Towards a computational account of egodystonia

Lucy Lai
University of California, San Diego, San Diego, California, United States

Tobias Hauser
University of Tübingen, Tübingen, Germany

Quentin Huys
University College London, London, United Kingdom

Abstract

Egodystonia refers to thoughts and behaviors that conflict with one's values or beliefs, which is often observed in psychiatric conditions like obsessive-compulsive disorder (OCD). While prior work has demonstrated dissociations between beliefs and actions (Vaghi et al., 2017, 2019), we lack a computational framework to explain the mechanisms underlying this mismatch. In a novel experiment combining behavior and subjective report, we induced egodystonic feelings in a healthy population with a range of obsessive-compulsive traits. Individuals scoring higher on the Obsessive-Compulsive Inventory (OCI-R) reported greater egodystonic experiences. Egodystonicity was not influenced by reward availability or action rate, but was driven by perceived consequences of inaction, as captured by a computational model of the task. This study provides the first experimental evidence of induced egodystonia and offers a foundation for theoretical advances in understanding this phenomenon.

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*