

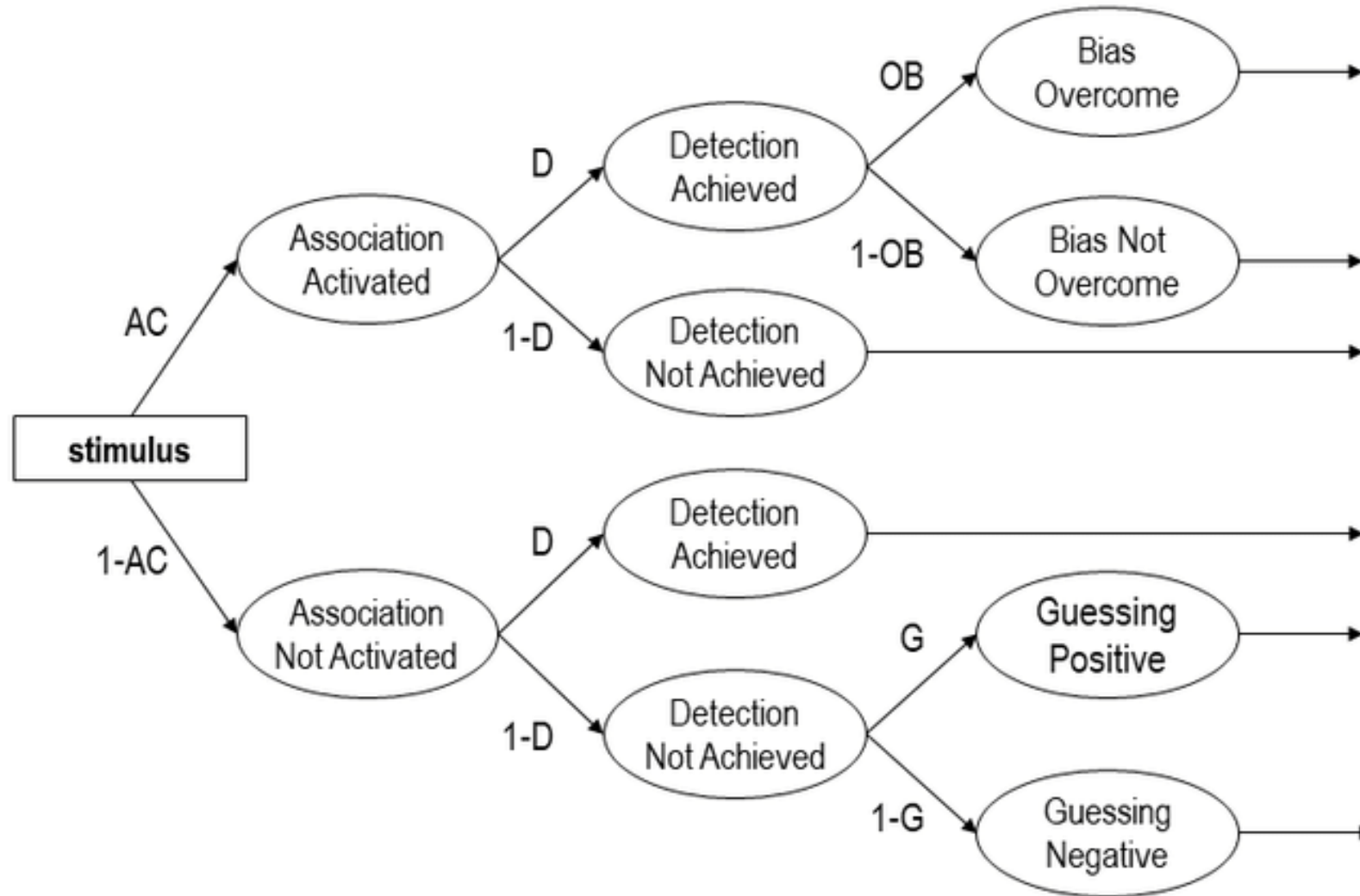
A Cognitive Atlas of Bias

We apply the Quadruple Process Model (Quad Model) to a large collection of data from the Project Implicit website

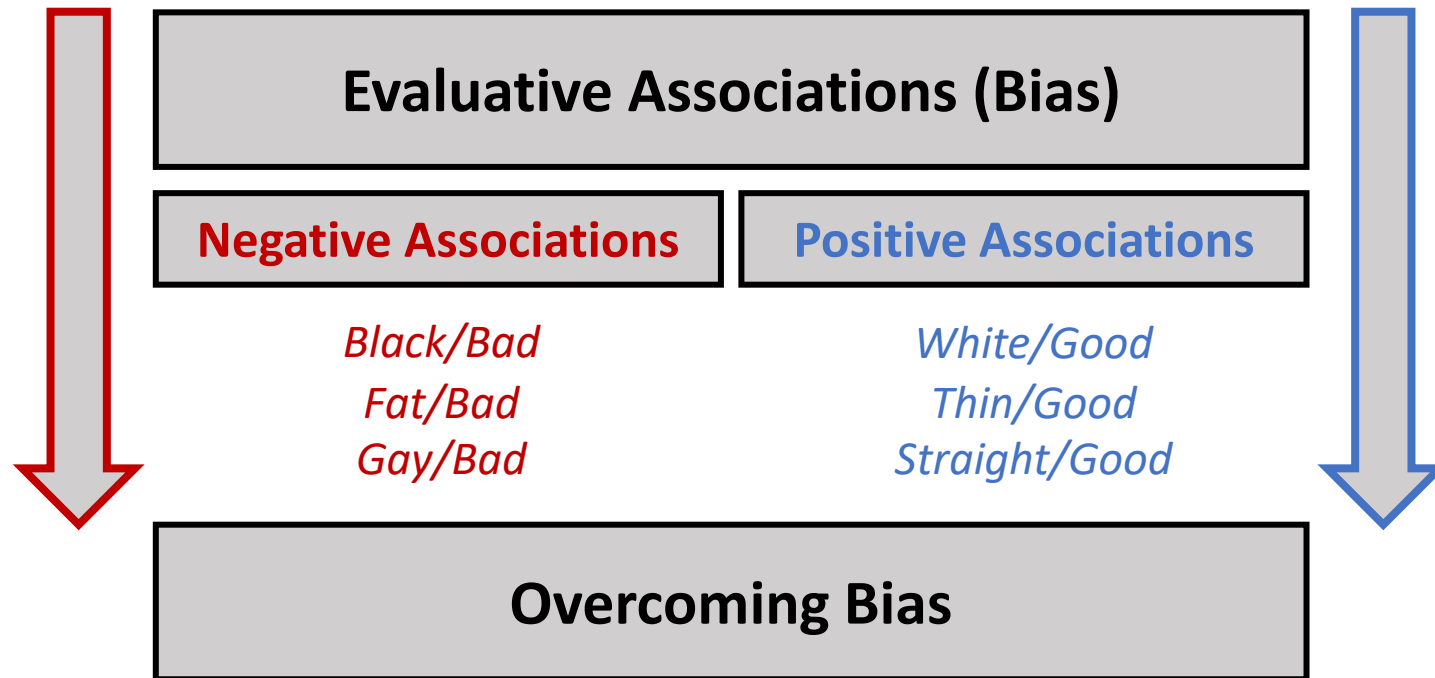
The dataset contains responses from variants of the Implicit Association Test, a categorization task designed to assess associations between social groups and evaluative attributes

The Quad Model is a multinomial processing tree model that decomposes responses on the IAT into multiple cognitive components

The Quadruple Process Model



The Quadruple Process Model



The QM assumes that the activation of biased associations can pressure responses on the IAT. However, if correct responses are also identified, these associations may not be expressed, as one may overcome them in favor of the correct response

The Quadruple Process Model: Parameters

Activation parameter (**AC**): extent to which evaluative information (e.g., Black-bad association) is activated

Detection parameter (**D**): extent to which correct response can be discriminated from incorrect response

Overcoming Bias parameter (**OB**): extent to which activated bias fails to influence responses

Guessing (**G**): Response bias in the absence of AC, OB, & D

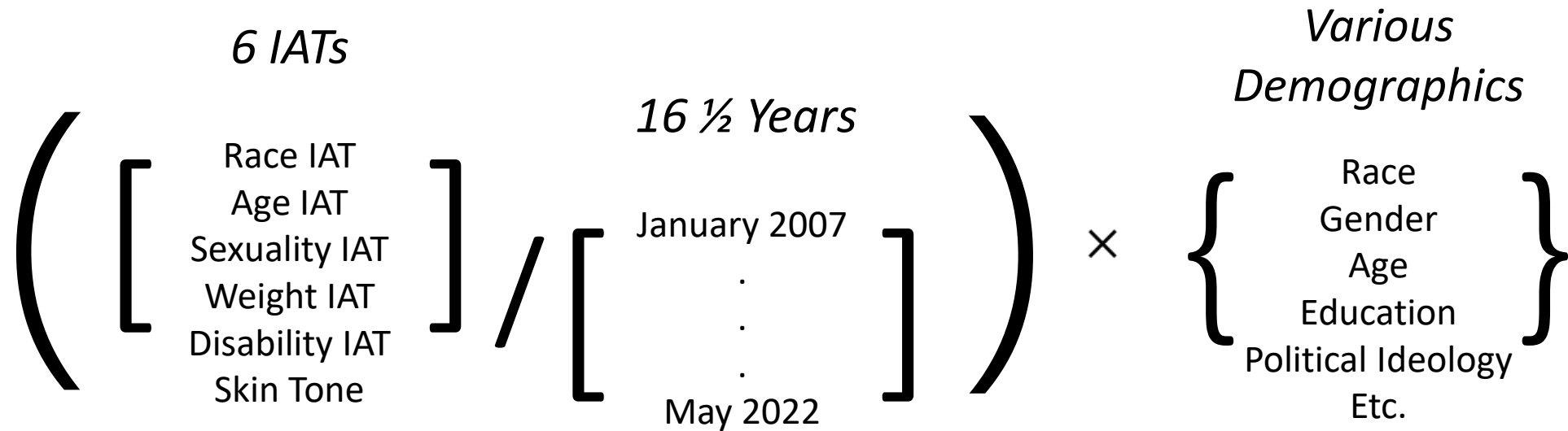
Relevant Papers

Sherman, J. W., Gawronski, B., Gonsalkorale, K., Hugenberg, K., Allen, T. J., & Groom, C. J. (2008). The self-regulation of automatic associations and behavioral impulses. *Psychological review*, *115*(2), 314.

Calanchini, J., Schmidt, K., Sherman, J. W., & Klein, S. A. (2022). The contributions of positive outgroup and negative ingroup evaluation to implicit bias favoring outgroups. *Proceedings of the National Academy of Sciences*, *119*(40), e2116924119.

Klein, S. A., & Sherman, J. W. (2022). Measuring and Modeling Implicit Cognition. *The Routledge Handbook of Philosophy and Implicit Cognition*.

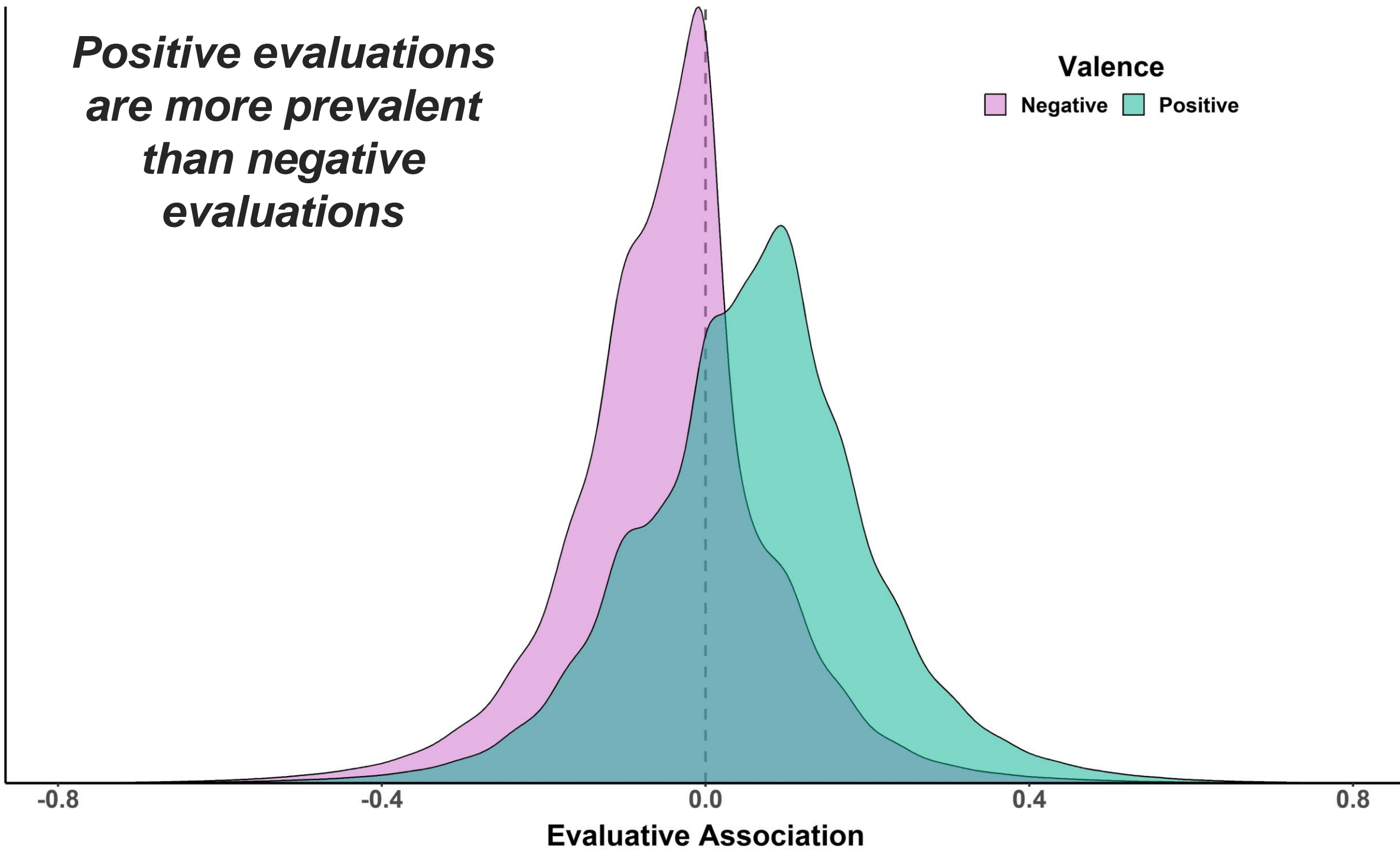
The Data (from Project Implicit)

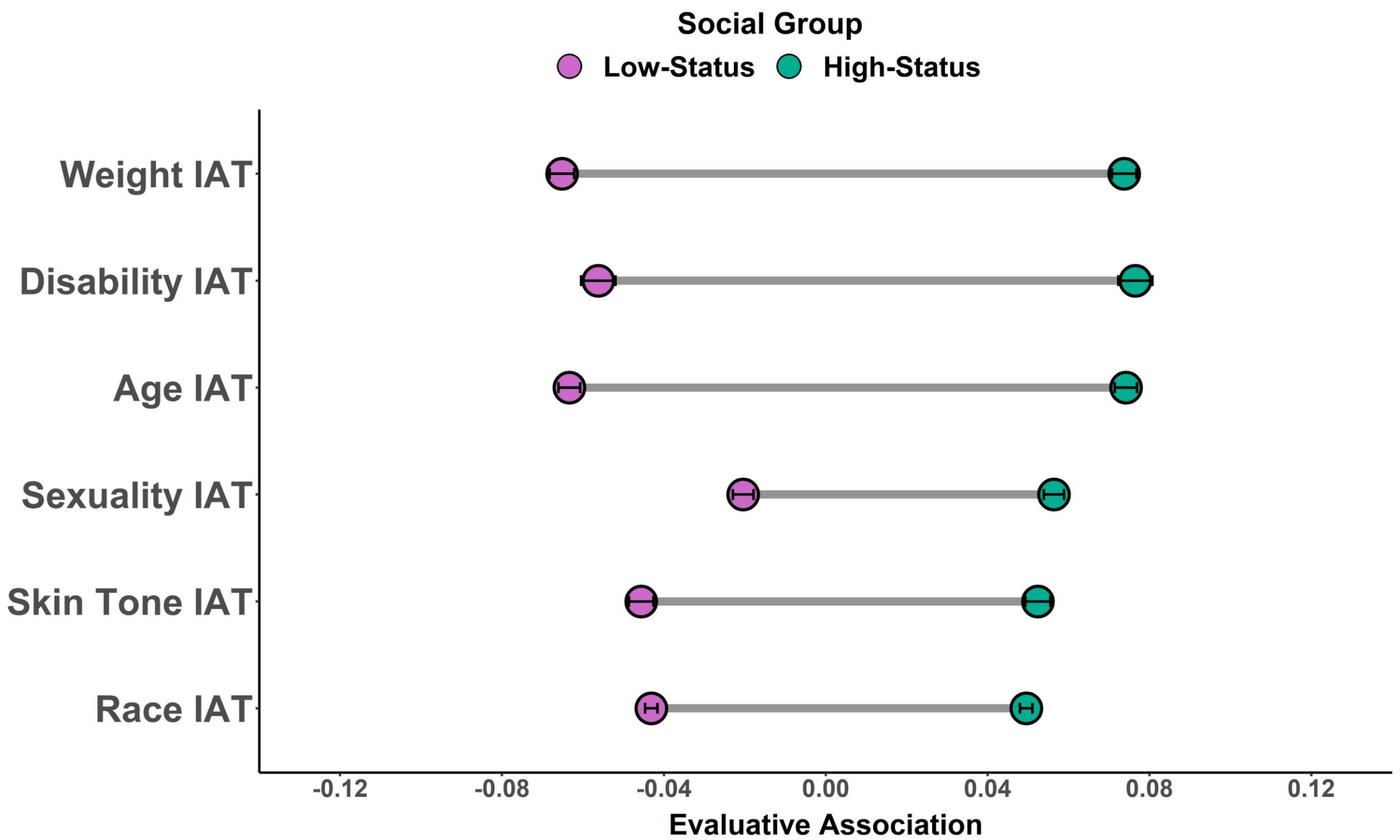


+8 ¼ Million IATs
+1 Billion Responses

***Positive evaluations
are more prevalent
than negative
evaluations***

Valence
Negative Positive





Convergence in Estimation Methods

Aggregate (complete-pooling approach): X # of IATs \rightarrow 1 model

Individual (no-pooling approach): X # of IATs \rightarrow X # of models

Rank-order between approaches appears quite stable in our data

 High rank correlations

Bias Measurement Comparison

We examined the association between bias measured in the Quad model and conventional measures of bias (D-scores, self-reported bias)

ρ (AC, D-Score)

ρ (AC, Self-Report)

Note: conventional measures cannot distinguish one group (e.g., Black) from another (e.g., White). For each IAT, the same values from each conventional measure are used to for correlations with both social groups (e.g., the same D-scores used to correlate with both estimated evaluative associations of Black and White)

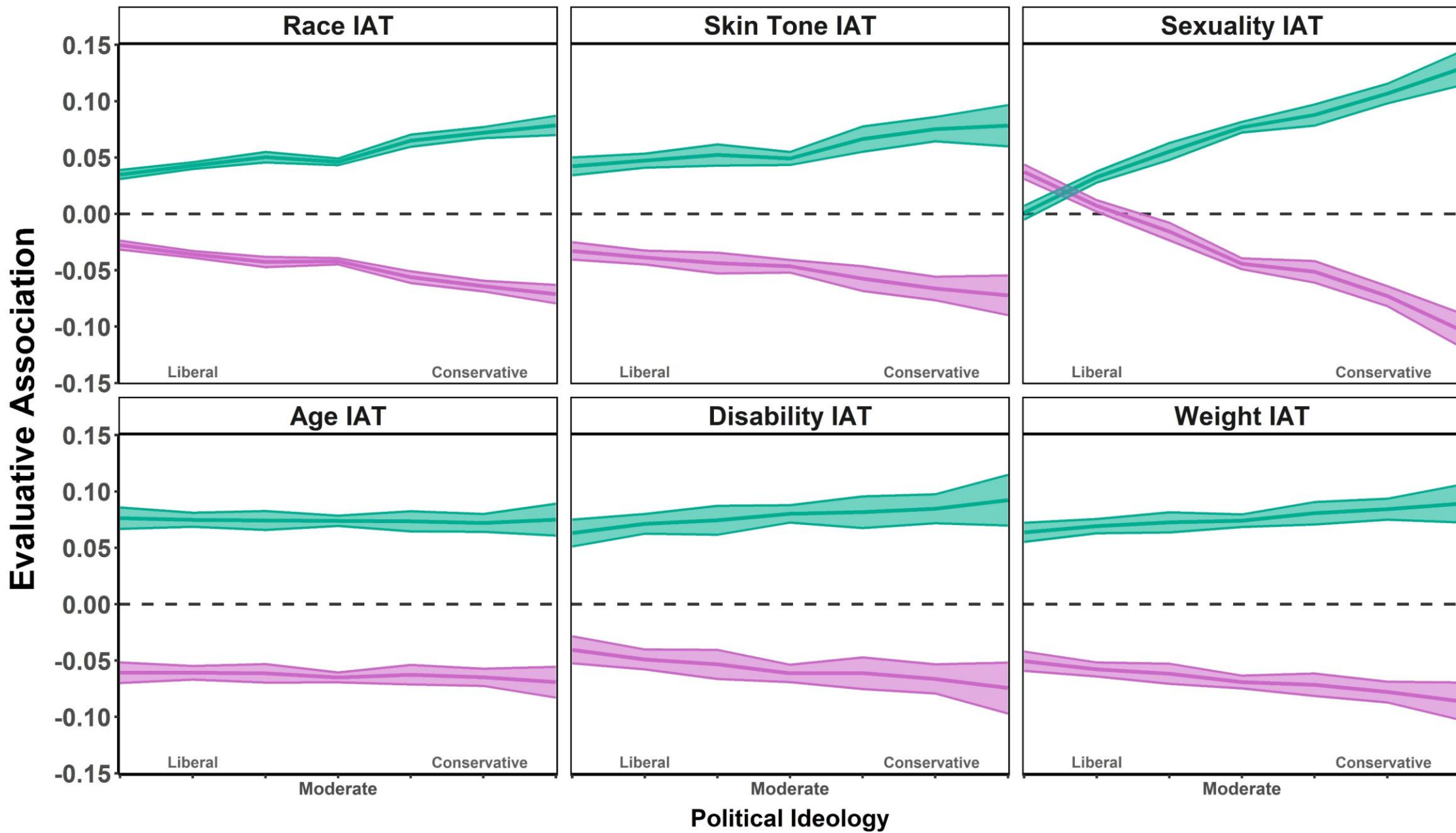
Social Group (per IAT)	IAT D-Score (Implicit Bias)		Self-Reported Preferences (Explicit Bias)	
	<i>rho</i>	<i>CI</i> _{95%}	<i>rho</i>	<i>CI</i> _{95%}
<i>Black</i>	-.277	[-.279, -.275]	-.289	[-.291, -.287]
<i>White</i>	0.264	[.262, .265]	.263	[.261, .265]
<i>Dark-Skinned</i>	-.262	[-.266, -.258]	-.209	[-.213, -.205]
<i>Light-Skinned</i>	.234	[.231, .238]	.194	[.190, .198]
<i>Gay</i>	-.424	[-.427, -.421]	-.454	[-.457, -.451]
<i>Straight</i>	.366	[.363, .368]	.366	[.363, .368]
<i>Old</i>	-.092	[-.096, -.088]	-.101	[-.105, -.097]
<i>Young</i>	.097	[.093, .101]	.092	[.089, .096]
<i>Disabled</i>	-.191	[-.197, -.185]	-.123	[-.129, -.117]
<i>Abled</i>	.216	[.210, .221]	.099	[.093, .104]
<i>Fat</i>	-.070	[-.075, -.066]	-.163	[-.167, -.159]
<i>Thin</i>	.079	[.075, .083]	.153	[.149, .157]



Political Ideology

Social Group

■ Low-Status ■ High-Status



Please Reach Out With Questions

This is merely a preview

I plan to update this with more information about the approach and our findings

It may take some time, though

If you'd like to know more before I've had time to update this, please reach out to me. I'll be happy to share more (e.g., other demographics, the OB & D parameters, motivation, etc.) and answer any questions

sawklein@ucdavis.edu