

Modeling Social Anxiety, Attentional Bias, and Metacognitive Beliefs with Supervised Learning Algorithms

Josep. Bosch Borrás^{1*}, Zarley. Mett², Marlenne. Reynoso-Calderon³


¹ Department of Educational Psychology, School of Information Sciences, University of Illinois at Urbana-Champaign, Champaign, IL, USA

² Department of Psychology, Clark University, Worcester, Massachusetts, USA



³ Department of Psychology, University of Toronto, Canada

* Corresponding author email address: boschborras-josep@gmail.com

Editor

Mehdi Rostami
Department of Psychology and
Counseling, KMAN Research
Institute, Richmond Hill, Ontario,
Canada
mehdirostami@kmanresce.ca

Reviewers

Reviewer 1: Daniela Gottschlich
Faculty of Health Sciences, Simon Fraser University, Vancouver, BC, Canada.
Email: danielagottschlich@wayne.edu
Reviewer 2: Sabine. Kraus
Department of Psychology, University of Nevada, Las Vegas, NV, USA.
Email: sabine_kraus10@gmail.com

1. Round 1

1.1. Reviewer 1

Reviewer:

The use of advanced supervised machine learning methods, particularly the identification of Gradient Boosting Regressor as the best-performing model, represents a notable methodological contribution. The inclusion of SHAP-based interpretability is especially commendable, because it allows the authors not only to optimize prediction but also to examine the relative contribution of individual predictors and their interaction patterns. This substantially strengthens the manuscript's analytical sophistication and makes the machine learning approach more transparent and psychologically meaningful.

The finding that negative metacognitive beliefs regarding the uncontrollability and danger of thoughts emerged as the strongest predictor of social anxiety severity is theoretically compelling and potentially impactful. This result is consistent with metacognitive models of anxiety and suggests that the manuscript makes a meaningful contribution to the literature by positioning these beliefs as a central computational driver of symptom severity. The additional finding that attentional bias was the second most important predictor further supports the value of integrating top-down and bottom-up cognitive processes in the study of social anxiety.

Concerns about generalizability are also substantial. The reliance on a web-based crowdsourcing platform and the restriction to adult participants residing in the United States may introduce sampling biases and limit the applicability of the findings across clinical, developmental, and cultural contexts. Because social anxiety, attentional processing, and metacognitive beliefs may vary meaningfully across age groups and sociocultural environments, the authors should temper broad claims and emphasize that external validation in clinically diagnosed, culturally diverse, and younger populations is necessary before strong general conclusions can be drawn.

Authors revised and uploaded the document.

1.2. Reviewer 2

Reviewer:

One of the most interesting aspects of the paper is the reported non-linear interaction between attentional bias and negative metacognitive beliefs, whereby elevated attentional bias appears to have a substantially greater predictive effect among individuals with severe negative metacognitive beliefs. This is an important insight that goes beyond additive models and may help explain why some individuals exhibit much more severe social anxiety despite apparently similar cognitive vulnerabilities. I encourage the authors to expand the explanation of this interaction in more accessible conceptual language so that its clinical and theoretical significance is fully appreciated by readers outside machine learning specialties.

Despite these strengths, the manuscript would benefit from much clearer and more complete reporting of methodological details. Based on the retrieved text, key information such as the exact sample size, the names of the psychometric instruments used, the structure of the attentional bias task, preprocessing procedures, model training workflow, hyperparameter tuning strategy, and validation approach is either not sufficiently visible or not explicitly stated. These details are essential for reproducibility, for evaluating the robustness of the findings, and for allowing readers to judge whether the predictive performance may reflect overfitting or sound model generalization.

The reporting of model performance should be substantially improved. Although the manuscript states that the Gradient Boosting Regressor explained an “exceptionally high” proportion of variance, the retrieved text does not clearly show the exact coefficient of determination or other performance metrics such as RMSE, MAE, cross-validation scores, or confidence intervals. For a machine learning paper, especially one making strong claims about superiority over traditional linear regression, precise and transparent reporting of all evaluation metrics is indispensable. Without those details, the magnitude and practical significance of the reported predictive advantage remain difficult to assess rigorously.

The cross-sectional design is a major limitation and should be discussed even more cautiously in relation to the manuscript’s theoretical interpretations. While the authors acknowledge that causality cannot be established, some of the wording still risks implying developmental or mechanistic directionality that the present data cannot support. The manuscript would be stronger if it more explicitly distinguished predictive association from causal explanation and avoided language that could be read as demonstrating the etiological primacy of metacognitive beliefs or attentional bias in the absence of longitudinal or experimental data.

Authors revised and uploaded the document.

2. Revised

Editor’s decision after revisions: Accepted.

Editor in Chief’s decision: Accepted.