Negative performance evaluation in the imposter phenomenon

Reference:
Gadsby Stephen, Hohwy Jakob.- Negative performance evaluation in the imposter phenomenon
Current psychology: research and reviews - ISSN 1936-4733 - New york, Springer, 43:10(2024), p. 9300-9308
Full text (Publisher's DOI): https://doi.org/10.1007/S12144-023-05030-0
To cite this reference: https://hdl.handle.net/10067/1991940151162165141

Negative performance evaluation in the imposter phenomenon

Stephen Gadsby [1,2] ✉ **and Jakob Hohwy** [2]

[1]Centre for Philosophical Psychology, Department of Philosophy, University of Antwerp; [2]Cognition and Philosophy Laboratory, School of Philosophical, Historical and International Studies, Monash University

Abstract

The imposter phenomenon (IP) is associated with a bias towards negative evaluation of one's own performances. This study employs an online problem-solving task to investigate this bias. Participants (graduate students from the UK, US, and Europe; n = 163) solved reasoning problems and subsequently evaluated their performance. Participants high in IP evaluated their performance more negatively than participants low in IP. This pattern was observed both during the task and after completion. It was also observed in objective assessments (estimates of accuracy) and comparative assessments (estimates of rank amongst participants). Performance evaluation bias was not associated with a bias in the selection of feedback about performance nor was it mediated by depression or self-esteem.

✉ **For correspondence:**
gadsby.st@gmail.com

Present address: Centre for Philosophical Psychology, Department of Philosophy, University of Antwerp, Antwerp, Belgium.

Data availability: The dataset and code for this project is available on the Open Science Framework (OSF) page: https://osf.io/3n96e/.

Funding: This study was funded by an Australian Government Research Training Program (RTP) Scholarship (S.G), a Monash University Postgraduate Publications Award (S.G.), a Fonds Wetenschappelijk Onderzoek (FWO) Postdoctoral Fellowship (1267022N) (S.G), and the Faculty of Arts at Monash University (J.H).

Competing interests: The authors declare no competing interests.

Introduction

In the imposter phenomenon (IP), successful and intelligent people believe that they are less competent than their peers, and fear being exposed as such (***Clance and Imes, 1978***). It has been most extensively researched using observational methods, which have uncovered various negative correlates of the condition (***Bravata et al., 2019***; ***Stone-Sabali et al., 2023***). For example, IP is associated with higher levels of anxiety and depression (***Cozzarelli and Major, 1990***; ***McGregor et al., 2008***) and negative work outcomes, such as work-related stress (***Rohrmann et al., 2016***) and burnout (***Sakulku and Alexander, 2011***). While IP was initially assumed to exclusively affect women (***Clance and Imes, 1978***), later research revealed that men also experience it, and several studies show no difference in prevalence between genders (***Bravata et al., 2019***).

Negative performance evaluation is a crucial feature of, and driving force behind, IP (***Clance, 1985***). People high in IP expect themselves to perform poorly in upcoming tasks (***Cozzarelli and Major, 1990***), and, after successfully completing tasks, misattribute their success to luck or hard work, rather than intelligence (***Ibrahim et al., 2022***). By misattributing their success, individuals high in IP maintain their belief that they are incompetent (***Clance, 1985***). By expecting themselves to perform poorly, they exacerbate their fear of failure (***Cozzarelli and Major, 1990***). They may also react to these expectations with harmful preparatory behaviours, such as procrastinating or overworking (***Cozzarelli and Major, 1990***). Given these negative consequences, various studies have employed problem solving tasks to investigate the relationship between IP and performance evaluation (***Badawy et al., 2018***; ***Brauer and Proyer, 2022***; ***Gadsby and Hohwy, 2022***; ***Ibrahim et al., 2022***; ***Thompson et al., 2000***). Nevertheless, there are several unexplored features of this relationship.

An important feature of negative performance evaluation amongst individuals high in IP relates to the time point at which it occurs. Negatively biased performance evaluation amongst high IP participants has been uncovered both before and after performance (***Cozzarelli and Major, 1990***; ***Gadsby and Hohwy, 2022***). To our knowledge, however, no studies have explored performance evaluation during tasks. If the negativity of those high in IP persists during task performance, then this may represent another way in which negative evaluation influences their performance, either distracting them or affecting their motivation while carrying out the task (***Norem, 2008***). More generally, discovering that negativity persists during task performance can help clinicians in understanding when their clients suffer from performance evaluation bias and, therefore, how to teach their clients to counteract such bias (***Zanchetta et al., 2020***).

Research into performance evaluation bias in IP has predominately focused on objective judgments related to performance, for example, whether success is attributed to intellect or hard work (***Thompson et al., 1998***) or what they estimate their result to be in an upcoming exam (***Cozzarelli and Major, 1990***; ***Leary et al., 2000***). However, these forms of performance evaluation are importantly distinct from the socially comparative type that the clinical literature has focused on, where those high in IP specifically downplay their intellect in relation to their peers (***Clance, 1985***). Some studies have uncovered socially comparative performance evaluation bias amongst participants high in IP (***Ibrahim et al., 2022***). However, none have explored the relationship between this bias and judgments of objective performance. There are two routes through which socially comparative misjudgements could occur. One could misjudge their comparative performance due to misjudging how well their peers performed. Alternatively, one could misjudge their comparative performance due to misjudging how well they have (objectively) performed. To distinguish between these hypotheses, studies must involve both objective and comparative performance evaluations.

A common observation from the clinical literature is that those high in IP pay inordinate attention to evidence of failure, while ignoring evidence of success (***Clance and O'Toole, 1987***). Researchers have yet to experimentally investigate this phenomenon and its relationship to negative performance evaluation. Nevertheless, there is a clear theoretical link between the two phenomena. If one was to predominately seek out and attend to negative evidence about their performances then their beliefs about such performances would become negatively biased (***Gadsby, 2022***). This biased searching of evidence thus offers a potential explanation for the biased performance evaluation associated with IP.

Finally, previous research has shown a link between IP, self-esteem, and depression, wherein IP negatively correlates with self-esteem and positively correlates with depression (***Cozzarelli and Major, 1990***; ***McGregor et al., 2008***). However, these traits have not been investigated in relationship to the link between IP and performance evaluation bias. High levels of depression and low levels of self-esteem are both associated with a bias towards negative self-evaluation (***Campbell and Fairey, 1985***; ***Kovacs and Beck, 1978***). This suggests that they may mediate the relationship between IP and performance evaluation bias.

Our study had four aims. The first aim was to test for performance evaluation bias during task performance. The second aim was to test for performance evaluation bias of a socially comparative nature and (if uncovered) test whether it was exclusively driven by misestimation of objective performance. The third aim was to test for an association between IP and search bias. The fourth aim was to explore the relationship between IP, performance evaluation bias, self-esteem, and depression.

In pursuit of these aims, we employed a paradigm wherein participants solved a set of reasoning problems and evaluated their own performance. After solving each problem, participants were required to rate their confidence in the accuracy of their solution. After solving the problem set, participants were allowed to select between one and four of their solutions to receive feedback on. After receiving this feedback, participants were required to estimate their total accuracy on the set of problems, as well as the percent of other participants that they outperformed.

Based on our study design, we calculated three measures of performance evaluation bias: Ongoing evaluation bias was based on participants' confidence in their solutions, during the task; Objective evaluation bias was based on participants' estimates of their total accuracy on the set of problems; Comparative evaluation bias was based on participants' estimates of the percent of other participants that they outperformed. We also calculated a measure of Search Bias, which was determined by the problems participants chose to receive feedback on. If participants opted for feedback on problems where they had lower confidence in their solutions, their feedback was more likely to be negative. Consequently, this was categorized as a negative search bias. Following previous research on the link between IP and performance evaluation (***Cozzarelli and Major, 1990***; ***Ferrari and Thompson, 2006***; ***Thompson et al., 1998***, ***2000***), we compared a group of participants with high levels of IP against a group with low levels.

Our first aim was to test for performance evaluation bias during task performance. Given prior findings of biased performance evaluation at other time points (retrospectively and prospectively), we hypothesised that during the task, participants with high levels of IP would demonstrate less confidence in their performance, compared to participants with low levels of IP (H1).

The second aim was to explore the relationship between IP and socially comparative performance evaluation bias. Given the emphasis of comparative performance evaluation bias in the clinical literature, and prior findings of objective evaluation bias, we hypothesised that participants with high levels of IP would demonstrate more negative performance evaluation compared to participants with low levels of IP, in both objective and comparative forms (H2).

The third aim was to test for an association between IP and search bias. Based on clinical reports of negatively biased selection of feedback amongst individuals with high levels of IP, we constructed our third hypothesis: IP would be associated with a negative search bias, such that participants with high levels of IP would select more negative feedback regarding their performance compared to participants with low levels of IP (H3).

The fourth aim was to explore the relationship between IP, depression, self-esteem, and performance evaluation bias. To achieve this, we assessed participants' levels of depression and self-esteem and examined the correlations between these variables and the three forms of performance evaluation bias. Consistent with prior research illustrating a link between IP, depression, and self-esteem, we hypothesised that IP would positively correlate with depression and negatively correlate with self-esteem (H4). Given the link between depression, self-esteem, and biased self-evaluation, we constructed two further hypotheses: depression would mediate the relationship between IP and performance evaluation bias (in all three forms) (H5) and self-esteem would mediate the relationship between IP and performance evaluation bias (in all three forms) (H6).

Methods & Materials

Participants

201 participants were recruited online, through Prolific (https://www.prolific.co/). We recruited participants between the ages of 18 and 65, who were currently enrolled in graduate studies (e.g., MA/MSc/MPhil/PhD) and residing in the United Kingdom (14%), the United States of America (3%), and continental Europe (83%). The European countries represented in the sample were Austria (n = 1), Belgium (n = 2), Bulgaria (n = 1), Czech Republic (n = 1), Estonia (n = 2), Finland (n = 1), France (n = 4), Germany (n = 8), Greece (n = 12), Hungary (n = 6), Italy (n = 47), the Netherlands (n = 6), Norway (n = 1), Poland (n = 15), Portugal (n = 15), Spain (n = 8), Sweden (n = 3), and Switzerland (n = 2). Using one-way ANOVAs, we tested for a main effect of geographical area (UK vs Continental Europe vs US) on IP, Ongoing Evaluation Bias, Objective Evaluation Bias, Comparative Evaluation Bias, and Search Bias. However, no significant main effects were found ($p > 0.05$). Consequently, all participants were aggregated into a single group for further analysis.

Graduate students were selected as participants due to their tendency to exhibit high levels of IP and engage in social comparisons with peers. Given our focus on socially comparative performance evaluation, we restricted the geographic location of participants to increase the chance that the participants would view each other as peers and competitors. All participants reported fluency in English and no current or prior diagnosis of a neurological or psychiatric condition.

Participants received a base rate of £1.70 as compensation for their participation. Additionally, a bonus incentive was provided to encourage effort. Each participant had two of their answers randomly selected, and they were paid £0.80 for each correct answer. Taking into account the bonuses, the total payment per participant ranged between £1.70 and £3.30.

A total of 5 participants were excluded due to failing one of two effort measures: either scoring less than 3 out of 16 questions correctly or failing a question designed to assess effort ("What is the sixth month of the year?"). A further 31 participants were excluded for failing one of two measures of general interest, either choosing not to see their final score and the set of answers (*n* = 13) or selecting four answers in a row (e.g., 1, 2, 3, 4) when selecting feedback (*n* = 18). The rationale behind these exclusions was to minimize low-quality data commonly associated with online studies, even though some usable data may have been lost in the process. Ultimately, the final sample consisted of 163 participants (93 females, 70 males; *M* age = 25.4; *SD* = 4.4). The (pre-exclusion and post-exclusion) data sets for this project are available on the Open Science Framework page (https://osf.io/3n96e/).

Ethics

This study was approved by Monash University Human Research Ethics Committee (MUHREC Project ID: 25939). Participants were informed about the design and purpose of the study and provided informed consent before taking part.

Questionnaires

To measure IP, we employed the Clance Imposter Phenomenon Scale (CIPS) (***Clance, 1985***). The CIPS is the most commonly employed questionnaire for measuring IP and exhibits the strongest validity and reliability (α = .91) (***Holmes et al., 1993***). It contains 20 statements related to the IP construct e.g., "I can give the impression that I'm more competent than I really am," "At times, I feel my success was due to some kind of luck." Participants report their agreement with these statements on a five-point Likert scale ranging from 1 = *not at all true* to 5 = *very true*.

To measure self-esteem, we employed the Rosenberg Self-Esteem Scale (RSS), a widely used measure of global self-esteem (α = .88) (***Rosenberg, 1965***). It contains 10 items related to general feelings about oneself, e.g. "On the whole, I am satisfied with myself," "I certainly feel useless at times." Participants report their agreement with these statements on a four-point Likert scale ranging from 1 = *Strongly Agree* to 4 = *Strongly Disagree*.

To measure depression, we employed the Beck Depression Inventory (BDI-II), a commonly used scale for measuring depression (α = .93) (***Beck et al., 1996***). It contains 21 groups of statements designed to assess the severity of various symptoms of depression felt by participants during the past two weeks, including sadness, loss of interest in activities, changes in appetite or sleep, feelings of worthlessness or guilt, and suicidal thoughts. Participants are required to select a single statement related to each category, e.g., "I do not feel sad," "I am so sad or unhappy that I can't stand it."

Procedure

Data was collected through Qualtrics (http://www.qualtrics.com/). After providing informed consent, participants filled out a demographic questionnaire, completed the main task, and completed three questionnaires (CIPS, RSS, BDI-II, provided in that order). They were then given the opportunity to return directly to Prolific or view their total score and the complete set of answers.

In the main task, participants were required to complete 16 reasoning problems ("designed to test your intelligence"). Problems were taken from the international cognitive reasoning ability resource (***Condon and Revelle, 2014***) and the test of figural analogies (***Blum et al., 2016***). These included verbal reasoning, letter and number series, three-dimensional rotation, matrix reasoning, and figural analogy problems (for the full problem set, refer to the supplementary material).

Participants were allowed two minutes to solve each problem. After providing a solution, they were required to report their confidence in it ("How confident are you that you answered correctly?") on a scale from 0 ("not confident at all") to 100 ("completely confident").

After solving the entire problem set, participants were shown each of the problems, along with the confidence that they reported in their solutions. They were asked to select between one and four of the problems to receive feedback on: "You may now check whether some of your answers were correct. You will only be informed about whether each answer was correct, you will not be given the answers to the questions."

After receiving feedback for the problems that they selected, they were asked to evaluate their performance. First, they estimated how many problems (0–16) they had solved correctly. Second, they estimated the percent of other participants (presented as graduate students in the UK, US, and Europe) they outperformed (0%–100%).

Data Preparation

CIPS (α = .902), BDI-II (α = .901), and RSS (α = .881) scores were calculated by extracting total scores from their respective questionnaires. Participants' total scores on the reasoning problems were recorded, as well as their self-reported confidence in each solution and their retrospective performance evaluations. Four dependent variables were calculated based on these scores:

Ongoing Evaluation Bias was calculated by subtracting a participant's score (%) from the mean of their confidence ratings (obtained after each reasoning problem). Negative values represent a negative bias (being less confident in one's performance than warranted), while positive values represent a positive bias (being more confident than warranted);

Objective Evaluation Bias was calculated by subtracting a participant's score (%) on the reasoning problems from their estimated score (their estimate of how many problems they solved, converted into %). Negative values indicate a negative bias (underestimating one's own score), while positive values indicate a positive bias (overestimating it);

Comparative Evaluation Bias was calculated by subtracting the percent of the sample that a participant outperformed from the percent that they estimated themselves to have outperformed (see the OSF project for the syntax used to calculate rank). Negative values indicate a negative bias (underestimating one's comparative performance), while positive values indicate a positive bias (overestimating it);

Search Bias was calculated by subtracting the mean confidence (%) of a participant's searched answers from their average confidence (%). A negative value represents a negative bias (participants choosing feedback for solutions that they felt less confident in, i.e., believed were more likely to be inaccurate). A positive value represents a positive bias (participants choosing feedback for solutions that they felt more confident in, i.e., believed were more likely to be accurate).

Analysis

Analysis was conducted using IBM SPSS Statistics (27.0.0.0) and JASP (JASP Team, 2020). To test the first two hypotheses, between-group analyses were conducted which compared participants who scored high in IP against those who scored low in IP on all three measures of performance evaluation bias (H1 & H2) and Search Bias (H3). We split the data set into two groups, based on CIPS scores: the top-third of participants (high-IP group; n = 54; range = 70–91; M = 79.37, SD = 6.54) and the bottom-third (low-IP group; n = 54; range = 26–58; M = 49.91, SD = 6.67). These ranges are consistent with the suggested CIPS cut-off between impostors and non-impostors (61) (***Holmes et al., 1993***).

To test for between-group differences in performance evaluation bias, ANCOVAs were conducted on Ongoing Evaluation Bias and Comparative Evaluation Bias, with Group as the independent variable and (objective) Score as the covariate. For Objective Evaluation Bias, a general linear model was conducted, using Group as the independent variable and Score as the covariate. An interaction between Group and Score was found in this model and included. Normality was assessed using Shapiro-Wilk tests, and homogeneity of variance was assessed using Levene's tests. In cases of violation, the results of (non-parametric) Kruskal-Wallis (*KS*) tests were reported. Estimated marginal means of each group were compared against zero, and Bonferroni adjusted p-values were reported.

To test for between-group differences in CIPS, Search Bias, and Score, means of the high-IP and low-IP groups were compared. Shapiro-Wilk tests were used to assess normality, and Levene's tests were used to assess homogeneity of variance. Student t-tests were used for normally distributed data, and Mann-Whitney U-tests were used for non-normal data.

To further explore H1 and H2 with reference to the complete sample, partial correlations were computed between CIPS and Ongoing Evaluation Bias, Objective Evaluation Bias, and Comparative Evaluation Bias, controlling for Score. To further explore H3, a correlation between CIPS and Search Bias was computed.

To test whether CIPS positively correlates with BDI-II and negatively correlates with RSS (H4), correlations between the total scores on each questionnaire were computed. To test whether depression or self-esteem mediates the relationship between CIPS and performance evaluation bias (H5 & H6), structural equation modeling was used to test a regression model with CIPS as the predictor, BDI-II and RSS as mediators, Ongoing Evaluation Bias, Objective Evaluation Bias, and Comparative Evaluation Bias as outcomes, and Score as a background confounder. Confidence intervals and standard errors were computed using a bias-corrected bootstrap method.

Where appropriate, Bayes factors (*BF01*) were calculated to assess the strength of evidence in favour of the null hypothesis. The default priors set by JASP (Cauchy prior, $r = 0.707$) were used and evidence for the null hypothesis provided by each *BF01* was interpreted using the cut-offs suggested by Lee & Wagenmakers (2014): 1–3 = anecdotal evidence, 3–10 = moderate evidence, 10–30 = strong evidence.

Results

Between-group analysis

For Ongoing Evaluation Bias, the ANCOVA revealed a significant main effect of Group $F(1,105) = 9.411$, $p = .003$, $\eta 2p = .082$, *KS*: $p = .014$, indicating that the high-IP group (*EMM* = -1.735; *SE* = 1.882) reported more negatively biased confidence (during the task) compared to the low-IP group (*EMM* = 6.441; *SE* = 1.882) (Figure 1). The low-IP group showed a significant difference from 0 ($t = 3.422$, $p = .002$), indicating a positive bias, while the high-IP group did not show a significant difference from 0 ($t = -.922$, $p = .718$).

For Objective Evaluation Bias, the general linear model indicated a significant main effect of Group $F(1,104) = 9.914$, $p = .002$, $\eta 2p = .087$, *KS*: $p = .003$, indicating that the high-IP group (*EMM* = -8.148; *SE* = 1.709) provided more negatively biased estimates of objective performance compared to the low-IP group (*EMM* = -0.137; *SE* = 1.707) (Figure 2). The high-IP group exhibited a significant difference from 0 ($t = -4.768$, $p < .001$), indicating a negative bias, while the low-IP group did not show a significant difference from 0 ($t = -.080$, $p = 1.000$). Additionally, there was a significant interaction between Group and Score $F(1,104) = 5.779$, $p = .018$, $\eta 2p = .053$ (refer to the supplementary materials for the plot).

For Comparative Evaluation Bias, the ANCOVA revealed a significant main effect of Group $F(1,105) = 9.532$, $p = .003$, $\eta 2p = .083$, *KS*: $p = .015$, indicating that participants in the high-IP (*EMM* = -12.091; *SE* = 3.053) provided more negatively biased estimates of comparative performance compared to the low-IP group (*EMM* = 1.256; *SE* = 3.053) (see Figure 3.). The high-IP group exhibited a significant difference from 0 ($t = -3.960$, $p < .001$), indicating a negative bias, while the low-IP group did not show a significant difference from 0 ($t = .411$, $p = 1.000$). These effects remained significant even after including Objective Evaluation Bias as a covariate in the model (refer to the supplementary materials), indicating that the observed effects are not solely driven by differences in Objective

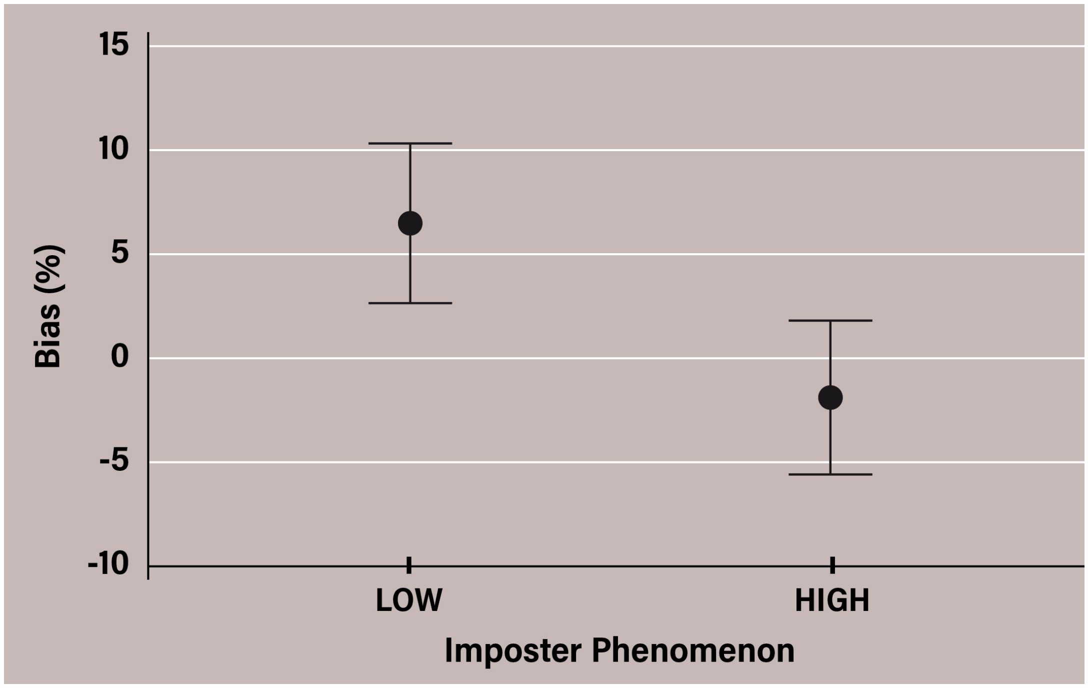

Figure 1. Estimated marginal means for Ongoing Evaluation Bias. Error bars represent 95% CI.

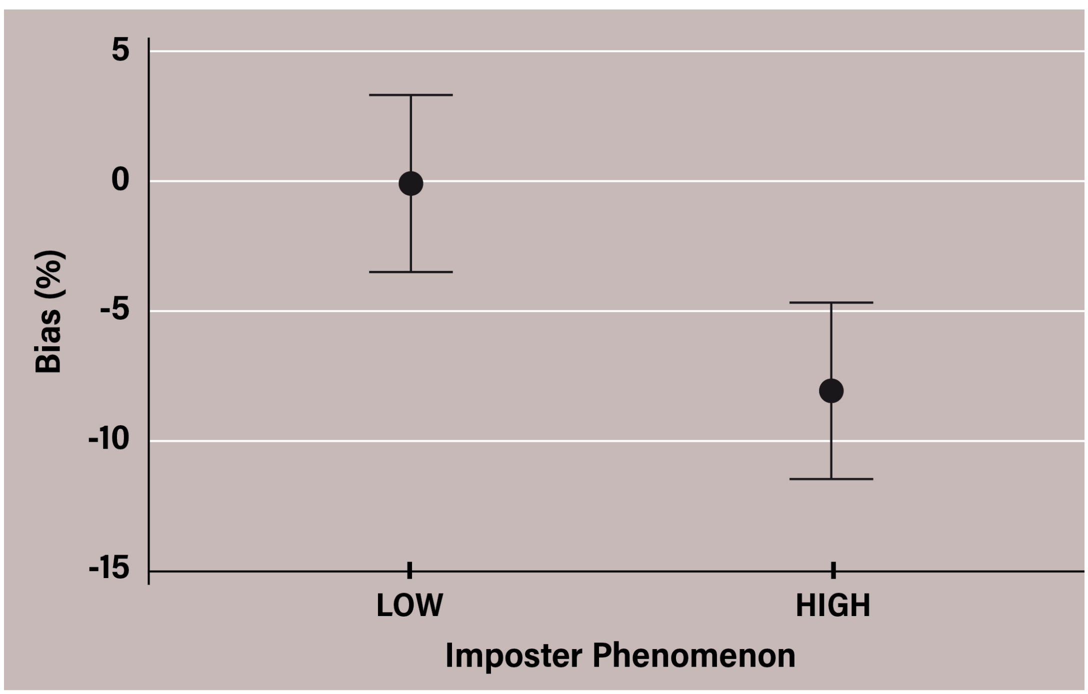

Figure 2. Estimated marginal means for Objective Evaluation Bias. Error bars represent 95% CI.

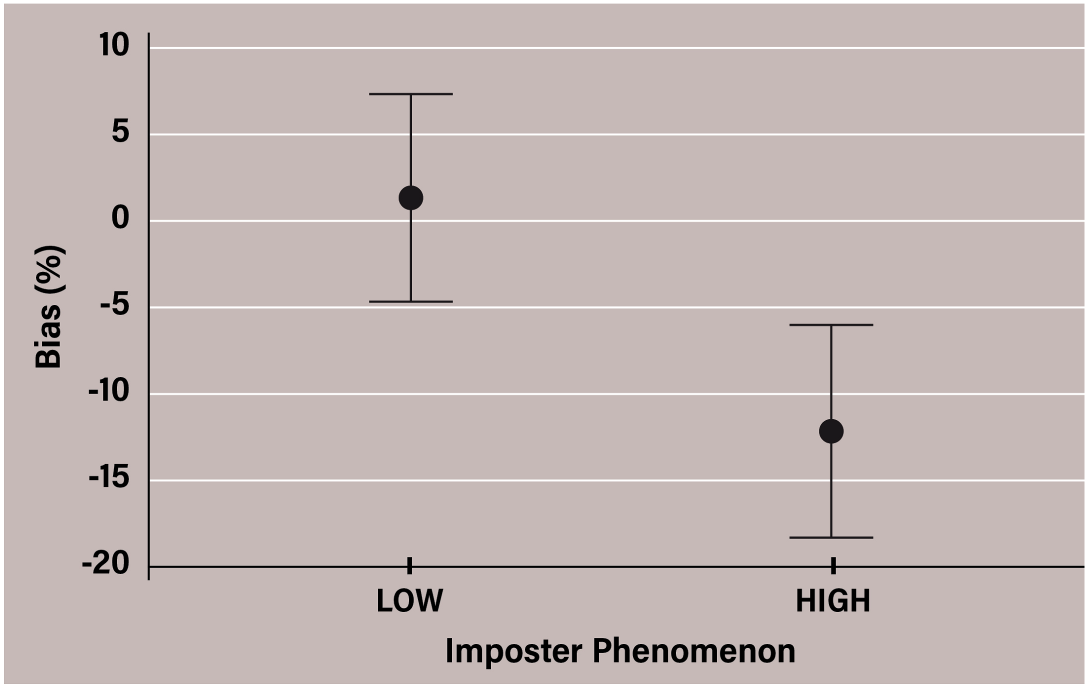

Figure 3. Estimated marginal means for Comparative Evaluation Bias. Error bars represent 95% CI.

Evaluation Bias.

A Mann-Whitney U-test revealed a significant difference in CIPS scores between the high-IP and low-IP groups ($U = 0$, $p < .001$, $d = 1.00$). Independent sample t-tests revealed no significant difference in Search Bias between the High-IP ($M = -12.64$, $SD = 24.40$) and Low-IP ($M = -13$, $SD = 23.6$) groups, $t(106) = -.08$, $p = .938$, $d = -0.015$, $BF01 = 4.894$. Similarly, there was no significant difference in Score between the high-IP ($M = -61.11$, $SD = 18.60$) and Low-IP ($M = -63.43$, $SD = 14.69$) groups, $t(106) = -0.718$, $p = .475$, $d = -0.138$, $BF01 = 3.897$. Bayesian analysis uncovered moderate evidence for both null hypotheses. Detailed descriptive statistics for each group can be found in the supplementary material.

IP, self-esteem, depression, and performance evaluation bias

Analysis of the complete data set uncovered a positive correlation between CIPS and BDI-II ($r = .57$, $p < .001$) and a negative correlation between CIPS and RSS ($r = -.67$, $p < .001$). Additionally, females ($n = 70$, $M = 68.83$) scored higher on the CIPS ($p = .001$) compared to males ($n = 93$, $M = 61.98$).

Consistent with the between-group findings, CIPS negatively correlated with Ongoing Evaluation Bias ($r = -.23$, $p = .002$), Objective Evaluation Bias ($r = -.21$, $p = .004$) and Comparative Evaluation Bias ($r = -.39$, $p = .001$), while controlling for Score. These correlations indicate that higher IP scores are associated with more negatively biased performance evaluation. The correlation between CIPS and Search Bias was not significant ($r = -.03$, $p = .377$, $BF01 = 13.247$). All significant results passed a bonferroni-corrected threshold ($p < .007$).

Structural equation modelling uncovered no significant mediating effects of BDI-II on the relationship between CIPS and Ongoing Evaluation Bias ($p = .100$), Objective Evaluation Bias ($p = .407$), or Comparative Evaluation Bias ($p = .568$). Similarly, there were no significant mediating effects of RSS on the relationship between CIPS and Ongoing Evaluation Bias ($p = .419$), Objective Evaluation Bias ($p = .587$), or Comparative Evaluation Bias ($p = .501$). For further information, see supplementary material.

Discussion

This study explored the relationship between IP and negative performance evaluation. In an online setting, participants solved a set of reasoning problems while evaluating each of their solutions, and, after receiving some self-selected feedback, estimated their own performance. The study tested five hypotheses. H1: the high-IP group would be more negative than the low-IP group, in Ongoing Evaluation Bias; H2: the high-IP group would be more negative than the low-IP group, in both Objective and Comparative Evaluation Bias; H3: the high-IP group would exhibit a more negative search bias than the low-IP group; H4: IP would positively correlate with depression and negatively correlate with self-esteem; H5: depression would mediate the relationship between IP and performance evaluation bias; H6: self-esteem would mediate the relationship between IP and performance evaluation bias.

The findings supported both H1 and H2. The high-IP group was more negative in their performance evaluation than the low-IP group, despite performing equally well. This occurred in relation to each form of performance evaluation. Consistent with these between-group findings, we discovered that, in the complete sample, IP negatively correlated with all three forms of bias. In the case of Comparative Evaluation Bias, this effect remained when we controlled for differences in Objective Evaluation Bias, suggesting that differences in comparative performance evaluation between high-IP and low-IP groups were not exclusively driven by misestimation of (objective) performance. Put differently, participants in the high-IP group did not underestimate how well they performed compared to others only because they underestimated their own score, they also overestimated how well other participants performed.

Our findings did not support H3. There was no significant difference in Search Bias between high-IP and low-IP groups, nor did IP correlate with Search Bias. Both null hypotheses were supported by Bayesian analysis. Instead, our results suggest that, on average, all participants selected feedback for questions that they were moderately unsure of. This suggests that participants' curiosity regarding the accuracy of their solutions may have guided their selection of feedback, potentially outweighing alternative motivations to engage in biased selection. Consequently, bias in the selection of feedback may be associated with IP, though the present study was unable to discover it.

Regarding H4, we uncovered a positive correlation between IP and depression and a negative correlation between IP and self-esteem. These findings are consistent with previous research (***Cozzarelli and Major, 1990***; ***McGregor et al., 2008***). However, we uncovered no evidence in support of H5 or H6, as neither depression nor self-esteem mediated the relationship between IP and (any form of) performance evaluation bias.

Theoretical Contributions and Future Directions

Prior evidence suggests that the performance evaluation biases associated with IP disappear under conditions of anonymity (***Leary et al., 2000***). In contrast, our findings suggest that even in anonymous online settings, participants high in IP are biased towards negative performance evaluation. Consequently, these results contribute to research illustrating the utility of anonymous online problem-solving tasks for studying IP (***Gadsby, 2022***; ***Ibrahim et al., 2022***). Our findings further indicate that participants high in IP do not only misestimate how well they will perform (***Cozzarelli and Major, 1990***) and misattribute the cause of their success (***Ibrahim et al., 2022***), but also misestimate how well they have performed, in objective and comparative terms. Future research should focus more closely on these forms of biased performance evaluation, and how they might contribute to IP.

A novel finding from this study relates to the time point at which negative performance evaluation occurs. We discovered that high IP participants were less confident in their performance than low IP participants during the task. Researchers have suggested different negative outcomes associated with prospective and retroactive negative performance evaluations but have not discussed the effects of low confidence during the task itself (***Cozzarelli and Major, 1990***). Low confidence during tasks may affect participants' motivation to exert effort, either reducing their effort (because the task is seen as futile) or increasing their effort (to overcome their perceived shortcomings) (***Gadsby, 2022***). This represents another important avenue for researchers and clinicians to explore.

The absence of a between-group difference in search bias speaks to the strength of the bias in retrospective performance estimates, amongst high IP participants. Participants were only required to answer 16 questions and could receive feedback for up to four of them. It was therefore relatively simple to estimate one's score, which, on average, low IP participants were able to do accurately. Nevertheless, despite selecting the kind of feedback that ought to have facilitated an accurate estimate, high IP participants still provided negatively biased estimates. Negative performance evaluation in our study appears to have been underpinned by strategies distinct from biased selection of feedback. Future research should focus on uncovering these strategies. For example, those high in IP may have exhibited biased memory recall, disproportionately remembering problems that they could not solve (***Zimmermann, 2020***). This suggestion is consistent with clinical descriptions of individuals with high levels of IP selectively remembering negative experiences (***Clance, 1985***). Future research should explore alternative strategies that might underpin the performance evaluation bias associated with the IP.

Limitations

This study has several limitations. First, participants were aware that they were taking part in a study about the IP, which may have affected their behaviour. Second, we included participants from a broad geographic area (UK, US, and Europe), which may have obscured important cultural differences in the way the IP presents itself. Finally, we administered the questionnaires at the end of the study (after the main task), thus we cannot rule out an effect of doing the intelligence test on participants' answers to the questions.

Conclusion

Our results showed that individuals high in IP exhibit a bias towards negative performance evaluation, in an online setting. This bias extends to different forms of performance evaluation (objective and comparative) and different time points (during a task and retrospectively). However, this bias was not associated with a bias in the selection of feedback on performance. Future research should evaluate the nature and consequences of low confidence during task performance. It should also explore alternative mechanisms underpinning negative performance evaluation, such as biased memory.

Acknowledgment

Thanks to Manja Engel and members of the Monash Cognition and Philosophy lab for useful feedback. Thanks to Andrew Gadsby for figure design. This preprint was created using the LaPreprint template (https://github.com/roaldarbol/lapreprint) by Mikkel Roald-Arbøl.

Supplementary

Supplementary materials are available from: https://osf.io/w439a

References

Badawy RL, Gazdag BA, Bentley JR, Brouer RL. Are all impostors created equal? Exploring gender differences in the impostor phenomenon-performance link. Personality and Individual Differences. 2018; 131:156–163. doi: 10.1016/j.paid.2018.04.044.

Beck AT, Steer RA, Brown GK. Beck Depression Inventory Manual (2nd ed.). San Antonio, TX: The Psychological Corporation Harcourt Brace & Company; 1996.

Blum D, Holling H, Galibert MS, Forthmann B. Task difficulty prediction of figural analogies. Intelligence. 2016; 56:72–81. doi: 10.1016/j.intell.2016.03.001.

Brauer K, Proyer RT. The Impostor Phenomenon and causal attributions of positive feedback on intelligence tests. Personality and Individual Differences. 2022; 194:111663. doi: 10.1016/j.paid.2022.111663.

Bravata DM, Watts SA, Keefer AL, Madhusudhan DK, Taylor KT, Clark DM, Nelson RS, Cokley KO, Hagg HK. Prevalence, Predictors, and Treatment of Impostor Syndrome: a Systematic Review. Journal of General Internal Medicine. 2019; p. 1–24.

Campbell JD, Fairey P. Effects of self-esteem, hypothetical explanations, and verbalization of expectancies on future performance. Journal of Personality and Social Psychology. 1985; 48(5):1097.

Clance PR. The impostor phenomenon: Overcoming the fear that haunts your success (kindle version ed.). Peachtree; 1985.

Clance PR, Imes SA. The imposter phenomenon in high achieving women: Dynamics and therapeutic intervention. Psychotherapy: Theory, Research & Practice. 1978; 15(3):241. doi: 10.1037/h0086006.

Clance PR, O'Toole MA. The imposter phenomenon: An internal barrier to empowerment and achievement. Women & Therapy. 1987; 6(3):51–64. doi: 10.1300/J015V06N03_05.

Condon DM, Revelle W. The International Cognitive Ability Resource: Development and initial validation of a public-domain measure. Intelligence. 2014; 43:52–64. doi: 10.1016/j.intell.2014.01.004.

Cozzarelli C, Major B. Exploring the validity of the impostor phenomenon. Journal of Social and Clinical Psychology. 1990; 9(4):401–417. doi: 10.1521/jscp.1990.9.4.401.

Ferrari JR, Thompson T. Impostor fears: Links with self-presentational concerns and self-handicapping behaviours. Personality and Individual Differences. 2006; 40(2):341–352.

Gadsby S. Imposter Syndrome and Self-Deception. Australasian Journal of Philosophy. 2022; 100(2):247–261. doi: 10.1080/00048402.2021.1874445.

Gadsby S, Hohwy J. Incentivising accuracy reduces bias in the imposter phenomenon. Current Psychology. 2022; p. 1–9. doi: 10.1007/s12144-022-03878-2.

Holmes SW, Kertay L, Adamson LB, Holland C, Clance PR. Measuring the impostor phenomenon: A comparison of Clance's IP Scale and Harvey's IP Scale. Journal of Personality Assessment. 1993; 60(1):48–59. doi: 10.1207/s15327752jpa6001_3.

Ibrahim F, Göddertz D, Herzberg PY. An experimental study of the non-self-serving attributional bias within the impostor phenomenon and its relation to the fixed mindset. Current Psychology. 2022; p. 1–10. doi: 10.1007/s12144-022-03486-0.

Kovacs M, Beck AT. Maladaptive cognitive structures in depression. American Journal of Psychiatry. 1978; 135(5):525–533.

Leary MR, Patton KM, Orlando AE, Wagoner Funk W. The impostor phenomenon: Self-perceptions, reflected appraisals, and interpersonal strategies. Journal of Personality. 2000; 68(4):725–756.

McGregor LN, Gee DE, Posey KE. I feel like a fraud and it depresses me: The relation between the imposter phenomenon and depression. Social Behavior and Personality: An International Journal. 2008; 36(1):43–48.

Norem J. The positive power of negative thinking. Basic Books; 2008.

Rohrmann S, Bechtoldt MN, Leonhardt M. Validation of the impostor phenomenon among managers. Frontiers in Psychology. 2016; 7:821.

Rosenberg M. Society and the adolescent self-image. Princeton University Press; 1965.

Sakulku J, Alexander J. The impostor phenomenon. The Journal of Behavioral Science. 2011; 6(1):75–97.

Stone-Sabali S, Bernard DL, Mills KJ, Osborn PR. Mapping the evolution of the impostor phenomenon research: A bibliometric analysis. Current Psychology. 2023; p. 1–13.

Thompson T, Davis H, Davidson J. Attributional and affective responses of impostors to academic success and failure outcomes. Personality and Individual Differences. 1998; 25(2):381–396. doi: 10.1016/s0191-8869(98)00065-8.

Thompson T, Foreman P, Martin F. Impostor fears and perfectionistic concern over mistakes. Personality and Individual Differences. 2000; 29(4):629–647. doi: 10.1016/s0191-8869(99)00218-4.

Zanchetta M, Junker S, Wolf AM, Traut-Mattausch E. "Overcoming the fear that haunts your success"–the effectiveness of interventions for reducing the impostor phenomenon. Frontiers in Psychology. 2020; 11:405.

Zimmermann F. The dynamics of motivated beliefs. American Economic Review. 2020; 110(2):337–361.