**Mining Transdiagnostics Symptoms in Social Media Data**

**Abstract**

Mining social media data to predict mental health condition and psychological traits have increasingly attracted attention in the clinical psychology domain. Instead of using social media data to predict a diagnosis specific criteria, we adopted a disorder-independent approach by investigating factors that contribute to a variety of psychological disorders. Treatments that targets these factors are called transdiagnostic treatment, which has been widely employed to tackle anxiety and depression disorders. Transdiagnostic symptoms predispose an individual to a variety of mental disorders. We leverage FB data from 77 users who participated in the myPersonality project who completed a self-reported depression measurement back in 2011. We labeled negative emotion and two transdiagnositc components - reasoning bias and negative thinking among 5000 Facebook posts. Our study include detecting the some of the predisposing factors to depression and predicting depression by transdiagnostic components and user profile on Facebook. We used a simple linear regression model to predict self-reported depression score. A significant regression equation was found (F( , ) \_\_ p<0.001), with an of .. Users with more transdiagnostic symptoms have lower satisfaction with life, high in neuroticism and ...

Keywords: cognitive distortion, emotion, negative thinking, transdiagnostics, depression, social media, Facebook

**Introduction**

In the recent years, there is a surging amount of studies attempting to use social media data to predict psychopathology diagnosis. Various attempts on predicting depression has achieved good performance (citation). What makes these prediction tasks challenging at the moment is comorbidity is very common in psychopathology. According to the literature, about 60% - 70% (Brown, Antonu, & Barlow, 1995) of individuals diagnosed with anxiety disorder also meet some of the criteria in depressive and affective disorders.

The traditional conceptual structure approach to understand psychological disorder is to provide a diagnosis of a specific disorder. However, there is increasing recognition that criteria diagnosis are of less value because a majority of patients have coexisting disorders, which is also called comorbidity. In light of the challenge, psychologists are shifting towards a transdiagnositc approach in the recent years. Instead of giving multiple diagnosis to a patient with comorbidity, transdiagnostic approach focuses on common psychological processes underlie the syndromes, which provides a better explanation to the high rate of comorbidity observed in clinical practice. Treatments targeting transdiagnostic symptoms have been found to be more effective than ..treatments in anxiety and depressive disorders (citation).

**Investigate Transdiagnositic Symptoms on Social Media**

We explore some of the transdiagnostic symptoms in social media posts. People often posts threads or updates about their opinions, emotions and daily life activities. Data capturing these information provide researchers a platform to study their longitudinal behaviors (citation). Motivated by the fact that most of the studies look at whether social media behaviors reflect mental health symptoms focus on a specific diagnosis criteria (citation), which did not consider high comorbidity rate in many disorders. We consider using social media posts to detect some of the transdiagnostics symptoms.

The components of transdiagnostic treatment or research include attention, memory, reasoning, thought and behaviour. Reasoning refers to thinking that involves deducing conclusions, generating judgements and testing hypotheses logically. Biased reasoning often draws in conclusion different from the reality (Allison, Book). Reasoning bias is in parallel with cognitive distortion in cognitive behavior therapy. Assessing cognitive distortion is an index of the improvements in behaviours and emotional resiliency (citation). Later we will explain the details of identifying cognitive distortion.

Negative thinking …..

We also identify the negative emotion…

We compare the transdiagnostic symptoms with user’s self-reported depression score. Reasoning bias and negative thinking are predisposing factors in developing psychopathology. Individuals who develop reasoning bias and thought might or might not have depression, but these individuals have higher risks to develop psychological disorders if they encounter some important life event, such as lost or failure (citation). Therefore, we won’t use self-reported score of psychological disorder as a benchmark, instead, we see how the transdiagnostic symptoms contribute to depression, influence satisfaction with life and how it related to personality. Meanwhile, we also look at the profile information to highlight the predisposing factors of depression. Finally, we will investigate the characteristics of users who have more transdiagnostic symptoms.

**BACKGROUND AND RELATED WORK**

**The Cognitive Behavior Therapy (CBT)**

Cognitive models of psychopathology proposed that pathological behaviors and emotions are often the consequences of cognitive biases or distortion, which is the inadequate interpretation of situations. Beck’s cognitive model of psychopathology emphasizes the role of cognitive distortion in the the maintenance of anxiety, depression and other mental disorders (Beck, 1967; Beck, 2011). Since cognitive distortion is central in some of the disorders, a common goal of cognitive behavioral therapy is to help an individual to adjust these biases. The process through which is called “cognitive restructuring’. Cognitive restructuring modifies the clients’ problematic ways of thinking about themselves, their world and their future (Beck, 1976). To identify these biases, they look at thoughts that might contain cognitive distortions and investigate the schema that generate the distortive thoughts (Cory F. Newman, 2015).

Cognitive assessment include three levels: 1) the automatic thoughts, which are spontaneous and compulsive thoughts that are not based on reflection or deliberate thinking. 2) Intermediate beliefs, which are underlying assumptions that cause the tendency of particular type of distortion. It is one’s general assumptions regarding oneself, others, the world and the future (citation). 3) Underlying personal schema. The underlying assumptions and schemas are rigid, and contribute to the development of depression episodes or anxiety (citation).

Automatic thoughts can be classified according to the cognitive distortion in its content. For example, mind reading, personalization, labeling and all-or-nothing thinking, etc. (citation). These thoughts can be true or dismissive to the reality. For example, “My boyfriend doesn’t like me anymore.” This statement may be true or based on mind reading. The cognitive distortions are a part of mood and anxiety disorders, and the content often indicate the types of disorder. For instance, individuals with social anxiety disorder engages in mind reading “No one wants me to be around with them” or catastrophizing “It will be a disaster if I say something wrong in the group”. Depressed individuals engages in a wide range of cognitive distortions - labeling “I am the black sheep of the family”, fortune telling “I can never be happy without you.” (Back & Haigh, 2014).

**Assessment of Cognitive Distortion**

The most widely applied method for assessing cognitive distortion is the cognitive distortion checklist (citation). The list has been validated in experimental and clinical work (citation). However, the measurement of cognitive distortion was not given much attention in the past. The Cognitive Distortion Questionnaire (CD-Quest; de Oliveira et al., 2014) is a 15-item questionnaire that assess the frequency and intensity of cognitive distortion. It is administered before the therapy session to help a client to keep track on their thinking errors thus enabling them to aware of the change over time as the therapy goes on.

The limitation of using a survey approach is that cognitive distortion is highly context based. Therapists identify the cognitive distortions via a semi-structural or non-structural interview with the client instead of administer a questionnaire to the patients. CD-Quest only provide a few contextualized examples as references for a client to evaluate their cognitive distortions over a varieties of situations (table 1) . Based on the few examples provided on a questionnaire, it could be difficult for a client to identify the distortion without sufficient trainings. In this study, we adopted the CD-Quest in our annotation guideline as a criteria for annotators to identify cognitive distortion based on the context information provided in the Facebook posts.

Table 1.

Example of CD-Quest

1. Dichotomous thinking (also called all-or-nothing, black-and-white, or

polarized thinking): I view a situation, a person, or an event in “either-or”

terms, fitting them into only two extreme categories instead of on a continuum.

EXAMPLES: “I made a mistake; therefore my performance was a failure.” “I ate

more than I planned, so I blew my diet completely.”

**Negative thinking**

Nolen-Hoeksema (1991) defined rumination as thoughts and behaviours that repetitively focus an individual’s attention on his or her negative feelings, and the nature and implications of these feelings (including the causes, meanings, and consequences of the feelings). (Allison Book)

**Negative Emotion and Psychopathology**

In addition to cognitive distortion, the cognitive model ascribes exaggerated negative response to negative life events as a manifestation of vulnerability. In psychotherapy, it is essential to distinguish between events, thoughts and feelings. An event is a specific situation occur in the past, present and future. It can be physical sensation, factual or anticipated (Robert, Leahy). Whereas, feelings or emotions are the result of how an individual think about an event. The frequency and intensity of emotion are indexes of psychopathology. Intense distressed emotion is painful and dysfunctional because it interferes an individual to think clearly and effectively (citation). Patients with psychological disorder often experience intense emotion that can be excessive to the situation, therefore, the level of distress also distinguish the patients with non-patients.

Studies show that negative emotion is related to depression.

# A Depression Detection Model Based on Sentiment Analysis in Micro-blog Social Network.

Pennebeker LIWC and depression, share task depression,

**Method and Materials**

**Data**

This corpus consists of 5000 Facebook posts from individuals who participated in the myPersonality project from January 2009 to December 2011. Our methods were carried out in accordance with the approved guidelines from myPersonality. myPersonality was a Facebook-based application collecting psychometric tests from users. Participants opt to allow myPersonality to collect their account information and public Facebook posts. collection of myPersonality complied with the terms of Facebook service. All data are anonymized and gathered with opt-in consents for research purposes. The sample used in our study contains 301 participants who have completed the CES-D scale, Satisfaction with Life Scale, Big-5 Personality Scale and Schwartz Value Survey.

**Sampling Approach**

To ensure we have enough posts to conduct a longitudinal study. We only include regular posters in our sample. We define regular posters as individuals who posted twice per week or more. We estimated this using the average post count per day during the sampling frame. If an individual had a post count per day of 0.3, this individual made around 109.5 posts in 365 days, which was roughly equivalent to an average of 2.1057692 posts per week. In our sample, 122 out of 301 participants were regular posters. To make sure our sampling approach was conducted under a standard sampling framework, we included 91 regular posters whose last post obtained by myPersonality was less than a week before they completed the CES-D scale. Then we obtain a sample of 4696 posts that were produced two months before CES-D score was obtained. We future eliminate 14 posters who produced less than 20 posts during the two months and posts that are not written in English. Eventually we yield a sample of 4145 posts from 77 users.

**Annotation**

The annotation guideline was developed using 4362 Facebook posts to illustrate negative emotion and cognitive distortions. The extracted posts were first annotated by a specific trained psychologist according to the annotation guideline. Another two trained annotators labeled 10% of all the posts for agreement. Two clinical psychologists checked all the posts that contain cognitive distortion. In scoring the cognitive distortion, annotators are given specific cues - Cognitive Distortion Checklist as reference of the measurement of cognitive distortion, but are also instructed to rely on their linguistic intuition. In addition, posts from quotes, lyrics, and repost are labeled as non-original posts.

Negative emotion…

Negative thinking…

Annotators are trained by following instructions and sets of practice examples on the annotation guideline, the discussing deviant scoring and scoring additional items until interrater reliability of xx is reached.

**Self-reported measurement scale**

We now presents a number of user characteristics lists in their profile and four self-reported scales they have filled out.

We define X different user characteristics that measure a user’s nature of activity on Facebook.

Center for Epidemiologic Studies Depression Scale (CES-D Scale)

CES-D is a self-reported scale designed to measure depression symptoms in the general population(citation). The scale consists of 20 depression symptoms associated items. The scale has been tested in psychiatric settings across various cultures over the years. It was found to have high internal consistency and test-retest reliability (citation). Its validity was assessed via correlation with clinical diagnosis of depression and other self-reported depression measurement (citation).

Five Factor Model of Personality

The five factor personality model was established in an attempt to understand the description of traits. The dimensions composing the 5-factor models are extraversion, agreeableness, conscientiousness, neuroticism and openness to experience. The five factor structure has been proved to be robust in both self and peer ratings (McCrae & Costa, 1987), children and adult (citation) and across different cultures (McCrae & Costa,1997; Pulver, Allik, Pulkkinen, & Hamalainen, 1995; Salgado, 1997). Early literature found that big-5 is relatively stable over time (Costa & McCrae,1992a,1988). However, recently literature found the opposite (citation). Neuroticism was found to have strong correlation with a bunch of psychological disorders, such as anxiety and depression (citation). Individuals who score high on neuroticism tend to experience negative mood frequently and physical symptoms. They are also more likely to be affected by negative life events (Suls, Green, & Hills, 1998). Recent studies found that social media data can predict the 5-factor model of personality (personality and social media data study).

Satisfaction with Life Scale (SWLS)

The 5-item satisfaction with life scale was developed to measure global life satisfaction. The SWLS has been tested across different cultures and age groups (citation) and has been found to have high internal consistency and temporal reliability (Diener et at., 1985). Its validity was assessed by correlation with other measures of subjective well-being and specific personality dimension.

Values are cognitive representations of desirable, abstract goals (e.g., security, justice) (Rokeach, 1973; Schwartz, 1992). Similar to needs, motives, and goals, values motivate actions (Rohan, 2000; Seligman, Olson, & Zanna, 1996). Values differ from specific goals (Emmons, 1989; King, 1995; Roberts & Robins, 2001; Winnel, 1987) because values are transsituational. Unlike needs and motives (Bilsky, 1998; McClelland, 1985), values are inherently desirable and must be represented cognitively in ways that enable people to communicate about them. Explicating the relations of personality traits to values will deepen our understanding of both

Neuroticism. Individuals high on Neuroticism tend to be anxious, depressed, angry, and insecure. Those low on Neuroticism tend to be calm, poised, and emotionally stable. We anticipate no positive associations between value priorities and Neuroticism. Neuroticism is not likely to facilitate the attainment of the motivational goal of any type of value. Moreover, as Bilsky and Schwartz (1994) reasoned, “The depression characteristic of people high on neuroticism might result from failure to attain the desired level of any one of the ten values” (p. 171).

**Results**

**Transdiagnostic labels**

Among 4145 posts, 804 of them reflect negative emotion of the author, 36 of them contain a mix of positive and negative emotion. Among 840 posts contain negative emotion, only 41 show cognitive distortion, 111 show negative thinking (85 worry, 26 rumination). 3 posts shows both cognitive distortion and negative thinking. Cognitive distortion is rare, it only occurs in 1% of the posts in this sample. We aggregate a per day negative emotion score by summing up the number of negative post from each user and divide it by 60 days. We use the same approach to generate a per day distortion score and per day negative thinking score for each user. Table 1 shows the statistics of the per day score from the transdiagnostic components and their correlation with self-reported CES-D score. Although posts show cognitive distortion only account for 1% of all the posts but it shows a medium correlation with self-reported depression score whereas negative emotion and negative thinking are not significantly correlated with CES-D.

Table 1

|  |  |  |  |
| --- | --- | --- | --- |
|  | mean | sd | Cor. with CES-D |
| Negative emotion | 0.182 | 0.197 | 0.192 |
| Cognitive distortion | 0.009 | 0.016 | 0.300\*\* |
| Negative thinking | 0.018 | 0.032 | 0.110 |

\* p<0.05, \*\*p<0.01, \*\*\*p<0.001

**Dataset Statistics**

We now subset a set of individuals with both of their negative emotion day score higher than the mean and the cognitive distortion day score higher than the mean, which yields a sample of 16 individuals. We also subset another sample (n = 43) in which individuals have lower than average negative emotion day score and cognitive distortion day score.

In Table 1, we present transdiagnostic symptoms scores from the two groups together with user characteristics and their self-reported big-5 personality score, satisfaction with life score and depression score. Two users didn’t report their age on their profiles, here we assign the mean age to the them. Our observation indicate that users psychological characteristics distinguish their transdiagnostics symptoms. Specifically, users with more symptoms tend to post more posts (nearly twice more than the low symptom group), and younger in age. They reported significantly more depression symptoms (28% higher than low symptom users), less open-minded and extroverted.

Table 1. Psychological characteristics among individuals with high and low transdiagnostic symptoms

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | All | | High Trans | | Low Trans | | p | Cohen’s d |
|  | mean | sd | mean | sd | mean | sd |  |  |
| num. of posts | 53.940 |  | 99.25 |  | 33.49 |  | \*\*\* | L |
| age | 23.120 |  | 21.450 |  | 24.400 |  | \* | S |
| negative emotion | 0.181 |  | 0.437 |  | 0.081 |  | \*\*\* | L |
| cognitive distortion | 0.009 |  | 0.031 |  | 0.000 |  | \*\*\* | L |
| negative thinking | 0.018 |  | 0.035 |  | 0.006 |  | \*\* | L |
| SWL | 4.221 |  | 3.675 |  | 4.353 |  |  | M |
| CES-D | 23.860 |  | 28.19 |  | 22.05 |  | \* | M |
| ope | 4.166 |  | 3.939 |  | 4.207 |  | \* | M |
| con | 3.183 |  | 3.237 |  | 3.171 |  |  |  |
| ext | 3.101 |  | 2.701 |  | 3.175 |  | \* | M |
| agr | 3.539 |  | 3.230 |  | 3.613 |  |  | M |
| neu | 3.022 |  | 3.224 |  | 2.890 |  |  | S |

\* p<0.05, \*\*p<0.01, \*\*\*p<0.001

effect size: 0.8 = large(L); 0.5= moderate(M); 0.2 = small(S)

num. of posts: Number of posts in two months; SWL: Satisfaction with Life score

CES-D: Center for Epidemiological Studies Depression (CESD); ope: openness; con: conscientiousness; ext: extraversion; agr: agreeableness; neu: neuroticism.

We further divide users according to their demographic characteristics (gender, marital status, relationship status and relationship with parents), and observe their differences in transdiagnostic symptoms. Users missing some of the characteristics information or are assigned under the category ‘other’. P values shows whether the group mean is different from the mean of the whole sample.

Our observation shows that there is no gender difference in transdiagnostics symptoms. It seems to appear that married individuals have significantly lower transdiagnostics symptom and depressive symptoms. However, we find that married individual with children, all of them are female, have very high depressive symptoms. Having a partner to provide mental support seems to be a protective factor in developing transdiagnostic symptom, whereas, no difference is found among people being in a relationship. Therefore, the result can be interpreted the other way around, people with less transdiagnostic symptoms are more likely to get married.

Individuals with parents who are together tend to have lower negative emotion (42% lower than the mean) , cognitive distortion (67% lower than the mean) and negative thinking (72% lower than the mean). Whereas, those with separated parents and no frequent contact with non-resident mothers appear to have a lot more depressive symptoms (55% higher than the mean). Their negative emotion and negative thinking are more than twice higher than the group mean and cognitive distortion four fold of the group mean.

Table 2. Day score for different groups of people

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | n | Negative emotion | | Cognitive distortion | | Negative thinking | | CESD |
|  |  | mean | sd | mean | sd | mean | sd |  |
| All participants | 77 | 0.182 | 0.197 | 0.009 | 0.016 | 0.018 | 0.032 | 23.86 |
| gender | | | | | | | |  |
| male | 24 | 0.173 | 0.143 | 0.010 | 0.020 | 0.018 | 0.035 | 22.88 |
| female | 53 | 0.186 | 0.219 | 0.008 | 0.014 | 0.019 | 0.031 | 24.30 |
| Parents together | | | | | | | |  |
| No\\ and not in frequent contact with non-resident father | 10 | 0.162 | 0.080 | 0.008 | 0.009 | 0.009 | 0.22 | 27.1 |
| No\\ and not in frequent contact with non-resident mother | 3 | 0.433 | 0.377 | 0.039 | 0.042 | 0.047 | 0.046 | 37\*\*\* |
| No\\ but in frequent contact with non-resident mother | 1 | 0.067 | NA | 0 | NA | 0 | NA | 28 |
| No\\ but in frequent contact with non-resident father | 10 | 0.25 | 0.178 | 0.005 | 0.008 | 0.021 | 0.032 | 20.89 |
| Parents are together | 19 | 0.105 | 0.076 | 0.003 | 0.006 | 0.005 | 0.013 | 25.32 |
| other | 34 | 0.197 | 0.239 | 0.011 | 0.018 | 0.025 | 0.039 | 19.63 |
| Relationship status | | | | | | | |  |
| single | 49 | 0.188 | 0.217 | 0.009 | 0.017 | 0.021 | 0.035 | 24.31 |
| In a relationship | 20 | 0.208 | 0.171 | 0.01 | 0.016 | 0.018 | 0.027 | 22.25 |
| married | 8 | 0.079\*\*\* | 0.074 | 0\*\*\* | 0 | 0.003\*\*\* | 0.009 | 25.13 |
| Married with child/children | 3 | 0.15 | 0.100 | 0.006 | 0.010 | 0.011 | 0.019 | 33.33 |

\* p<0.05, \*\*p<0.01, \*\*\*p<0.001

Late night posts

Sleep disturbance is one of the major symptoms in depression. We attempt to find out if there’s any relationship between sleep disturbance and cognitive distortion. We count the number of posts written from midnight until 6:00 in the morning. We find that the number of late night posts and user’s negative emotion (r = 0.754, p < 0.001) is highly correlated, cognitive distortion (r = 0.474, p < 0.001) and negative thinking (r = 0.395, p <0.001) are moderately correlated with number of late night post. Our result shows that individuals with high

Linguistics styles

Regression model

**Discussion**

By combining cognitive distortion, emotions and its intensity, we gain better insight of an individual’s predisposing factors to mental disorder.

Problem in annotation

1. Automatic thoughts are sometimes expressed in the form of a question, making evaluation difficult.

‘Will I pass the test?’

‘Why did this happen to me’

1. Overgeneralizing and dichotomous thinking is sometimes confusing
2. Some of the thinkings does not fit into any of these categories but are obviously negative

E.g. is stressed enough to pull out her hair.

feels like screaming.

**(Label Negative automatic thought as first layer)**

1. A lot of the status shows the emotion or thoughts but didn’t show the reason why, that’s what makes cognitive evaluation difficult

E.g. At times, I really hate myself

I’m angry today

**(Label Negative automatic thought as first layer)**

1. Missing information

e.g

is in shock and worried because of what people did and can't wait to go to EC one Week

5. Sometimes we can’t tell whether it’s a fact or distorted cognitive perception

**(Label Negative automatic thought as first layer)**

E.g

feeling depressed, and fat.

(Is the person really fat or self-perceived fat?)

6. There are times that neither the author nor the annotator is able to

identify the affect. People are often confused about what they feel, for example, whether

it is one affect or another or some combination of both. One will experience oscillation between anger and sadness when one is disappointed in love.

E.g.

is wishing that she could stop feeling like shit. she doesn't know why and she doesn't want to really know.

Just woke up from a doze to suddenly start crying, what the hell is wrong with me?

Sentistrengh misclassify some of the tweets, do we need to remove them in the training sample?

Is it better to recruit psychology students to do labeling

Find the entries in the list that can be identified in the posts and ask participants to label those e

Start with longer posts

5000 on both end of the scale

Preotiuc-Pietro, D., Sap, M., Schwartz, H. A., & Ungar, L. H. (2015, June). Mental Illness Detection at the World Well-Being Project for the CLPsych 2015 Shared Task. In *CLPsych@ HLT-NAACL* (pp. 40-45).

**Abstract**

In this paper, we construct a annotated corpus -- the Negative Automatic Thoughts, a corpus of consists of xx Facebook posts. This corpus has been annotated for negative automatic thoughts in a three tier hierarchical format. The first tier of label identifies posts with negative automatic thoughts; the second tier identify the category of the thoughts. We also evaluate the emotion and situation of the thoughts.

**Introduction**

There is a growing consensus that social media data provide help us to gain insights on how different aspects of mental health issues are manifested in language. A growing number of mental health disorder classification models have been developed based on language features and user behavior patterns on social media. (Mitchell et al., 2015; Coppersmith et al., 2015; Munmun De Choudhury, 2015; Chen et al., 2017). Most of these models make binary classification on mental health disorders. However, symptoms of mental health disorders often overlapped. Single or multiple depressed episodes, low activity level are universal symptoms in both depressive and bipolar disorder (citation); mood swing and psychotic symptoms are characteristics in both bipolar and schizophrenia (citation). Therefore, it is hard for machine learning classifiers to distinguish disorders with a lot of overlapped symptoms, which is the case for most of the mental disorders. For instance, a depression classifier in Shared task have good performance in identifying depression (true positive rate 0.8) but poor performance in distinguishing depression and Post-traumatic stress disorder (true positive rate 0.4) (Preotiuc-Pietro et al., 2015).

Moreover, the baseline of these mental health classifiers are often self-reported diagnosis. Participants fill out a self-reported measurement scale online (Munmun De Choudhury, 2015). For the shared task studies, data was retrieved by keyword search, such as “I am diagnosed with…” (Preotiuc-Pietro et al., 2015). It is rare to have patients who are willing to openly talk about their mental health on social media and such a sample might be very biased in the language they use.

Mental health classifier based on social media data has its limitation. Based on the above findings, we investigate if social media data can identify predisposing factor, precipitation factor, perpetuating factor and protective factor for mental illness (citation). Some of the psychological factors, such as rumination, worry, impaired reasoning and memory occur across many mental health disorders. These components show very similar characteristics across a range of disorders (citation). We aim at investigating how a transdiagnostic component - repetitive negative thinking, manifested in social media data.

Repetitive negative thinking refers to the repetitive thought about one or more negative topics that is experienced as difficult to to control (Ehring & Watkins, 2008). RNT has been found to be a transdiagnostic component in at least 13 disorders, including major depressive disorder (MDD) (r = 0.53) and general anxiety disorder (GAD) (r = 0.34, <0.01) (Spinhoven et al., 2015); Ehring et al. 2011; Harvey et al., 2004), PTSD (Clohessy & Ehlers, 1999), social phobia (Abbott & Rapee, 2004), obsessive-compulsive disorder (Amir, Cashman, & Foa 1997), eating disorder (Nolen-Hoeksema et al., 2007) and other maladaptive cognitive processes (e.g., cognitive avoidance; McEvoy et al. 2013). RNT also constitutes a vulnerability factor for comorbid depression and anxiety disorder (Kessler, Chiu, Demler & Walter, 2005). RNT may also lead to sense of entrapment and hopelessness that may contribute to suicidal ideation (Law et al., 2017)

Studies compare the RNT across different disorders and found them share three characteristics: repetitive, passive and/or uncontrollable, focus on negative content). However, the content of RNT are disorder-specific, in a way that reflects individual’s current concern (Thomas Ehring and Edward Watkins, 2008). For example, GAD patients have catastrophic concern about what might go wrong in the future (Vasey and Borkovec 1992; Davey and Levy 1998). Patients with panic disorder are more concern with how they feel about their bodies (Clark 1993).

**Negative Automatic Thoughts**

People often post their thoughts on social media, some of these thoughts are elaborative, **whereas most of them seem to be spontaneous and compulsive. It’s not based on reflection or deliberate thinking. These thoughts are identified as ‘automatic thoughts’ in cognitive behavior model. Sometimes these thoughts are expressed in exclamation words but the meaning of these thoughts can be easily interpreted in a context.** For instance, a person working in the office and the boss handover a stack a document to him, when the person has the thought ‘Oh, no!’, the meaning was he is tired of the workload. **Whereas, some of the thoughts are the expressions of an idea. For instance**, “I’m stupid to be late to the exam today.”

**Automatic thoughts can be evaluated according to their validity and utility (Judith). Automatic thoughts affect the individual response to the situation or events appropriately. There are three types of automatic thoughts. The first type of automatic thought is distorted to the fact despite there is objective evidence to support the contrary.** For instance, “all of my friends told me that I am already slim but I don’t think so.”(invalid)  **The second type of thought is accurate to the fact but the conclusion from the individual may be distorted.** For example, “I’m bad at the exam (valid thought), therefore, I’m hopeless (distorted conclusion). ”

**The third type of thought is also valid, but dysfunctional by decisions that would triggered more problems**. For example, “I am so tired because I worked until 3 A.M. last night. I spent my daytime playing games because I don’t feel like to work in the morning.” This thought is accurate, but the individual suffers from decrease of motivation and concentration because of inappropriate time schedule decided by himself.

**Automatic thoughts are often brief, they can be in verbal or imagination form. People usually think their automatic thoughts are true without evaluation. For social media data, we are not able to identify validity of all the thoughts. But we can identify whether they are automatic by the linguistics features. Whether the thoughts are dysfunctional or the conclusion from the author is distorted.**

**To investigate repetitive negative thinking, we start from identifying negative automatic thoughts, analysing the emotion and situation associate with the thoughts and investigate if some of them demonstrate a repetitive pattern.**

1. **1. Does the thinking of the author not correspond to the facts/reality? If yes, AT. If no or can’t tell, go to 2 and 3**

**According to your knowledge, is the thinking counter to the fact?**

**E.g. Homework over break kinda defeats the purpose of the break.**

**Not correspond to the facts. You can have a good break with proper time schedule**

**If I fail everything in my life, at least my bedroom is clean**

**Not correspond to the facts. One fails a specific task, not everything in life.**

**E.g. I fail the exam, my life is hopeless. (the conclusion ‘my life is hopeless’ is distorted)**

1. **Is the thinking negative and compulsive (hard to control)? If yes, AT, if no or not sure go to 2 and 3**

**(this category is usually for posts with strong emotions or feelings only, without showing any reflective thinking)**

**E.g. Grrr...I am really irritated. (seems compulsive, it’s something you can’t control)**

**Oh, shit.**

1. **Is the thinking/behavior creates more problems in life or interpersonal relationships? If yes, AT**

**(e.g. blaming everything to a person, persistent bad habits)**

**i am waking up at four tomorrow for last minute black friday plans. because i am crazy girl who overestimates her powers of staying awake.**

**E.g. boring boring boring boring**

GAAAAHHH!!!! HATE MOOD SWINGS!!! >.< ONE MOMENT YOU WANT TO HIDE IN CORNER, THE NEXT YOU WANT TO GO UP TO SOME RANDOM PERSON'S DOOR AND DANCE THE POKA BECAUSE YOU HAVE TO MUCH ENERGY!!!

Negative automatic thought is a crucial component in the cognitive models of behavior.

Cognitive models of psychopathology proposed that pathological behaviors and emotions are often the consequences of cognitive biases or distortion, which is the inadequate interpretation of situations. The common goal of cognitive behavioral therapy is to help an individual to adjust these biases. The process through which is called “cognitive restructuring’. Cognitive restructuring modifies the clients’ problematic ways of thinking about themselves, their world and their future (Beck, 1976). To identify these biases, they look at thoughts that might contain cognitive distortions and investigate the schema that generate the distortive thoughts (Cory F. Newman, 2015).

Automatic thoughts is at the ‘surface’ level of identifying cognitive distortion. It is one’s reaction towards what’s happening in the environment, recollections, ruminations and so on. Such thoughts come and go quickly and closely tie to one’s emotion, however, they do not necessarily reflect the reality accurately. Dysfunctional thoughts contribute to the intermediate beliefs, which are one’s general assumptions regarding oneself, others, the world and the future. At the deepest level are the core beliefs, which is the fundamental beliefs in people’s lives. For example, an individual with an abandoned schema tend to think all the people in their lives will leave them eventually no matter what they did (citation). Therefore, collecting the automatic thoughts is the first step of understanding the schema.

**Situation and Salience of Automatic Thoughts**

Automatic thoughts could be abstract without any context (e.g. It’ll probably be a waste of time) and most of the automatic thoughts on social media are of this kind. Identifying these thoughts is a beginning step in cognitive behavior therapy (CBT), because one of the goals of the therapist is to teach clients to specify their thoughts (e.g. It’ll probably be a waste of time if I go to this lecture) so that they can identify if these thoughts are appropriate to the situation (citation). **The automatic thoughts can be categorized by the types of situation that trigger the thoughts. The situation can be external events, stream of thoughts, emotion, behavior and physiological or mental experience (citation).** Social media posts various in length, short posts often do not imply any situation and context, whereas, situation can sometimes be identified in long post.

Another aspect of the automatic thoughts is the emotions that goes with it. Intense distressed emotion is painful and dysfunctional because it interferes an individual to think clearly and effectively (citation). Patients with psychological disorder often experience intense emotion that can be excessive to the situation, therefore, the level of distress distinguish the patients with non-patients. It is important to identify whether the content of the automatic thoughts match with the emotion.

**Thoughts on social media**

Social media data contain thoughts on a variety of situations (citation and my paper). A lot of these thoughts are spontaneous, temporary and repetitive. Some of these thoughts are tie to a specific situation. For example, … (external event); … (behavior); …(physiological); ….(mental experience). Some of them does not imply any context. For instance, …. (emotion); … (stream of thoughts). Therefore, one of our goals is to identify whether these thoughts are linked to a specific situation.

In this paper, we build a binary classifier on automatic thoughts and a binary classifier to see if the thoughts are tied to a situation. Then we analyze the emotions and situation tie to these thoughts.

**There is an accumulating evidence shows that expressive writing about an upsetting and emotional event produce long-term beneficial effect in both mental and physical health** (Greenberg and Stone 1992; Pennebaker 1993; Spera el al. 1994; Littrell 1998; Smyth 1998)

**Annotation Method**

Our tagging process include two stages. **At stage one**, the posts are assigned sentiment score automatically by Sentistrengh. (Introduction of Sentistrengh) and then briefly corrected by human annotation. Section x present the results of a comparison between manually corrected tagging and automatic assigned tagging.

**At stage two**, we labeled posts with a situation. These situations include external events, behavior, physiological and mental experience.

**Types of Negative Automatic Thoughts**

**Negative automatic thoughts may pertain to self (“I am angry”) or others (“He is not a nice person”).** The content of negative automatic thoughts can be categorised according to the source that trigger the thoughts. Some of the posts might contain information from more than one categories. For example, “I Just woke up from a doze to suddenly start crying, what the hell is wrong with me?” The above example shows a behavior (suddenly cried), and stream of thought (What is going on with me?).

**External event**

An event that is associate with the change of affect.

**Behavior**

Behavior that results on change of affect

E.g. is stressed enough to pull out her hair.

**Physiological**

Statements that indicate physical complain

**Emotion**

Emotions that does not link to any situation

E.g.

feels like screaming

I’m angry today

**Stream of thoughts**

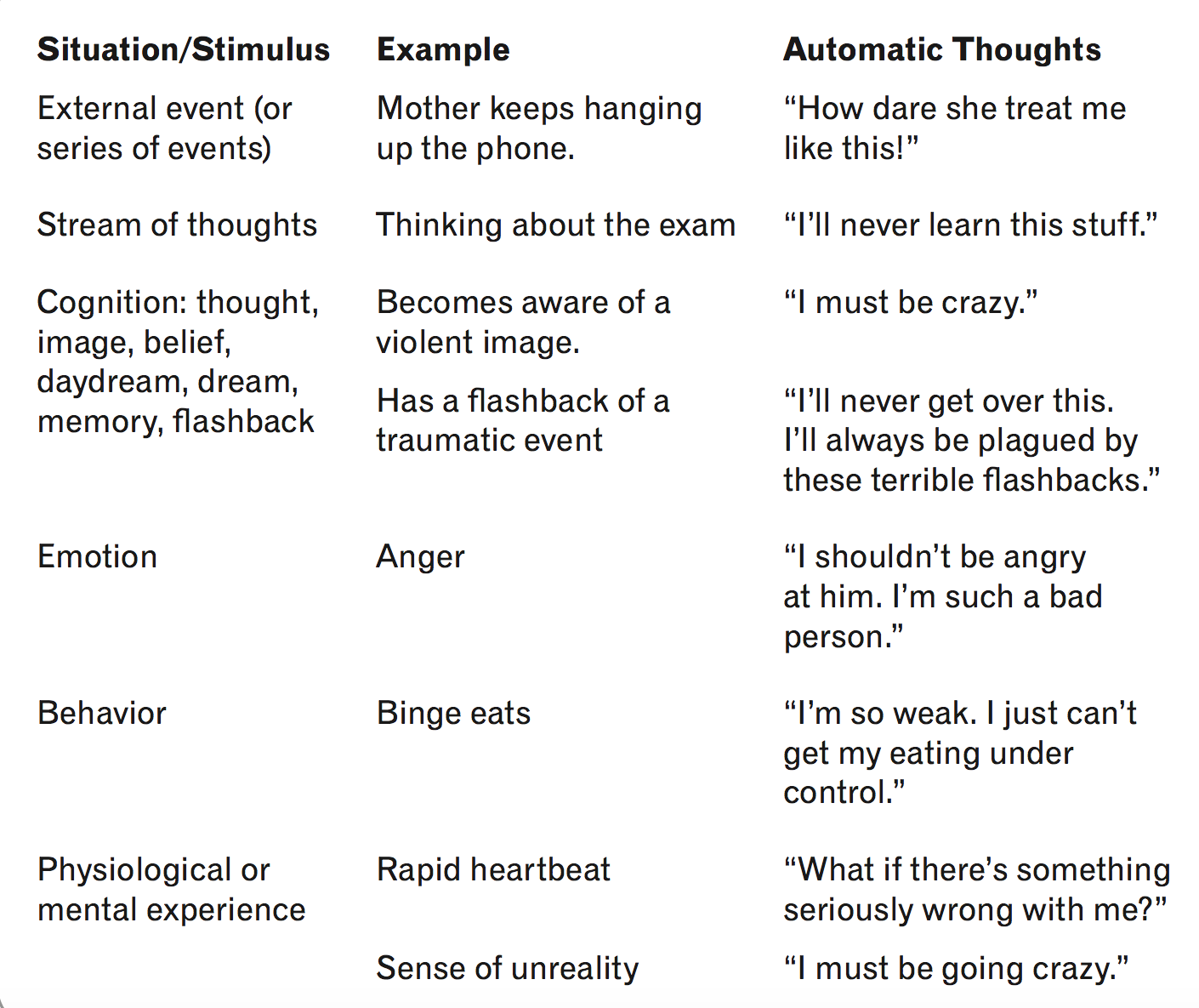
Stream of thoughts that does not indicate any situation. Stream of thoughts are sometimes expressed in the form of question.

E.g.

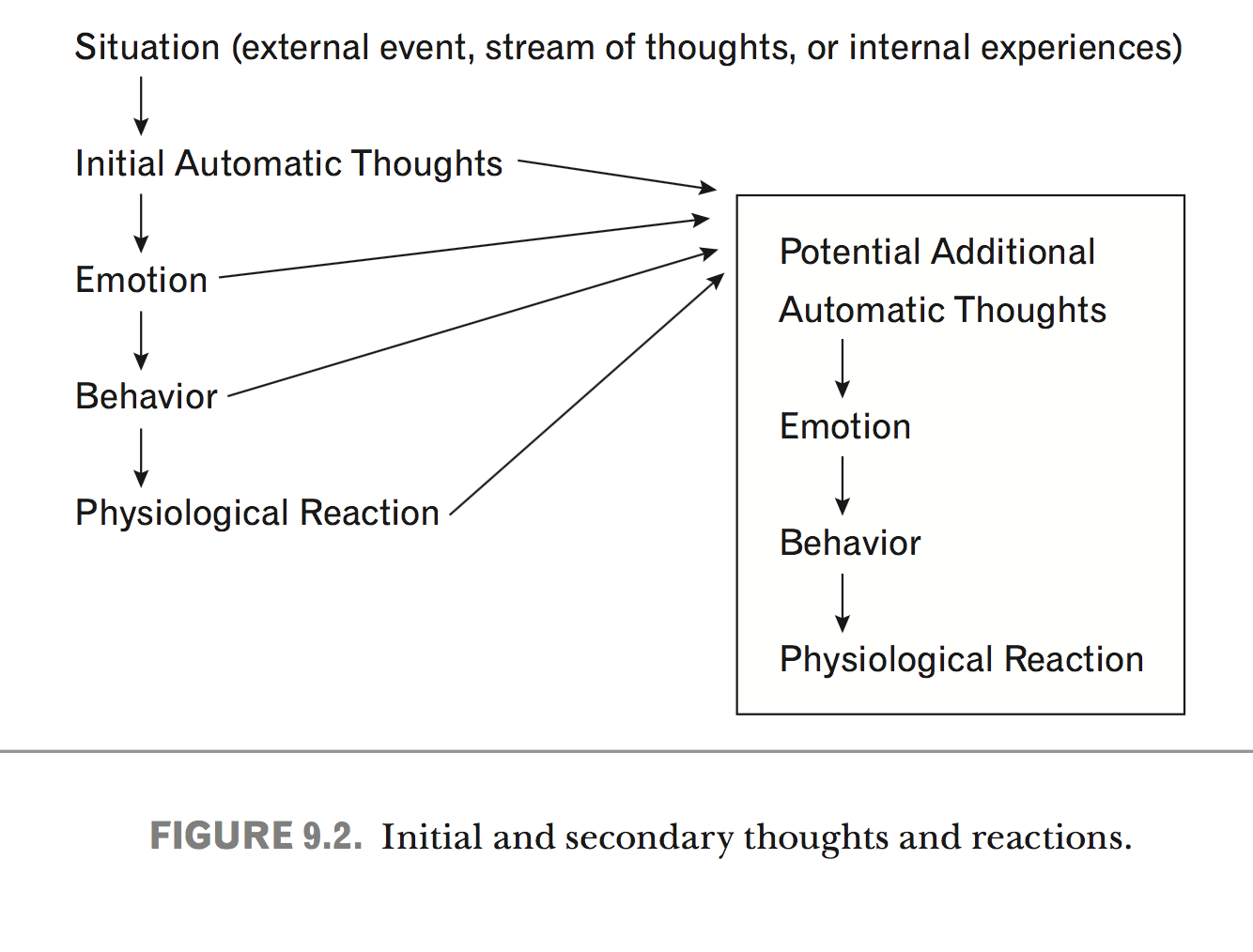
‘Will I pass the test?’

‘Why did this happen to me’

At times, I really hate myself



￼



Problem in annotation

1. Automatic thoughts are sometimes expressed in the form of a question, making evaluation difficult.

‘Will I pass the test?’

‘Why did this happen to me’

1. Overgeneralizing and dichotomous thinking is sometimes confusing
2. Some of the thinkings does not fit into any of these categories but are obviously negative

E.g. is stressed enough to pull out her hair.

feels like screaming.

**(Label Negative automatic thought as first layer)**

1. A lot of the status shows the emotion or thoughts but didn’t show the reason why, that’s what makes cognitive evaluation difficult

E.g. At times, I really hate myself

I’m angry today

**(Label Negative automatic thought as first layer)**

1. Missing information

e.g

is in shock and worried because of what people did and can't wait to go to EC one Week

5. Sometimes we can’t tell whether it’s a fact or distorted cognitive perception

**(Label Negative automatic thought as first layer)**

E.g

feeling depressed, and fat.

(Is the person really fat or self-perceived fat?)

6. There are times that neither the author nor the annotator is able to

identify the affect. People are often confused about what they feel, for example, whether

it is one affect or another or some combination of both. One will experience oscillation between anger and sadness when one is disappointed in love.

E.g.

is wishing that she could stop feeling like shit. she doesn't know why and she doesn't want to really know.

Just woke up from a doze to suddenly start crying, what the hell is wrong with me?

Sentistrengh misclassify some of the tweets, do we need to remove them in the training sample?

Is it better to recruit psychology students to do labeling

Find the entries in the list that can be identified in the posts and ask participants to label those e

Start with longer posts

5000 on both end of the scale

Preotiuc-Pietro, D., Sap, M., Schwartz, H. A., & Ungar, L. H. (2015, June). Mental Illness Detection at the World Well-Being Project for the CLPsych 2015 Shared Task. In *CLPsych@ HLT-NAACL* (pp. 40-45).