

## **Beyond Self-Report: Direct Detection of Depressive Symptoms from Neural Activity**

Apoorva Vallampati<sup>1</sup> ([avalla01@tufts.edu](mailto:avalla01@tufts.edu)), Victoria Sharpe<sup>1</sup>, Arim Choi-Perrachione<sup>1</sup>, Julia Klein<sup>1</sup>, Maggie Wargo<sup>1</sup>, Bethany Bracken<sup>2</sup>, Spencer Lynn<sup>2</sup>, Gina R. Kuperberg<sup>1,3</sup>

<sup>1</sup>Tufts University, Medford, MA, USA

<sup>2</sup>Charles River Analytics, Cambridge, MA, USA

<sup>3</sup>Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

Depression is typically assessed through clinical interviews and self-report questionnaires. These approaches, however, are prone to biases from conscious deliberation and self-presentation. Can symptoms instead be inferred directly from patterns of brain activity produced as individuals read depression-consistent and -inconsistent statements? To address this question, we developed 160 self-relevant vignettes designed to evaluate various domains of depression, including emotional (e.g., mood, guilt), cognitive (e.g., concentration, suicidal ideation), behavioral (e.g., motivation, withdrawal), and physiological (e.g., energy, appetite), as informed by rating scales and the DSM-5. Each vignette was carefully controlled to isolate specific symptoms from the valence of a sentence-final critical word (e.g., "...Lately my mind hasn't been that/has been quite foggy/clear"). We recorded event-related potentials as 39 participants with a range of scores on the Beck Depression Inventory read these vignettes word-by-word. In participants with higher depression scores, depression-consistent (versus inconsistent) critical words elicited a sustained positivity between 300-1000ms. This effect was highly selective: participants with lower depression scores showed no such differences, and neither group exhibited differences in activity when reading the same critical words in a matched set of non-self-relevant vignettes. These findings demonstrate that depressive symptoms can be directly detected from neural activity recorded during natural reading, which reflects the earliest stages of access to meaning (300–500ms) and later evaluative processing (500ms onward). These targeted, self-relevant vignettes provide a promising foundation for developing a measure to objectively evaluate depressive symptomatology directly from neural activity, without the need for explicit behavioral responses.