Forbes et al. (2008)

If the study has a broad focus and this data extraction focuses on just one component of the study, please specify this here

☒ Not applicable (whole study is focus of data extraction)

☐ Specific focus of this data extraction (please specify)

Study aim(s) and rationale

Was the study informed by, or linked to, an existing body of empirical and/or theoretical research?

Please write in authors' declaration if there is one. Elaborate if necessary, but indicate which aspects are reviewers' interpretation.

☒ Explicitly stated (please specify)

☐ Implicit (please specify)

☐ Not stated/unclear (please specify)

- Social neuroscience
- Stereotype threat
- Psychological disengagement
- devaluing and discounting
- anterior cingulate of the prefrontal cortex (ACC)
    - ACC is impicated in monitoring one's environment for goal conflic and in alerting areas in the prefrontal cortex of attributional discrepancies.

Do authors report how the study was funded?

☒ Explicitly stated (please specify)

☐ Implicit (please specify)

☐ Not stated/unclear (please specify)

Study research question(s) and its policy or practice focus

What is/are the topic focus/foci of the study?

- In this study, two distinct forms of chronic disengagement, devaluing the academic domain and discounting intellectual tests, were examined as predictors of performance monitoring processes of stigmatized minorities taking a putative intelligence test.
- The present study employed a cognitive neuroscience paradigm to assess activity in the anterior cingulate of the prefrontal cortex (ACC) in response to errors.
- The present study investigated how different form of psychological disengagement predict online performance monitoring processes.

What is/are the population focus/foci of the study?

- minority participants

What is the relevant age group?

☐ Not applicate (focus not learners)

☐ 0 - 4

☐ 5 - 10

☐ 11 - 16

☐ 17 - 20

☐ 21 and over

☒ Not stated/unclear

What is the sex of the population focus/foci?

☐ Not applicate (focus not learners)

☐ Female only

☐ Male only

☐ Mixed sex

☒ Not stated/unclear

What is/are the educational setting(s) of the study?

☐ Community centre

☐ Correctional institution

☐ Government department

☐ Higher education institution

☐ Home

☐ Independent school

☐ Local education authority

☐ Nursery school

☐ Other early years setting

☐ Post-compulsory education institution

☐ Primary school

☐ Residential school

☐ Secondary school

☐ Special needs school

☐ Workplace

☐ Other educational setting

In Which country or cuntries was the study carried out?

☐ Explicitly stated (please specify)

☐ Not stated/unclear (please specify)

Please describe in more detail the specific phenomena, factors, services, or interventions with which the study is concerned

What are the study reserach questions and/or hypotheses?

Research questions or hypotheses operationalise the aims of the study. Please write in authors' description if there is one. Elaborate if necessary, but indicate which aspects are reviewers' interpretation.

☒ Explicitly stated (please specify)

☐ Implicit (please specify)

☐ Not stated/unclear (please specify)

- We hypothesized that valuing would predict greater ERN amplitudes when the task was described as a measure of intelligence as this would activate goal-relevant motivations and more vigilant early stage error monitoring.

- Hypotheses for the Pe were more tentative given that relatively less is understood about the cognitive correlated of this potential

- If discounting requires a later evaluation of errors to cue an external attribution in response to perceived threat, however, then discounting might predict greater Pe amplitudes to errors when the task is described as a measure of intelligence.

- In addition, the role of disengagement tendencies and diagnosticity information on performance (e.g., errors, posterror slowing) and task construals (e.g. perceived difficulty, self-doubt) was examined.

- If the value placed on academics has implications for motivation, then valuing should predict performance-related variables such as fewer total errors and a tendency to slow down following an error.

- In contrast, if discounting has implications for how the situation is constructed, then discounting should predict more negative perceptions of a task linked to intelligence.

Methods - Design

Which variables or concepts, if any, does the study aim to measure or examine?

☒ Explicitly stated (please specify)

☐ Implicit (please specify)

☐ Not stated/unclear (please specify)

- devaluing measure

- discounting measure

- EEG data

- flankers task

Study timing

Please indicate all that apply and give further details where possible.

If the study examines one or more samples, but each at only one point in time it is cross-sectional.
If the study examines the same samples, but as they have changed over time, it is retrospective, provided that the interest is in starting at one timepoint and looking backwards over time.
If the study examines the same samples as they have changed over time and if data are collected forward over time, it is prospective provided that the interest is in starting at one timepoint and looking forward in time.

☒ Cross-sectional

☐ Retrospective

☐ Prospective

☐ Not stated/unclear (please specify)

If the study is an evaluation, when were measurements of the variable(s) used for outcome made, in relation to the intervention?

If at least one of the outcome variables is measured both before and after the intervention, please use the before and after category.

☐ Not applicable (not an evaluation)

☐ Before and after

☐ Only after

☐ Other (please specify)

☐ Not stated/unclear (please specify)

Methods - Groups

If comparisons are being made between two or more groups, please specify the basis of any divisions made for making these comparisons.

Please give further details where possible.

☐ Not applicable (not more than one group)

☒ Prospecitive allocation into more than one group (e.g. allocation to different interventions, or allocation to intervention and control groups)

☐ No prospective allocation but use of pre-existing differences to create comparison groups (e.g. receiving different interventions, or characterised by different levels of a variable such as social class)

☐ Other (please specify)

☐ Not stated/unclear (please specify)

How do the groups differ?

☐ Not applicable (not more than one group)

☒ Explicityly stated (please specify)

☐ Implicit (please specify)

☐ Not stated/unclear (please specify)

- diagnosticity condition (control condition vs. diagnostic of intelligence [DIQ] condition -> stereotype threat)

Number of groups

For instance, in studies in which comparisons are made between groups, this may be the number of groups into which the dataset is divided for analysis (e.g. social class, or form size), or the number of groups allocated to, or receiving, an intervention.

☐ Not applicable (not more than one group)

☐ One

☒ Two

☐ Three

☐ Four or more (please specify)

☐ Other/unclear (please specify)

Was the assignment of participants to interventions randomised?

☐ Not applicable (not more than one group)

☐ Not applicate (no prospective allocation)

☒ Random

☐ Quasi-random

☐ Non-random

☐ Not stated/unclear (please specify)

Where there was prospective allocation to more than one group, was the allocation sequence concealed from participants and those enrolling them until after enrolment?

Bias can be introduced, consciously or otherwise, if the allocation of pupils or classes or schools to a programme or intervention is made in the knowledge of key characteristics of those allocated. For example: children with more serious reading difficulty might be seen as in greater need and might be more likely to be allocated to the 'new' programme, or the opposite might happen. Either would introduce bias.

☐ Not applicable (not more than one group)

☐ Not applicable (no prospective allocation)

☒ Yes (please specify)

☐ No (please specify)

☐ Not stated/unclear (please specify)

Apart from the experimental intervention, did each study group receive the same level of care (that is, were they treated equally)?

☒ Yes
☐ No
☐ Can't tell

Study design summary

In addition to answering the questions in this section, describe the study design in your own words. You may want to draw upon and elaborate the answers you have already given.

1. pretest (devaluation and discounting measures)
2. EEG preparation by a white male experimenter
3. EEG recording + baseline version of the Eriksen-Flankers task
4. Random assignment to diagnosticity condition (control vs. DIQ)
5. Demographic items (including race/ethnicity in the DIQ condition)
6. EEG recording + second flankers task
7. final questionnaire while hooked up to physiological equipment
8. Debriefing and compensation

Methods - Sampling strategy

Are the authors trying to produce findings that are representative of a given population?

Please write in authors' description. If authors do not specify please indicate reviewers' interpretation.

☐ Explicitly stated (please specify)

☒ Implicit (please specify)

☐ Not stated/unclear (please specify)

- minority participants

Which methods does the study use to identify people or groups of people to sample from and what is the sampling frame?

e.g. telephone directory, electoral register, postcode, school listing, etc. There may be two stages – e.g. first sampling schools and then classes or pupils within them.

☐ Not applicable (please specify)

☒ Explicitly stated (please specify)

☐ Implicit (please specify)

☐ Not stated/unclear (please specify)

- Undergraduates

Which methods does the study use to select people or groups of people (from the sampling frame)?

e.g. selecting people at random, systematically - selecting for example every 5th person, purposively in order to reach a quota for a given characteristic.

☐ Not applicable (no sampling frame)

☒ Explicitly stated (please specify)

☐ Implicit (please specify)

☐ Not stated/unclear (please specify)

- University course/undergraduates

Planned sample size

If more than one group please give details for each group separately.

☐ Not applicable (please specify)
☐ Explicitly stated (please specify)
☐ Not stated/unclear (please specify)

Methods - Recruitment and consent

Which methods are used to recruit people into the study?

>   *e.g. letters of invitation, telephone contact, face-to-face contact.*

☐ Not applicable (please specify)

☒ Explicitly stated (please specify)

☐ Implicit (please specify)

☐ Not stated/unclear (please specify)

- University course/undergraduates

Were any incentives provided to recruit people into the study?

☐ Not applicable (please specify)

☒ Explicitly stated (please specify)

☐ Not stated/unclear (please specify)

- course credit or $20

Was consent sought?

>   *Please comment on the quality of consent if relevant.*

☐ Not applicable (please specify)
☐ Participant consent sought
☐ Parental consent sought
☐ Other consent sought
☐ Consent not sought
☒ Not stated/unclear (please specify)

Are there any other details relevant to recruitment and consent?

☐ No

☒ Yes (please specify)

- Eligible participants were permanent US residents and had no disabilities that would impair performance.

Methods - Actual sample

What was the total number of participants in the study (the actual sample)?

>   *If more than one group is being compared please give numbers for each group.*

☐ Not applicable (e.g. study of policies, documents, etc)

☒ Explicitly stated (please specify)

☐ Implicit (please specify)

☐ Not stated/unclear (please specify)

- Participants were 57 minority undergraduates (46 Latino, 11 African American) who participated for credit or \$20.

- 14 participants were excluded (3 from all analyses, 11 from the ERP analyses)

- Final sample of 43 participants (35 Latino, 9 African American) for the main ERP hypotheses.

What is the proportion of those selected for the study who actually participated in the study?

Please specify numbers and percentages if possible.

☐ Not applicable (e.g. study of policies, documents, etc)

☒ Explicitly stated (please specify)

☐ Implicit (please specify)

☐ Not stated/unclear (please specify)

- 43 out of 57

Which country/countries are the individuals in the actual sample from?

If UK, please distinguish between England, Scotland, N. Ireland, and Wales if possible. If from different countries, please give numbers for each. If more than one group is being compared, please describe for each group.

☐ Not applicable (e.g. study of policies, documents, etc)

☐ Explicitly stated (please specify)

☒ Implicit (please specify)

☐ Not stated/unclear (please specify)

- permanent US residents

What ages are covered by the actual sample?

Please give the numbers of the sample that fall within each of the given categories. If necessary, refer to a page number in the report (e.g. for a useful table). If more than one group is being compared, please describe for each group. If follow-up study, age at entry to the study.

☐ Not applicable (e.g. study of policies, documents, etc)
☐ 0 to 4
☐ 5 to 10
☐ 11 to 16
☐ 17 to 20

☐ 21 and over
☒ Not stated/unclear (please specify)

What is the socio-economic status of the individuals within the actual sample?

If more than one group is being compared, please describe for each group.

☐ Not applicable (e.g. study of policies, documents, etc)
☐ Explicitly stated (please specify)
☐ Implicit (please specify)
☒ Not stated/unclear (please specify)

What is the ethnicity of the individuals within the actual sample?

If more than one group is being compared, please describe for each group.

☐ Not applicable (e.g. study of policies, documents, etc)

☒ Explicitly stated (please specify)

☐ Implicit (please specify)

☐ Not stated/unclear (please specify)

- 35 Latino, 9 African American in the final sample for the main ERP hypotheses.

What is known about the special educational needs of individuals within the actual sample?

e.g. specific learning, physical, emotional, behavioural, intellectual difficulties.

☐ Not applicable (e.g. study of policies, documents, etc)
☐ Explicitly stated (please specify)
☐ Implicit (please specify)
☒ Not stated/unclear (please specify)

Is there any other useful information about the study participants?

☐ Not applicable (e.g. study of policies, documents, etc)

☒ Explicitly stated (please specify no/s.)

☐ Implicit (please specify)

☐ Not stated/unclear (please specify)

- One participants suspicious of the hypotheses and two participants who failed to follow instructions were excluded from analyses.

- Due to equipment malfunction, an additional 11 participants were excluded from ERP analyses, yielding a final sample of 43 participants for the main ERP hypotheses.

How representative was the achieved sample (as recruited at the start of the study) in relation to the aims of the sampling frame?

Please specify basis for your decision.

☐ Not applicable (e.g. study of policies, documents, etc)
☐ Not applicable (no sampling frame)
☐ High (please specify)
☒ Medium (please specify)
☐ Low (please specify)
☐ Unclear (please specify)

If the study involves studying samples prospectively over time, what proportion of the sample dropped out over the course of the study?

If the study involves more than one group, please give drop-out rates for each group separately. If necessary, refer to a page number in the report (e.g. for a useful table).

☐ Not applicable (e.g. study of policies, documents, etc)
☒ Not applicable (not following samples prospectively over time)
☐ Explicitly stated (please specify)
☐ Implicit (please specify)
☐ Not stated/unclear

For studies that involve following samples prospectively over time, do the authors provide any information on whether and/or how those who dropped out of the study differ from those who remained in the study?

☐ Not applicable (e.g. study of policies, documents, etc)
☒ Not applicable (not following samples prospectively over time)
☐ Not applicable (no drop outs)
☐ Yes (please specify)
☐ No

If the study involves following samples prospectively over time, do authors provide baseline values of key variables such as those being used as outcomes and relevant socio-demographic variables?

☐ Not applicable (e.g. study of policies, documents, etc)
☒ Not applicable (not following samples prospectively over time)
☐ Yes (please specify)
☐ No

Methods - Data collection

Please describe the main types of data collected and specify if they were used (a) to define the sample; (b) to measure aspects of the sample as findings of the study?

☐ Details

- demographic data -> a (and b in the threat condition)

- devaluing and discounting measures -> b

- EEG data -> b

- flankers task -> b

Which methods were used to collect the data?

Please indicate all that apply and give further detail where possible.

☐ Curriculum-based assessment
☐ Focus group
☐ Group interview
☐ One to one interview (face to face or by phone)
☐ Observation
☐ Self-completion questionnaire
☐ Self-completion report or diary
☐ Exams
☐ Clinical test
☐ Practical test
☐ Psychological test
☐ Hypothetical scenario including vignettes
☐ School/college records (e.g. attendance records etc)
☐ Secondary data such as publicly available statistics
☐ Other documentation
☐ Not stated/unclear (please specify)

Details of data collection methods or tool(s).

Please provide details including names for all tools used to collect data and examples of any questions/items given. Also please state whether source is cited in the report.

☒ Explicitly stated (please specify)

☐ Implicit (please specify)

☐ Not stated/unclear (please specify)

- Eriksen-Flankers task (Eriksen and Eriksen, 1974; Gehring et al., 1993)

- devaluing (five items, e.g. "Being good at academics is an important part of who I am" - reversed) and discounting (four items, e.g. "Most intelligence tests do not really measure what they are supposed to do"; Major and Schmader, 1998), items were rated from 1 (strongly disagree) to 7 (strongly agree)

- EEG was recorded from 32 tin electrodes embedded within a stretch-lycra cap, with a Cz reference and a ground lead located anterior to Fz on the mid-line of participant's scalps

- final questionnaire: participants reted (on a 7-point scale) the extent to which they felt doubtful, foolish, inferior, insecure and unsure while completing the flankers task

Who collected the data?

Please indicate all that apply and give further detail where possible.

☐ Researcher
☐ Head teacher/Senior management
☐ Teaching or other staff
☐ Parents
☐ Pupils/students
☐ Governors
☐ LEA/Government officials
☐ Other education practitioner
☐ Other (please specify)
☐ Not stated/unclear

Do the authors describe any ways they addressed the reliability of their data collection tools/methods?

e.g. test-retest methods (Where more than one tool was employed please provide details for each.)

☐ Details

Do the authors describe any ways they have addressed the validity of their data collection tools/methods?

e.g. mention previous validation of tools, published version of tools, involvement of target population in development of tools. (Where more than one tool was employed please provide details for each.)

☐ Details

Was there concealment of study allocation or other key factors from those carrying out measurement of outcome – if relevant?

Not applicable – e.g. analysis of existing data, qualitative study. No – e.g. assessment of reading progress for dyslexic pupils done by teacher who provided intervention. Yes – e.g. researcher assessing pupil knowledge of drugs - unaware of pupil allocation.

☐ Not applicable (please say why)
☐ Yes (please specify)
☐ No (please specify)

Where were the data collected?

e.g. school, home.

☐ Explicitly stated (please specify)

☐ Implicit (please specify)
☐ Unclear/not stated (please specify)

Are there other important features of data collection?

e.g. use of video or audio tape; ethical issues such as confidentiality etc.

☐ Details

Methods - Data analysis

Which methods were used to analyse the data?

Please give details e.g. for in-depth interviews, how were the data handled? Details of statistical analysis can be given next.

☒ Explicitly stated (please specify)
☐ Implicit (please specify)
☐ Not stated/unclear (please specify)

Final questionnaire: - responses were averaged to form a composite measure of self-doubt

EEG recording and data reduction: - All impedances were reduced below 10 $k\Omega$ prior to recording EEG activity - All EEG signals were filtered online from 0.05 to 200 Hz, amplified by a factor of 500 with Synamps digital amplifiers, and sampled at 1000 Hz - Offline, EEG signals were filtered with a 3003 point finite impulse response filter that passed signals from 1.5 to 15 Hz - Artefacts other than blinks were manually rejected off-line, and the effect of blinks was corrected using an ocular artefact regression correction procedure (Semlitsch et al., 1986) - EEG signals were then epoched off-line using Neuroscan software, response locked to the participants incorrect or correct responses, extending from 500 ms prior to the response onset and 1500 ms postresponse - Epochs were baseline corrected by subtracting the average value of EEG 50 ms before the response from the entire epoch - Following procedures typical of this paradigm (Hajcak et al., 2004), error-specific activity was isolated by subtracting average waveforms for correct responses from the average waveforms for error responses. - From these difference waveforms, the ERN was defined as the most negative deflection at site Fz between 50 and 130 ms after the response, and the Pe was defined as the most positive deflection at site Pz between 200 and 500 ms after the error (Falkenstein et al., 2000).

Descriptive statistics: - means and correlations for all variables involved in analyses

General error effects on brain electrical activity: - A general ERN pattern was established by a repeated measures analysis on premanipulation early stage amplitudes as a function of site (Fz, Cz, Pz) and accuracy (correct, error) - The Pe was established by a repeated measures analysis on premanipulation later stage amplitudes that revealed two main effect, qualified by a significant interaction

ERP results: - In a series of hierarchical regression analyses, devaluing and then discounting were tested as moderators of diagnosticity effects on postmanipulation ERN

amplitude (recorded during the letters task at site Fz) and Pe amplitudes (at site Pz) - Step 1 of these analyses included site-specific baseline activity recorded during the arrows task as a covariate. - The mean centred disengagement variable (devaluing or discounting) and diagnosticity condition were entered on Step 2, with their interaction entered on Step 3 (Aiken and West, 1991).

ERN: devaluing as a moderator - Simple slope analyses

Posterror slowing: - An index of posterror slowing was created by subtracting mean reaction times on correct trails following a correct response from mean reaction times on trails following an error (higher number equal more slowing) - simple slopes analyses

Which statistical methods, if any, were used in the analysis?

☐ Details

see above

What rationale do the authors give for the methods of analysis for the study?

e.g. for their methods of sampling, data collection, or analysis.

☐ Details

For evaluation studies that use prospective allocation, please specify the basis on which data analysis was carried out.

'Intention to intervene' means that data were analysed on the basis of the original number of participants as recruited into the different groups. 'Intervention received' means data were analysed on the basis of the number of participants actually receiving the intervention.

☐ Not applicable (not an evaluation study with prospective allocation)
☐ 'Intention to intervene'
☐ 'Intervention received'
☐ Not stated/unclear (please specify)

Do the authors describe any ways they have addressed the reliability of data analysis?

e.g. using more than one researcher to analyse data, looking for negative cases.

☐ Details

Do the authors describe any ways they have addressed the validity of data analysis?

e.g. internal or external consistency; checking results with participants.

☐ Details

Do the authors describe strategies used in the analysis to control for bias from confounding variables?

☐ Details

Please describe any other important features of the analysis.

☐ Details

Please comment on any other analytic or statistical issues if relevant.

☐ Details

Results and Conclusions

How are the results of the study presented?

e.g. as quotations/figures within text, in tables, appendices.

☐ Details

- in text

- table

- figures

What are the results of the study as reported by authors?

Please give details and refer to page numbers in the report(s) of the study where necessary (e.g. for key tables).

☐ Details

Descriptive statistics: - Devaluing and discounting were uncorrelated and did not vary by distnosticity condition - These two variables were distinct, and minority students generally reported valuing academics but discounting intelligence tests

General error effects on brain electrical activity: - ERN pattern revealed two main effects, qualified by a significant interaction - Planned contrasts revealed that ERN differences between error and correct trails were maximal at Fz and Cz compared to Pz - Pe revealed two main effects, qualified by a significant interaction - Consistent with the more centro-parietal scalp distribution of the Pe (Nieuwenhuis et al., 2001), planned contrasts indicated that Pe differences between error and correct trails were larger at Pz and Cz compared to Fz - For simplicity, ERN and Pe results were analysed for sites Fz and Pz, respectively

RN: devaluing as a moderator: - The analyses of dvaluing as a moderator of distnosticity on ERN amplitudes yielded the predicted interaction - Simple slope analyses revealed that devaluing predicted smaller ERN amplitudes in the DIQ condition, but not in the control condition - The diagnosticity effect was significant at 1 SD below the mean of devaluing but not at 1 SD above the mean - Because more negative ERN amplitudes indicate increased error monitoring, these results suggest that minority students exhibited more error monitoring to the degree that they valued the academic domain but only when

the task was linked to intelligence.
- The same analysis using discounting as a moderator yielded no significant effects.

Pe: discounting as a moderator: - Discounting did moderate diagnosticity effects on Pe amplitudes at Pz - Unexpectedly, discounting predicted lower Pe amplitudes when the task was neutral - But when the task was linked to intelligence, this association tended to be positive although non-significant - Linking the task to intelligence led to smaller Pe amplitudes compared to control among participants low in discounting, whereas this pattern tended to reverse among participants high in discounting - These results suggest that a tendency to discount academic performance predicts less error salience among minority students only when the task is not explicitly framed as a measure of intelligence. - When devaluing was tested as a moderator of diagnosticity effects on Pe amplitudes at Pz, no effects were significant, except for a devaluing main effect, suggesting that errors were generally more salient to academic values.

Error analyses: - On average, participants made 16.72% (SD = 37) of errors on the postmanipulation flankers task - Analyses of the number of errors made on the task yielded only a devaluing by diagnosticity interaction on total errors - Indicative of the idea that devaluing would entail lower motivation, devaluing predicted more errors when the task was linked to intelligence but not when it was described neutrally - No other effects (including those with discounting) were significant.

Posterror slowing: - There were no effects on overall reaction time, however, consistent with past evidence of posterror slowing (Rabbitt, 1981; Luu et al., 2000), reaction times for trails following errors were significantly slower than reaction times on error trails - After first accounting for posterror slowing at baseline, devaluing and discounting were examined as moderators of diagnosticity effects on posterror slowing, yielding a devaluing by diagnosticity interaction - Simple slopes revealed that, unexpectedly, devaluing predicted more posterror slowing when the test was linked to intelligence, but not when it was described neutrally - When paired with effects on the ERN and errors, these results suggest that among minorities who value academics, stereotype threat cues an automatic vigilance towards errors, and a tendency to make fewer errors and actually speed up on trails following errors

Self-reported difficulty and self-doubt: - Analyses of perceived difficulty and self-doubt yielded only discounting by diagnosticity interactions - When the task was linked to intelligence, discounting predicted greater perceived difficulty and a non-significant trend toward greater self-doubt - In the control condition, these patterns were reveresed but non-significant - Thus, discounting, but not devaluing, predicted greater subjective evaluations of threat in the diagnostic condition.

Was the precision of the estimate of the intervention or treatment effect reported?

- CONSIDER:
    - Were confidence intervals (CIs) reported?
  ☐ Yes
  ☒ No
  ☐ Can't tell

Are there any obvious shortcomings in the reporting of the data?

☐ Yes (please specify)
☒ No

Do the authors report on all variables they aimed to study as specified in their aims/research questions?

This excludes variables just used to describe the sample.

☒ Yes (please specify)
☐ No

Do the authors state where the full original data are stored?

☐ Yes (please specify)
☒ No

What do the author(s) conclude about the findings of the study?

Please give details and refer to page numbers in the report of the study where necessary.

☐ Details

The present results suggest that the answer lies in the extent to which stigmatized minorities tend to devalue the academic domain or discount intellectual tests. Specifically, devaluing and discounting were distinguished as two forms of disengagement that have different motivational and interpretative implications for minority students' performance monitoring processes in intellectually threatening environments.

When performance was framed in terms of intelligence, valuing predicted early stage motivational processes in performance monitoring (ERN amplitudes as an indicator of error detection) and behavior (faster reaction times following an error and fewer errors overall). These findings suggest that stigmatized minorities who value academics respond to stereotype threatening cues by becoming implicitly vigilant for performance relevant stimuli and more efficient in responding to them. These results highlight the increased motivation to excel that situations of stereotype threat elicit in domain identified targets (Jamieson and Harkins, 2007) and help delineate the processes by which threat-induced motivation translates into more efficient responding on tasks that are simple or well-learned (O'Brien and Crandall, 2003; Ben-Zeev et al., 2005).

It is noteworthy that high levels of valuing predicted larger ERN amplitudes under threat compared to control, but there was no evidence that devaluing related to smaller ERN amplitudes under threat compared to control. This pattern is consistent with the proposal that factors of the person (valuing the domain) and the situation (cues to domain relevance) might both be necessary to cue implicit processing of goal relevant stimuli (Klinger and Cox, 2004).

Whereas valuing moderated diagnosticity effects on the ERN, discounting moderated diagnosticity effects on Pe amplitudes. Unexpectedly, discounting predicted smaller Pe amplitudes when the task was described neutrally.

It is worth noting that the Pe has received relatively little attention, and research is still trying to isolate its psychological correlates (Overbeek et al., 2005). The interpretation we provide here is consistent with the view that this neural response might underscore evaluative salience of an error (Amodio et al., 2006).

Additional evidence that discounters were more threatened by their errors in the diagnostic condition was found in their posttask ratings of difficulty and self-doubt. Whereas discounting was unrelated to task construals in the neutral condition, it predicted greater perceived difficulty and relatively more self-doubt when the task was linked to intelligence.

Quality of the study - Reporting

Is the context of the study adequately described?

Consider your answer to questions: Why was this study done at this point in time, in those contexts and with those people or institutions? (Section B question 2) Was the study informed by or linked to an existing body of empirical and/or theoretical research? (Section B question 3) Which of the following groups were consulted in working out the aims to be addressed in the study? (Section B question 4) Do the authors report how the study was funded? (Section B question 5) When was the study carried out? (Section B question 6)

☐ Yes (please specify)
☐ No (please specify)

Are the aims of the study clearly reported?

Consider your answer to questions: What are the broad aims of the study? (Section B question 1) What are the study research questions and/or hypotheses? (Section C question 10)

☐ Yes (please specify)
☐ No (please specify)

Is there an adequate description of the sample used in the study and how the sample was identified and recruited?

Consider your answer to all questions in Methods on 'Sampling Strategy', 'Recruitment and Consent', and 'Actual Sample'.

☐ Yes (please specify)
☐ No (please specify)

Is there an adequate description of the methods used in the study to collect data?

Consider your answer to the following questions in Section I: Which methods were used to collect the data? Details of data collection methods or tools Who collected the data? Do the authors describe the setting where the data were collected? Are there other important features of the data collection procedures?

☐ Yes (please specify)
☐ No (please specify)

Is there an adequate description of the methods of data analysis?

Consider your answer to the following questions in Section J: Which methods were used to analyse the data? What statistical methods, if any, were used in the analysis? Who carried out the data analysis?

☐ Yes (please specify)
☐ No (please specify)

Is the study replicable from this report?

☐ Yes (please specify)
☐ No (please specify)

Do the authors avoid selective reporting bias?

(e.g. do they report on all variables they aimed to study as specified in their aims/research questions?)

☐ Yes (please specify)
☐ No (please specify)

Quality of the study - Methods and data

Are there ethical concerns about the way the study was done?

Consider consent, funding, privacy, etc.

☐ Yes, some concerns (please specify)
☐ No concerns

Were students and/or parents appropriately involved in the design or conduct of the study?

☐ Yes, a lot (please specify)
☐ Yes, a little (please specify)
☐ No (please specify)

Is there sufficient justification for why the study was done the way it was?

☐ Yes (please specify)
☐ No (please specify)

Was the choice of research design appropriate for addressing the research question(s) posed?

☐ Yes (please specify)
☐ No (please specify)

To what extent are the research design and methods employed able to rule out any other sources of error/bias which would lead to alternative explanations for the findings of the study?

e.g. (1) In an evaluation, was the process by which participants were allocated to or otherwise received the factor being evaluated concealed and not predictable in advance? If not, were sufficient substitute procedures employed with adequate rigour to rule out any alternative explanations of the findings which arise as a result? e.g. (2) Was the attrition rate low and if applicable similar between different groups?

☐ A lot (please specify)
☐ A little (please specify)
☐ Not at all (please specify)

How generalisable are the study results?

☐ Details

Weight of evidence - A: Taking account of all quality assessment issues, can the study findings be trusted in answering the study question(s)?

In some studies it is difficult to distinguish between the findings of the study and the conclusions. In those cases please code the trustworthiness of this combined results/conclusion. **Please remember to complete the weight of evidence questions B-D which are in your review specific data extraction guidelines.**

☐ High trustworthiness (please specify)
☐ Medium trustworthiness (please specify)
☐ Low trustworthiness (please specify)

Have sufficient attempts been made to justify the conclusions drawn from the findings so that the conclusions are trustworthy?

☐ Not applicable (results and conclusions inseparable)
☐ High trustworthiness
☐ Medium trustworthiness
☐ Low trustworthiness

Wells et al. (2014)

CASE CONTROL STUDIES

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Exposure categories. A maximum of two stars can be given for Comparability.

Selection

Is the case definition adequate?

- a) yes, with independent validation

- b) yes, e.g., record linkage or based on self reports
- c) no description

Representativeness of the cases

- a) consecutive or obviously representative series of cases *
- b) potential for selection biases or not stated

Selection of Controls

- a) community controls *
- b) hospital controls
- c) no description

Definition of Controls

- a) no history of disease (endpoint) *
- b) no description of source

Comparability

Comparability of cases and controls on the basis of the design or analysis

- a) study controls for _____ (Select the most important factor.) *
- b) study controls for any additional factor * (This criterion could be modified to indicate specific control for a second important factor.)

Exposure

Ascertainment of exposure

- a) secure record (e.g., surgical records) *
- b) structured interview where blind to case/control status *
- c) interview not blinded to case/control status
- d) written self report or medical record only
- e) no description

Same method of ascertainment for cases and controls

- a) yes *
- b) no

Non-Response rate

- a) same rate for both groups *
- b) non respondents described
- c) rate different and no designation

COHORT STUDIES

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Outcome categories. A maximum of two stars can be given for Comparability.

Selection

Representativeness of the exposed cohort

- a) truly representative of the average _____ (describe) in the community *
- b) somewhat representative of the average _____ in the community *
- c) selected group of users, e.g., nurses, volunteers
- d) no description of the derivation of the cohort

Selection of the non exposed cohort

- a) drawn from the same community as the exposed cohort *
- b) drawn from a different source
- c) no description of the derivation of the non exposed cohort

Ascertainment of exposure

- a) secure record (e.g., surgical records) *
- b) structured interview *
- c) written self report
- d) no description

Demonstration that outcome of interest was not present at start of study

- a) yes *
- b) no

Comparability

Comparability of cohorts on the basis of the design or analysis

- a) study controls for _____ (select the most important factor) *
- b) study controls for any additional factor * (This criterion could be modified to indicate specific control for a second important factor.)

Outcome

Assessment of outcome

- a) independent blind assessment *
- b) record linkage *
- c) self report
- d) no description

Was follow-up long enough for outcomes to occur

- a) yes (select an adequate follow up period for outcome of interest) *
- b) no

Adequacy of follow up of cohorts

- a) complete follow up - all subjects accounted for *
- b) subjects lost to follow up unlikely to introduce bias - small number lost - > _____ % (select an adequate %) follow up, or description provided of those lost) *
- c) follow up rate < _____% (select an adequate %) and no description of those lost
- d) no statement

University of Glasgow (n.d.)

DOES THIS REVIEW ADDRESS A CLEAR QUESTION?

Did the review address a clearly focussed issue?

- Was there enough information on:
  - The population studied
  - The intervention given
  - The outcomes considered
☐ Yes
☐ Can't tell
☐ No

Did the authors look for the appropriate sort of papers?

- The 'best sort of studies' would:
  - Address the review's question
  - Have an appropriate study design
☐ Yes
☐ Can't tell
☐ No

ARE THE RESULTS OF THIS REVIEW VALID?

Do you think the important, relevant studies were included?

- Look for:
  - Which bibliographic databases were used
  - Follow up from reference lists
  - Personal contact with experts
  - Search for unpublished as well as published studies
  - Search for non-English language studies
☐ Yes
☐ Can't tell
☐ No

Did the review's authors do enough to assess the quality of the included studies?

- The authors need to consider the rigour of the studies they have identified. Lack of rigour may affect the studies results.

☐ Yes

☐ Can't tell

☐ No

If the results of the review have been combined, was it reasonable to do so?

- Consider whether:
    - The results were similar from study to study
    - The results of all the included studies are clearly displayed
    - The results of the different studies are similar
    - The reasons for any variations are discussed

☐ Yes

☐ Can't tell

☐ No

WHAT ARE THE RESULTS?

What is the overall result of the review?

- Consider:
    - If you are clear about the review's 'bottom line' results
    - What these are (numerically if appropriate)
    - How were the results expressed (NNT, odds ratio, etc)

How precise are the results?

- Are the results presented with confidence intervals?

☐ Yes

☐ Can't tell

☐ No

WILL THE RESULTS HELP LOCALLY?

Can the results be applied to the local population?

- Consider whether:
    - The patients covered by the review could be sufficiently different from your population to cause concern
    - Your local setting is likely to differ much from that of the review

☐ Yes

☐ Can't tell

☐ No

Were all important outcomes considered?

☐ Yes
☐ Can't tell
☐ No

Are the benefits worth the harms and costs?

- Even if this is not addressed by the review, what do you think?

☐ Yes
☐ Can't tell
☐ No

References

Critical Appraisal Skills Programme. (2018). CASP Systematic Review Checklist [Organization]. In *CASP - Critical Appraisal Skills Programme*. https://casp-uk.net/casp-tools-checklists/.

EPPI-Centre. (2003). *Review guidelines for extracting data and quality assessing primary studies in educational research* (Guidelines Version 0.9.7). Social Science Research Unit.

Forbes, C. E., Schmader, T., & Allen, J. J. B. (2008). The role of devaluing and discounting in performance monitoring: A neurophysiological study of minorities under threat. *Social Cognitive and Affective Neuroscience*, *3*(3), 253–261. https://doi.org/10.1093/scan/nsn012

University of Glasgow. (n.d.). *Critical appraisal checklist for a systematic review* [Checklist]. Department of General Practice, University of Glasgow.

Wells, G., Shea, B., O'Connell, D., Robertson, J., Welch, V., Losos, M., & Tugwell, P. (2014). The newcastle-ottawa scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. *Ottawa Health Research Institute Web Site*, *7*.