

Supplementary Information for

Facebook Language Predicts Depression in Medical Records

Johannes C. Eichstaedt*, Robert J. Smith*, Raina M. Merchant, Lyle H. Ungar, Patrick Crutchley,
Daniel Preotiuc-Pietro, David A. Asch, H. Andrew Schwartz

*These authors contributed equally to this work

Corresponding Author:

Johannes C. Eichstaedt

johannes.penn@gmail.com

This PDF file includes:

Table S1

Figs. S1 to S2

Table S1.

LIWC Dictionaries Associated with Depression (including non-Benjamini-Hochberg-significant associations).

Positively associated	β	Negatively associated	β
Pronouns		Personal Concerns	
1st pers singular (I, me)	.19 ***BH	Work	-.10 *
3rd pers singular (s/he)	.09 *	Leisure	-.08 *
Emotions		Other	
Feel (perceptual process)	.15 ***BH	Power (drives)	-.11 **
Negative Emotions	.14 **BH	Achievement (drives)	-.09 *
Sadness	.17 ***BH	Hear (perceptual process)	-.10 *
Anxiety	.08 *		
Cognitive Processes			
Discrepancy	.12 **BH		
Tentative	.10 *		
Other			
Health	.11 **BH		
Female references	.11 *		
Present focus	.10 **		

Note. Shown here are all pronoun and psychological process LIWC 2015 dictionaries significantly associated with future depression status controlling for demographics, with strengths of associations given as standardized regression coefficients. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; BH $p < 0.05$ after Benjamini-Hochberg correction for multiple comparisons.

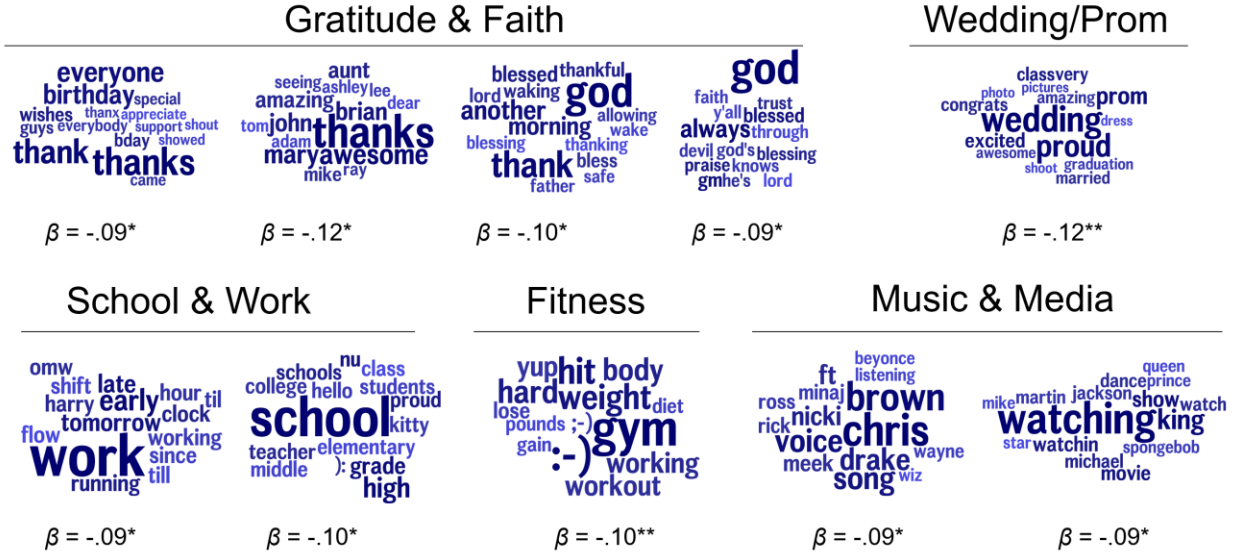


Figure S1. Ten language topics most negatively associated with a future depression diagnosis controlling for demographics. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; ^{†††} $p < 0.05$ after Benjamini-Hochberg correction for multiple comparisons. Font size reflects relative prevalence of words within topics. Color shade aids readability and carries no meaning.

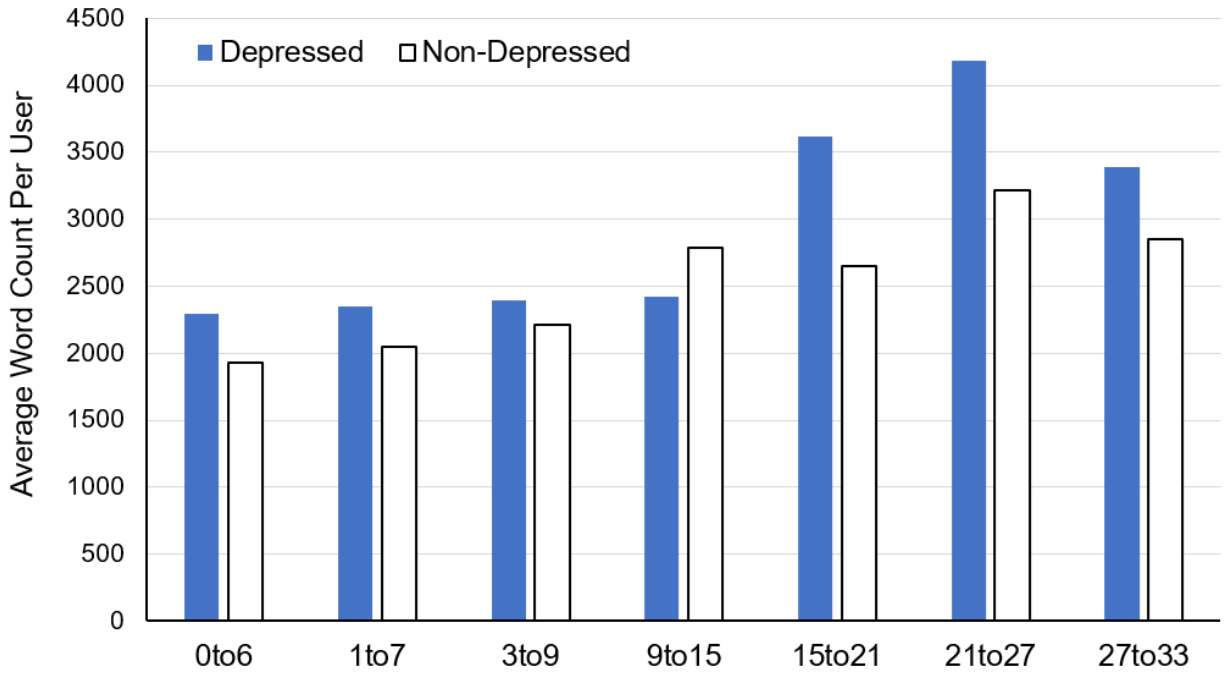


Figure S2. Average word counts for 56 depressed and 251 non-depressed users across the 6-month temporal windows. All users have at least 20 words in all temporal windows.