

Human ratings for validating a Suicide Risk Lexicon

You will see a list of **tokens** (i.e., single words and short phrases such as "lonely", "no one cares about me") under a given **construct** (e.g., Loneliness & Isolation, Active Suicidal Ideation).

Please rate each token as to whether it is a prototypical expression of the construct as used in everyday language (prototypical means among the most typical examples that come to mind when you think of the construct). Our main use case would be if you want to flag a text document (clinical interview transcript, EHR, survey response) if it contains that token for further inspection of that construct in a mental health domain. For instance, if you're trying to measure the construct Eating Disorder, would you want to flag a document containing "anorexia"? "I don't eat much"?

Scale:

3 Definitely prototypical of the construct (i.e., I would want to flag it)

2 More yes than no

1 More no than yes

0 Not prototypical of the construct (i.e., very likely to be about something else so often that I would not want to flag it)

Note:

- False positive: flagging a document not about the construct
- False negatives: missing a document about the construct.
- "Depression" is more commonly used for mental health than economy or geology; whereas "alarm" would likely create many false positives for Panic (e.g., no Panic in these phrases: "The report alarmed us", "The alarm went off").
- So the goal is to balance reducing expected false positives as well as false negatives

Example

Construct: Eating disorders

Prompt name (the construct name given to GPT-4 to create some of these tokens): eating disorders and related behaviors

Definition: Eating disorders are mental health conditions characterized by unhealthy and obsessive behaviors related to food, eating, and body image. Common types include anorexia nervosa, bulimia nervosa, and binge-eating disorder.

Tokens	Eating disorders_include (your scores)	Eating disorders_add	Notes for annotators
I'm fat	3		While this may not be an eating disorder, it is very prototypical of one. 2 would also be a reasonable response.
Weight	1		It's very common for people to mention "weight" without

			describing a likely eating disorder
I'm not eating	2 or 3		Follow your intuition
I'm restricting	2		Note: in the context of eating disorders, this would definitely be prototypical; however, without context it is likely to be used when not talking about eating disorders, so probably best to put 2
anorexia	3		Very prototypical
anorexia nervosa	3		
appetite	0 or 1	never have an appetite	"appetite" is very likely to be used when not talking about eating disorders
binge	2		
binge eating	3		
bulimia	3		

Notes:

- **Add tokens if you want** if you think should be included in the <construct>_add column next to each construct's list (in any row). We'll consider it 3 unless you specify 2 next to the word. Don't worry about morphology ("don't eat much" vs. "doesn't eat much") or punctuation (binge-eating disorder vs binge eating disorder). We might edit your addition to reflect most colloquial uses.
- **It's okay if a token appears under multiple constructs** (e.g., "pills" under the active suicidal ideation construct and under the lethal means and treatment; or "can't sleep" under constructs for Sleep issues as well as for Bipolar Disorder).
- Remove if you think it will be a false positive too often. You can decide whether to remove a token (assigning it 0 or 1) if you think it will too often be associated with a different construct. For instance "crime" is related to sexual abuse but there are so many more types of crimes than sexual abuse that including it as "sexual abuse" may create too many false positives. Also, while certain words might characterize suicidal notes ("just", "irreversible"), they are not prototypical of suicide, and would create too many false positives.
- **You can leave a comment anywhere** with anything you want me to take into account in the <construct>_add column (e.g., suggesting a change in a definition, incorrect example token).
- **Search for how words are used in suicide notes or other subreddits if you want by googling:** reddit r/suicidewatch "<token>" and looking at the Reddit r/SuicideWatch entries.



Reddit · r/SuicideWatch

5 comments · 5 years ago



I don't want to be a part of this society. : r/SuicideWatch

Go to school, go to college, get a job, live a boring life and wallow in your **unimportance**. What's the point ? Society is not made for ...

-
- Or for instance, googling: reddit r/EatingDisorders "<token>"
- Sometimes tokens seem like they're missing some words or a final letter. A final letter or word may be missing to capture multiple possible suffixes or a word may be missing: for instance, "skin" in "scratch my" in Direct Self-Injury. For instance, "quick and painless" would be more prototypical if it were "quick and painless death" or "I want a quick and painless death"; however, this token can appear with multiple contexts, so good to keep it minimal a priori. But feel free to score it low and add the broader phrase or word in the add column.
- At the bottom you'll see OLD TOKENS: DO NOT CODE, this is just in case you want to see some tokens that were not found in the large datasets or the current tokens are substrings of them and so they're redundant