The accuracy of individual implicit measure scores

Jamie Cummins1 & Ian Hussey2

1Ghent University, Belgium

2Ruhr University Bochum, Germany

**Author note**

JC, Department of Experimental Clinical and Health Psychology, Ghent University, & IH, Faculty of Psychology, Ruhr University Bochum. JC was conducted with the support of Grant BOF16/MET\_V/002 to Jan De Houwer. IH was supported by Ghent University grant 01P05517 and the META-REP Priority Program of the German Research Foundation (#464488178). Correspondence should be sent to [jamie.cummins@UGent.be](mailto:jamie.cummins@UGent.be).

**Introduction**

The use of implicit measures has become part of the bedrock of social psychology’s efforts to investigate mental processes assumed to mediate human behavior. The narrative for how these measures came about is well-repeated in the introduction section of many such articles: social psychologists were interested in

**Method**

This study uses openly available data acquired from Project Implicit, originally collected by Bar-Anan and Nosek (2014; OSF LINK). Data, processing scripts, and analysis scripts can be found on the Open Science Framework repository for this project (XXXX).

**Participants**

Detailed information regarding the collection of these data can be found in Bar-Anan and Nosek (2014). The data consisted of a total of 21060 participants in total who completed, and met the screening criteria for, at least one measure in the overall study. Within this, 6902 participants completed the IAT, 7238 completed the AMP, 6039 completed the BIAT, 6795 completed the EPT, 6529 completed the GNAT, and 6626 completed the ST-IAT. These completions were divided approximately evenly across the three domains of race, politics, and self-esteem.

**Measures**

Herein we briefly describe each of the measurement procedures used within the study. For more detailed descriptions, see Bar-Anan and Nosek (2014).

**Implicit Association Test.**

**Brief Implicit Association Test.**

**Single-Target Implicit Association Test.**

**Affect Misattribution Procedure.**

**Go-No Go Association Task.**

**Evaluative Priming Task.**

**Procedure**

For all participants, each session lasted approximately 15 minutes. Within each session, participants were presented with two “long-duration” and two “short-duration” measures (the implicit measures were divided across these two categories; see Bar-Anan and Nosek, 2014). There were no constraints on participants in terms of the measures they would receive beyond the fact that the same exact measure/domain combination could not be presented twice in one session.

**Results**

**Analytic Strategy**

The first key aspect of our analytic strategy was the estimation of confidence intervals around each participants scores on a given implicit measure. We follow the strategy originally deployed by Hussey (2021); namely the use of the BCA method with 5000 resamples. The implicit measures compared here have typically use different methods and metrics for scoring. The IAT, ST-IAT, and B-IAT tend to use the *D* score based on response times (REF); the AMP tends to use proportion of prime-consistent evaluative responses; the GNAT and EPT tend to be scored based on differential response latencies (although the GNAT can also be scored based on accuracy differentials). These different methods of scoring, and the corresponding differences in error variances associated with them, make direct comparisons between the measures somewhat difficult. As such, we opted to instead score every measure using the same analytic method: namely, using probabilistic index (PI) scores. PI scores have previously been used as a method of scoring several implicit measures (De Scryver et al., 2018; Cummins et al., 2021) and tend to exhibit superior psychometric properties to other methods of scoring, particularly in relation to their robustness to outliers (REF). In short, PI

**Results**

**Discussion**