Weight Salience and Reasoning Performance among African Americans: Insights from A Randomized Field Experiment

Hieu Nguyen
Department of Economics & Business, Southwestern University

Abstract

I test whether experimentally induced weight salience lowers cognitive reasoning among Black/African American adults (n = 500), a population facing dual weight- and race-based stereotypes yet understudied in salience-performance research. Participants complete either a weight-reflection writing task (treatment) or neutral reading task (control) before the 4-item Cognitive Reflection Test (CRT). The manipulation successfully raises salience (d = 1.0, p < 0.01), but elevated salience does not significantly reduce CRT performance (d = -0.08, p > 0.10). These findings establish important boundary conditions for stereotype-threat effects, suggesting that acute salience alone may be insufficient to disrupt reasoning in low-stakes, non-evaluative contexts.

Keywords: obesity stigma; weight salience; cognition; health disparities; field experiment
JEL Codes: I14; I24; D91; J15
Abbreviations: BMI: body mass index; CRT: Cognitive Reflection Test; OLS: ordinary least squares; 2SLS: two-stage least squares.
Conflict of interest: None.
Correspondence: Hieu Nguyen, Economics & Business Department, Southwestern University, 1001 E University Ave, Georgetown, TX 78626. Email: nguyen23@southwestern.edu.
Disclosure statement: This research involved anonymous data collection on human subjects, and received approval from the Southwestern University Institutional Review Board as "IRB-Exempt" (reference #: FA25_21).
Funding statement: Partial financial support from the Sam Taylor Fellowship (United Methodist General Board of Higher Education and Ministry) and Southwestern University was gratefully received. I declare no other funding sources and/or competing interests.

1. Introduction

Obesity rates in the United States remain high, affecting 41.9% of adults overall, but are disproportionately concentrated among the non-Hispanic African American population, where the prevalence of obesity has reached approximately 50% (CDC, 2024). As a stigmatized trait, the body weight shapes one's self-image, social identity, and economic opportunity. Tomiyama et al.'s (2018) comprehensive review finds that weight *stigma*, independent of weight itself, drives adverse health outcomes through both psychological and physiological pathways. This finding is in line with extensive reviews in public health and social psychology documenting that devaluation and negative stereotyping affect overweight individuals' subjective well-being (Puhl & Heuer, 2009). Evidence also shows that stress-induced stigma can worsen self-efficacy and lead to complete avoidance of overall health services (Puhl & Heuer, 2010). These consequences are disproportionately experienced by individuals at intersections of stigmatized group identities, including racial minorities with higher average weight status such as Black/African American adults (Petersen et al., 2019; CDC, 2024).

Empirical evidence suggests that weight-based discrimination undermines perceived competence and cognitive functioning. Adults who report weight discrimination exhibit poorer numeric reasoning scores (Sutin et al., 2020), and children confronted with weight-related stereotypes display impaired working memory (Guardabassi & Tomasetto, 2020). In the workplace, people with higher weight may experience discrimination in employment, education, and healthcare (Andreyeva et al., 2008; Puhl & Heuer, 2009), resulting in lasting reductions in wages and productivity (Cawley, 2004). Recently, economists have started observing persistent labor-market disadvantages associated with obesity, including lower wages and employment

probabilities, even conditional on education and job characteristics (Andreyeva et al., 2008). One potential (structural) explanation for these equilibrium results is that weight stigma generates biased perceptions of productivity and reduces performance on cognitive assessments that are used in skill evaluation. Major et al. (2012) show that making weight visible during an evaluative task (a speech about dating potential) significantly increased blood pressure and impaired Stroop-task performance among higher-BMI women. Hunger et al. (2015) further develop a weight-based social identity threat model, arguing that anticipated stigma activates chronic-stress pathways that undermine both mental and physical health.

A growing literature in behavioral economics demonstrates that cognitive performance, along with the productive labor it underpins, is highly sensitive to psychological stressors and identity-based threat (Akerlof & Kranton, 2000; Schilbach et al.., 2016). Indeed, experimental economics research has shown that *priming* ethnic identity shifts economic preferences among Asian American participants (Benjamin et al., 2010). Likewise, Hoff and Pandey (2006) find a significant reduction in cognitive performance among low-caste Indian schoolchildren exposed to caste-identity salience. When individuals are reminded of negatively stereotyped identities, mental resources may be diverted toward monitoring the threat rather than toward cognitive control (Steele & Aronson, 1995; Spencer et al., 1999). This diversion may be especially consequential for critical-thinking tasks where reasoning performance requires suppressing certain intuitive impulses. At the other end of the spectrum, cognitive flexibility and reasoning ability are often viewed as key drivers of improvements in productivity and decision-making. If identity salience causally impacts cognitive effort choices, weight-related cues, even if unintentional, could impair reasoning exactly in scenarios where accuracy matters most, such as

assessments or selection processes. This raises an important equity concern: Should institutions redesign tasks to reduce salience of stigmatized traits?

I seek to address this question with a novel, light-touch, and scalable field experiment conducted on a historically marginalized group to explore the nexus between acute weight salience and critical thinking performance. Specifically, I experimentally test whether momentary salience of body weight causally reduces cognitive reflection among Black/African American adults. This population faces a dual burden of weight- and race-based stereotypes (Himmelstein et al., 2017), yet remains understudied in salience-performance research. The results have direct implications for public policies directed at ensuring equitable screening in education, hiring, and other contexts.

2. Literature Review

While Akerlof & Kranton (2000) are often touted as a key reference for economists' understanding of identity, a compelling theoretical basis for studying weight salience and performance dates back to psychological models of identity-relevant threat. Stereotype threat theory explains how awareness of negative stereotypes imposes psychological costs and reduces performance on cognitively demanding tasks. Classic experiments in social psychology find that participants from stigmatized groups perform worse when a diagnostic stereotype is invoked, e.g., African Americans in academic testing and women in math-intensive tasks (Steele & Aronson, 1995; Spencer et al., 1999). Since individuals typically function in social settings and possess multiple group identities, the direction and strength of identity-salience effects may be determined by which specific identity is primed. For instance, Shih et al. (1999) show that Asian-American women performed better on math when their Asian identity was salient but worse

when their female identity was salient. The integrated process model of Schmader et al. (2008) explains these stereotype-threat effects through three mechanisms: physiological stress responses, active monitoring of performance, and efforts to suppress negative thoughts, all of which consume one's brain resources. Consistent with this framework, Inzlicht et al. (2006) show that being the target of prejudice depletes self-regulatory resources, impairing performance on subsequent tasks requiring executive control. Beilock and Carr (2005) provide complementary evidence that pressure-induced worry reduces working memory capacity and impairs math performance among high-ability individuals who would otherwise perform well.

Recent research has begun to apply stereotype-threat theory specifically to weight-based identity. Some studies show strong associations between weight-based stereotype threat and acute impairments in numerical reasoning and attentiveness (Guardabassi & Tomasetto, 2018, 2020). Puhl et al. (2018) estimate that 40% of U.S. adults internalize weight stigma, with *higher* rates among underrepresented minorities, underscoring the prevalence and relevance of weight-based identity threat in racially diverse populations. Although they do not isolate weight-specific mechanisms, Barnes et al. (2012) find that perceived discrimination is associated with worse cognitive performance among *older* African Americans. These mechanisms remain understudied among African Americans, who already face systemic barriers to human-capital accumulation.

Despite substantial evidence that weight stigma is harmful and that salience-based threat undermines cognitive control, the causal role of momentary weight salience in reasoning performance among Black adults remains largely unexplored. A major methodological shortcoming of previous correlational studies is that they cannot disentangle whether weight discrimination causes cognitive deficits, or whether unobserved heterogeneity drives both. Even

among causality-focused studies, very few pair experimental manipulations with a structural econometric model that interprets the behavioral mechanism, which is often of great interest to policymakers, in a straightforward manner. This study addresses these gaps in the literature by randomizing weight salience in a moderately sized sample of African American adults and estimating its effect on cognitive control within a transparent IV framework. Understanding this pathway can help improve assessment policies relying on measured reasoning ability, or at least illuminate how seemingly innocuous design elements may disproportionately disadvantage individuals living with weight-based stereotypes.

For the main experimental task, I make use of psychologists' Cognitive Reflection Test, a standard measure of reflective reasoning (Frederick, 2005). CRT performance has been shown to predict a wide range of heuristics-and-biases tasks (Toplak, West, & Stanovich, 2011), and to be correlated with economic decision-making, patience, and accuracy in evaluative settings (Toplak, West, & Stanovich, 2014). Importantly, because the CRT isolates controlled reasoning from general knowledge and demands impulse control rather than extensive computation, it is particularly sensitive to the so-called "mental bandwidth" taxes. Thomson and Oppenheimer (2016) validate alternative CRT items, showing that the construct can be reliably measured with different item sets, which motivates my use of a four-item version.

3. Conceptual Framework and Testable Hypotheses

To fix ideas, I develop a simple micro-founded static model of cognitive reasoning where CRT performance results from an optimal deliberative effort choice under momentary psychological burdens. One can think of the proposed mechanism as akin to the notion of "mental bandwidth"

or cognitive taxation in Schilbach et al. (2016). Specifically, let $e_i \geq 0$ denote individual *i*'s exerted effort when responding to analytical questions such as those on the Cognitive Reflection Test. Exerting effort generates benefits proportional to improved performance (i.e., higher accuracy), but it also imposes both *psychological* and *attentional* costs. These costs are assumed to rise with the degree to which the individual's body weight is salient in mind at the moment of decision. This assumption is grounded in theories of stereotype threat and identity-based stress, which from the above literature review suggest that negative identity cues draw cognitive resources away from executive functioning. The utility of exerting effort consists of a linear performance benefit and a convex cost term that increases in weight-related salience. Formally, the individual selects e_i to solve:

$$\max_{e_i \geq 0} \left\{ a e_i - \frac{b}{2} e_i^2 - (\lambda_0 + \lambda_1 m_i) e_i \right\} \qquad (1),$$

where $a > 0$ represents the marginal value of improved task performance (i.e., correct reasoning), $b > 0$ ensures convexity of cognitive cost, and $\lambda_0 > 0$ captures baseline cognitive cost (burden). Importantly, $\lambda_1 > 0$ measures the additional marginal cost imposed by momentary weight salience m_i, reflecting the degree to which the agent is consciously attending to their body weight at that moment. When weight is psychologically more intrusive, effort becomes more costly, resulting in lower optimal performance-enhancing effort.

Solving the first-order condition yields the unique optimal effort expenditure function of:

$$e_i^*(m_i) = \frac{a - \lambda_0 - \lambda_1 m_i}{b} \qquad (2),$$

which implies a negative equilibrium effect of salience,

$$\frac{\partial e_i^*}{\partial m_i} = -\frac{\lambda_1}{b} < 0 \qquad (3).$$

Thus, the theory predicts that heightened attention to one's own weight diminishes optimal cognitive effort expenditure. Because the Cognitive Reflection Test requires suppression of intuitive but incorrect answers, diminished effort translates into poorer measured performance.

Observed performance on the CRT can therefore be modeled as a function of (latent) optimal effort plus some random noise as follows:

$$y_i = e_i^*(m_i) + \varepsilon_i = \beta_0 + \beta_m m_i + \varepsilon_i \qquad (4),$$

where y_i denotes the total number of correct CRT responses, ε_i captures unpredictability in performance, and the empirical parameters are defined as $\beta_0 = \frac{a-\lambda_0}{b}$ and $\beta_m = -\frac{\lambda_1}{b} < 0$. The negative sign of β_m follows directly from $\lambda_1 > 0$ and $b > 0$ (greater weight salience should, in theory, decrease CRT scores). In econometric terms, $\beta_m < 0$ is the key testable inequality restriction, quantifying the salience-induced increase in marginal cognitive cost.

However, salience is inherently endogenous due to unobserved anxiety. One can imagine that individuals who are chronically dissatisfied with their weight (or more anxious about stereotypes in general) may exhibit higher m_i even in the absence of the prime, and that those same latent traits may also depress cognitive performance directly. A naïve regression of y_i on m_i would therefore bias the OLS estimate of β_m. To overcome this endogeneity, I employ a randomized intervention that exogenously elevates weight salience. Let Z_i equal one for participants assigned

to the salience writing task (treatment) and zero for those assigned to a neutral, unframed reading activity (control). Salience production can be modeled using the first-stage relationship:

$$m_i = \phi_0 + \theta Z_i + \phi' X_i + \eta_i \qquad (5),$$

where we expect $\theta > 0$ to represent the causal effect of the prime on salience and X_i denotes the baseline controls (including demographic and BMI information). Substituting this expression back into the performance equation yields a reduced-form relationship:

$$y_i = \gamma_0 + \tau Z_i + \gamma' X_i + \nu_i \qquad (6),$$

where $\tau < 0$ captures the intent-to-treat effect. Combining the first stage and reduced form implies that the structural effect of salience on performance is identified through 2SLS as:

$$y_i = \beta_0 + \beta_m \widehat{m_i} + \beta' X_i + \varepsilon_i \qquad (7),$$

with $\beta_m = \tau/\theta < 0$. This quantity is both economically meaningful and policy-relevant. It offers a transparent prediction: the increase in marginal cognitive cost per unit of salience (scaled by $1/b$) should translate into lower CRT performance. The theory and empirical evidence give rise to the following testable hypotheses.

H1 (First Stage). Assignment to the weight-reflection writing task increases momentary salience of body weight relative to the neutral reading control. Econometrically, we expect $\theta > 0$.

Inspired by previous research showing that identity-focused tasks heighten self-consciousness (Steele & Aronson, 1995; Puhl & Heuer, 2009), the writing prompt is designed to make body-

related characteristics salient, thereby shifting attention toward the stigmatized identity. The randomization should therefore produce a *positive* first-stage effect on the salience measure.

H2 (Reduced Form). Participants in the salience condition perform worse on the CRT than those in the neutral condition. Econometrically, we would expect: $\tau < 0$.

If the writing task successfully raises weight salience, and if higher salience lowers reflective reasoning, then participants randomly assigned to the salience condition should, on average, perform worse on the CRT than those in the control condition. This direct intention-to-treat effect reflects the overall cognitive burden imposed by making weight identity salient immediately prior to a reasoning task.

H3 (Structural Relationship). Greater weight salience reduces CRT performance by elevating the marginal cost of controlled reasoning. Econometrically, we expect $\beta_m < 0$.

From the optimal effort equation derived above, $e_i^*(m_i) = b^{-1}(a - \lambda_0 - \lambda_1 m_i)$, it can be seen that as salience increases, $\lambda_1 m_i$ increases, thereby lowering optimal effort and hence expected accuracy. Intuitively, increases in salience decrease the incentive to exert cognitive control, leading to less frequent suppression of intuitive but incorrect responses. This prediction aligns with studies demonstrating that weight-based identity threat impairs working memory and numerical reasoning, both of which are essential components of CRT performance (Frederick, 2005; Guardabassi & Tomasetto, 2018, 2020).

Together, these three hypotheses provide a coherent causal mechanism. Random assignment heightens momentary salience, which increases the marginal cost of cognitive effort, and this increased cost manifests as reduced performance on the reasoning task.

H4 (Exploratory). There will be some heterogeneity.

If the intensity of identity-related burden depends on baseline vulnerability (e.g., objective BMI, internalized stigma, etc.), then the structural parameter may exhibit heterogeneity. Gender differences in stigma are also likely to matter here. Prior work shows that women report higher rates of weight-related social evaluation and body dissatisfaction (Puhl & Heuer, 2009). Therefore, among other exploratory exercises, I will explore whether β_m is more negative among women than among men, and whether effects are larger among participants with higher BMI or those perceiving themselves as overweight.

4. Experimental Design and Procedures

This study employs a randomized online field experiment conducted on the Prolific platform. Recruitment is restricted to adults who self-identify as Black/African American, ensuring that the sample reflects a population for whom both weight stigma and racial stereotype pressures are socially relevant.

Participants and Sample Size. Based on a prospective power analysis, I target a sample of 500 complete responses. Prior research provides some guidance on expected effect sizes. For instance, Major et al. (2012) found that weight-visible conditions reduced Stroop performance by approximately Cohen's d = 0.35-0.45 among higher-BMI women, and Guardabassi and Tomasetto (2018, 2020) found working memory impairments of d = 0.30-0.40 under stereotype threat conditions. Given that my manipulation is lighter-touch, i.e., a brief writing task rather than a videotaped speech in a standard lab setting, I assume a conservative effect size of d = 0.25 for the intention-to-treat effect. For a two-group comparison with equal allocation, $\alpha = 0.05$ (two-

tailed), and 80% power, the required sample size is approximately 253 per group, yielding a total of 506. Should the true effect be larger (d = 0.30), power would exceed 90%; should it be smaller (d = 0.20), power would fall to approximately 65%, still sufficient for detecting meaningful effects and establishing boundary conditions. Participant compensation is $3.00 per completed response, yielding an hourly rate of approximately $18. This figure aligns with both the Federal minimum wage and experimental economics norms. In addition, 5% of participants are randomly selected to receive a $1.00 bonus, independent of their answers. This lottery is designed to discourage over-expenditure of effort on the CRT items and reduce incentives for cheating (i.e., consulting with external resources while completing the Qualtrics survey).

To maximize participant availability and data quality, I deploy both the pilot study and the actual study on a Tuesday morning at 10:00 AM Eastern Time. This reflects both platform-related considerations and specific characteristics of the target population. Previous studies on online labor markets suggest that participant activity on Prolific peaks during *weekday mid-mornings*, when a larger and more diverse pool of respondents is typically available (Peer et al., 2022). Deploying on Tuesday rather than Monday avoids the potential backlog effects that can occur as participants return to the platform after the weekend. Besides, a mid-morning deployment (in the Eastern time zone) accommodates participants across all US time zones during reasonable waking hours, reducing selection bias toward respondents with non-standard schedules. This timing strategy therefore enhances the efficiency of data collection and enables me to obtain a more representative Prolific sample of Black/African American adults.

Procedures. After providing informed consent, participants complete baseline demographic questions including age, race/ethnicity, and sex at birth, as well as self-reported anthropometric

measures for height and weight. These measures are necessary for the computation of Body Mass Index (BMI), which will inform my testing of the exploratory hypotheses. Participants also report perceived weight status on a five-point scale (very underweight to very overweight).

Treatment Assignment. Participants are then randomly assigned with equal probability (50/50) to one of two experimental conditions:

(1) Treatment (Weight-Salience Writing Task): Participants in this group complete a short writing task asking them to describe in 4-5 sentences how their body weight or body shape influences their daily life. This design induces attention to weight, and is intended to increase the momentary salience of this stigmatized attribute.

(2) Control (Neutral Reading Task): Participants assigned to this group will read a brief paragraph about the extinction of dinosaurs. The content of this reading is non-affective in nature, and deliberately unrelated to personal identity: "Paleontologists have argued for a long time that the demise of the dinosaurs was caused by climatic alterations associated with slow changes in the positions of continents and seas resulting from plate tectonics. Off and on throughout the Cretaceous (the last period of the Mesozoic era, during which dinosaurs flourished), large shallow seas covered extensive areas of the continents. Data from diverse sources, including geochemical evidence preserved in seafloor sediments, indicate that the Late Cretaceous climate was milder than today's. The days were not too hot, nor the nights too cold. The summers were not too warm, nor the winters too frigid. The shallow seas on the continents probably buffered the temperature of the nearby air, keeping it relatively constant."

This reading passage is adapted from the publicly available official TOEFL iBT practice materials published by the Educational Testing Service (ETS, 2023). This choice offers several methodological advantages. First, TOEFL reading passages undergo rigorous professional validation to ensure appropriate difficulty levels for adult readers and emotional neutrality across diverse populations. Second, the passage describes paleontological evidence about dinosaur extinction, an engaging topic that is entirely unrelated to body weight, physical appearance, health behaviors/outcomes, or social evaluation. Third, because standardized testing materials are designed to minimize extraneous cognitive interference, using such content reduces the risk that the control condition (inadvertently) primes any identity-relevant constructs. Finally, the passage length and complexity are fairly comparable to the treatment condition's writing task. To further ensure equivalent time-on-task across conditions, I impose a minimum page time of 90 seconds within the neutral reading block.

Right after the randomization, all participants respond to a single-item measure quantifying weight salience using an 11-point slider scale: "At this very moment, how much are you thinking about your body weight or your body shape?" (0 = Not at all on my mind, 10 = Very much on my mind). This measure directly operationalizes the endogenous variable m_i in the structural model, and provides a great manipulation check for the proposed psychological mechanism.

Once the salience measure is recorded, all participants complete a four-item, numeric-response-only version of the Cognitive Reflection Test (CRT). Backward navigation is prevented. Prior to presenting the items, I provide participants with a brief instruction: "Next, you will answer a few short, everyday problem-solving questions. Each question has a single best numeric answer. Please answer without looking up answers online, using calculators, or consulting other

resources. Please answer each question in order; do not skip ahead. All participants will receive the base payment for completing the study, and 5% of participants will be randomly selected for a separate bonus. The bonus lottery is completely independent of the responses in this section. By continuing, you confirm you will rely only on your own reasoning. You will move on automatically once this part is finished." Following the CRT, participants complete an integrity check asking whether they used external resources (search engines, calculators, or other tools) during the problem-solving section. The study concludes with a brief "debriefing message" re-explaining the research purpose and providing researcher contact information. This design provides a transparent mapping to the proposed theoretical and econometric models. Because the random assignment precedes measurement and does not alter incentives for CRT success, I argue that it satisfies the exclusion restriction required for valid IV estimation.

5. Estimation Strategy

In order to test the hypotheses, I adopt a three-stage empirical strategy. The *first stage* estimates the effect of random assignment on the salience measure corresponding to equation (5). I then estimate the effect of random assignment on CRT performance via the *reduced form* in equation (6). Finally, to document the causal effect of weight salience on performance, I perform two-stage least squares (2SLS) *structural estimation* of equation (7) using the treatment assignment as an instrument for weight salience. Throughout estimation I employ heteroskedasticity-robust standard errors. The first-stage F-statistic will allow me to assess instrument strength and interpret effect sizes in standardized units. Because CRT scores are bounded and have relatively low variance, as one of the robustness checks I also carry out the same analyses using a standardized outcome y_i^z (mean zero, SD one) to facilitate interpretation across settings.

6. Results

6.1. Study 1 (Pilot)

Before deploying the full study, I conducted a small pilot study with around 60 Black/African American adults recruited via Prolific (excluding invalid or incomplete submissions) to stress-test the survey instruments, validate the experimental manipulation, and obtain preliminary effect-size estimates. As can be seen, the pilot sample was fairly balanced across conditions (n_{Ctrl} = 32 and n_{Treat} = 30). The balance tests showed no statistically significant differences in baseline characteristics across conditions (all $p > 0.10$), giving me reassurance that the randomization was successful. Importantly, the first-stage results indicated that treated participants reported an average salience score 3.446 points higher (on the 0-10 scale) than control participants (*salience* $mean_{Ctrl}$ = 2.19, SD_{Ctrl} = 2.89, *salience* $mean_{Treat}$ = 5.63, SD_{Treat}= 3.61), corresponding to a first-stage F-statistic of 17.05 and Cohen's d = 1.06. This validated the strength of the prime. The reduced-form results indicate that treatment assignment induces a 0.504-point reduction in CRT performance (SE = 0.367, p = 0.175), corresponding to Cohen's d = -0.35. Although not statistically significant at conventional levels, as expected given the pilot's limited power, the point estimate was directionally consistent with Hypothesis 2 and larger than the preregistered assumption of d = 0.25. Control participants averaged 2.44 correct responses on the CRT items (SD = 1.48), while treatment participants averaged 1.93 correct responses (SD = 1.41), representing a 21% reduction in accuracy. Finally, the 2SLS structural estimate showed that all

else equal, a one-point increase in momentary salience reduced CRT scores by 0.146 points.[1] Median completion time slightly exceeded 6 minutes.

These pilot findings directly inform the main study in three important ways. First, the robust first-stage F-statistic of 17.05 substantially exceeds the conventional threshold of 10, confirming that the manipulation is sufficiently strong for a valid instrumental variables estimation. Second, the observed reduced-form effect size of d = -0.35 exceeds my initial assumption of d = 0.25; at this effect size, the planned sample of n = 500 achieves statistical power exceeding 90%. Third, the pilot showed a standard deviation of 1.48 on CRT scores for the control group, providing a reasonable benchmark for interpreting effect magnitudes in the main study.

The pilot did reveal some technical issues (e.g., duplicate submissions, which might have to do with inadvertently refreshing browsers) that motivated me to make several small improvements for the main study that aimed to enhance the quality of research data. Specifically, I screened out participants who took part in the pilot study, participants with no previous Prolific submissions, and participants who had an approval rate of less than 90%. Within the survey builder, I enabled Qualtrics' in-survey multiple-submissions prevention to block repeated submissions and added a Prolific-generated completion code to my original "thank-you" message to ensure proper submission tracking. To further minimize partial exposure and restart behavior, respondents are also gently instructed at the outset not to refresh or reload the survey while completing it, as doing so may cause responses to restart or not be recorded. In addition, for one CRT item,

[1] Cohen's d is calculated as: $d = \frac{\overline{Y_1} - \overline{Y_0}}{SD_{pooled}} = \frac{\overline{Y_1} - \overline{Y_0}}{\sqrt{\frac{(n_1-1)s_1^2 + (n_0-1)s_0^2}{n_1+n_0-2}}}$. For the first-stage Cohen's d (salience score), plugging in $\overline{Y_1} = 5.63;\ \overline{Y_0} = 2.19;\ n_1 = 30$, $s_1 = 3.61$, $n_0 = 32$, and $s_0 = 2.89$ yields $d^{firststage} = 1.06$. For the reduced-form Cohen's d (CRT score), plugging in $\overline{Y_1} = 1.93;\ \overline{Y_0} = 2.44;\ n_1 = 30$, $s_1 = 1.41$, $n_0 = 32$, and $s_0 = 1.48$ yields $d^{reducedform} = -0.35$.

despite my explicitly asking for the answer to be in cents, some pilot respondents entered dollar-formatted responses (e.g., 0.05). In the main study, I therefore imposed a simple input validation requiring whole-number entries corresponding to cents only. I also embedded a simple subtle attention check asking participants to identify the last digit of the number 2026, placed immediately after the Prolific ID verification, and before any demographic or substantive questions. This early placement ensures that inattentive respondents are identified before randomization, avoids any potential confounds between the attention check and the experimental manipulation, and prevents post-treatment selection bias that could arise from excluding participants based on responses collected after treatment assignment. Based on these pilot results and survey refinements, I proceeded with the full study.

6.2. Study 2 (Actual Deployment)

For the main study, I recruited 500 Black/African American adults via Prolific, with participants randomized in near-equal proportion to the weight-salience treatment condition (n = 243) or the control condition (n = 257).[2] Table 1 presents balance tests for baseline characteristics, showing that no statistically significant differences between treatment and control groups on age, gender composition, BMI, or perceived weight status were found. This provided me with reassurance that internal validity was achieved, and that any observed outcome differences were attributable to the experimental manipulation.

[2] Although the survey was sent to an initial 500 respondents (encompassing both the treatment group and control group) at once, due to my quality enhancement settings a small number of respondents were permanently disqualified and their data not collected by Qualtrics. Therefore, I raised the Prolific target to 535 to achieve my target Qualtrics sample size of 500. All data collection was completed on the same day.

Table 2 summarizes the two key outcome variables separated by experimental condition. For the first-stage outcome (weight salience), control participants reported a mean salience score of 3.39 (SD = 3.28) on the 0-10 scale, while treatment participants reported a mean of 6.59 (SD = 2.94). This difference of 3.20 points closely replicates the pilot finding of 3.45 points and corresponds to a very large Cohen standardized effect size, with the treatment nearly doubling baseline salience levels. Figure 2 complements this by displaying both the estimates and 95% confidence intervals, visually confirming the clear difference between the two conditions.

Table 3 shows the results for the 2SLS estimation. First, in line with the summary statistics, the first-stage regression shows that assignment to treatment increases weight salience by 3.21 points (SE = 0.278, $p < 0.01$). The first-stage F-statistic of 132.87 far exceeds conventional "rules of thumb" for instrument strength, indicating that the intervention worked as intended. Recall from Table 2 that for the primary outcome (number of correctly answered CRT items out of 4 total), control participants averaged 1.99 correct responses (SD = 1.52), while treatment participants averaged 1.87 correct responses (SD = 1.50). This is confirmed in the reduced-form reduction of 0.12 points and visualized in Figure 1. The effect is not statistically significant (SE = 0.135, $p = 0.375$). The standardized effect size (Cohen's $d = -0.08$) is negligible compared to the pilot estimate of -0.35. Importantly, the narrow 95% confidence interval suggests that even if a true negative effect exists, it is likely smaller than 0.4 points on the four-item scale. Finally, the last column of Table 3 presents the 2SLS estimate of the structural parameter. Each one-point increase in salience is associated with a 0.037-point reduction in CRT performance (SE = 0.042), but given the null reduced form, this estimate is not statistically meaningful.

Taken together, Hypothesis 1 is strongly supported by the data, whereas Hypothesis 2 and Hypothesis 3 are not supported. Overall, the findings suggest that, in an anonymous online environment, briefly heightening weight salience does not acutely impair reflective reasoning among Black/African American adults. This calls into question the external validity of identity-salience effects previous documented in the social psychology literature.

6.3. Heterogeneity and robustness

To explore whether the cognitive effects of weight salience vary across observable characteristics, I estimate a series of interaction models of the form:

$$y_i = \mu + \rho Z_i + \delta G_i + \psi(Z_i \times G_i) + \varepsilon_i, \qquad (8)$$

where G_i is a binary indicator of a chosen subgroup. Table 4 summarizes the results for the heterogeneity exercise conducted across five dimensions, namely gender, BMI category, perceived overweight status, age, and obesity status. It is important to note that because these (exploratory) subgroup analyses involve multiple comparisons, the associated *p*-values should therefore be interpreted with caution rather being treated as confirmatory.

First, given prior evidence showing higher rates of weight-based social evaluation and body dissatisfaction among women, I test for possible gender differences in the treatment effect. Table 4 shows that the interaction is negative but are both statistically and economically insignificant (estimate = -0.179, SE = 0.262, $p = 0.495$). Second, I examine whether objective weight status moderates the main effect, as individuals with higher BMI may be more vulnerable to weight-based identity threat (Major et al., 2012). The interaction coefficients for the subgroups of Blacks with BMI of at least 25 and of Blacks with BMI of at least 30 both reverse sign and are

imprecisely estimated. Third, contrary to some previous studies highlighting the internalization of weight stigma independent of objective BMI, I fail to reject the hypothesis that perceived weight status (Overweight vs. Not Overweight) moderates the treatment effect (coefficient = 0.051, SE = 0.269, $p = 0.849$). Finally, I find that age (Age < 40 vs. Age ≥ 40) does not moderate the effect (coefficient = -0.086, SE = 0.273, $p = 0.754$).

In addition to interaction-based heterogeneity analyses, I conduct several robustness checks which apply the full 2SLS procedure to alternative samples and outcome definitions, including screens for implausible BMI values, standardized CRT outcomes, restrictions to respondents passing the attention check, and restrictions to respondents reporting no use of external resources). In every specification, the treatment produces a strong and precisely estimated first stage, with the first-stage F-statistics consistently in the 133-140 range, yet the corresponding 2SLS estimates remain small and statistically indistinguishable from zero. The results for the robustness exercise are suppressed to conserve space but are available upon request.

Both the heterogeneity and robustness exercises suggest that the null main effect is not driven by particular subgroups, measurement choices, or sample restrictions. This appears to contrast previous findings (Major et al., 2012, Puhl et al., 2018). However, in addition to the differing contexts, two methodological caveats warrant attention. First, although the sample ($n = 500$) as a whole provides adequate power to detect reasonable main effects, interaction models presumably require much larger samples. My prospective power analysis suggests that smaller (but potentially meaningful) heterogeneity could remain undetected. Second, the subgroup definitions employed here necessarily involve some degree of arbitrariness and exclusivity, preventing a definitive interpretation.

7. Discussion

Motivated by the dual obesity and racial-stereotype challenges facing the Black/African American community, I test whether experimentally induced weight salience translates into a reduction in cognitive performance among this population. Using a randomized online field experiment with 500 anonymous participants recruited through Prolific, I find that the weight-salience writing task powerfully increased momentary attention to body weight, almost doubling baseline salience levels with a standardized effect size exceeding one standard deviation. Despite the success of the first-stage experimental manipulation, however, I find no evidence that elevated salience impaired Blacks' performance on the Cognitive Reflection Test. In particular, the reduced-form effect is small (d = -0.08) and statistically insignificant. Exploratory heterogeneity analyses reveal no systematic variation across gender, age, or weight-related subgroups, and the planned robustness checks show that the null result is not driven by particular sample restrictions, control sets, or outcome definitions. My paper's contribution is therefore to document a novel set of 2SLS results and identify a clear boundary condition. Results show that even when strongly induced, weight salience does not appear to as disruptive to Black/African American adults' reasoning performance as previous work suggests.

These findings challenge psychological evidence documenting cognitive impairments (reduced Stroop-task performance and lower working memory) under weight-based identity threat (Major et al., 2012; Guardabassi and Tomasetto, 2018, 2020). An important contrast between my investigation and these studies is the social environment that embeds the identity cue. In Major et al.'s design, for instance, the research participants deliver a speech they believe will be judged by observers, creating a high-stakes evaluative context with explicit audience presence. On the other

hand, my manipulation involves *private, written* reflection, and the outcome task is completed anonymously without feedback (or social comparison). The current results therefore suggest that salience alone is likely not sufficient to induce psychological stress, thought suppression, and other mechanisms which have been previously theorized.

My heterogeneity results likewise refute the hypothesis that effects are concentrated among the most vulnerable subgroups. For example, despite evidence of the influence of internalized weight stigma on psychological well-being and the higher rates of weight-based social evaluation among women in prior work, I show that the aggregate result is neither driven by respondents with "overweight" status, nor suggestive of substantial gender differences (Puhl et al., 2018; Puhl and Heuer, 2009). These patterns hold for BMI categories, age, and obesity status as well, signifying the consistency of the null main effect across the sample. This conclusion is reinforced with a battery of robustness checks entailing different sample restriction rules and outcome definitions. Across specification, the experimentation consistently produces a strong first stage (with exceedingly high F-statistics), whereas the corresponding 2SLS estimates remain economically and statistically insignificant. The structural interpretation suggests very modest implied counterfactual gains from reducing salience.

At least three main limitations of the study merit attention. First, for both logistic and internal-validity reasons, the weight-salience manipulation is intentionally brief, private, and text-based. It therefore may not capture many real, high-stakes environments where identity-related cognitive effects prevail. Second, the outcome measure of most interest captures a single dimension of cognitive performance. In particular, the Cognitive Reflection Test only assesses one's propensity to engage in reflective reasoning, and the results may not readily generalize to

other domains such as working memory, processing speed, or sustained attention. Relatedly, while the non-incentivized nature of the study is a strength in itself (i.e., it helps cleanly distinguish effort from the tendency to override intuitive responses), it limits the study's applicability to longer, incentive-compatible tasks. Finally, although the overall sample size provides adequate power to detect reasonable average effects, interaction models require substantially larger samples, so smaller heterogeneous effects may remain undetected.

8. Concluding Remarks

This paper provides the first well-powered experimental test of whether acute weight salience causally impairs cognitive reflection among Black/African American adults. Using a Prolific experiment (n = 500), I show that in the first stage, salience strongly rises by more than one standard deviation, yet the reduced-form and 2SLS estimates on CRT performance are modest, and show no systematic heterogeneity across gender, age, or weight-related subgroups. The null effect is precisely estimated, robust across subgroups, and inconsistent with theoretical predictions derived from the stereotype-threat and weight-stigma literatures. Altogether, these findings challenge the assumption that brief identity-based primes work in online, low-stakes environments. They also suggest that context matters in interpreting earlier laboratory results.

Considering the aforementioned limitations, several implementable extensions follow naturally. First, future lab designs can explicitly embed audience evaluation or higher stakes to test whether salience interacts with the social environment. Second, future work can explore a broader range of outcomes beyond the CRT (including working memory, sustained attention, and incentive-compatible tasks), and study which cognitive margins respond the strongest to weight salience and under what conditions. Finally, as field experiments are gaining traction in most subfields of

economics, embedding the weight-salience manipulation within consequential assessments, such as job readiness evaluations or educational placement tests, would help translate the insights from this study into more policy-relevant settings. By providing novel experimental evidence on an understudied group, this study contributes to a better understanding of when and to whom identity-salience effects matter.

References

Akerlof, G. A., & Kranton, R. E. (2000). Economics and identity. *The Quarterly Journal of Economics*, *115*(3), 715-753.

Andreyeva, T., Puhl, R. M., & Brownell, K. D. (2008). Changes in perceived weight discrimination among Americans. *Obesity, 16*(5), 1129–1134.

Barnes, L. L., Lewis, T. T., Begeny, C. T., Yu, L., Bennett, D. A., & Wilson, R. S. (2012). Perceived discrimination and cognition in older African Americans. *Journal of the International Neuropsychological Society*, *18*(5), 856-865.

Beilock, S. L., & Carr, T. H. (2005). When high-powered people fail: Working memory and "choking under pressure" in math. *Psychological Science*, *16*(2), 101-105.

Benjamin, D. J., Choi, J. J., & Strickland, A. J. (2010). Social identity and preferences. *American Economic Review*, *100*(4), 1913-1928.

Cawley, J. (2004). The impact of obesity on wages. *Journal of Human Resources*, *39*(2), 451-474.

Centers for Disease Control and Prevention. (2024). *Adult obesity facts*. U.S. Department of Health and Human Services. Retrieved from: https://www.cdc.gov/obesity/adult-obesity-facts/?CDC_AAref_Val=https://www.cdc.gov/obesity/data/adult.html.

Deming, D. J. (2017). The growing importance of social skills in the labor market. *Quarterly Journal of Economics, 132*(4), 1593–1640.

Educational Testing Service. (2023). *TOEFL iBT Reading Practice Questions* [Practice test materials]. Retrieved from: https://www.ets.org/pdfs/toefl/toefl-ibt-reading-practice-sets.pdf.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, 19(4), 25–42.

Guardabassi, V., & Tomasetto, C. (2018). Does weight stigma reduce working memory? Evidence of stereotype threat susceptibility in adults with obesity. *International Journal of Obesity*, *42*(8), 1500-1507.

Guardabassi, V., & Tomasetto, C. (2020). Weight status or weight stigma? Obesity stereotypes—Not excess weight—Reduce working memory in school-aged children. *Journal of Experimental Child Psychology*, *189*, 104706.

Himmelstein, M. S., Puhl, R. M., & Quinn, D. M. (2017). Intersectionality: an understudied framework for addressing weight stigma. *American Journal of Preventive Medicine*, *53*(4), 421-431.

Hoff, K., & Pandey, P. (2006). Discrimination, social identity, and durable inequalities. *American Economic Review*, *96*(2), 206-211.

Hunger, J. M., Major, B., Blodorn, A., & Miller, C. T. (2015). Weighed down by stigma: How weight-based social identity threat contributes to weight gain and poor health. *Social and Personality Psychology Compass*, *9*(6), 255-268.

Inzlicht, M., McKay, L., & Aronson, J. (2006). Stigma as ego depletion: How being the target of prejudice affects self-control. *Psychological Science*, *17*(3), 262-269.

Major, B., Eliezer, D., & Rieck, H. (2012). The psychological weight of weight stigma. *Social Psychological and Personality Science*, *3*(6), 651-658.

Peer, E., Rothschild, D., Gordon, A., Evernden, Z., & Damer, E. (2022). Data quality of platforms and panels for online behavioral research. *Behavior Research Methods*, *54*(4), 1643-1662.

Petersen, R., Pan, L., & Blanck, H. M. (2019). Racial and ethnic disparities in adult obesity in the United States: CDC's tracking to inform state and local action. *Preventing Chronic Disease*, *16*, E46.

Puhl, R. M., & Heuer, C. A. (2009). The stigma of obesity: A review and update. *Obesity*, 17(5), 941–964.

Puhl, R. M., & Heuer, C. A. (2010). Obesity stigma: Important considerations for public health. *American Journal of Public Health*, 100(6), 1019–1028.

Puhl, R. M., Himmelstein, M. S., & Quinn, D. M. (2018). Internalizing weight stigma: prevalence and sociodemographic considerations in US adults. *Obesity*, *26*(1), 167-175.

Schilbach, F., Schofield, H., & Mullainathan, S. (2016). The psychological lives of the poor. *American Economic Review*, *106*(5), 435-440.

Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review*, *115*(2), 336.

Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science*, *10*(1), 80-83.

Spencer, S. J., et al. (1999). Stereotype threat and women's math performance. *Journal of Experimental Psychology: Applied, 5*(1), 4–28.

Steele, C. M., & Aronson, J. (1995). Stereotype threat. *Journal of Personality and Social Psychology, 69*(5), 797–811.

Sutin, A. R., Stephan, Y., Gerend, M. A., Robinson, E., Daly, M., & Terracciano, A. (2020). Perceived weight discrimination and performance in five domains of cognitive function. *Journal of Psychosomatic Research*, *131*, 109793.

Thomson, K. S., & Oppenheimer, D. M. (2016). Investigating an alternate form of the cognitive reflection test. *Judgment and Decision Making*, *11*(1), 99-113.

Tomiyama, A. J., Carr, D., Granberg, E. M., Major, B., Robinson, E., Sutin, A. R., & Brewis, A. (2018). How and why weight stigma drives the obesity 'epidemic'and harms health. *BMC Medicine*, *16*(1), 123.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition*, *39*(7), 1275-1289.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the Cognitive Reflection Test. *Thinking & Reasoning*, *20*(2), 147-168.

Table 1. Baseline Summary Statistics

		(1)		(2)		(1)-(2)
		Control		**Treatment**		Pairwise t-test
Variable	N	Mean/(SE)	N	Mean/(SE)	N	Mean difference
Age	257	38.809	243	39.770	500	-0.960
		(0.823)		(0.812)		
Female	257	0.560	243	0.519	500	0.042
		(0.031)		(0.032)		
BMI	257	27.155	243	27.693	500	-0.538
		(0.472)		(0.479)		
Perceived Weight	257	3.409	243	3.436	500	-0.028
		(0.046)		(0.046)		

Note: The table shows the differences in means across groups. ***, **, and * indicate significance at 0.01, 0.05, and 0.10. *Age* is a numeric variable indicating self-reported age in years. *Female* is binary indicator of gender assigned at birth (Female = 1, and Male = 0). *BMI* is a numeric variable derived from self-reported weight in pounds and self-reported height in inches according to the formula $703 \times \frac{Weight}{Height^2}$. *Perceived Weight* is an ordinal variable with five categories (1 – Very underweight, 2 – Underweight, 3 – About the right weight, 4 – Overweight, and 5 – Very overweight).

Table 2. Outcome Summary Statistics Separated by Treatment Status

Variable	Condition	N	Mean	SD	Min	Max
CRT Score	Control	257	1.992	1.518	0	4
(0-4)	Treatment	243	1.872	1.501	0	4
Weight Salience	Control	257	3.385	3.280	0	10
(0-10)	Treatment	243	6.593	2.939	0	10

Note: The table reports summary statistics for the two primary outcome variables by treatment status. *CRT Score* is the total number of correct responses on the four-item Cognitive Reflection Test. *Weight Salience* is a self-reported measure of momentary attention to body weight on a 0-10 scale. Control and Treatment refer to assignment to the neutral reading task and the weight-related reflective writing task, respectively.

Table 3. Main Instrumental Variable Estimation Results

	(1)	(2)	(3)
	First Stage	**Reduced Form**	**2SLS**
Treatment	3.207***	-0.120	
	(0.278)	(0.135)	
Salience			-0.037
			(0.042)
Observations	500	500	500
R-squared	0.210	0.002	.
First-stage *F*-stat	132.868		

Note: Column (1) reports the first-stage regression of salience on treatment assignment. Column (2) reports the reduced-form regression of CRT performance on treatment assignment. Column (3) reports the two-stage least squares (2SLS) estimate instrumenting salience with treatment assignment. The first-stage F-statistic is reported from column (1). Robust standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 4. Exploratory Heterogeneity Analyses

	(1)	(2)	(3)	(4)	(5)
	Female	BMI ≥ 25	Perceived Overweight	Young (<40)	BMI ≥ 30
Treatment	-0.055	-0.222	-0.140	-0.069	-0.148
	(0.191)	(0.212)	(0.180)	(0.205)	(0.159)
Treatment × *Subgroup*	-0.179	0.210	0.051	-0.086	0.089
	(0.262)	(0.275)	(0.269)	(0.273)	(0.279)
p-value (interaction)	0.495	0.444	0.849	0.754	0.750
Observations	500	500	500	500	500

Note: Each column reports results from a separate regression of CRT score on the treatment indicator, a subgroup indicator, and their interaction. The coefficient on "Treatment" gives the effect for the omitted reference category, while the interaction term reports the differential effect for the subgroup. Robust standard errors are reported in parentheses.

Figure 1. Effect of Random Assignment on Reasoning

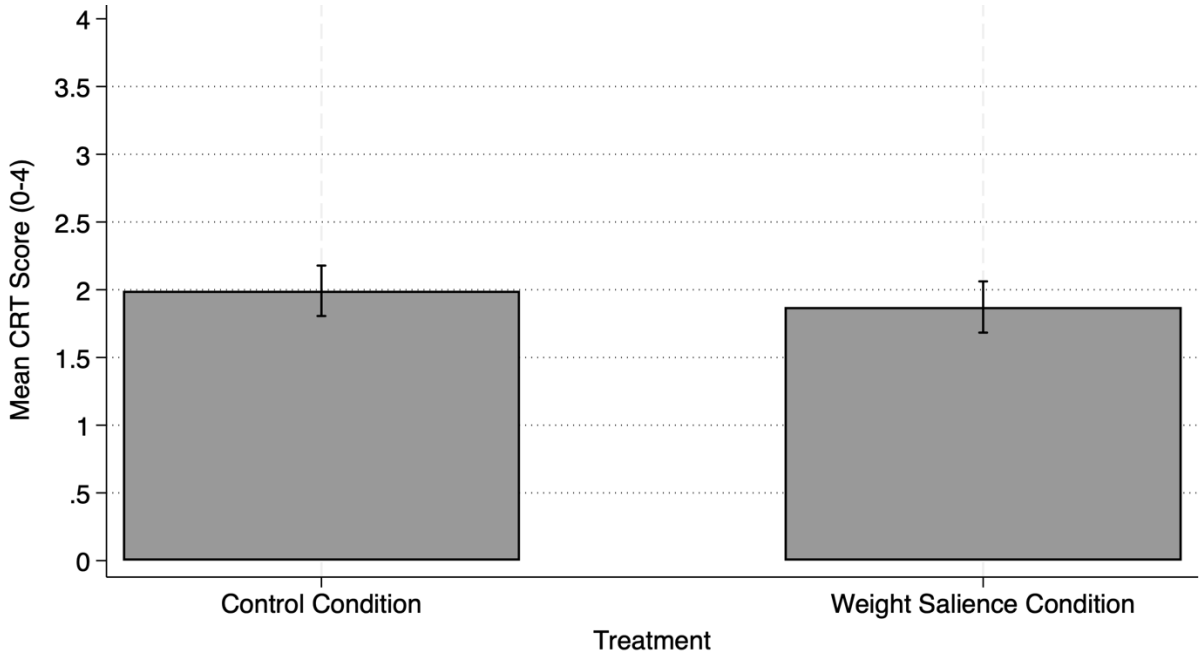
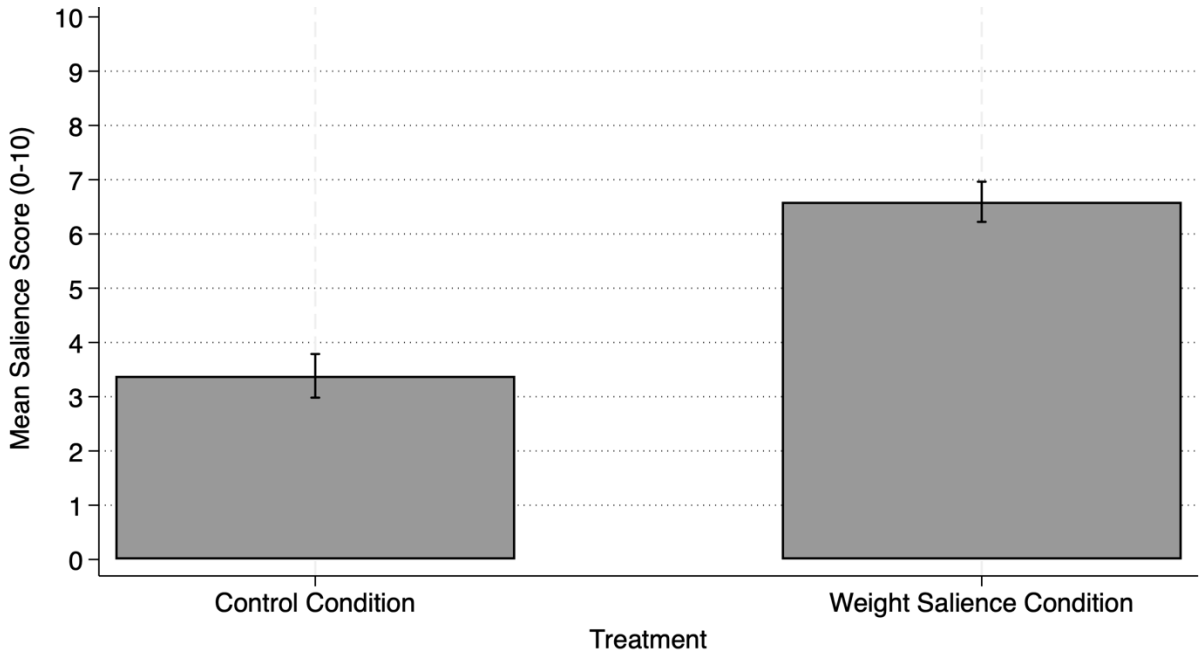


Figure 2. Effect of Randomization on Salience Score



Appendix A1. Baseline Summary Statistics (Pilot)

		(1)		(2)		(1)-(2)
		Control		**Treatment**		Pairwise t-test
Variable	N	Mean/(SE)	N	Mean/(SE)	N	Mean difference
Age	32	40.594	30	37.567	62	3.027
		(2.142)		(1.875)		
Female	32	0.469	30	0.533	62	-0.065
		(0.090)		(0.093)		
BMI	32	26.144	30	27.474	62	-1.330
		(1.328)		(1.404)		
Perceived Weight	32	3.375	30	3.300	62	0.075
		(0.133)		(0.109)		

Note: The table shows the differences in means across groups. ***, **, and * indicate significance at 0.01, 0.05, and 0.10. *Age* is a numeric variable indicating self-reported age in years. *Female* is binary indicator of gender assigned at birth (Female = 1, and Male = 0). *BMI* is a numeric variable derived from self-reported weight in pounds and self-reported height in inches according to the formula $703 \times \frac{Weight}{Height^2}$. *Perceived Weight* is an ordinal variable with five categories (1 – Very underweight, 2 – Underweight, 3 – About the right weight, 4 – Overweight, and 5 – Very overweight).

Appendix A2. Outcome Summary Statistics Separated by Treatment Status (Pilot)

Variable	Condition	N	Mean	SD	Min	Max
CRT Score	Control	32.0000	2.4375	1.4797	0	4
(0-4)	Treatment	30.0000	1.9333	1.4126	0	4
Weight Salience	Control	32.0000	2.1875	2.8898	0	10
(0-10)	Treatment	30.0000	5.6333	3.6149	0	10

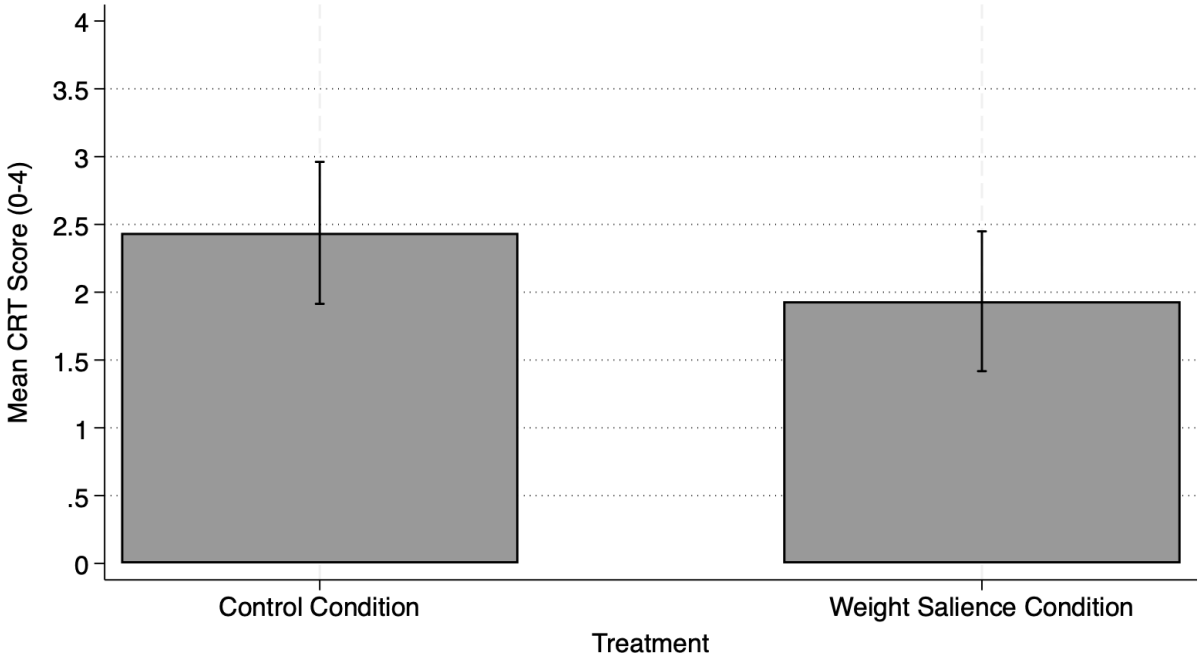
Note: The table reports summary statistics for the two primary outcome variables by treatment status. *CRT Score* is the total number of correct responses on the four-item Cognitive Reflection Test. *Weight Salience* is a self-reported measure of momentary attention to body weight on a 0-10 scale. Control and Treatment refer to assignment to the neutral reading task and the weight-related reflective writing task, respectively.

Appendix A3. Main Instrumental Variable Estimation Results (Pilot)

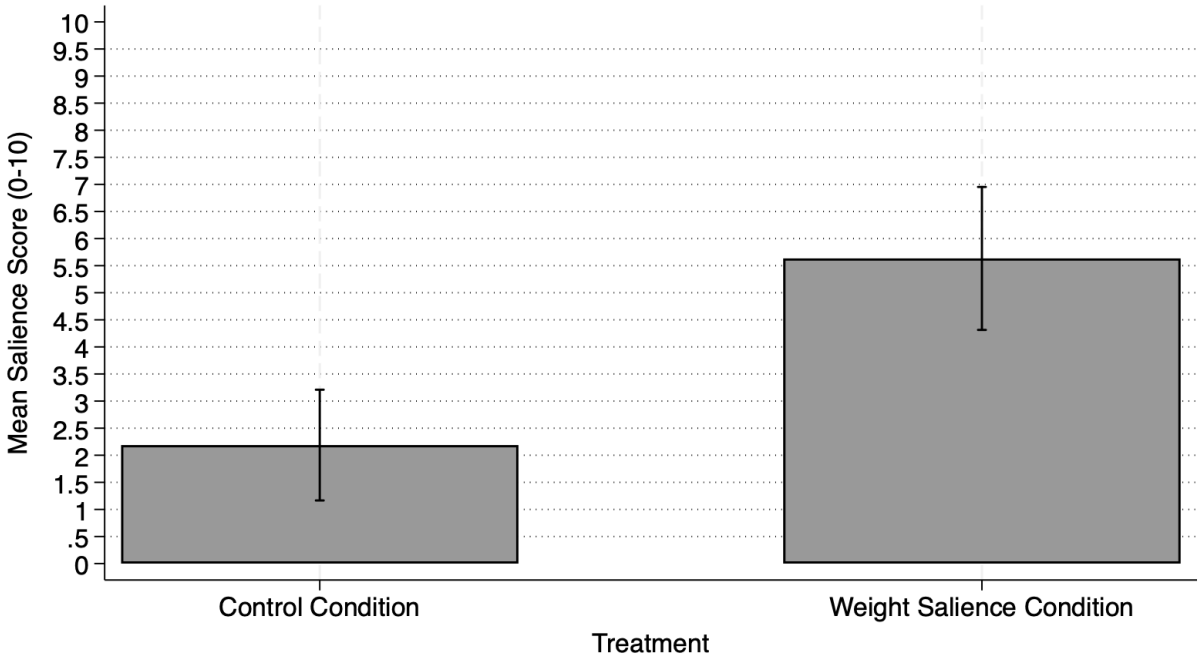
	(1)	(2)	(3)
	First Stage	**Reduced Form**	**2SLS**
Treatment	3.446***	-0.504	
	(0.834)	(0.367)	
Salience			-0.146
			(0.116)
Observations	62.000	62.000	62.000
R-squared	0.224	0.030	.
First-stage *F*-stat	17.052		

Note: Column (1) reports the first-stage regression of salience on treatment assignment. Column (2) reports the reduced-form regression of CRT performance on treatment assignment. Column (3) reports the two-stage least squares (2SLS) estimate instrumenting salience with treatment assignment. The first-stage F-statistic is reported from column (1). Robust standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Appendix A4. Effect of Random Assignment on Reasoning



Appendix A5. Effect of Randomization on Salience Score



Appendix A6: Full Survey Instruments (Actual Deployment)

Economics and Everyday Reasoning: The African American Experience

INVITATION TO PARTICIPATE:
Welcome. You are invited to participate in a research study about everyday thoughts and brief decision-making tasks. This study is conducted by Dr. Hieu M. Nguyen, Assistant Professor of Economics at Southwestern University. You are being recruited because you previously indicated on Prolific that you self-identify as Black or African American. Please read this form carefully and ask any questions you may have before agreeing to participate.

KEY INFORMATION ABOUT THIS RESEARCH STUDY:
The following is a short summary of this study to help you decide whether to participate; more details are provided below.

STUDY PURPOSE:
The purpose of this research is to better understand how individuals think about daily experiences and how these may relate to brief reasoning tasks.

PROCEDURES FOR THE STUDY:
If you choose to participate, you will complete an online survey that includes questions about yourself and a few short writing or reading activities, followed by several quick questions that involve thinking about everyday situations and providing numerical answers. The survey can be completed wherever you feel comfortable. No identifying information such as your name or contact details will be requested. Approximate duration: 10 minutes.

RISKS:
The risks associated with the study are no greater than those experienced in everyday life.

CONFIDENTIALITY:
The records of this study will be kept private, in accordance with Prolific's data protection and privacy policy (click here for more information). Your responses are anonymous. Any data downloaded by the researcher will contain only anonymized response codes. All information will be stored securely on a password-protected device accessible only to Dr. Hieu Nguyen. Data may be used in scholarly presentations or publications, but only in aggregate, and without any details that could reasonably identify you. An anonymized dataset may later be shared in a secure academic repository to support scientific transparency. Third-party platforms (e.g., Prolific, Qualtrics) have independent privacy policies outside the control of Southwestern University; however, we re-emphasize that the researcher will not receive personal identifying information.

BENEFITS OF TAKING PART IN THE STUDY:
There are no direct personal benefits to you beyond some monetary compensation. However, your participation contributes to a better understanding of decision-making processes, which may benefit research in behavioral economics and social psychology.

PAYMENT OR INCENTIVE:

You will receive $3.00 for completing the study. Additionally, a random 5% of participants will receive a $1.00 bonus. This bonus lottery is independent of your answers and not based on performance.

VOLUNTARY PARTICIPATION:
Taking part in this study is voluntary. You may choose not to take part or may leave the study at any time. Leaving the study will not result in any penalty. Your decision whether or not to participate in this study will not affect your current or future relations with Southwestern University or any of its representatives. You have the right to refuse to participate or to withdraw at any time without penalty or loss of benefits.

ALTERNATIVES TO TAKING PART IN THE STUDY:
If you do not want to be in the study, there are no other choices except not to take part in the study.

CONTACT INFORMATION:
If you have any questions, you may contact Dr. Hieu Nguyen at 412-218-6301 or nguyen23@southwestern.edu. You may have a copy of this form for your records. If you have any questions about your rights as a participant in this research or if you feel you have been placed at risk, contact the Southwestern University IRB at irb@southwestern.edu.

PARTICIPANT'S CONSENT:
- USE OF YOUR INFORMATION IN FUTURE RESEARCH

Your information collected for this study will NOT be used or shared for future research studies, even if we remove the identifiable information like your name or date of birth.

- ACKNOWLEDGEMENT OF CONSENT

By clicking the Consent button below, you acknowledge:

"I have read and understood the above explanations. The information that I provide will be used for research purposes only. I understand that my participation is voluntary and that I may withdraw at any time without penalty. Further, I acknowledge that I have been given the opportunity to ask questions, and my questions have been answered. If I have any concerns about my experience in this study (e.g., that I was treated unfairly or felt unnecessarily threatened), I may contact Dr. Hieu Nguyen or the Chair of the Institutional Review Board at Southwestern University regarding my concerns. I affirm that I am at least 18 years of age, I reside in the United States, and I voluntarily consent to participate in this research study."

Block 1: Consent and Basic Screening Questions
Q1 … *(full consent text)*
Q2 Are you 18 years of age or older?

Yes

No

Q3 How would you describe your race?

Black or African American (qualifies)

White

Hispanic or Latino/a/x

Asian

American Indian or Alaska Native

Native Hawaiian or Other Pacific Islander

Q4 Please enter your unique Prolific ID, which has exactly 24 alphanumeric characters. _____
Block 2: Basic Demographic Characteristics
Q5 Next question:
What is the last digit of 2026? Enter a single digit (0-9).
Q6 What is your age in years?
Q7 What is your sex assigned at birth?
Q8 What is your weight (in pounds)?
Q9 How would you describe your weight?

Very underweight

Underweight

About the right weight

Overweight

Very overweight

Q10 This question will ask you about your height. Please enter your height in two separate boxes below. Do not enter decimals or text.
Q10a For feet: Please enter a whole number between 4 and 7. ____
Q10b For inches: Please enter a whole number between 0 and 11. ____

[After this, participants will be randomized into either the Treatment block or the Control block.]

Block 3: Weight Salience Writing (Treatment)
Q11 Please write 4-5 sentences about how your body weight or your body shape affects your daily life. You may discuss any experiences that come to mind.

[Minimum character count: 150]

Block 4: Neutral Reading (Control)
Q12 Please read the short paragraph below. To ensure sufficient reading time, the 'Next' button will appear after a brief pause. Please do not refresh your browser.

Paleontologists have argued for a long time that the demise of the dinosaurs was caused by climatic alterations associated with slow changes in the positions of continents and seas resulting from plate tectonics. Off and on throughout the Cretaceous (the last period of the Mesozoic era, during which dinosaurs flourished), large shallow seas covered extensive areas of the continents. Data from diverse sources, including geochemical evidence preserved in seafloor sediments, indicate that the Late Cretaceous climate was milder than today's. The days were not too hot, nor the nights too cold. The summers were not too warm, nor the winters too frigid. The shallow seas on the continents probably buffered the temperature of the nearby air, keeping it relatively constant.

[Minimum page time: 90 seconds]

Block 5: Salience Measure
Q13 At this very moment, how much are you thinking about your body weight or your body shape?
(0 = Not at all on my mind, 10 = Very much on my mind)

Block 6: CRT Items
Next, you will answer a few short, everyday problem-solving questions. Each question has a single best numeric answer. Please answer without looking up answers online, using calculators, or consulting other resources. Please answer each question in order; do not skip ahead. All participants will receive the base payment for completing the study, and 5% of participants will be randomly selected for a separate bonus. The bonus lottery is completely independent of the responses in this section. By continuing, you confirm you will rely only on your own reasoning. You will move on automatically once this part is finished.

Q14 CRT Bat & Ball
A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball.
How much does the ball cost *in cents*? Please enter a whole number only (no dollar sign or decimal).
Q15 CRT Widgets
If 5 machines take 5 minutes to make 5 widgets, how many minutes would it take 100 machines to make 100 widgets? Please enter a whole number only.
Q16 CRT Race Position

You are running a race and pass the person in second place.
What place are you in now? Please enter a whole number only.
Q17 CRT Lily Pads
A patch of lily pads doubles in size every day.
If it takes 48 days to cover the entire pond, after how many days will it cover half the pond?
Please enter a whole number only.

Block 7: Integrity Check
Q18 During the previous problem-solving section, did you use any outside resources (such as a search engine, calculator, or other tools) to help answer the questions?
No. I did not use any outside sources.
Yes. I used outside sources.

Block 8: Debrief
Thank you for participating!
This study examines how attention to daily experiences may influence reasoning. The completion code is: _____ (unique study code generated by Prolific).
If you have questions, contact: nguyen23@southwestern.edu. Please click Next.