Human Connectome Project (HCP) dataset, which recruited a sample of young adults across a wide BMI range.

Our **Primary Aims** were the following:

Aim 1: **To examine the relationship between BMI and emotion recognition accuracy.**

Hypothesis 1: Higher BMI will be associated with greater difficulties in accurately recognizing emotions from facial expressions, after controlling for sociodemographic and clinical covariates, and executive functioning.

Aim 2: **To examine the role of emotion valence in the relationship between BMI and emotion recognition accuracy.**

Hypothesis 2: The relationship between BMI and accuracy in recognizing emotions from facial expressions will be moderated by emotion valence such that higher BMI will be associated with less accuracy in recognizing negative emotions (i.e., anger, sadness, and fear) but not positive emotions.

The study Open Science Foundation preregistration (French and Chen, 2022) can be found in Appendix A (10.17605/OSF.IO/ZG246).

CHAPTER 3

METHODS

The proposed study aims to utilize publicly available data provided by the Human Connectome Project, Young Adult sample, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University (van Essen et al., 2013). We are utilizing the "1200 Subjects Data Release" behavioral dataset which was released March 2017.

Data was collected at one time-period during a two-day procedure after an initial phone screen and diagnostic interview. During Day 1 participants completed questionnaires, drug screening, blood tests, structural and functional MRI sessions, and NIH toolbox tests (including cognitive functioning measures). On Day 2 participants completed diffusion and functional MRI sessions and non-NIH behavioral tasks (including the Penn Emotion Recognition Test). See Van Essen et al. (2013) and the "WU-Minn HCC 1200 Subjects Data Release Reference Manual" (WU-Minn Consortium, 2018) for more details.

Inclusion/Exclusion Criteria of Dataset (Criteria for Larger HCP Dataset).

The HCP study included 1,206 participants aged 22 to 35 years at the time of the phone screen. Additionally, this dataset included some participants that were twins. Participants were excluded for having a significant previously diagnosed psychiatric or substance abuse history, although some participants were identified as having depression or substance use disorders during study-related diagnostic assessments. Medical exclusions included a history of neurological disorders or events (e.g., stroke), cardiovascular disease, genetic disorders, significant head injuries, premature birth, chemotherapy or radiation, recent thyroid hormone treatment, diabetes, high blood pressure, or daily migraine medication use in the past month. Additionally, participants with as score below 25 on the Mini Mental State Exam (MMSE) or contraindications to MRI procedures (e.g., pregnancy, unsafe metals in body, etc.) were excluded. See Van Essen et al. (2013) for full details of the inclusion and exclusion criteria.

Additional Exclusions for Current Study.

In addition to the criteria needed for the broader HCP study, we added the following inclusion criteria for the subset of participants used in our analysis: women must have reported having (**1**) no history of thyroid disease (e.g., hyperthyroidism or hypothyroidism) as these conditions may affect weight status (Pearce, 2012). As thyroid disease is much more common in women, this question was only asked of female participants. We also excluded individuals with (**2**) a blood alcohol concentration (BAC) of .05% or higher or (**3**) who tested positive for screened acute drugs during study visits, as both alcohol and drug use can influence task performance (Castellano et al., 2015). Additionally, during the study, participants were screened for substance use, and those meeting Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (American Psychiatric Association, 2000) criteria for alcohol abuse or dependence disorder were excluded, as alcohol abuse is associated with poorer facial emotion recognition skills (Castellano et al., 2015). This left us with 832 participants.

Finally, besides these criteria, we excluded participants that did not complete the emotion recognition task (Penn Emotion Recognition Test) or the executive functioning tasks (from the NIH Toolbox Cognition Battery), as these were the primary variables of interest in this proposal. Only eight participants out of 832 did not complete one of these, leaving us with 824 participants that met our inclusion/exclusion criteria.

Measures.

Emotion Recognition

The **Penn Emotion Recognition Test** (Gur et al., 2002) was used to assess how well participants recognized different emotions from pictures of facial expressions. Forty images (five emotions x eight individuals) of faces were balanced for poser's gender, age, and race or ethnicity (Gur et al., 2010). Each facial image was presented individually, and participants were asked to choose which emotion expression was displayed out of the following options: “Happy”, “Sad”, “Angry”, “Scared”, and “No Feeling” (i.e., “Neutral”). Each of the emotion expressions were displayed 8 times. Half of the faces were males and half were females. The number of correct responses (Accuracy measures) was tallied for each emotion: “Happy”, “Sad”, “Angry”, “Scared”, and “No Feeling”. An overall accuracy score for all 5 emotions had been pre-calculated in the HCP dataset by tallying the total correct responses on the Penn Emotion Recognition Test. A Negative Emotion Accuracy score was calculated by tallying the amount of “Sad”, “Angry”, and “Scared” faces correctly recognized.

Sociodemographic Variables

Gender, age (in years), race, ethnicity, educational level (in years completed), in-school/work status, and household income were also measured in this study. Variables that were correlated with BMI, executive functioning, and/or emotion recognition (the outcome variable) will be included in the analysis as covariates.

Dichotomous variables were dummy coded with beta coefficients being reported for women for gender (reference group: 0=men), identifying as Hispanic for ethnicity (reference group: 0=non-Hispanic), and “Yes” to being a student or working full-time (reference group: 0=“No”). Race data was reported using beta coefficients for individuals that identified as: Black/African American, White/Caucasian, Asian/Native Hawaiian/Pacific Islander, Native American/Alaskan Natives, or multiracial (reference group: not reported/unknown).

Clinical variables

BMI was calculated using self-reported height and weight using the formula: 703 * (weight in lbs) /(height in.)^2, which converts weight to the metric of kg/meters^2.

Twin-status was also reported in the HCP dataset using self-report and genotyping for the majority of participants in the sample to confirm whether participants were monozygotic or dizygotic twins. This was also a non-dichotomous variable, with beta coefficients being reported for monozygotic twin-status, dizygotic twin-status, or unknown or unverified zygosity twin-status (reference group: individuals who were not twins).

Depression, anxiety, and ADHD symptoms. Other important clinical variables, including **depression, anxiety, and ADHD symptoms** are controlled for in this study using the age and gender adjusted T-scores from the Achenbach Adult Self-Report (ASR) form (Rescorla & Achenbach, 2004) for the depression, anxiety, & ADHD scales respectively. This measure has good internal and external validity as well as good reliability (Achenbach et al., 2017; Rescorla & Achenbach, 2004). Additionally, we examined lifetime history of a major depressive episode using DSM-IV criteria

(American Psychiatric Association, 2000), with the with beta coefficients being reported for past depressive episodes being present and dummy coded as 1 (reference group: dummy coded as 0 for no lifetime history of a major depressive episode).

Executive Functioning Variables

The HCP dataset measured executive functioning by using the NIH Toolbox Cognition Battery, which includes the **(1)** Dimensional Change Card Sort task (DCCS; measuring cognitive flexibility), **(2)** List Sorting Working Memory Test ("List Sorting"; measuring working memory), and **(3)** Flanker Inhibitory Control and Attention Test (measuring inhibition and attention). These tasks have been shown to have good test-retest reliability and validity in adult samples (Tulsky et al., 2014; Weintraub et al., 2014; Zelazo et al., 2014). Finally, we included the **(4)** Delayed Discounting test (Estle et al., 2006; Green & Myerson, 2004) in our analysis, which is a standalone test that is not part of the NIH Toolbox Cognition Battery.

Cognitive flexibility (also called "set shifting," "task shifting," or "switching") is the ability to switch between two different concepts, sets, or tasks (Archambeau & Gevers, 2018). The DCCS task displays target pictures on the screen that vary along two dimensions: shape and color. Participants are asked to match a series of pictures based on a target picture, first according to one dimension (e.g., shape) and then the other (e.g., color). The rules switch throughout the task, thus set shifting is needed. There are a total of 40 trials, and the task takes about 4 minutes to complete. Scores are standardized based on age (Weintraub et al., 2013).

Working memory. Working memory is the amount of information that can be held temporarily in mind and used to complete cognitive tasks (Cowan, 2014). We used the **List Sorting test** to measure working memory. In this task, a series of stimuli are presented as pictures on a screen with a paired sound clip orally dictating the word associated to it, and participants are instructed to repeat the word. Participants learn stimuli in two trials. In trial 1, participants must compare the size of one category of stimuli (food or animals) and order them by size. Then in trial 2, participants are given stimuli from two categories (food and animals) and must list back all stimuli in order of relative size. The task takes about 7 minutes to complete. Scores are standardized based on age (Weintraub et al., 2013).

Inhibition/attention. **Inhibition** is the ability to curtail a behavior or response. It requires **effortful attention** to control your response, as refraining from a behavior often requires more effort than exhibiting the behavior (Wong, 2013). **The Flanker Inhibitory Control and Attention test** measures this construct by requiring the participant to focus on target stimuli (i.e., arrows) on the center of a computer. A second similar "distractor" stimulus (i.e., arrows) is placed to the left or right of the target. The "distractor" arrows can either be pointing in the same ("congruent") or opposite ("incongruent") direction as the target stimuli. The participant must determine the direction of the target stimuli while inhibiting the distractor stimuli. Scoring is based on a combination of accuracy and reaction time and is standardized based on age. There are 40 trials, and it takes about 4 minutes to complete this task (Weintraub et al., 2013).

Another measure of inhibition, ***delayed discounting***, describes the process of undervaluing larger delayed rewards in favor of smaller immediate rewards (Green & Myerson, 2004; Myerson et al., 2001). Delayed reward amounts in this task were either $200 or $40,000. Immediate rewards amounts were adjusted on a trial-by-trial basis dependent on the participant's response in order to find the indifference point (Estle et al., 2006; Green et al., 2007). The area-under-the-curve is a validated summary measure of how steeply the participant discounts delayed rewards (Estle et al., 2006). In our study, if scores using $200 and $40,000 were highly correlated, we averaged scores on these two subtasks.

Descriptive Statistics and Data Analysis.

Analysis. All data was analyzed using R studio, version 4.1.3. Of note, the significance level for all tests was set to $p < 0.05$ (Schafer, 1999).

Descriptive statistics were gathered for all variables examined in our study (mentioned above for the 824 subjects meeting our inclusion/exclusion criteria). Preliminary correlations were used to examine the relationship between all identified variables in our study. Spearman's correlation was used between two continuous or ranked variables. Cramer's V was used between two categorical variables. A Point-Biserial correlation coefficient (using dummy coding) was calculated to assess the association between dichotomous and continuous variables. Finally, a Kruskal-Wallis test was used to examine the relationship between non-dichotomous nominal (e.g. race) and continuous variables. See *Table 1* for correlations of the sociodemographic variables with BMI, emotion recognition, and executive functioning variables. Extreme values or outliers that affect the results and/or are non-representative of the population were assessed using descriptive statistics. Based on the correlations, we found that the following variables correlated with one or more of our predictors (executive functioning variables or BMI) or our outcome variable (emotion recognition accuracy): age, gender, educational attainment, household income, race, ethnicity, ASR ADHD scale score, ASR

Depression scale score, history of major depressive episodes, and twin status. Therefore, these variables were added in our final model as controls.

Missing data. Missing data can affect statistical inferences. According to Schaffer (Schafer, 1999), data with a missing rate of less than five percent is not likely to influence the results of an analysis. In our sample, there were 25 out of 824 participants (3.03%) that had missing data for any measure included in our analysis. Therefore, we used listwise deletion. Our final sample size was 799 participants. Of note, our final dataset was similar to the full HCP dataset in all examined variables, except for three statistical differences (p's<0.05): 1) years of completed education attainment (final dataset: M=15.04, SD=1.74; full dataset: M=14.86, SD=1.82), 2) Adult Self Report, ADHD scale (final dataset: M=54.30, SD=5.26; full dataset: M=54.91, SD=5.81), and 3) gender (final dataset=percentage female: 61.20%; full dataset=percentage female: 54.39%). No other statistical differences were found.

Testing assumptions of a hierarchal linear regression. Linearity between each of the main predictors (BMI and the executive functioning variables) and the outcome variable (emotion recognition accuracy) was assessed by testing linear versus nonlinear relationships between predictors and outcome variables (quadric, cubic, and quartic relationships compared to linear relationships in regression modeling) and visual inspection of residual plots. The normal distribution of the residuals was assessed through visual inspection (e.g., Q-Q plot and histogram), although the Central Limit Theorem may protect against violations in a large dataset.

Outliers of residuals were assessed using the dfbeta values to check if there were observations significantly influencing the regression models. If there were observations above a sample-adjusted cut-off score (Belsley et al., 1980), which indicate observations that shift the regression coefficient estimate considerably, we performed a sensitivity analysis using the final model results from our original analysis, comparing it to the results of a Robust Linear Regression.

Multicollinearity between predictors was assessed with the Variance Inflation Factor (VIF) & tolerance scores, with scores of above 5 and under 0.20, respectively used to define multicollinearity (O'brien, 2007). Heteroscedasticity was assessed using the Breusch-Pagan Test, and the Durbin-Watson test was used to assess the independence of the residuals. Non-significance of the Breusch-Pagan Test and the Durbin-Watson test suggest that these variables are respectively not significantly heteroscedastic and that residuals are not autocorrelated.

Testing the primary aim. We conducted a hierarchical regression with emotion recognition accuracy as the dependent variable, entering sociodemographic variables and clinical variables first, executive functioning variables second, and BMI third. We chose a hierarchical regression to allow for an assessment of the amount of variance contributed by the predictors entered at each step (i.e., change in R^2). Additionally, by comparing models, we were able to determine when variables entered in each step significantly explained variance (R^2).

Testing the secondary aim. Analyses of our behavioral meta-analysis (French & Chen, 2021) showed that on emotion recognition tasks assessing negative emotions, overweight and obese participants performed significantly more poorly than healthy weight control participants. Given this, we predicted that higher BMI would impact emotion recognition of negative emotions and not positive emotions. Therefore, we planned on testing differences between emotion recognition accuracy scores between emotions measured. If there were no significant differences between emotion type, we would not conduct separate analyses. If all negative emotion variables were correlated with each other, we would combine negative emotions into a composite measure and conduct separate analyses for emotion recognition accuracy for positive ("Happy") and negative emotions ("Sad," Anger," and "Fear").

Power. To our knowledge, no study has been conducted that looked at the relationship between BMI, emotion recognition accuracy, and executive functioning in a young adult sample. In a meta-analysis, we found that emotion recognition abilities varied widely (French & Chen, 2021). We conservatively assumed that the true effect size of BMI predicting emotion recognition skills was small (f^2=0.02; Cohen, 1988). Our "main predictor variables" of emotion recognition accuracy were: **(1)** BMI, **(2)** DCCS task, (**3**) Listing Sorting test, **(4)** Flanker Inhibitory Control and Attention Test, and **(5)** Delayed Discounting averaged score (as scores on the two subtasks were highly correlated with each other, r=0.722).

In our sample, the following sociodemographic variables correlated with one or more of our predictor variables and were identified as "covariates": **(1)** age, **(2)** gender, **(3)** educational attainment, **(4)** income, **(5)** race, **(6)** ethnicity, **(7)** ASR ADHD scale age and gender adjusted T-scores, **(8)** ASR Depression scale score age and gender adjusted T-scores, and **(9)** history of major depressive episodes, and (**10**) twin-status. See *Table 1* for a correlation matrix of variables examined in our analysis.

Using G*Power 3.1.9.4, the following parameters in a sample size collection for a "linear multiple regression: Fixed model, R^2 increase" statistical test type was entered: (1)

f^2=0.02, which assumed a small effect size, (2) α=0.05, (3) power=0.80, (4) the five "main predictors" variables, as above, and (5) 15 total variables, totaling the "control" variables and "main predictors" listed above . A total sample size of 648 participants gives 80% power, when all 15 variables are added. Since our sample exceeds this amount, we have sufficient power to determine a small effect size in our analyses.

CHAPTER 4

RESULTS

Preliminary Testing

Overall Emotion Recognition Analysis; see *Table 2* for sample characteristics.

Assumption testing. Normality of the residuals was within normal limits upon visual inspection, although there was some kurtosis at the tail ends of the data for the overall emotion recognition variable. Given the large sample size, the Central Limit Theorem suggests that departures from normality are less impactful (Lumley et al., 2002). Examination of dfbeta values revealed 453 out of 17,578 values were influential outliers (2.58% of the values). There was no evidence of multicollinearity (all VIF's < 2 and tolerance values > 0.65), heteroscedasticity (Breusch–Pagan (21)=27.17, p=0.165), or autocorrelation (Durbin-Watson=2.01, p=0.546). There was also no evidence of any predictor variables having a nonlinear relationship with the outcome variable.

Sensitivity analysis. As some influential outliers were identified, we conducted a sensitivity analysis of the final model, comparing the original ordinary least squares (OLS) hierarchical linear model results with a Robust Linear Model (RLM) using an MM-estimator; this estimator was chosen as it was less sensitive to outliers and has high efficiency (Wilcox, 2012; Yohai,1987). Both models identified the same significant predictors and had similar R^2 values (OLS=0.074; RLM=0.060). For the sake of broader interpretability, we report results of the original OLS model. See Appendix A, *Supplemental Table 1* for results of the RLM.

Negative Emotion Recognition

Assumption testing. When examining valence of emotion recognition accuracy, a Kruskal-Wallis comparison showed that emotional recognition accuracy was significantly different between different emotion types, $H(4)=940$, $p<0.001$. Additionally, a Spearman's correlation matrix showed accuracy scores for each negative emotion (sadness, fear, and anger) were positively correlated with each other (see Appendix C, *Supplemental Table 2*, for the correlation matrix). Therefore, as planned, the accuracy scores for the three negative emotion scores were combined into a composite score.

After looking at the dfbeta values, we identified 449 outliers out of the 17,578 dfbeta values (2.55 % of the values). Normality of the residuals was within normal limits upon visual inspection, although some kurtosis was present. There was no evidence of multicollinearity (all VIF's < 2 and tolerance values > 0.65), heteroscedasticity (Breusch–Pagan (21)=26.39, $p=0.192$), or autocorrelation (Durbin-Watson=1.99, $p=0.423$).

Sensitivity analysis. We also carried out a sensitivity analysis comparing the original OLS model results with a Robust Linear Model (RLM) for negative emotion recognition accuracy, using the MM-estimator (Yohai, 1987). The OLS identified gender, ethnicity, DCCS, and the List Sorting test as significant predictors. RLM also identified these predictors, except that List Sorting and gender were marginally significant (*p's* =0.073 for both). The RLM also identified history of depressive episodes and Asian Americans/Pacific Islander as significant predictors. The R^2 values between both models was similar (OLS=0.058; RLM=0.060). Given the results were similar between analyses, we decided to report the results of the OLS for ease of interpretability. See Appendix B, *Supplemental Table 1* for RLM results of the final model.

Preliminary Testing for Positive Emotion Recognition – Analysis Not Conducted

In examining influential outliers, we identified 208 out of 17,578 dfbeta values (1.18% of values). There was no evidence of multicollinearity (all VIF's < 2 and tolerance values > 0.65), heteroscedasticity (Breusch–Pagan (21)=13.35, *p*=0.896), or autocorrelation (Durbin-Watson=2.05, *p*=0. 0.733). Upon visual inspections of the residuals for the regression model, data appeared nonlinear. Of note, there was very little variance in positive emotion accuracy scores, with most participants getting a perfect score. See Appendix D, *Supplemental Table 3*, for the HLR model results for positive emotion recognition accuracy, which produced a nonsignificant model.

Primary Analysis

HLR for Overall Emotional Recognition Accuracy

Sociodemographic and clinical variables. *See Table 3* for the HLR model results for overall emotional accuracy. Sociodemographic and clinical variable predictors for emotion recognition produced a significant overall model fit. Ethnicity was a predictor of emotion recognition, such that individuals identifying as Hispanic had higher emotion recognition accuracy scores than non-Hispanic individuals.

Sociodemographic, clinical, and executive functioning variables. When executive functioning variables were added to model the overall model was significantly improved, Δ Adjusted R^2=0.031, $F(4)$= 6.56, p<0.001. The DCCS task (measuring cognitive flexibility) and the List Sorting test (assessing working memory both significantly contributed to the model, and ethnicity remained a significant predictor.

Final model – sociodemographic, clinical, executive functioning, and BMI variables. After adding BMI, our final model demonstrated a significant overall model fit

[F(21, 777)=2.96, p <0.001]. Ethnicity, cognitive flexibility, and working memory remained significant. The BMI beta coefficient was not significant (p=0.088) and adding BMI to the model did not improve overall model fit, Δ Adjusted R^2=0.002, F(1)=2.92, p= 0.088), indicating BMI was not a predictor of overall emotion recognition accuracy when controlling for other variables in the model.

HLR for Negative Emotion Recognition Accuracy

Sociodemographic and clinical variables. *See Table 4* for the HLR model results for negative emotional accuracy responses. Sociodemographic and clinical variables predictors for negative emotion recognition accuracy produced a significant model fit. Ethnicity was a predictor of negative emotion recognition, such that individuals identifying as Hispanic had higher accuracy of identifying negative emotions. Additionally, higher education levels also predicted better emotion recognition accuracy.

Sociodemographic, clinical, and executive functioning variables. After adding the executive functioning variables, the model remained significant. Ethnicity remained a significant predictor, while education level was no longer significant (p=0.070). Gender also became a significant predictor, such that women were more accurate at recognizing negative emotions than men. DCCS task performance and the List Sorting test performance both significantly contributed to the model. Additionally, adding the executive functioning variables significantly improved the model, ΔR^2=0.016, F(4)=4.31, p=0.002, supporting that cognitive flexibility and working memory predict emotion recognition accuracy.

Final model – BMI added to model with sociodemographic, clinical, and executive functioning variables. After adding BMI, our final model was still significant.

Ethnicity, gender, the DCCS task, and the List Sorting test also remained significant predictors. The BMI beta coefficient was not significant (p=0.357) and adding BMI to the model did not improve overall model fit, Δ Adjusted R^2<0.001, $F(1)$= 0.85, p=0.357), indicating BMI was not an independent predictor of negative emotion recognition accuracy.

CHAPTER 5

DISCUSSION

Contrary to our hypothesis, BMI was not a significant predictor of accuracy in recognizing emotions or recognizing negative emotions from facial expressions when controlling for socio-demographic, clinical, and executive functioning variables. Instead, executive functioning and certain sociodemographic variables contributed to differences in emotion recognition of facial expressions. In the final model, better overall emotion recognition accuracy for facial expressions was predicted by Hispanic identity status, greater cognitive flexibility (i.e., the DCCS task), and better working memory (i.e., the List Sorting test). Similarly, greater accuracy in recognizing negative emotions from facial expressions was predicted by Hispanic identity status, greater cognitive flexibility and better working memory. Being female also predicted better accuracy identifying negative facial emotions.

BMI was not significant in predicting emotion recognition accuracy*, *over and above socio-demographic and clinical variables, and executive functioning. Contrary to our hypothesis and previous literature, we did not find a positive relationship between BMI and emotion recognition. One reason for the difference in findings may be that most studies examining the relationship between BMI and accuracy in recognizing emotions have not controlled for executive functioning or other clinical variables. For example, the only study we could find of young adults showed that the high BMI group performed more poorly on an emotion recognition task than a healthy-weight BMI group. However, this study did not control for differences in executive functioning, nor did it examine clinical variables other than medical conditions (Balter et al., 2021).

Other studies have controlled for clinical covariates and executive functioning covariates and still found significant differences between BMI groups. For example, another study found that participants (aged 16 to 40 years) with BMIs in the obese range performed more poorly on an emotion recognition test than participants with healthy range BMIs, even when controlling for depressive symptoms (i.e., as measured by the Hospital Anxiety and Depression Scale), cognitive flexibility (i.e., the Wisconsin Card Sorting Test (WCST), and verbal IQ (Caldú et al., 2019). Of note, both of these studies treated BMI as a dichotomous variable (Balter et al., 2021; Caldú et al., 2019), and it is unclear if these results would be different if BMI was treated as a continuous variable, as we did in the current study.

Our findings are more similar to studies that have included individuals with high BMIs, such as bariatric patients, that covary with clinical, sociodemographic, and executive functioning variables. In a large mixed sample of healthy adults and bariatric candidates (BMI: M=34.91; SD=10.98), BMI was not found to predict emotion recognition accuracy after controlling for executive functioning and other sociodemographic and clinical variables (Miller, 2015). Similar to our findings, Miller (2015) found that a composite executive functioning variable was correlated with emotion recognition accuracy. Another study of bariatric patients showed that performance on an executive functioning task (i.e., the Austin Maze) but not BMI predicted emotion recognition accuracy (Manderino et al., 2015). Although findings from these studies are difficult to generalize to a population of BMI-diverse healthy young adults, when taken together with our study, they suggest that BMI does not independently predict emotion recognition accuracy.

Our lack of findings for the relationship between BMI and emotion recognition accuracy reflect the equivocal findings of our behavior meta-analysis (French & Chen, 2021). This meta-analysis across 12 studies with N=1,259 showed that there was a non-significant but medium-sized association between poor overall emotion recognition accuracy in obese or overweight BMI range relative to individuals in the healthy BMI range. There was also a small but significant association between poorer negative emotion recognition in individuals in the overweight and obese range relative to the healthy weight group. Differences between our current study using the HCP dataset and our behavioral meta-analysis may be related to the use of different emotion recognition tasks and the use of a dataset of young adults. However, while the study of the relationship between BMI and emotion processing is still nascent, it is important to note that the size of our current study using the HCP dataset was eight times the average size of the studies included in our behavioral meta-analysis. This gives us reasonable confidence in being able to reject the hypothesis that higher BMI is associated with poorer emotion recognition of facial expressions.

Greater cognitive flexibility and larger working memory capacity, rather than BMI, is associated with enhanced overall and negative emotion. We found that cognitive flexibility (i.e., the DCCS task) and working memory (on the List Sorting test) were both significant predictors of overall and negative emotion recognition. Recognizing emotions from facial expressions requires skill in being able to switch attention and focus from one stimulus to the next. Working memory capacity may be important in comparing recalled facial expressions with the facial task stimuli presented in an emotion recognition task. Working memory may be more related to emotion recognition when attentional sources are limited (Lee & Cho, 2019). Unfortunately, there are few studies measuring cognitive flexibility, working memory, emotion recognition, and BMI in healthy adult samples. One study of adults with a healthy BMI and adults with obesity (bariatric surgery patients) found a significant correlation between poorer working memory (digit span and another spatial memory test) and higher emotion recognition accuracy (Miller, 2015). Although this study found that poorer executive functioning (including cognitive flexibility) was associated with poorer emotion recognition, they did not break down their results by specific tasks, making it difficult to fully understand emotion recognition's relationship with cognitive flexibility (Miller, 2015).

Imaging and behavioral studies of patients with lesions offer some insights into the relationship between executive functioning and emotional processing. For example, Ouerchefani et al. (2022) found that in patients with prefrontal damage, facial emotion recognition was correlated with cognitive flexibility and inhibition but not working memory. They also conducted a voxel-based lesion symptoms mapping conjunction analysis, which showed that overlapping regions in the ventromedial and dorsomedial prefrontal regions were associated with both negative emotion recognition problems and poorer cognitive flexibility and inhibition (Ouerchefani et al., 2022). Similarly, another prefrontal lesion study found an association between verbal fluency and emotion recognition on behavioral tasks (Martins et al., 2011). Although these studies suggest that executive functioning problems may contribute to emotion recognition abilities, most studies have not examined this relationship or looked at the additional impact of BMI

(Callahan et al., 2011; Hopkins et al., 2002; Hornak et al., 1996, 2003; Tsuchida & Fellows, 2012).

Although few imaging studies directly look at associations between emotion recognition and executive functioning, overlapping brain regions can be identified by examining meta-analyses for emotion recognition (Xu et al., 2021) and executive functioning (Zhang et al., 2021). Although the precise overlap of brain regions between emotion recognition and executive functioning is difficult to confirm without the report of the coordinates of significant regions (Xu et al., 2021), qualitative comparison suggests that the insula is involved in both negative emotion recognition processing and cognitive flexibility (Xu et al., 2021; Zhang et al., 2021). Regions involved in both negative emotion recognition and working memory appear to share activation in the ventrolateral prefrontal cortex or inferior frontal gyrus, the pars opercularis, and the pars triangularis (Xu et al., 2021; Zhang et al., 2021). The inferior frontal gyrus is regarded as part of the ventral attention network but has overlap with the salience network particularly in its connectivity with the fronto-insular region, with the anterior insula being important in the salience network (Menon & D'Esposito, 2022).

Another large-scale imaging study of 759 children and young adults, ages 8- to 23-years-old in the Philadelphia Neurodevelopmental Cohort, has confirmed the importance of the insula as part of the salience network in differential activation to a facial emotion recognition task (Zhang et al., 2019). However, this study and other meta-analyses (Fusar-Poli et al., 2009; Kober et al., 2008) did not find that the inferior frontal gyrus (part of the ventral attention network) was involved in facial emotion recognition. Instead, this study's findings implicate the medial prefrontal regions such as the ventromedial and dorsomedial prefrontal regions in facial emotion recognition – regions that have been found to be more associated with differences in BMI (Chen et al., 2018, 2020; García-García et al., 2019; Herrmann et al., 2019). The meta-analysis by Zhang et al. (2019) did not assess the relationship between the brain circuits associated with facial emotion recognition and executive functions.

Ethnicity and both overall and negative emotion recognition. Hispanic identity was associated with emotion recognition accuracy, such that individuals that identified as Hispanic scored better on both overall and negative emotion accuracy measures. While some research suggests that participants are better at identifying emotions in facial expressions of their own race (Elfenbein et al., 2002), there is little research examining the role of cultural identity on emotion recognition accuracy in Latinx samples. Some research suggests that differences between ethnic groups may be due to cultural factors that discourage or encourage the expression of certain emotions (Engelmann et al., 2013). In a meta-analysis examining the effects between emotion recognition and BMI-status, only 3 out of 15 papers reported racial or ethnic demographic information (French & Chen, 2021). As only a relatively small amount of participants in our study identified as Hispanic, further research is needed to confirm the association and should also explore the role of culture and other covariates that may mediate this result.

Gender and negative emotion recognition. Consistent with most studies, gender predicted negative emotion recognition, with women more accurately recognizing negative emotions. Other studies also found that women may score higher on emotion recognition tests (Abbruzzese et al., 2019; Miller, 2015; Olderbak et al., 2019; Quintana et al., 2012). Past research shows that women tend to attend to negative emotion stimuli more than men and have increased brain activity during the processing of negatively valenced stimuli (Abbruzzese et al., 2019).

Strengths, limitations, and future research. One limitation to our study was that the version of the emotion recognition task used may have produced a ceiling effect. The average participant (median score) correctly identified 90% (36 out of 40) of all the emotional stimuli and 87.5% (21 out of 24) of the negative emotional stimuli. In the future, studies should use more ambiguous /complex emotional stimuli (e.g. subthreshold facial emotions) to increase the difficulty in detecting emotions and thus reduce these ceiling effects (Johnston et al., 2003). Additionally, future studies may want to focus on complex emotional expressions, as there is a wider range of abilities for identifying complex emotions in healthy samples (Baron-Cohen et al., 2001). Although the HCP dataset (Barch et al., 2013) does not include tasks assessing theory of mind using facial expressions of emotion, future use of the HCP dataset may examine use the Frith-Happe animation task (White et al., 2011), which examines other aspects of theory of mind as an outcome (e.g., Yang et al., 2018). This may be important to explore as accurate recognition of emotions requires an understanding of mental and emotional states in oneself and in others, and this accurate emotion recognition has shown to implicate the prefrontal cortex (Laillier et al., 2019; Premack & Woodruff, 1978).

Our study examined emotion recognition in a large, healthy young adult sample while controlling for executive functioning variables. Additionally, we had the power to detect small effect sizes while adding and controlling for a number of predictors in our model. To our knowledge, our study is the largest to examine the relationship between emotion recognition and BMI in youth. Given this, we have confidence in rejecting the hypotheses that higher BMI is associated with greater difficulties in accurately recognizing emotions generally and negative emotions, particularly. However, our findings suggest that executive functioning, rather than BMI, explains some of the differences in emotion recognition accuracy in a sample of relatively young and healthy adults. Future studies examining the relationship between the brain regions associated with emotion recognition, cognitive flexibility, and working memory are needed. Additionally, because this study is cross-sectional, it cannot provide information on whether earlier challenges in executive functioning and weight gain predict problems in emotion and social functioning later in life, including emotion recognition. Therefore, more longitudinal studies across the lifespan are needed to examine the relationship between weight gain and obesity, executive functioning, and emotion recognition.

BIBLIOGRAPHY

Abbruzzese, L., Magnani, N., Robertson, I. H., & Mancuso, M. (2019). Age and Gender Differences in Emotion Recognition. *Frontiers in Psychology*, *10*, 2371. https://doi.org/10.3389/fpsyg.2019.02371

Achenbach, T. M., Ivanova, M. Y., & Rescorla, L. A. (2017). Empirically based assessment and taxonomy of psychopathology for ages 1½–90+ years: Developmental, multi-informant, and multicultural findings. *Comprehensive Psychiatry*, *79*, 4–18. https://doi.org/10.1016/j.comppsych.2017.03.006

Albano, G., Rowlands, K., Baciadonna, L., Coco, G. lo, & Cardi, V. (2019). Interpersonal difficulties in obesity: A systematic review and meta-analysis to inform a rejection sensitivity-based model. *Neuroscience & Biobehavioral Reviews*, *107*, 846–861. https://doi.org/10.1016/j.neubiorev.2019.09.039

American Psychiatric Association. (2000). Diagnostic and statistical manual of mental disorders (4th ed., text rev.). Washington, DC: Author.

Archambeau, K., & Gevers, W. (2018). *Chapter 16 - (How) Are Executive Functions Actually Related to Arithmetic Abilities?* (A. Henik & W. B. T.-H. of F. in N. C. Fias, Eds.; pp. 337–357). Academic Press. https://doi.org/10.1016/B978-0-12-811529-9.00016-9

Balter, L. J., Raymond, J. E., Aldred, S., Higgs, S., & Bosch, J. A. (2021). Age, BMI, and inflammation: Associations with emotion recognition. *Physiology and Behavior*, *232*, 113324. https://doi.org/10.1016/j.physbeh.2021.113324

Barch, D. M., Burgess, G. C., Harms, M. P., Petersen, S. E., Schlaggar, B. L., Corbetta, M., Glasser, M. F., Curtiss, S., Dixit, S., Feldt, C., Nolan, D., Bryant, E., Hartley, T., Footer, O., Bjork, J. M., Poldrack, R., Smith, S., Johansen-Berg, H., Snyder, A. Z., … WU-Minn HCP Consortium. (2013). Function in the human connectome: Task-fMRI and individual differences in behavior. *NeuroImage*, *80*, 169–189. https://doi.org/10.1016/j.neuroimage.2013.05.033

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The “Reading the mind in the eyes” Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry*, *42*(2), 241–251. https://doi.org/10.1111/1469-7610.00715

Barrett, L. F. (1998). Discrete Emotions or Dimensions? The Role of Valence Focus and Arousal Focus. *Cognition and Emotion*, *12*(4), 579–599. https://doi.org/10.1080/026999398379574

Belsley, D. A., Kuh, E., & Welsch, R. E. (1980). Detecting Influential Observations and Outliers. In *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity* (pp. 6-84). John Wiley & Sons, Inc.

Bora, E., & Berk, M. (2016). Theory of mind in major depressive disorder: A meta-analysis. *Journal of Affective Disorders*, *191*, 49–55. https://doi.org/10.1016/j.jad.2015.11.023

Caldú, X., Ottino-González, J., Sánchez-Garre, C., Hernan, I., Tor, E., Sender-Palacios, M.-J., Dreher, J.-C., Garolera, M., & Jurado, M. Á. (2019). Effect of the catechol-O-methyltransferase Val 158Met polymorphism on theory of mind in obesity. *European Eating Disorders Review*, *27*(4), 401–409. https://doi.org/10.1002/erv.2665

Callahan, B. L., Ueda, K., Sakata, D., Plamondon, A., & Murai, T. (2011). Liberal bias mediates emotion recognition deficits in frontal traumatic brain injury. *Brain and Cognition*, *77*(3), 412–418. https://doi.org/10.1016/j.bandc.2011.08.017

Castellano, F., Bartoli, F., Crocamo, C., Gamba, G., Tremolada, M., Santambrogio, J., Clerici, M., & Carrà, G. (2015). Facial emotion recognition in alcohol and substance use disorders: A meta-analysis. *Neuroscience & Biobehavioral Reviews*, *59*, 147–154. https://doi.org/10.1016/j.neubiorev.2015.11.001

Centers for Disease Control and Prevention. (2022). *Defining Adult Overweight & Obesity*. Obesity. Retrieved from https://www.cdc.gov/obesity/basics/adult-defining.html

Centers for Disease Control and Prevention. (2021). *Nutrition, Physical Activity, and Obesity: Data, Trends and Maps*. https://www.cdc.gov/nccdphp/dnpao/data-trends-maps/index.html

Chen, E. Y., Eickhoff, S. B., Giovannetti, T., & Smith, D. v. (2020). Obesity is associated with reduced orbitofrontal cortex volume: A coordinate-based meta-analysis. *NeuroImage. Clinical*, *28*, 102420. https://doi.org/10.1016/j.nicl.2020.102420

Chen, E. Y., Murray, S., Giovannetti, T., & Smith, D. v. (2018). Reduced gray matter volume in the orbitofrontal cortex is associated with greater body mass index: a coordinate-based meta-analysis. *BioRxiv*, 359919. https://doi.org/10.1101/359919

Cheng, H. L., Medlow, S., & Steinbeck, K. (2016). The Health Consequences of Obesity in Young Adulthood. *Current Obesity Reports*, *5*, 30–37. https://doi.org/10.1007/s13679-016-0190-2

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences, 2nd ed.* Lawrence Erlbaum Associates, Publishers. https://doi.org/10.1234/12345678

Cowan, N. (2014). Working Memory Underpins Cognitive Development, Learning, and Education. *Educational Psychology Review*, *26*, 197–223. https://doi.org/10.1007/s10648-013-9246-y

Cserjési, R., Vermeulen, N., Lénárd, L., & Luminet, O. (2011). Reduced capacity in automatic processing of facial expression in restrictive anorexia nervosa and obesity. *Psychiatry Research*, *188*(2), 253–257. https://doi.org/10.1016/j.psychres.2010.12.008

da Matta, A., Gonçalves, F. L., & Bizarro, L. (2012). Delay discounting: Concepts and measures. *Psychology & Neuroscience*, *5*(2), 135–146. https://doi.org/10.3922/j.psns.2012.2.03

Dalili, M. N., Penton-Voak, I. S., Harmer, C. J., & Munafò, M. R. (2015). Meta-analysis of emotion recognition deficits in major depressive disorder. *Psychological Medicine*, *45*(6), 1135–1144. https://doi.org/10.1017/S0033291714002591

Damasio, A. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. AVON BOOKS.

Demenescu, L. R., Kortekaas, R., den Boer, J. A., & Aleman, A. (2010). Impaired Attribution of Emotion to Facial Expressions in Anxiety and Major Depression. *PLoS ONE*, *5*(12), e15058. https://doi.org/10.1371/journal.pone.0015058

Dietz, W. H. (2017). Obesity and Excessive Weight Gain in Young Adults: New Targets for Prevention. *JAMA*, *318*(3), 241–242. https://doi.org/10.1001/jama.2017.6119

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, *6*(3–4), 169–200. https://doi.org/10.1080/02699939208411068

Ekman, P., Sorenson, E. R., & Friesen, W. v. (1969). Pan-Cultural Elements in Facial Displays of Emotion. *Science*, *164*(3875), 86 LP – 88. https://doi.org/10.1126/science.164.3875.86

Elfenbein, H. A., Marsh, A. A., & Ambady, N. (2002). Emotional intelligence and the recognition of emotion from facial expressions. In *The wisdom in feeling: Psychological processes in emotional intelligence.* (pp. 37–59). The Guilford Press.

Elks, C., den Hoed, M., Zhao, J. H., Sharp, S., Wareham, N., Loos, R., & Ong, K. (2012). Variability in the Heritability of Body Mass Index: A Systematic Review and Meta-Regression. *Frontiers in Endocrinology*, *3*. https://www.frontiersin.org/articles/10.3389/fendo.2012.00029

Engelmann, J. B., Pogosyan, M., Roberson, D., Da, S., & Ramos, S. (2013). *FOCUSED REVIEW Emotion perception across cultures: the role of cognitive mechanisms*, *4*. https://doi.org/10.3389/fpsyg.2013.00118

Erikson, E. H. (1982). *The life cycle completed*. Norton.

Estle, S. J., Green, L., Myerson, J., & Holt, D. D. (2006). Differential effects of amount on temporal and probability discounting of gains and losses. *Memory & Cognition*, *34*, 914–928. https://doi.org/10.3758/BF03193437

Fabio, R. A., Esposito, S., Carrozza, C., Pino, G., & Caprì, T. (2020). Correlations between facial emotion recognition and cognitive flexibility in autism spectrum disorder. *Advances in Autism*, *6*(3), 195–204. https://doi.org/10.1108/AIA-02-2019-0005

French, M. N., & Chen, E. Y. (2021). The association of Body Mass Index and Emotion Recognition: A Meta-Analysis. *[Unpublished Manuscript]. Temple University*.

Fusar-Poli, P., Placentino, A., Carletti, F., Landi, P., Allen, P., Surguladze, S., Benedetti, F., Abbamonte, M., Gasparotti, R., Barale, F., Perez, J., McGuire, P., & Politi, P. (2009). Functional atlas of emotional faces processing: a voxel-based meta-analysis of 105 functional magnetic resonance imaging studies. *Journal of Psychiatry & Neuroscience : JPN*, *34*(6), 418–432.

García-García, I., Michaud, A., Dadar, M., Zeighami, Y., Neseliler, S., Collins, D. L., Evans, A. C., & Dagher, A. (2019). Neuroanatomical differences in obesity: meta-analytic findings and their validation in an independent dataset. *International Journal of Obesity (2005)*, *43*, 943–951. https://doi.org/10.1038/s41366-018-0164-4

Global BMI Mortality Collaboration, Di Angelantonio, E., Bhupathiraju, S., Wormser, D., Gao, P., Kaptoge, S., Berrington de Gonzalez, A., Cairns, B., Huxley, R., Jackson, C., Joshy, G., Lewington, S., Manson, J., Murphy, N., Patel, A., Samet, J., Woodward, M., Zheng, W., Zhou, M., … Hu, F. (2016). Body-mass index and all-cause mortality: individual-participant-data meta-analysis of 239 prospective studies in four continents. *Lancet*, *388*(10046), 776–786. https://doi.org/10.1016/S0140-6736(16)30175-1

Green, L., & Myerson, J. (2004). A discounting framework for choice with delayed and probabilistic rewards. *Psychological Bulletin*, *130*(5), 769–792. https://doi.org/10.1037/0033-2909.130.5.769

Green, L., Myerson, J., Shah, A. K., Estle, S. J., & Holt, D. D. (2007). Do adjusting-amount and adjusting-delay procedures produce equivalent estimates of subjective value in pigeons? *Journal of the Experimental Analysis of Behavior*, *87*(3), 337–347. https://doi.org/10.1901/jeab.2007.37-06

Gur, R. C., Richard, J., Hughett, P., Calkins, M. E., Macy, L., Bilker, W. B., Brensinger, C., & Gur, R. E. (2010). A cognitive neuroscience-based computerized battery for efficient measurement of individual differences: Standardization and initial construct validation. *Journal of Neuroscience Methods*, *187*(2), 254–262. https://doi.org/10.1016/j.jneumeth.2009.11.017

Gur, R. C., Sara, R., Hagendoorn, M., Marom, O., Hughett, P., Macy, L., Turner, T., Bajcsy, R., Posner, A., & Gur, R. E. (2002). A method for obtaining 3-dimensional facial expressions and its standardization for use in neurocognitive studies. *Journal of Neuroscience Methods*, *115*(2), 137–143. https://doi.org/10.1016/S0165-0270(02)00006-7

Herrmann, M. J., Tesar, A.-K., Beier, J., Berg, M., & Warrings, B. (2019). Grey matter alterations in obesity: A meta-analysis of whole-brain studies. *Obesity Reviews*, *20*(3), 464–471. https://doi.org/10.1111/obr.12799

Hopkins, M. J., Dywan, J., & Segalowitz, S. J. (2002). Altered electrodermal response to facial expression after closed head injury. *Brain Injury*, *16*(3), 245–257. https://doi.org/10.1080/02699050110103346

Hornak, J., Bramham, J., Rolls, E. T., Morris, R. G., O’Doherty, J., Bullock, P. R., & Polkey, C. E. (2003). Changes in emotion after circumscribed surgical lesions of the orbitofrontal and cingulate cortices. *Brain*, *126*(7), 1691–1712. https://doi.org/10.1093/brain/awg168

Hornak, J., Rolls, E. T., & Wade, D. (1996). Face and voice expression identification in patients with emotional and behavioural changes following ventral frontal lobe damage. *Neuropsychologia*, *34*(4), 247–261. https://doi.org/10.1016/0028-3932(95)00106-9

Hruby, A., Manson, J. E., Qi, L., Malik, V. S., Rimm, E. B., Sun, Q., Willett, W. C., & Hu, F. B. (2016). Determinants and Consequences of Obesity. *American Journal of Public Health*, *106*(9), 1656–1662. https://doi.org/10.2105/AJPH.2016.303326

Johnston, P. J., McCabe, K., & Schall, U. (2003). Differential susceptibility to performance degradation across categories of facial emotion--a model confirmation. *Biological Psychology*, *63*(1), 45–58. https://doi.org/10.1016/s0301-0511(03)00026-7

Kim, Y.-T., Kwon, D.-H., & Chang, Y. (2011). Impairments of facial emotion recognition and theory of mind in methamphetamine abusers. *Psychiatry Research*, *186*(1), 80–84. https://doi.org/10.1016/j.psychres.2010.06.027

Kober, H., Barrett, L. F., Joseph, J., Bliss-Moreau, E., Lindquist, K., & Wager, T. D. (2008). Functional grouping and cortical–subcortical interactions in emotion: A meta-analysis of neuroimaging studies. *NeuroImage*, *42*(2), 998–1031. https://doi.org/10.1016/j.neuroimage.2008.03.059

Koch, A., & Pollatos, O. (2015). Reduced facial emotion recognition in overweight and obese children. *Journal of Psychosomatic Research*, *79*(6), 635–639. https://doi.org/10.1016/j.jpsychores.2015.06.005

Kostandov, E. A., Cheremushkin, E. A., Petrenko, N. E., & Yakovenko, I. A. (2016). A cognitive hypothesis of the development of differentiative cortical inhibition in humans. *Human Physiology*, *42*(4), 351–360. https://doi.org/10.1134/S0362119716040083

Kostandov, E. A., Cheremushkin, E. A., Yakovenko, I. A., & Ashkinazi, M. L. (2012). The Role of the Context of Cognitive Activity in the Recognition of Facial Emotional Expressions. *Neuroscience and Behavioral Physiology*, *42*, 293–301. https://doi.org/10.1007/s11055-012-9565-3

Kouklari, E.-C., Thompson, T., Monks, C. P., & Tsermentseli, S. (2017). Hot and Cool Executive Function and its Relation to Theory of Mind in Children with and without Autism Spectrum Disorder. *Journal of Cognition and Development*, *18*(4), 399–418. https://doi.org/10.1080/15248372.2017.1339708

Laillier, R., Viard, A., Caillaud, M., Duclos, H., Bejanin, A., de La Sayette, V., Eustache, F., Desgranges, B., & Laisney, M. (2019). Neurocognitive determinants of theory of mind across the adult lifespan. *Brain and Cognition*, *136*, 103588. https://doi.org/10.1016/j.bandc.2019.103588

Lavagnino, L., Arnone, D., Cao, B., Soares, J. C., & Selvaraj, S. (2016). Inhibitory control in obesity and binge eating disorder: A systematic review and meta-analysis of neurocognitive and neuroimaging studies. *Neuroscience & Biobehavioral Reviews*, *68*, 714–726. https://doi.org/10.1016/j.neubiorev.2016.06.041

Lee, H. J., & Cho, Y. S. (2019). Memory facilitation for emotional faces: Visual working memory trade-offs resulting from attentional preference for emotional facial expressions. *Memory & Cognition*, *47*, 1231–1243. https://doi.org/10.3758/s13421-019-00930-8

Leppänen, J. M., & Nelson, C. A. (2006). The development and neural bases of facial emotion recognition. *Advances in Child Development and Behavior*, *34*, 207–246. https://doi.org/10.1016/S0065-2407(06)80008-X

Lindquist, K. A., Satpute, A. B., Wager, T. D., Weber, J., & Barrett, L. F. (2016). The Brain Basis of Positive and Negative Affect: Evidence from a Meta-Analysis of the Human Neuroimaging Literature. *Cerebral Cortex*, *26*(5), 1910–1922. https://doi.org/10.1093/cercor/bhv001

Lumley, T. & Diehr, P., Emerson, S., & Chen, L. (2002). The Importance of the Normality Assumption in Large Public Health Data Sets. *Annual review of public health*. *23*, 151-69. https://doi.org/10.1146/annurev.publhealth.23.100901.140546

Liu, M., Zhou, L., Wang, X., & Ye, B. (2020). Sad expressions during encoding attenuate recognition of facial identity in visual working memory: behavioural and electrophysiological evidence. *Cognition and Emotion*, *34*(6), 1271–1283. https://doi.org/10.1080/02699931.2020.1726291

Luppino, F. S., de Wit, L. M., Bouvy, P. F., Stijnen, T., Cuijpers, P., Penninx, B. W. J. H., & Zitman, F. G. (2010). Overweight, Obesity, and Depression: A Systematic Review and Meta-analysis of Longitudinal Studies. *Archives of General Psychiatry*, *67*(3), 220–229. https://doi.org/10.1001/archgenpsychiatry.2010.2

Manderino, L., Spitznagel, M. B., Strain, G., Devlin, M., Cohen, R., Crosby, R. D., Mitchell, J. E., & Gunstad, J. (2015). Cognitive dysfunction predicts poorer emotion recognition in bariatric surgery candidates. *Obesity Science & Practice*, *1*, 97–103. https://doi.org/10.1002/osp4.9

Martins, A. T., Faísca, L., Esteves, F., Muresan, A., Justo, M. G., Simão, C., & Reis, A. (2011). Traumatic brain injury patients: Does frontal brain lesion influence basic emotion recognition? *Psychology & Neuroscience*, *4*(3), 377–384. https://doi.org/10.3922/j.psns.2011.3.011

Mazur, J. E. (1987). An adjusting procedure for studying delayed reinforcement. In M. L. Commons, J. E. Mazur, J. A. Nevin, & H. Rachlin (Eds.), *The effect of delay and of intervening events on reinforcement value.* Lawrence Erlbaum Associates, Inc.

McCubbin, J. A., Merritt, M. M., Sollers, J. J. I. I. I., Evans, M. K., Zonderman, A. B., Lane, R. D., & Thayer, J. F. (2011). Cardiovascular-Emotional Dampening: The Relationship Between Blood Pressure and Recognition of Emotion. *Psychosomatic Medicine*, *73*(9), 743–750. https://doi.org/10.1097/PSY.0b013e318235ed55

Meinhardt-Injac, B., Kurbel, D., & Meinhardt, G. (2020). The coupling between face and emotion recognition from early adolescence to young adulthood. *Cognitive Development*, *53*, 100851. https://doi.org/10.1016/j.cogdev.2020.100851

Menon, V., & D'Esposito, M. (2022). The role of PFC networks in cognitive control and executive function. *Neuropsychopharmacology*, *47*, 90–103. https://doi.org/10.1038/s41386-021-01152-w

Miller, A. N. R. (2015). The role of body mass index and its covariates in emotion recognition. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, *75*(8-B(E)).

Munt, A. E., Partridge, S. R., & Allman-Farinelli, M. (2017). The barriers and enablers of healthy eating among young adults: a missing piece of the obesity puzzle: A scoping review. *Obesity Reviews*, *18*(1), 1–17. https://doi.org/10.1111/obr.12472

Myerson, J., Green, L., & Warusawitharana, M. (2001). Area under curve as a measure of discounting. *Journal of the Experimental Analysis of Behavior*, *76*(2), 235–243. https://doi.org/10.1901/jeab.2001.76-235

Neel, R. (2014). "Expert in the language of fear": Stigmatized targets' perception of others' emotion-specific prejudice. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, *74*(8-B(E)).

Niedenthal, P. M., & Brauer, M. (2011). Social Functionality of Human Emotion. *Annual Review of Psychology*, *63*, 259–285. https://doi.org/10.1146/annurev.psych.121208.131605

O'brien, R. M. (2007). A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Quality & Quantity*, *41*, 673–690. https://doi.org/10.1007/s11135-006-9018-6

Olderbak, S., Wilhelm, O., Hildebrandt, A., & Quoidbach, J. (2019). Sex differences in facial emotion perception ability across the lifespan. *Cognition and Emotion*, *33*(3), 579–588. https://doi.org/10.1080/02699931.2018.1454403

Ouerchefani, R., Ouerchefani, N., ben Rejeb, M. R., & le Gall, D. (2022). Impaired Perception of Unintentional Transgression of Social Norms after Prefrontal Cortex Damage: Relationship to Decision Making, Emotion Recognition, and Executive Functions. *Archives of Clinical Neuropsychology*, *37*(2), 249–273. https://doi.org/10.1093/arclin/acab078

Pearce, E. N. (2012). Thyroid hormone and obesity. *Current Opinion in Endocrinology, Diabetes and Obesity*, *19*(5), 408–413. https://doi.org/10.1097/MED.0b013e328355cd6c

Pratt, L. A., & Brody, D. J. (2014). Depression and obesity in the U.S. adult household population, 2005-2010. *NCHS Data Brief*, *167*, 1–8.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(4), 515–526. https://doi.org/DOI: 10.1017/S0140525X00076512

Puhl, R., & Suh, Y. (2015). Health Consequences of Weight Stigma: Implications for Obesity Prevention and Treatment. *Current Obesity Reports*, *4*, 182–190. https://doi.org/10.1007/s13679-015-0153-z

Quintana, D. S., Guastella, A. J., Outhred, T., Hickie, I. B., & Kemp, A. H. (2012). Heart rate variability is associated with emotion recognition: Direct evidence for a relationship between the autonomic nervous system and social cognition. *International Journal of Psychophysiology*, *86*(2), 168–172. https://doi.org/10.1016/j.ijpsycho.2012.08.012

Rabinovici, G. D., Stephens, M. L., & Possin, K. L. (2015). Executive dysfunction. *Continuum*, *21*(3 Behavioral Neurology and Neuropsychiatry), 646–659. https://doi.org/10.1212/01.CON.0000466658.05156.54

Rapcsak, S. Z. (2019). Face Recognition. *Current Neurology and Neuroscience Reports*, *19*(41). https://doi.org/10.1007/s11910-019-0960-9

Rescorla, L. A., & Achenbach, T. M. (2004). The Achenbach System of Empirically Based Assessment (ASEBA) for Ages 18 to 90 Years. In *The use of psychological testing for treatment planning and outcomes assessment: Instruments for adults, Volume 3, 3rd ed.* (pp. 115–152). Lawrence Erlbaum Associates Publishers.

Rolls, E. T. (2013). *Emotion and Decision Making Explained*. Oxford University Press.

Rubino, F., Puhl, R. M., Cummings, D. E., Eckel, R. H., Ryan, D. H., Mechanick, J. I., Nadglowski, J., Ramos Salas, X., Schauer, P. R., Twenefour, D., Apovian, C. M., Aronne, L. J., Batterham, R. L., Berthoud, H.-R., Boza, C., Busetto, L., Dicker, D., de Groot, M., Eisenberg, D., … Dixon, J. B. (2020). Joint international consensus statement for ending stigma of obesity. *Nature Medicine*, *26*, 485–497. https://doi.org/10.1038/s41591-020-0803-x

Schafer, J. L. (1999). Multiple imputation: a primer. *Statistical Methods in Medical Research*, *8*(1), 3–15. https://doi.org/10.1177/096228029900800102

Tang, J., Chrzanowski-Smith, O. J., Hutchinson, G., Kee, F., & Hunter, R. F. (2019). Relationship between monetary delay discounting and obesity: a systematic review and meta-regression. *International Journal of Obesity*, *43*, 1135–1146. https://doi.org/10.1038/s41366-018-0265-0

Tsuchida, A., & Fellows, L. K. (2012). Are You Upset? Distinct Roles for Orbitofrontal and Lateral Prefrontal Cortex in Detecting and Distinguishing Facial Expressions of Emotion. *Cerebral Cortex*, *22*(12), 2904–2912. https://doi.org/10.1093/cercor/bhr370

Tulsky, D. S., Carlozzi, N., Chiaravalloti, N. D., Beaumont, J. L., Kisala, P. A., Mungas, D., Conway, K., & Gershon, R. (2014). NIH Toolbox Cognition Battery (NIHTB-CB): list sorting test to measure working memory. *Journal of the International Neuropsychological Society : JINS*, *20*(6), 599–610. https://doi.org/10.1017/S135561771400040X van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E. J., Yacoub, E., & Ugurbil, K. (2013). The WU-Minn Human Connectome Project: An overview. *NeuroImage*, *80*, 62–79. https://doi.org/10.1016/j.neuroimage.2013.05.041

Weintraub, S., Dikmen, S. S., Heaton, R. K., Tulsky, D. S., Zelazo, P. D., Bauer, P. J., Carlozzi, N. E., Slotkin, J., Blitz, D., Wallner-Allen, K., Fox, N. A., Beaumont, J. L., Mungas, D., Nowinski, C. J., Richler, J., Deocampo, J. A., Anderson, J. E., Manly, J. J., Borosh, B., … Gershon, R. C. (2013). Cognition assessment using the NIH Toolbox. *Neurology*, *80*(11 Suppl 3), S54–S64. https://doi.org/10.1212/WNL.0b013e3182872ded

Weintraub, S., Dikmen, S. S., Heaton, R. K., Tulsky, D. S., Zelazo, P. D., Slotkin, J., Carlozzi, N. E., Bauer, P. J., Wallner-Allen, K., Fox, N., Havlik, R., Beaumont, J. L., Mungas, D., Manly, J. J., Moy, C., Conway, K., Edwards, E., Nowinski, C. J., & Gershon, R. (2014). The cognition battery of the NIH toolbox for assessment of neurological and behavioral function: validation in an adult sample. *Journal of the International Neuropsychological Society: JINS*, *20*(6), 567–578. https://doi.org/10.1017/S1355617714000320

White, S. J., Coniston, D., Rogers, R., & Frith, U. (2011). Developing the Frith-Happé animations: A quick and objective test of Theory of Mind for adults with autism. *Autism Research*, *4*(2), 149–154. https://doi.org/10.1002/aur.174

Wilcox, R. (2012). Chapter 10 - Robust Regression. In *Introduction to Robust Estimation and Hypothesis Testing* (pp. 471–532). Academic Press. https://doi.org/10.1016/b978-0-12-386983-8.00010-x

Wong, C. (2013). *Behavioral Inhibition BT - Encyclopedia of Behavioral Medicine* (M. D. Gellman & J. R. Turner, Eds.; pp. 191–193). Springer New York. https://doi.org/10.1007/978-1-4419-1005-9_1093

WU-Minn Consortium, H. C. P. (2018). *WU-Minn HCC 1200 Subjects Data Release Reference Manual*. https://www.humanconnectome.org/study/hcp-young-adult/document/1200-subjects-data-release

Xu, H., Cupples, L. A., Stokes, A., & Liu, C.-T. (2018). Association of Obesity With Mortality Over 24 Years of Weight History: Findings From the Framingham Heart Study. *JAMA Network Open*, *1*(7), e184587. https://doi.org/10.1001/jamanetworkopen.2018.4587

Xu, P., Peng, S., Luo, Y., & Gong, G. (2021). Facial expression recognition: A meta-analytic review of theoretical models and neuroimaging evidence. *Neuroscience & Biobehavioral Reviews*, *127*, 820–836. https://doi.org/10.1016/j.neubiorev.2021.05.023

Yang, Y., Shields, G. S., Guo, C., & Liu, Y. (2018). Executive function performance in obesity and overweight individuals: A meta-analysis and review. *Neuroscience and Biobehavioral Reviews*, *84*, 225–244. https://doi.org/10.1016/j.neubiorev.2017.11.020

Yohai, V. J. (1987). High Breakdown-Point and High Efficiency Robust Estimates for Regression. *The Annals of Statistics*, *15*(2), 642–656. https://doi.org/10.1214/aos/1176350366

Zelazo, P. D., Anderson, J. E., Richler, J., Wallner-Allen, K., Beaumont, J. L., Conway, K. P., Gershon, R., & Weintraub, S. (2014). NIH Toolbox Cognition Battery (CB): Validation of Executive Function Measures in Adults. *Journal of the International Neuropsychological Society*, *20*(6), 620–629. https://doi.org/10.1017/S1355617714000472

Zhang, Y., Padmanabhan, A., Gross, J. J., & Menon, V. (2019). Development of Human Emotion Circuits Investigated Using a Big-Data Analytic Approach: Stability, Reliability, and Robustness. *The Journal of Neuroscience*, *39*(36), 7155–7172. https://doi.org/10.1523/JNEUROSCI.0220-19.2019

Zhang, Z., Peng, P., Eickhoff, S. B., Lin, X., Zhang, D., & Wang, Y. (2021). Neural substrates of the executive function construct, age-related changes, and task materials in adolescents and adults: ALE meta-analyses of 408 fMRI studies. *Developmental Science*, *24*(6), e13111. https://doi.org/10.1111/desc.13111

Table 1: Correlations between sociodemographic and clinical variables with body mass index, emotion recognition, and executive functioning in the sample (*N*=824).

	BMI‡	**ER‡ - Overall**	**ER‡ - Neg Emo‡**	**ER‡ - Pos Emo‡**	**DCCS‡ test**	**List Sorting test**	**Flanker Inhib/ Attn‡ test**	**DD‡**
BMI‡ (kg/m^2)	-	-0.076*	-0.049	-0.019	-0.170***	-0.001	-0.096**	-0.061†
Age (years)	0.081*	-0.039	-0.001	0.027	-0.015	-0.074*	0.019	-0.001
Ethnicity	0.012	0.087*	0.089*	-0.0030	-0.007	-0.039	-0.025	-0.023
Gender	-0.126***	0.004	0.075*	0.034	-0.056	-0.165***	-0.061†	-0.101***
Race								
- Black/AA‡	0.180***	-0.053	-0.060†	-0.047	-0.048	-0.195***	-0.091**	-0.276***
- White/Caucasian	-0.038	0.033	0.038	0.013	0.050	0.165	0.050	0.159***
- AAPI‡	-0.223***	0.040	0.036	0.019	0.047	0.028	0.097**	0.102**
- NA/AN‡	0.002	0.040	0.027	0.006	0.011	-0.016	0.015	0.044
- Multiracial	-0.006	-0.061†	-0.034	0.023	-0.063†	-0.031	-0.025	0.109**
- unknown/not reported	0.077*	0.007	-0.005	0.022	-0.068†	-0.038	-0.088*	-0.102***
Education attainment (years)	-0.182***	0.081*	0.093**	0.031	0.216***	0.043	0.117***	0.168***
Household Income	-0.152***	0.075*	0.035	0.067†	0.155***	0.070*	0.156***	0.164
History of MDE‡	-0.019	0.040	0.064†	-0.054	0.117***	-0.031	-0.038	-0.030
ASR‡ DSM‡ Depression T-score	0.089*	-0.020	0.004	-0.075*	-0.050	0.003	-0.045	-0.034
ASR‡ DSM‡ Anxiety T-score	0.052	-0.057†	-0.039	-0.058†	-0.055	-0.049	-0.049	-0.073†
ASR‡ DSM‡ ADHD T-score	0.075*	-0.056	-0.030	-0.104**	-0.012	0.039	-0.053	-0.046
Twin-status								
- Not a twin	0.047	-0.012	0.006	-0.069*	-0.016	0.075*	0.030	0.004
- Monozygotic twin	-0.027	0.008	-0.016	0.047	-0.010	-0.044	-0.002	0.025
- Dizygotic twin	0.007	0.035	0.037	0.021	0.014	-0.012	-0.014	-0.026
- twin, unverified typed	-0.044	-0.035	-0.030	0.017	0.024	-0.041	-0.030	-0.010

Notes. Spearman's correlation was used between two ranked or continuous variables. A point-biserial correlation was conducted between a dichotomous and a continuous variable.

‡***Abbreviations***. AA=African American; AAPI=Asian/Asian Americans and Pacific Islanders (including Native Hawaiians); ASR=Adult Self Report; BMI=Body mass index; DCCS=Dimensional Change Card Sort; DD=Delayed Discounting (averaged score for $200 and $40K immediate awards); DSM=Diagnostic and Statistical Manual of Mental Disorders; ER=Emotion recognition; Inhib/Attn=Inhibitory Control and Attention; MDE=Major depressive episode; NA/AN=Native American/Alaskan Natives; Neg Emo=Negative emotion; Pos Emo=Positive emotion

Significance. † $p < 0.1$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 2. Sociodemographic and clinical characteristics of the dataset (*N*=799).

	Mean	SD	Median	Min	Max
BMI (kg/m^2) ‡	27.05	5.88	25.82	16.48	48.25
Age (years)	28.90	3.71	29	22.00	36
Education attainment (years)	15.04	1.74	16	11.00	17
ASR‡ DSM‡ Depression T-score	53.58	5.40	51	50.00	81
ASR‡ DSM‡ Anxiety T-score	52.98	4.97	51	50.00	80
ASR‡ DSM‡ ADHD T-score	54.30	5.26	52	50.00	80
Delayed Discounting - AUC†	0.39	0.23	0.36	0.02	0.98
DCCS‡ test	101.93	9.86	102.23	66.55	122.65
List Sorting	103.50	13.23	102.63	60.09	140.86
Flanker Inhib/Attn‡ test	101.56	10.08	101.72	69.31	123.56
ER‡ accuracy	35.61	2.49	36	24	40
ER‡ anger accuracy	6.75	1.04	7	2	8
ER‡ fear accuracy	6.92	1.16	7	0	8
ER‡ sadness accuracy	6.82	1.11	7	1	8
ER‡ happiness accuracy	7.96	0.22	8	5	8
ER‡ neutral accuracy	7.15	1.19	8	1	8
Household income bracket	NA	NA	$50,000-$74,999	<$10,000	≥100,000
	N	(%)			
Gender - Female	489	61.20%			
Race					
- Black/African American	118	14.77%			
- Caucasian	597	74.72%			
- AAPI‡	56	7.01%			
- NA/AN‡	1	0.13%			
- Multi-racial	14	1.75%			
- Race not reported/unknown	13	1.63%			
Twin status					
- Not a twin	357	44.68%			
- Monozygotic twin	217	27.16%			
- Dizygotic twin	134	16.77%			
- Twin, unverified type	91	11.39%			
Ethnicity (Hispanic/Latinx) §	62	7.76%			
School/work status (in school or working full-time)	643	80.48%			
History of MDEs‡ (Yes)	65	8.14%			

‡***Abbreviations***. AAPI=Asian/Asian Americans and Pacific Islanders (including Native Hawaiians); ASR=Adult Self Report; AUC=area under the curve (summary sore, for averaged responses for delayed rewards of $200 & $40K); BMI=Body mass index; DCCS= Dimensional Change Card Sort; DSM=Diagnostic and Statistical Manual of Mental Disorders; ER=Emotion recognition; Inhib/Attn=Inhibitory Control and Attention; MDE= Major depressive episode; NA/AN=Native Americans/Alaskan Natives
§***Hispanic Ethnicity***. 62 participants identified as Hispanic, 727 as non-Hispanic, and 10 were not reported out of data subset.

Table 3. Hierarchal linear regressions (Ordinary Least Squares) for overall emotion recognition (*N*= 799).

Category	Variable	Covariates only		Covariates + EF		Final: Covariates, EF, & BMI	
		β	**SE**	**β**	**SE**	**β**	**SE**
Sociodemographic & Clinical Variables (Covariates)	Age (years)	-0.038	0.03	-0.038	0.03	-0.032	0.03
	Gender (female)	0.045	0.19	0.136	0.19	0.097	0.19
	Education attainment (years)	0.086	0.05	0.068	0.05	0.058	0.05
	Income	0.087†	0.05	0.059	0.05	0.052	0.05
	Race						
	- Black/African American	-0.138	0.26	-0.057	0.27	-0.002	0.27
	- White/Caucasian	0.497	0.36	0.536	0.36	0.438	0.36
	- AAPI‡	2.126	2.49	2.361	2.46	2.378	2.45
	- NA/AN‡	-0.889	0.68	-0.506	0.67	-0.523	0.67
	- Multiracial	-0.565	0.76	-0.492	0.76	-0.421	0.76
	Ethnicity (Hispanic/Latinx)	1.022**	0.37	1.054**	0.37	1.038**	0.36
	ASR‡ DSM‡ ADHD T-score	-0.031	0.02	-0.023	0.02	-0.024	0.02
	ASR‡ DSM‡ Depression T-score	0.002	0.02	-0.001	0.02	<0.001	0.02
	History of MDEs‡ (Yes)	0.273	0.34	0.384	0.34	0.389	0.34
	Twin status (reference = not a twin)						
	- Monozygotic twin-status	0.277	0.22	0.369†	0.22	0.343	0.22
	- Dizygotic twin-status	0.287	0.26	0.335	0.26	0.325	0.26
	- Unverified twin-status	-0.125	0.30	-0.045	0.29	-0.082	0.29
Executive functioning (EF)	DCCS‡ Test - Cognitive flexibility			0.027*	0.01	0.025*	0.01
	List Sorting - Working memory			0.023***	0.01	0.023***	0.01
	Flanker–test - Inhibition & attention			0.002	0.01	0.003	0.01
	Delayed Discounting - Inhibition			-0.537	0.40	-0.614	0.41
BMI‡	BMI (kg/m2) ‡					. -0.027†	0.02
	R^2	0.039		0.070		0.074	
	Adjusted R^2	0.020		0.047		0.049	
	F	1.99*		2.95***		2.96***	

‡***Abbreviations***. AAPI=Asian/Asian Americans and Pacific Islanders (including Native Hawaiians); ASR=Adult Self Report scale; BMI=Body mass index; DSM=Diagnostic and Statistical Manual of Mental Disorders; DCCS=Dimensional Change Card Sort; MDE=Major depressive episode; NA/AN=Native Americans/Alaskan Natives
Significance. † $p < 0.1$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 4. Hierarchal linear regressions (Ordinary Least Squares) for negative emotion recognition (*N*=799).

Category	Variable	Covariates only		Covariates + EF		Final: Covariates, EF, & BMI	
		β	**SE**	**β**	**SE**	**β**	**SE**
Sociodemographic & Clinical Variables (Covariates)	Age (years)	-0.006	0.023	-0.005	0.023	-0.002	0.02
	Gender (female)	0.309†	0.166	0.394*	0.167	0.375*	0.17
	Education attainment (years)	0.109*	0.048	0.087†	0.048	0.083†	0.05
	Income	0.010	0.041	-0.010	0.041	-0.014	0.04
	Race						
	- Black/African American	-0.241	0.23	-0.134	0.24	-0.107	0.24
	- White/Caucasian	0.386	0.31	0.379	0.31	0.332	0.32
	- AAPI‡	1.490	2.18	1.548	2.16	1.555	2.16
	- NA/AN‡	-0.507	0.59	-0.296	0.59	-0.304	0.59
	- Multiracial	-0.769	0.67	-0.626	0.67	-0.592	0.67
	Ethnicity (Hispanic/Latinx)	0.875**	0.324	0.894**	0.322	0.886**	0.32
	ASR‡ DSM‡ ADHD T-score	-0.015	0.018	-0.009	0.018	-0.009	0.02
	ASR‡ DSM‡ Depression T-score	-0.001	0.017	-0.004	0.017	-0.004	0.02
	History of MDEs‡ (Yes)	0.308	0.296	0.361	0.296	0.364	0.30
	Twin status (reference = not a twin)						
	- Monozygotic twin-status	0.013	0.195	0.070	0.195	0.057	0.20
	- Dizygotic twin-status	0.157	0.228	0.193	0.227	0.188	0.23
	- Unverified twin-status	-0.149	0.260	-0.092	0.259	-0.110	0.26
Executive functioning (EF)	DCCS‡ Test - Cognitive flexibility			0.019*	0.009	0.018*	0.01
	List Sorting - Working memory			0.015*	0.006	0.015*	0.01
	Flanker test - Inhibition & attention			0.007	0.009	0.008	0.01
	Delayed Discounting - Inhibition			-0.023	0.356	-0.060	0.36
BMI	BMI (kg/m2) ‡					-0.013	0.01
	R^2	**0.036**		**0.057**		**0.058**	
	Adjusted R^2	**0.016**		**0.032**		**0.032**	
	F	**1.81***		**2.34*****		**2.26****	

‡***Abbreviation****s*. AAPI=Asian/Asian Americans and Pacific Islanders (including Native Hawaiians); ASR DSM=Adult Self Report scale; BMI=BMI=Body mass index; DSM=Diagnostic and Statistical Manual of Mental Disorders; DCCS=Dimensional Change Card Sort; NA/AN=Native Americans/Alaskan Natives; MDE=Major depressive episode
Significance. † $p < 0.1$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

APPENDIX A

OPEN SCIENCE FRAMEWORK PREREGISTRATION

Study Information

Hypotheses. Aim 1: Examine the relationship between BMI and emotion recognition accuracy. Hypothesis 1: Higher BMI will be associated with emotion recognition inaccuracy, after controlling for sociodemographic and clinical covariates, and executive functioning. Aim 2: Examine the role of emotion valance on the relationship between BMI and emotion recognition accuracy Hypothesis 2: The relationship between BMI and emotion recognition accuracy will be moderated by emotion valence such that higher BMI will be negatively associated with emotion recognition inaccuracy for negative emotions (i.e. anger, sadness, and fear) but not positive emotions.

Design Plan

Study type. Observational Study - Data is collected from study subjects that are not randomly assigned to a treatment. This includes surveys, “natural experiments,” and regression discontinuity designs.

Blinding

No blinding is involved in this study.

Is there any additional blinding in this study? No blinding

Study design

The proposed study aims to utilize publicly available data provided by the Human Connectome Project, Young Adult sample, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems

Neuroscience at Washington University (Van Essen et al., 2013). We are utilizing the “1200 Subjects Data Release” behavioral dataset which was released March 2017. Data was collected at one period time during a two-day visit process, after an initial phone screening and diagnostic interview. During Day 1, participants completed some questionnaires, drug screening, blood tests, structural and functional MRI sessions, and NIH toolbox behavioral tests (including cognitive functioning measures). On Day 2, participants completed diffusion and functional MRI sessions and non-NIH behavioral tasks (including the Penn Emotion Recognition Test). See Van Essen et al (2013) and the “WU-Minn HCC 1200 Subjects Data Release Reference Manual” (WU-Minn Consortium, 2018) for more details.

Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E. J., Yacoub, E., & Ugurbil, K. (2013). The WU-Minn Human Connectome Project: An overview. NeuroImage, 80, 62–79. https://doi.org/10.1016/j.neuroimage.2013.05.041

WU-Minn Consortium, H. C. P. (2018). WU-Minn HCC 1200 Subjects Data Release Reference Manual. Retrieved from https://www.humanconnectome.org/study/hcp-young-adult/document/1200-subjects-data-release

Randomization - Not applicable

Sampling Plan

Existing Data. Registration prior to analysis of the data

Explanation of existing data. Hypothesis were made a priori before looking at the dataset based on a meta-analysis and literature review the first author conducted. Descriptive statistics were only made to summarize the data after hypotheses have been made.

Data collection procedures. Inclusion/exclusion criteria of dataset For the Humane Connectome Project - Participants needed to be aged 22 to 35 years at the time of the phone screen. Additionally, this dataset included participants with twins and non-twin siblings. Participants were excluded for having a significant psychiatric, substance abuse, neurological, or cardiovascular disease as indicated by reported a clinical diagnosis by a physician, being hospitalized for any condition for two days or longer, or receiving pharmacologic or behavioral treatment for 12 months or longer (other than for childhood ADHD). Other exclusions included epilepsy, genetic disorders, neurological conditions or events (e.g. stroke, cerebral palsy, or a brain tumor), significant head injuries, premature birth, history of chemotherapy or radiation, recent thyroid hormone treatment, diabetes or high blood pressure, daily migraine medication in past month, low score on the Mini Mental State Exam (MMSE), and contraindications to MRI procedures (e.g. pregnancy, unsafe metals in body, and moderate to severe claustrophobia). See Van Essen et al. (2013) for full details of the inclusion and exclusion criteria. Additional exclusions for proposal Additionally, women must have reported having (1) no history of thyroid disease (e.g., hyperthyroid or hypothyroid) as these conditions may affect weight status (Pearce, 2012). As thyroid disease is much more common in females, this question was only asked to female participants. We also excluded for having a (2) blood alcohol concentration (BAC) of .05% or higher or (3) testing positive for screened drugs during study visits, as both alcohol and drug use can influence task performance (Castellano et al., 2015). Additionally, participants who met the Diagnostic and statistical manual of mental disorders, 4th ed (DSM-IV; American Psychiatric Association, 2000) criteria for alcohol abuse or dependence disorder were excluded, as substance abuse is negatively associated with worse facial emotion recognition skills (Castellano et al., 2015). This left us with a sample of 832 participants that met our inclusion criteria. Besides these criteria, we excluded participants that did not complete the emotion recognition or the executive functioning tasks, as these were the primary variables of interest in this proposal.

Sample size. 824 participants from the Human Connectome Project dataset met our inclusion/exclusion criteria. We plan to exclude missing data from the analysis.

Sample size rationale. In our sample, the following sociodemographic variables correlated with one or more of our predictor variables and are potential "control" variables: (1) age, (2) gender, (3) educational attainment, (4) income, (5) race, (6) ethnicity, (7) in-school/work status, (8) ASR ADHD scale score, (9) ASR Depression scale score, and (10) history of major depressive episodes. Using G*Power 3.1.9.4, the following parameters in a sample size collection for a "linear multiple regression: Fixed model, R2 increase" statistical test type: (1) f2 = 0.02, which assumes a small effect size, (2) α = .05, (3) power = 0.80, (4) five predictor variables, and (5) 15 total predictors. A total sample size of 648 participants gives 80% power, if all predictors are added. Since our sample exceeds this amount, we have sufficient power to determine a small effect size in our primary analyses.

Stopping rule. N/A

Variables

Manipulated variables. N/A - this is not experimental.

Measured variables. Emotion Recognition = The Penn Emotion Recognition Test (outcome) Main predictors of interest (IVs) - BMI - Executive functioning measures: - (1) Delayed Discounting test - (2) Dimensional Change Card Sort (measuring cognitive flexibility NIH Toolbox) - (3) Flanker Inhibitory Control and Attention Test (measuring inhibition and attention from NIH Toolbox) - (4) List Sorting Working Memory Test (measuring working memory). Sociodemographic & clinical variables (control variables) - (1) Gender - (2) age - (3) race - (4) ethnicity - (5) educational level - (6) household income - (7) Achenbach Adult Self Report (ASR) for Depression - (8) ASR for Anxiety - (9) ASR for ADHD - (10) DSM-IV history of past major depressive episodes See the Humane Connectome Project manual and citations for full details on these measures:

van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E. J., Yacoub, E., & Ugurbil, K. (2013). The WU-Minn Human Connectome Project: An overview. *NeuroImage*, *80*, 62–79. https://doi.org/10.1016/j.neuroimage.2013.05.041

WU-Minn Consortium, H. C. P. (2018). WU-Minn HCC 1200 Subjects Data Release Reference Manual. Retrieved from https://www.humanconnectome.org/study/hcp-young-adult/document/1200-subjects-data-release

Indices. If two subtasks our high correlated (such as the Delay Discounting tasks), average scores will be calculated. Additionally, if negative emotion sub-scores (for fear, anger, and sadness) on the Penn Emotion Recognition Test are correlated with each other, a composite score will be calculated by adding these sub-scores

Analysis Plan

Statistical models - t**esting assumptions of a hierarchal linear regression**. Linearity between each of the main predictors (BMI and the executive functioning variables) with the outcome variable (emotion recognition accuracy) will be assessed using a Spearman's correlations. Predictors that do not have a significant linear relationship with the outcome variable will be dropped from the model. Normal distribution of the residuals will be assessed through visual inspection (e.g., Q-Q plot and histogram), although the Central Limit Theorem may protect again some violations in this large dataset. Outliers of residuals will be assessed using the dfbeta values to check if there are observations significantly influencing the regression models. If there are observations above a sample size-adjusted cut-off score representing observations that shift the regression coefficient estimate considerably, we will perform a sensitivity analysis to compare the results from our original analysis to the results of a Robust Linear Regression. If results are similar between the RLM and the HLR, we will report the results of the HLR model. If results differ, we will report the results of the RLM. Multicollinearity between predictors will be assessed with the Variance Inflation Factor (VIF) & tolerance scores, with cut-off scores of above 5 and under 0.20, respectively. Heteroscedasticity will be assessed using the Breusch-Pagan Test. Finally, the Durbin-Watson test will be used to assess the independence of the residuals. Testing the Aim 1. We will conduct a hierarchical regression with emotion recognition accuracy as the dependent variable, entering sociodemographic variables and clinical variables in the first step, each executive functioning variables in the second step, and BMI in the third step. We chose a hierarchical regression to allow for an assessment of the amount of variance contributed by the predictors entered at each step (i.e., change in R^2). Additionally, by comparing model steps, we will be able to determine when variables entered in each step significantly explained variance (R^2). Testing the Aim 2. Finally, we predict that BMI will impact emotion recognition of negative emotions more than positive emotions. Therefore, we plan on testing differences between emotion recognition accuracy scores between emotions measured. If there are no significant differences between emotion type, we will not conduct separate analyses. If significant differences exist, and all negative emotion variables are correlated with each other, we will combine negative emotions into a composite measure and conduct separate analyses for emotion recognition accuracy for positive ("Happy") and negative emotions ("Sad," Anger," and "Fear").

Transformation. Dichotomous variables will be dummy coded with beta coefficients being reported for women for gender (reference group: men), identifying as Hispanic for ethnicity (reference group: non-Hispanic), and "Yes" to being a student or working full-time (reference group: "No"). If we have a categorical group that is not dichotomous, we will transform this data.

Inference criteria. We will use the standard $p<0.05$ criteria of all statistical tests determining significance, including for regression models, beta-coefficient of predictor, and correlation values

Data exclusion. Extreme values or outliers affecting the results and/or are non-representative of the population will be assessed using descriptive statistics. Outliers of residuals will be assessed using the dfbeta values to check if there are observations significantly influencing the regression models. If there are observations above a sample size-adjusted cut-off score representing observations that shift the regression coefficient estimate considerably, we will perform a sensitivity analysis to compare the results from our original analysis to the results of a Robust Linear Regression. If results are similar between the RLM and the HLR, we will report the results of the HLR model. If results differ, we will report the results of the RLM.

Missing data. If less than 5% participants have missing data for any variables, we will use a completer-only analysis. Participants with missing data for the variables used in the regression analysis will not be included.

Exploratory analysis - We do not have exploratory analyses planned.

Other [optional] - *Other* - *[No response]*

APPENDIX B

Supplemental Table 1. Hierarchal linear regression (Robust Linear Model) – final model for overall and negative emotion recognition (*N*=799).

		Overall ER - Final Model		Negative ER - Final Model	
Predictor category	**Variable name**	**β**	**SE**	**β**	**SE**
Socio-demographic & Clinical Variables (Covariates)	Age (in years)	-0.018	0.03	0.001	0.02
	Gender (female)	-0.102	0.18	0.298†	0.17
	Education attainment (years completed)	0.059	0.05	0.084†	0.05
	Income	0.013	0.05	-0.026	0.04
	Race (reference = not reported/unknown)				
	- Black/African American	-0.087	0.26	-0.189	0.23
	- White/Caucasian	0.260	0.34	0.189	0.25
	- AAPI‡	1.959	2.32	1.330***	0.29
	- Native American/Alaskan Native	-0.831	0.64	-0.477	0.57
	- Multi-racial	-0.205	0.73	-0.691	0.75
	Ethnicity (Hispanic/Latinx)	1.029**	0.35	0.877**	0.27
	ASR‡ DSM‡ ADHD T-score	-0.033†	0.02	-0.016	0.02
	ASR‡ DSM‡ Depression T-score	-0.008	0.02	-0.008	0.02
	History of MDEs‡ (Yes)	0.479	0.31	0.578*	0.24
	Twin status (reference = not a twin)				
	- Monozygotic twin-status	0.149	0.21	-0.033	0.18
	- Dizygotic twin-status	0.390	0.25	0.241	0.22
	- Unverified twin-status	0.055	0.29	-0.116	0.27
Executive Functioning	DCCS‡ Test - Cognitive flexibility	0.025*	0.01	0.019*	0.01
	List Sorting test - Working memory	0.014*	0.01	0.011†	0.01
	Flanker test - Inhibition & attention test	-0.005	0.01	0.002	0.01
	Delayed Discounting - Inhibition	-0.754†	0.39	-0.283	0.36
BMI	BMI (kg/m2) ‡	-0.022	0.02	-0.015	0.01
	R^2	**0.060**		**0.060**	

‡***Abbreviations***. AAPI=Asian/Asian Americans and Pacific Islanders (including Native Hawaiians); ASR=Adult Self Report scale; BMI=Body mass index (kg/m2); DSM=Diagnostic and Statistical Manual of Mental Disorders; DCCS=Dimensional Change Card Sort; ER=Emotion recognition (accuracy); MDE=Major depressive episode
Significance. † $p < 0.1$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

APPENDIX C

Supplemental Table 2. Correlation between emotion recognition variables (*N*=799).

	1	2	3	4	5
1. ER‡ - Overall	-				
2. ER‡ - Sad	0.549***	-			
3. ER‡ - Anger	0.548***	0.167***	-		
4. ER‡ - Fear	0.553***	0.105**	0.133***	-	
5. ER‡ - Happy	0.171***	0.083*	0.056	0.029	-
6. ER‡ - Neutral	0.403***	-0.077*	-0.013	0.011	0.024

‡***Abbreviations***. ER=Emotion recognition (accuracy)

Significance. † $p < 0.1$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

APPENDIX D

Supplemental Table 3. Hierarchal linear regression (Ordinary Least Squares) for positive emotion recognition (*N*=799).

		Final: Covariates, EF, & BMI	
Predictor category	**Variable name**	**β**	**SE**
Sociodemographic & Clinical Variables (Covariates)	Age (in years)	-0.001	0.00
	Gender (female)	0.009	0.02
	Education attainment (years completed)	0.002	0.00
	Income	0.003	0.00
	Race (reference = not reported/unknown)		
	- Black/African American	-0.018	0.02
	- White/Caucasian	-0.005	0.03
	- AAPI‡	0.038	0.22
	- NA/AN‡	0.059	0.06
	- Multi-racial	0.049	0.07
	Ethnicity (Hispanic/Latinx)	0.010	0.03
	ASR‡ DSM‡ ADHD T-score	-0.001	0.00
	ASR‡ DSM‡ Depression T-score	-0.001	0.00
	History of MDEs‡ (Yes)	-0.009	0.03
	Twin status (reference = not a twin)		
	- Monozygotic twin-status	0.036†	0.02
	- Dizygotic twin-status	0.032	0.02
	- Unverified twin-status	0.034	0.03
Executive Functioning (EF)	DCCS‡ Test - Cognitive flexibility	-0.001	0.00
	List Sorting test - Working memory	0.001*	0.00
	Flanker test - Inhibition & attention test	0.000	0.00
	Delayed Discounting - Inhibition	-0.030	0.04
BMI	BMI (kg/m2) ‡	0.001	0.00
	R^2	**0.023**	
	Adjusted R^2	**-0.003**	
	F	**0.88**	

‡**Abbreviation**s. AAPI=Asian/Asian Americans and Pacific Islanders (including Native Hawaiians); ASR DSM=Adult Self Report scale; BMI=Body mass index (kg/m2); DSM=Diagnostic and Statistical Manual of Mental Disorders; DCCS=Dimensional Change Card Sort; ER=Emotion recognition (accuracy); MDE=Major depressive episode; NA/AN=Native Americans/Alaskan Natives
Significance. † $p < 0.1$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.00$