**Data description:**

**Participants**

Data was collected from adolescents aged 13-18, recruited from three publicly funded secondary schools in the UK (n = 822, [450 females] Mean age = 14.84, SD = 1.45). School one was in Northamptonshire (n = 207 [127 females], Mean age = 15.02, SD = 1.49), and schools two and three were in Wiltshire (n = 248 [138 females], Mean age = 14.72, SD = 1.50) and school 3 (n = 367 [185 females], Mean age = 14.81) respectively. In this study we included only adolescents who scored above the clinical cut off for depression (i.e., a score of 27 or above on the MFQ), and those who scored within the ‘healthy’ range (i.e., a score of 12 or below on the MFQ) The final sample was 371 ‘healthy’ adolescents (58.2% males) with a mean age of 14.85 years (SD = 1.52), and 178 adolescents with elevated symptoms of depression (27.5% males) with a mean age of 14.82 years (SD = 1.29).

**Task**Self-concept was assessed by the Twenty Statements Test (TST; Kuhn & McPartland, 1954). The TST provided 20 unfinished statements each beginning “I am…” Participants are asked to describe themselves by responding to the question “Who am I?” by completing as many statements as possible. They were told verbally, not to think too much about their responses, and not worry about the order/importance of statements when completing the measure, as they were to write statements in the order they came to them.

**Variables in Excel Spreadsheets**

|  |
| --- |
| **Excel spreadsheet: Difference scores for descriptors**  |
| Descriptor | The word generated by adolescents |
| Difference score | The % of adolescents in the elevated group who generated word A - the % of adolescents in healthy group who generated word A). A difference score of a positive value highlights that these words were generated more often by adolescents with elevated symptoms of depression |
| **Excel spreadsheet: Most common descriptors split by depression group**  |
| Descriptor  | The word generated by adolescents |
| % of adolescents from each group who used each descriptor (i.e., healthy and elevated symptoms of depression)  | For each descriptor a proportion scores was calculated (e.g., count of word A/sample size of the healthy group; count of word A/sample size of elevated group). These scores reflect how often each group (i.e., healthy and those with elevated symptoms of depression) used this word. The most common descriptors appear first.  |