

# Towards Personalized Linguistic Anxiety Recognition

*Laurens Rook, Maria Chiara Mazza, Iulia Lefter, & Frances Brazier*

June 17, 2021

# What is Generalized Anxiety Disorder (GAD)?

- ❖ **GAD (definition)** = extreme, uncontrollable, and persistent worry and anxiety, which increases in intensity with age
- ❖ Assumed to affect 8.9 million people in the European Union, with major repercussions on social and occupational functioning
- ❖ Behavioral symptoms and mechanisms are poorly understood (difficult to recognize)

# Research questions

- ❖ Can GAD be predicted from the linguistic characteristics of expressive writings?
- ❖ Does GAD prediction accuracy increase if we also account for individual differences in avoidance and approach motivation?

# Theoretical agreement on GAD

- ❖ People suffering from GAD struggle with emotional cues, and seek to **avoid** them -->
- ❖ This is a maladaptive coping response, which triggers periodic episodes of excessive worrying behavior that are difficult to control
- ❖ This (worrying) is a **verbal linguistic phenomenon**

# The unique characteristics of the individual

- ❖ The intensity of GAD in an individual (self-reported via **GAD-7**; Spitzer et al., 2006--> *What is the amount of anxiety you experienced in the previous weeks?* (4-point scale)
- ❖ Individual differences in avoidance (or approach) motivation (self-reported via **the BIS/BAS Scales**; Carver & White, 1994)
- ❖ High scorers on BIS are more vulnerable to anxiety disorders, incl. GAD (Maack et al., 2012; Vervoort et al., 2010)

# Linguistic nature of GAD expression

- ❖ The way in which a threat is perceived should influence how anxiety/worry is expressed in words
- ❖ To find out, people could be invited to write vividly about a stressful or traumatic autobiographical event (the “**expressive writing paradigm**” from social & health psychology; Pennebaker, 1997)
- ❖ Those texts could be assessed on (psycho)linguistic properties (among others via LIWC; Pennebaker et al., 2015)

# Method

- ❖ Pre-survey: Administered the GAD-7 and BIS/BAS Scales
- ❖ Task: Expressive writing exercise: “*Recollect an anxious experience in your university life*”
- ❖ Post-survey: Assessed demographics and debriefing

# Sample and descriptives

- ❖ Final sample: 142 participants (56 men and 86 women,  $M_{\text{age}} = 23.33$  yrs.,  $SD = 1.96$ )
- ❖ Average text length = 165 words
- ❖ Summed GAD score  $M = 8.95$  ( $SD = 5.79$ ) vs.  $GAD \geq 10$ . In our sample, this applied to 55 participants (38.73%)
- ❖ Good – very good scale reliability: GAD ( $\alpha = .90$ ), BIS ( $\alpha = .87$ ), and BAS ( $\alpha = .82$ )

# Correlations I

	Abbrev	GAD	BIS-sensitivity		BAS-sensitivity	
			low GAD	high GAD	low GAD	high GAD
Words/sentence	WPS	-.03	-.05	.04	.14	-.04
Words > 6 letters	SIX	-.18*	-.03	-.09	.14	.04
<b>Linguistic dimensions</b>						
Personal pronoun	PPron	.19*	.11	.18	-.14	.27
1 <sup>st</sup> singular	I	.10	.08	.17	-.08	.19
1 <sup>st</sup> plural	We	.08	.05	-.03	-.09	.12
Adverbs	adverb	-.06	-.01	-.01	.09	-.13
Negations	Negate	.30**	.18	.02	.03	-.07

# Correlations II

## Psychological processes

Affective processes	Affect	-.04	.02	.12	-.29**	.16
Positive emotions	Posemo	-.20*	.05	-.02	-.13	.10
Negative emotions	Negemo	.18*	-.06	.17	-.26**	.12
Anxiety	Anx	-.12	-.09	.12	-.19	.28*
Anger	Anger	.26**	.01	.05	-.08	.01
Sadness	Sad	.30**	.00	.08	-.13	-.16
Social processes	Social	.23**	.19	.11	.01	.08
Family	Family	.25**	.08	.09	.22*	-.08
Friends	Friends	-.01	.17	-.07	-.02	.00
Certainty	Certain	-.08	-.14	.05	.01	.22
Past focus	Focuspast	-.09	.10	-.16	.06	.07
Present focus	Focuspres	.18*	.01	.21	-.13	.07
Future focus	Focusfut	-.17*	-.04	-.07	-.17	-.06
Time	Time	-.14	-.11	-.19	.06	.01
Work	Work	-.14	.11	-.14	.09	.09
Leisure	Leisure	.04	.03	.13	.16	-.23
Home	Home	.09	-.07	.13	-.12	-.23
Money	Money	.05	-.18	-.01	-.05	-.01
Religion	Relig	.10	-.10	.13	-.32	-.06
Death	Death	.09	-.08	-.06	.21	-.10
Swear words	Swear	.09	.01	.06	.06	.12

# GAD prediction

Feature	Classifier	Acc	U_Prec	U_Rec	U_F1	W_Prec	W_Rec	W_F1
	LIWC	0.7	0.69	0.67	0.67	0.7	0.7	0.69
SVM	LIWC+BB	0.75	0.76	0.72	0.72	0.77	0.75	0.74
	BB	0.6	0.59	0.58	0.56	0.61	0.6	0.59
	LIWC	0.63	0.61	0.57	0.55	0.62	0.63	0.59
RF	LIWC+BB	0.71	0.72	0.67	0.67	0.73	0.71	0.69
	BB	0.67	0.67	0.63	0.63	0.68	0.67	0.65
	LIWC	0.66	0.62	0.62	0.6	0.64	0.66	0.63
NB	LIWC+BB	0.72	0.71	0.71	0.7	0.73	0.72	0.72
	BB	0.65	0.66	0.66	0.65	0.68	0.65	0.65
	LIWC	0.68	0.65	0.64	0.63	0.67	0.68	0.66
LR	LIWC+BB	0.72	0.73	0.7	0.69	0.74	0.72	0.71
	BB	0.64	0.63	0.62	0.61	0.65	0.64	0.63

# Feature importance for GAD classification

LIWC	Imp	LIWC_BISBAS	Imp	BISBAS	Imp
L_negate	0.18	BIS Overall	0.18	BIS Overall	0.20
L_negemo	0.16	L_negate	0.08	BAS Overall	0.20
L_sad	0.14	L_negemo	0.08	BIS_24	0.18
L_social	0.14	L_sad	0.06	BIS_2	0.16
L_WPS	0.10	L_focuspast	0.06	BAS_5	0.14
L_adverb	0.10	L_work	0.04	BAS_3	0.10
L_affect	0.10	BIS_2	0.04	BAS_10	0.08
L_posemo	0.10	L_WPS	0.02	BAS_15	0.08
L_certain	0.10	L_Sixltr	0.02	BAS_9	0.06
L_focusfuture	0.10	L_posemo	0.02	BIS_13	0.06

# Main conclusion

- ❖ GAD can, indeed, be predicted from the linguistic characteristics of expressive writings:
  - Negatively valenced emotion words and text entries on social processes suffice for GAD recognition from written text, especially when also the writer's BIS is accounted for
- ❖ GAD prediction accuracy increases if we also account for individual differences in avoidance (not approach) motivation

# Future work

- ❖ *Is BAS sensitivity needed for accurate GAD prediction models?*
  - Our linguistic analysis suggests not vs. GAD classification results suggest it to be an important feature --> the BAS-GAD linkage could be better explained via **comorbidity with depression** scores
  
- ❖ *How trustworthy are diary entries from people with high GAD scores – especially during episodes of emotion dysregulation?*
  - --> **affective computing** applications for tracing emotion regulation journeys during a sequence of expressive writing exercises

**Thank you!**