

# What Motivates You? Benchmarking Automatic Detection of Basic Needs from Short Posts

Sanja Štajner<sup>1</sup>, Seren Yenikent<sup>1</sup>, Bilal Ghanem<sup>2</sup>, Marc Franco-Salvador<sup>2</sup>

<sup>1</sup>Symanto Research, Nuremberg, Germany

<sup>2</sup>Symanto Research, Valencia, Spain

{sanja.stajner, seren.yenikent, marc.franco}@symanto.com  
bilalhgm@gmail.com

## Abstract

According to the self-determination theory, the levels of satisfaction of three basic needs (competence, autonomy and relatedness) have implications on people's everyday life and career. We benchmark the novel task of automatically detecting those needs on short posts in English, by modelling it as a ternary classification task, and as three binary classification tasks. A detailed manual analysis shows that the latter has advantages in the real-world scenario, and that our best models achieve similar performances as a trained human annotator.

## 1 Introduction

Motivation is one of the most crucial aspects of human behaviour with implications ranging from daily life to career and educational contexts. Self-determination theory (SDT) provides a meta-framework for understanding the broad, as well as specific, nutrients of the function and application of the concept of motivation (Deci and Ryan, 2000; Ryan and Deci, 2017a).

SDT differs from the other motivational theories from the psychology literature in two substantial aspects (Ryan and Deci, 2000; Rigby and Ryan, 2018): (1) Unlike the drive theories that explain motivation as a function of its deficit (e.g. people are motivated by success to compensate its deficit), SDT focuses on growth and constructivism (e.g. people are naturally and universally motivated by success), thus giving the theory a more realistic understanding of the human behaviour, and making it applicable to wider contexts; and (2) Due to the applicability advantage, SDT is based on strong behavioural evidence and is thus not only a well-validated model but also sustainable and actionable.

The SDT framework is supported by a body of cross-cultural studies strengthening the universality of the theory. Studies conducted in diverse countries showed that the basic needs are essentially

represented across cultures (Chen et al., 2015; Jang et al., 2009). Although universal, the SDT framework is also able to point out the impact of sociocultural environment on the variations of basic needs in different cultures. For example, a study conducted in 11 countries showed that the need for competence was more linked to school performance in Eastern cultures than in the West (Nalipay et al., 2019).

One of the central pillars of SDT are three basic psychological needs that drive the initiation of a behaviour and the maintenance of motivation:

- **Autonomy:** the basic need to be the owner and controller of one's decisions and behaviours.
- **Competence:** the basic need to feel competent, effective and master-like.
- **Relatedness:** the basic need to belong, bond and connect with others.

According to SDT, those three needs are universal and their importance does not change across individuals and situations. However, different contexts and time periods would require different support and resources for the maintenance of the motivations. For instance, cultivating autonomy need in students creates more engagement and willingness, thus leading to higher academic performance, lower dropouts, and more self-esteem in the long run (Ryan and Deci, 2020). Similarly, the SDT framework is used to increase levels of employee satisfaction and engagement, supportive leadership and parenting skills, healthier relationships, satisfactory consumer experience and better designed digital media and well-being tools (Slemp et al., 2018; Rigby and Ryan, 2018; Ryan and Deci, 2017b; Knee et al., 2002; Gilal et al., 2019; Peters et al., 2020; Peng et al., 2012).

Need	Post
Autonomy	Just treated myself to a Roland TB-3. Should arrive this evening. #excited
Autonomy	One thing’s for sure, I will not let you ruin my dreams, HIV. #determined
Competence	What an achievement. Finally getting some credit. #Fury #SPOTY
Competence	I fell asleep with socks on... I disgust myself.
Relatedness	I’m so lucky to have my best friend and boyfriend rolled into one! #soppy #proud
Relatedness	You know what I feels like to be #ALONE in this cold world?

Table 1: Annotated examples from the dataset (either satisfied or unsatisfied need).

Traditionally, basic motivations are assessed via questionnaires which provide intensity-based scores for each dimension. The scores represent the degree to which that particular dimension is satisfied (Deci and Ryan, 2000). Although these questionnaires were developed and validated via laboratory and field studies which provide a strong empirical basis, they could suffer from biases commonly observed in questionnaire respondents such as social desirability bias (Krumpal, 2011) and the reference-group effect (Heine et al., 2002). The basic motivations can also be revealed in a more implicit way, by collecting subjects’ narratives while showing them pictures and images (Murray, 1943; McClelland, 1979). Although being more expensive and time-consuming, as it requires the inclusion of trained assessors, this method shows that implicit motivations can be assessed from texts. A few studies attempted at automatic detection of basic motivations on the basis of their linguistic aspects from such narratives (Pennebaker and King, 1999; Johannssen and Biemann, 2019).

To the best of our knowledge, our study is the first that attempts to automatically detect the three basic needs from short posts. In this study, we:

- Benchmark the task of automatic detection of basic needs from English Twitter data using several architectures on an already existing manually annotated dataset.
- Provide a manual analysis which shed light on the complexity of the task and its usability.
- Discuss the limitations of the existing dataset, and suggest better annotation strategies.

## 2 Dataset

For our experiments, we used the first two layers of the Basic Psychological Needs Corpus (Alharthi et al., 2017), which is publicly available.<sup>1</sup> The

<sup>1</sup>We obtained the original dataset directly from the authors.

corpus contains Twitter posts annotated with five layers of annotation as the intention was to provide freely available multilayered annotated corpus for a wide range of applications (Alharthi et al., 2017). The manual annotation was performed by three annotators in three stages, encompassing thorough training sessions and detailed annotation guidelines, one round of collectively labelling tweets, one round of independently labelling the same posts for calculating inter-annotator agreement (IAA), and the final round of independently labelling the rest of the posts. The average pairwise agreement and the Fleiss Kappa ( $\kappa$ ) were 90% and 0.815 for whether or not the post contains enough content for assigning one of the three basic needs (autonomy, relatedness, or competence), and 89% and 0.819, respectively, for the assigned label (Alharthi et al., 2017).

The final dataset with manual annotations of basic needs was already pre-filtered for non-emotional posts and those that do not contain enough signal (Alharthi et al., 2017). It contains 6334 posts with the following distribution of the labels: 1229 posts labelled with *competence*, 1771 with *autonomy*, and 3334 with *relatedness* label. In our experiments, we used this dataset and only the labels of the second layer of annotation (basic needs). Several examples are given in Table 1. Here is important to note that the original dataset also contains, in the third layer, the annotation for the satisfaction level (satisfied, dissatisfied, neutral) of the assigned basic need. We acknowledge that the combination of the basic needs and their level of satisfaction are often used together, e.g. as indicators of person’s well-being (Deci and Ryan, 2011), violence and conflict possibility (Christie, 1997), stress and coping (Ntoumanis et al., 2008; Weinstein and Ryan, 2011). However, we opted for discarding these additional labels for three reasons: (1) because the inter-annotator agreement was significantly lower for this annotation layer

(the average pairwise agreement was 75% and the  $\kappa$  was 0.640); (2) so that we do not increase the total number of classes (to nine instead of three) and therefore significantly lower the number of instances in each class; (3) because this task appears similar to the task of assigning the sentiment polarity of the post (Alharthi et al., 2017), and therefore might be modelled with various other datasets.

### 3 Experimental Setup

#### 3.1 Preprocessing

The instances were already cleaned in the original dataset by removing all usernames (@username) and URLs, while preserving emoticons, punctuation marks, social acronyms and abbreviations, which might contain psycholinguistic signals (Alharthi et al., 2017). Furthermore, the dataset does not contain any duplicated instances, tweets with less than three words, or tweets with more than three hashtags (Alharthi et al., 2017). We noticed that for this particular task, the hashtags may help the models, e.g. #proud usually signalizes *competence*, #relationship signalizes *relatedness*. To better assess how well the models would perform on a different type of texts, we experimented with two versions of the dataset: WITHOUT HASHTAGS and WITH HASHTAGS.

#### 3.2 Data Splits

We randomly choose 15% of the instances for testing, and then 15% from the rest of the data for development, while maintaining the class ratio (Table 2). During our experiments, we found that applying upsampling on the minority classes (*competence* and *autonomy*) slightly improved the performances of some models, and had no change on others. Thus, we only report the results obtained by using upsampling.

#### 3.3 Task Definition

We approached the problem of detecting basic needs with two different scenarios: (1) as a ternary classification problem (assigning one of the three possible basic needs to each post), and (2) as three binary classification tasks (for each basic need, assigning either *yes* or *no* label). The ternary classification is a more natural choice for this particular dataset, as all instances were annotated with only one of the three basic needs. However, according to the SDT, each person have at all times the all three needs just with different intensities and sat-

Need	ORIGINAL			REPORTED		
	TRAIN	DEV	TEST	TRAIN	DEV	TEST
Autonomy	1248	228	290	2416	404	290
Competence	868	168	204	2416	404	204
Relatedness	2416	404	508	2416	404	508

Table 2: Dataset splits.

isfaction levels (Section 1). It is thus reasonable to assume that some posts will also contain signals of multiple basic needs. Therefore, we also performed three binary tasks which would allow us to model each basic need separately. By using three binary classifiers instead of one ternary, posts could be automatically labelled with none of, or any combination of, basic needs.

#### 3.4 Evaluation Metrics

For both types of classification tasks (binary and ternary), we used the per-class precision, recall, and F<sub>1</sub>-score, and the macro-averaged F<sub>1</sub>-score for evaluating the performances of the models.

#### 3.5 Architectures

In order to assess the importance of both lexical and semantic aspects of texts, we tested various approaches that use different text representations:

- **BOW**: word unigrams and bigrams model with the TF-IDF weighting scheme (Salton and Buckley, 1988) using a Support Vector Machines (Chang and Lin, 2011) classifier with a linear kernel.<sup>2,3</sup>
- **Char-CNN**: a Convolutional Neural Networks (CNN) architecture similar to the one proposed in (Zhang et al., 2015) but using a trainable character embedding layer as input.
- **BiLSTM**: a bidirectional Long Short-Term Memory (BiLSTM) (Hochreiter and Schmidhuber, 1997) neural network that uses Fast-Text word embeddings (Bojanowski et al., 2017) to represent texts. The BiLSTM hidden states are fed to an attention layer (Yang et al., 2016), and then the attention output is processed with a fully connected layer. As an output, a softmax layer is used to obtain the final classification.

<sup>2</sup>We also explored logistic regression, random forest, Naive Bayes, and support vector machines, with different kernels, during the prototyping phase.

<sup>3</sup>Character *n*-grams were also tested but as they did not lead to better performances, we do not report their results.

Model	WITHOUT HASHTAGS									WITH HASHTAGS										
	Autonomy			Competence			Relatedness			F <sub>1</sub> (macro)	Autonomy			Competence			Relatedness			F <sub>1</sub> (macro)
	P	R	F <sub>1</sub>	P	R	F <sub>1</sub>	P	R	F <sub>1</sub>		P	R	F <sub>1</sub>	P	R	F <sub>1</sub>	P	R	F <sub>1</sub>	
BiLSTM	.54	.54	.54	.48	.45	.47	.74	.76	.75	.59	.65	.47	.55	.47	.63	.54	.80	.82	.81	.63
Char-CNN	.50	.64	.56	.58	.24	.34	.73	.79	.76	.55	.61	.61	.61	.54	.54	.54	.82	.81	.82	.66
BOW	.61	.49	.54	.49	.54	.52	.76	.82	.79	.62	.67	.57	.62	.57	.54	.55	.79	.87	.83	.67
BERT	.62	.54	.58	.55	.61	.58	.84	.87	<b>.86</b>	.67	.70	.69	<b>.69</b>	.73	.52	.61	.83	.93	<b>.88</b>	.72
BERT+BiLSTM	.62	.60	<b>.61</b>	.56	.66	<b>.60</b>	.87	.82	.85	<b>.69</b>	.71	.61	.66	.69	.62	<b>.65</b>	.83	.93	<b>.88</b>	<b>.73</b>
Trained human	.78	.70	<b>.74</b>	.69	.88	<b>.77</b>	.88	.72	.79	<b>.77</b>	.78	.70	<b>.74</b>	.69	.88	<b>.77</b>	.88	.72	.79	<b>.77</b>
BERT+BiLSTM	.78	.70	<b>.74</b>	.79	.65	.71	.73	.93	<b>.81</b>	.75	.73	.75	<b>.74</b>	.81	.62	.70	.77	.93	<b>.84</b>	.76

Table 3: Results of the ternary classification task. The last two rows present the results on a subset of the test set that was annotated by a trained human annotator and contains 40 instances of each class.

- **BERT**: the neural language model, well-known for providing text representations that show leading performances on several natural language processing benchmarks (Devlin et al., 2019). We fine-tune BERT and use its hidden representation of the special [CLS] token to represent the full input text and feed it to a softmax output layer.
- **BERT+BiLSTM**: this model combines the previous two approaches. Instead of FastText word representation, the fine-tuned BERT embeddings are post-processed by the BiLSTM architecture defined above. We observed that such architectures help BERT to adapt to the target task and obtain better classification results in scenarios with small training datasets.

## 4 Results and Discussion

### 4.1 Ternary Classification

All models performed noticeably better on the original than on the cleaned dataset, thus supporting our hypothesis that the presence of the hashtags leads to better model performances (Table 3). As expected, the models that are based on transfer learning (BERT and BERT+BiLSTM) performed best. Interestingly, the non-neural model (BOW) outperformed the BiLSTM and Char-CNN models on the *competence* class using the cleaned dataset (F<sub>1</sub>-score of 0.52 against 0.47 and 0.34, respectively).

In all models, most misclassifications were observed between the *competence* and *autonomy* classes. A possible reason for this might lie in the SDT theory, as autonomy and competence are self-originated needs, whereas relatedness includes both self and others (Vansteenkiste et al., 2020).

This might lead to theme/topic overlaps between autonomy and competence due to the self-focus, while relatedness might be easier to distinguish due to including self and the others.

### 4.2 Human Performance and Error Analysis

To assess the expected performance ceiling, we hired a psychologist, well-versed in SDT, provided the annotation guidelines with several examples, and asked to annotate randomly selected 150 instances from the cleaned test set (50 from each class). The annotator was allowed to assign as many classes as needed to each post.

Our guidelines were based on a thorough review of psychology research by Ryan and Deci (2020, 2017a,b, 2000) who studied observable behavioural outcomes. We selected the following cues for each basic need:

- **Autonomy**: focus of initiative, ownership of self-actions, feelings of restriction by any type of external control.
- **Competence**: focus on behaviours associated with mastery, achievements, success, and growth (both positive and negative), search for personal or contextual challenges, well-structured environments, and positive feedback.
- **Relatedness**: focus on spending and appreciating time with significant others, search for community and connection, sense of nurturing and caring for others.

The annotator assigned two classes in 14 cases (9.3%). Some of those were the cases in which our best system (BERT+BiLSTM) made ‘wrong’ prediction, which turned out to be the same as one of the classes assigned by the human annotator

Gold	Predicted	Post
Relatedness	Autonomy	Wishing I was home this Christmas, maybe next year #homesick #holidays
Relatedness	Competence	I work with an amazing team. They work so hard and are so dedicated. Truly a top comms team #proud

Table 4: Examples of penalized predictions which actually caught the secondary signal. For those examples, the human annotator assigned both classes (the gold and the predicted one).

Task	Yes			No			F <sub>1</sub> (macro)
	P	R	F <sub>1</sub>	P	R	F <sub>1</sub>	
Autonomy	.74	.48	.58	.81	.93	.87	.73
Competence	.69	.63	.66	.91	.93	.92	.79
Relatedness	.85	.92	.88	.91	.83	.87	.87

Table 5: Results of the binary classification tasks on the datasets WITH HASHTAGS.

Task	Yes			No			F <sub>1</sub> (macro)
	P	R	F <sub>1</sub>	P	R	F <sub>1</sub>	
Autonomy	.61	.49	.54	.81	.87	.84	.69
Competence	.55	.64	.59	.90	.87	.88	.74
Relatedness	.83	.86	.85	.85	.82	.84	.84

Table 6: Results of the binary classification tasks on the datasets WITHOUT HASHTAGS.

(Table 4). Therefore, we took 120 instances for which the human annotator assigned only one class, and additionally ran our best model on that portion of the test set, to fairly compare its performance with the human performance (the last two rows in Table 3).

### 4.3 Binary Classifications

The results of the best performing architecture (BERT+BiLSTM) on the binary tasks using the datasets WITH HASHTAGS and WITHOUT HASHTAGS are presented in Tables 5 and 6.

To assess the performance of those systems in the real-world scenario, we took 100 random new tweets and ran all three models on them. At the same time, we asked the psychologist to annotate each post (without showing the obtained automatic predictions) by assigning one of the three labels (*no*, *low*, *high*) for each basic need. For example, “@matchbox\_sized Wait, you’ve seen it already? Thought it aired on Sunday nights?” was annotated as *low* for *relatedness*, *high* for *autonomy*, and *no* for *competence*. For the same example, the three best binary models assigned the following probabilities to each of the corresponding classes:

$p(\textit{autonomy}) = 0.88$ ,  $p(\textit{relatedness}) = 0.70$ , and  $p(\textit{competence}) = 0.30$ .

We further investigated whether or not the class probabilities obtained by the binary models were related to the labels assigned by the annotator. On those 100 examples, we found that the manually assigned label *no* corresponds to the  $p(\textit{yes}) \in [0, 0.5)$  (obtained by the models) in 90% of the cases, the manually assigned label *low* to the  $p(\textit{yes}) \in [0.5, 0.75)$  (obtained by the models) in 100% of the cases, and manually assigned label *high* to the  $p(\textit{yes}) \in [0.75, 1]$  (obtained by the models) in 98% of the cases. These findings indicate that it might be possible to use the binary models in a more general setup, i.e. on the posts which are not pre-filtered for containing emotions or needs signals, and on posts that reflect more than one need. Furthermore, it seems that those models could capture the intensity of the signals.

## 5 Conclusions

In this study, we benchmarked the automatic detection of basic motivations on short (Twitter) posts in English, framing the problem as a ternary classification task, as well as three binary classification tasks. On the ternary classification task, our BERT+BiLSTM model performed almost equally well as a trained human annotator.

We showed that modelling this problem as three binary classification tasks, instead of modelling it as one ternary classification task, allows for better applicability of the models. The proposed setup with three binary models assigns none of the basic motivations to those posts without any signal (all three models assign a *no* class), and multiple basic motivations to those posts with signals from multiple motivations (more than one model assigns a *yes* class), achieving a high agreement with the human annotator. We also found a high association between the class probabilities of the binary models and the human-perceived motivation intensities.

## 6 Ethics/Impact Statement

### 6.1 Intended Use

The goal of our experiments was to investigate if there is a possibility to automatically detect basic needs from short posts, and to benchmark this novel NLP task. As we do not have any demographic information in the dataset used, and we did not thoroughly investigate performances of our models on different text types, demographic groups, and in different contexts, we do not encourage the use of these particular models in real-world applications. Instead, the contribution of our study lies in setting the ground for future models of automatic detection of basic needs from short texts, by benchmarking the task with various machine learning architectures on a specific dataset, experimenting with both ternary and binary setups, providing performance ceiling estimation via human annotations, and discussing the usability of presented approaches. Our study thus provides the foundations for future models which, if trained on carefully sampled data (representative data with strict bias control), have the potential to speed up and provide additional quality checks for traditional questionnaire-based basic needs estimation procedures, which are already widely used for: (1) providing supportive information about the user in organizational contexts such as leadership style and team building processes (Rigby and Ryan, 2018); and (2) prompting learner perspectives in educational contexts such as designing motivation-supportive settings and activities (Schneider et al., 2018).

### 6.2 Failure Modes

To try to estimate how the model would perform if trained on different type of data, i.e. non-Twitter data, we evaluated models trained on posts with hashtags and models trained on the same posts but after removing all hashtags. However, it is not certain how would the reported models perform on different types of data, neither whether training models with different data sources would lead to similar results or not. On the used Twitter datasets, we found most misclassifications between *autonomy* and *competence* classes.

### 6.3 Biases

Given that we do not have any demographic information about the authors of the posts in the used dataset, and that the dataset was prefiltered for emotional and needs signals (Alharthi et al., 2017),

the presented models might suffer from various algorithmic biases. Furthermore, it is known that certain age groups or socio-economic groups are more present in Twitter than others (Tufekci, 2014; Morstatter et al., 2014), and that certain personality types are more active on particular media platforms (Goby, 2006).

### 6.4 Misuse Potential

Using automatic detection of basic needs in decision-making processes during hiring and placement could lead to a potential misuse and unfair decisions due to: (1) algorithmic biases and imperfections of the models; (2) giving too much weight to the estimation of basic needs instead of taking it only as one of many aspects of the employee (e.g. personality, educational background) and team work.

Basic needs could be used in combination with other psychological variables (e.g. personality) for marketing and consumer targeting purposes. Tailoring marketing materials for different personalities can be beneficial for consumers by leading them to spend their money on personality-matching items (Matz et al., 2016). However, it can also be misused by leading people to act against their best interests, e.g. by persuading them to gamble (Matz et al., 2016).

### 6.5 Potential Harm to Vulnerable Populations

As any other psychological modelling, when combined with demographic characteristics (e.g. age, gender, socio-economic background), machine learning models could potentially harm vulnerable groups such as immigrants or people with mental health issues. The models could potentially detect people who suffer from psychological and emotional instability, as it is highly likely that those people may be unsatisfied about their basic needs. To avoid such unintended harms, special attention should be given to carefully collecting a representative sample for any intended use (Williams et al., 2018).

## References

Rajwa Alharthi, Benjamin Guthier, Camille Guertin, and Abdulmotaleb El Saddik. 2017. A dataset for psychological human needs detection from social networks. *IEEE Access*, 5:9109–9117.

- Piotr Bojanowski, Edouard Grave, Armand Joulin, and Tomas Mikolov. 2017. Enriching Word Vectors with Subword Information. *Transactions of the Association for Computational Linguistics*, 5:135–146.
- Chih-Chung Chang and Chih-Jen Lin. 2011. Libsvm: A library for support vector machines. *ACM transactions on intelligent systems and technology (TIST)*, 2(3):1–27.
- Beiwen Chen, M. Vansteenkiste, W. Beyers, L. Boone, E. Deci, J. Kaap-Deeder, B. Duriez, W. Lens, Lennia Matos, Athanasios Mouratidis, R. Ryan, K. Sheldon, B. Soenens, S. Petegem, and Joke Verstuyf. 2015. Basic psychological need satisfaction, need frustration, and need strength across four cultures. *Motivation and Emotion*, 39:216–236.
- Daniel J. Christie. 1997. Reducing direct and structural violence: The human needs theory. *Peace and Conflict: Journal of Peace Psychology*, 3(4):315–332.
- Edward L. Deci and Richard M. Ryan. 2000. The “What” and “Why” of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry*, 4(11):227–268.
- Edward L. Deci and Richard M. Ryan. 2011. Levels of analysis, regnant causes of behavior and well-being: The role of psychological needs. *Psychological Inquiry*, 22(1):17–22.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186.
- Faheem Gul Gilal, Jian Zhang, Justin Paul, and Naeem Gul Gilal. 2019. The role of self-determination theory in marketing science: An integrative review and agenda for research. *European Management Journal*, 1(37):29–44.
- Valerie Priscilla Goby. 2006. Personality and online/offline choices: MBTI profiles and favored communication modes in a Singapore study. *CyberPsychology & Behavior*, 9:5–13.
- Steven J. Heine, Darrin R. Lehman, Kaiping Peng, and Joe Greenholtz. 2002. What’s wrong with cross-cultural comparisons of subjective likert scales?: The reference-group effect. *Journal of Personality and Social Psychology*, 82(6):903–918.
- Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural computation*, 9(8):1735–1780.
- Hyunghim Jang, Johnmarshall Reeve, Richard M. Ryan, and Ahyoung Kim. 2009. Can Self-Determination Theory Explain What Underlies the Productive, Satisfying Learning Experiences of Collectivistically Oriented Korean Students? *Journal of Educational Psychology*, 101:644–661.
- Dirk Johannssen and Chris Biemann. 2019. Neural classification with attention assessment of the implicit-association test and prediction of subsequent academic success. In *Proceedings of the 15th Conference on Natural Language Processing, KONVENS*. German Society for Computational Linguistics & Language Technology.
- C. Raymond Knee, Heather Patrick, Nathaniel A. Vitor, Aruni Nanayakkara, and Clayton Neighbors. 2002. Self-Determination as Growth Motivation in Romantic Relationships. *Personality and Social Psychology Bulletin*, 5(28):609–619.
- Ivar Krumpal. 2011. Determinants of social desirability bias in sensitive surveys: a literature review. *Quality & Quantity*, 47(4).
- Sandra C. Matz, Joe J. Gladstone, and David Stillwell. 2016. Money buys happiness when spending fits our personality. *Psychological science*, 27(5):715–725.
- David C. McClelland. 1979. Inhibited power motivation and high blood pressure in men. *Journal of Abnormal Psychology*, 88:182–190.
- Fred Morstatter, Jürgen Pfeffer, and Huan Liu. 2014. When is it biased? assessing the representativeness of twitter’s streaming api. In *Proceedings of the 23rd International Conference on World Wide Web, WWW ’14 Companion*, page 555–556, New York, NY, USA. Association for Computing Machinery.
- Henry A. Murray. 1943. *Thematic Apperception Test*. Harvard University Press, Cambridge, MA.
- Ma Jenina N. Nalipay, Ronnel B. King, and Yuyang Cai. 2019. Autonomy is equally important across East and West: Testing the cross-cultural universality of self-determination theory. *Journal of adolescence*, 78:67–72.
- Nikos Ntoumanis, Jemma Edmunds, and Joan L. Duda. 2008. Understanding the coping process from a self-determination theory perspective. *British Journal of Health Psychology*, 14(2):249–260.
- Wei Peng, Jih-Hsuan Lin, Karin A Pfeiffer, and Brian Winn. 2012. Need satisfaction supportive game features as motivational determinants: An experimental study of a self-determination theory guided exergame. *Media Psychology*, 2(15):175–196.
- James W. Pennebaker and Laura A. King. 1999. Linguistic styles: Language use as an individual difference. *Journal of personality and social psychology*, 77(6):1296–1312.
- Dorian Peters, Naseem Ahmadpour, and Rafael A Calvo. 2020. Tools for wellbeing-supportive design: Features, characteristics, and prototypes. *Multimodal Technologies and Interaction*, 3(4).

- C. Scott Rigby and Richard M. Ryan. 2018. Self-determination theory in human resource development: New directions and practical considerations. *Advances in Developing Human Resources*, 2(20):133–147.
- Richard M. Ryan and Edward L. Deci. 2000. The darker and brighter sides of human existence: Basic psychological needs as a unifying concept. *Psychological inquiry*, 4(11):319–338.
- Richard M. Ryan and Edward L. Deci. 2017a. *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford Publications, New York, USA.
- Richard M. Ryan and Edward L. Deci. 2017b. *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford Publications.
- Richard M. Ryan and Edward L. Deci. 2020. Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology*, (101860).
- Gerard Salton and Christopher Buckley. 1988. Term-weighting approaches in automatic text retrieval. *Information processing & management*, 24(5):513–523.
- Sascha Schneider, Steve Nebel, Maik Beege, and Günter Daniel Rey. 2018. The autonomy-enhancing effects of choice on cognitive load, motivation and learning with digital media. *Learning and Instruction*, 58:161–172.
- Gavin R. Slemp, Margaret L. Kern, Kent J. Patrick, and Richard M. Ryan. 2018. Leader autonomy support in the workplace: A meta-analytic review. *Motivation and emotion*, 5(42):706–724.
- Zeynep Tufekci. 2014. Big questions for social media big data: Representativeness, validity and other methodological pitfalls. *Proceedings of the International AAAI Conference on Web and Social Media*, 8(1).
- Maarten Vansteenkiste, Richard M. Ryan, and Bart Soenens. 2020. Basic psychological need theory: Advancements, critical themes, and future directions. *Motivation and Emotion*, 44:1–31.
- Netta Weinstein and Richard M. Ryan. 2011. A self-determination theory approach to understanding stress incursion and responses. *Stress & Health*, 27(1):4–17.
- Betsy Anne Williams, Catherine F Brooks, and Yotam Shmargad. 2018. How algorithms discriminate based on data they lack: Challenges, solutions, and policy implications. *Journal of Information Policy*, 8:78–115.
- Zichao Yang, Diyi Yang, Chris Dyer, Xiaodong He, Alex Smola, and Eduard Hovy. 2016. Hierarchical Attention Networks for Document Classification. In *Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: human language technologies*, pages 1480–1489.
- Xiang Zhang, Junbo Zhao, and Yann LeCun. 2015. Character-level Convolutional Networks for Text Classification. *Advances in neural information processing systems*, 28:649–657.