A Computational Linguistic Approach to Study Border Theory at Scale

Keywords: border theory, work-life balance, micro-blogs, linguistic analysis, topic modelling

Extended Abstract

Introduction. Social scientists have long been interested in studying the interface between work and life. This stems from the belief that increasing demands of work have begun to dominate areas of life, creating a sense of work-life imbalance. Many theories have been developed to study the work-life interface (Hobfoll 1989; Ashforth, Kreiner, and Fugate 2000; Clark 2000). Among them, *work-family border theory* (Clark 2000) has emerged as the most prominent. This theory explains how employees maintain balance by managing demarcations, or *borders*, between work and non-work life. According to Clark (2000), borders can be physical, psychological, or temporal, and are characterized by their *permeability* (the degree to which work and life domains come in contact) and *flexibility* (the degree to which domains expand and shift overtime). In times of crisis, border permeability and flexibility can change, affecting work-life balance and having implications for well-being. However, these changes cannot be studied using interviews or diaries as these are not at the scale required when societal disruptions occur. We propose a scalable computational methodology for operationalizing border theory. This methodology leverages Natural Language Processing techniques applied to readily available micro-blogging data to automatically detect borders and changes to permeability and flexibility over time. We apply this operationalization to a case study of around half a million tweets posted by nearly a hundred teachers in the UK from the beginning of 2019 to the end of 2022. We longitudinally study and compare the changes in borders between work and life before, during, and after COVID-19 lockdown periods.

Methodology. Instead of interviews of diaries, we collect data from readily available micro-blogging platforms, namely Twitter. We start by collecting all tweets of users selected as part of the study throughout the entire duration of the study period. Subsequently, we preprocess the data by extracting nouns, tokenizing and lemmatizing the tweets. To investigate whether we could identify tweets expressing work and life, we first apply a topic modeling technique known as BERTopic (Grootendorst 2022). BERTopic is a topic modeling framework that uses clustering algorithms to automatically obtain dense topics in a collection of documents, such as tweets. BERTopic computes tweet embeddings for a given input of tweets, before reducing their dimensions and finding dense areas of similar tweets in the vector space using clustering algorithms. We use mini-batch k-means clustering to fix the number of tweet clusters to 5, which we found was a reasonable number to discern patterns of work and life from the input tweets. After fitting the BERTopic model and generating topics, we evaluate the topics both quantitatively and qualitatively. These approaches helped us choose a meta-label representing the key theme found in each topic. Once meta-labels have been chosen, topics are grouped into either work or life categories based on the nature of the label. Having identified topics as either work or life, we develop a method for defining borders. In this study, we focus on operationalizing temporal borders, deferring an analysis of spatial and psychological borders for future research. We determine temporal borders of work and life domains collectively for all users in the following way. First, for each day in a week (Monday to Sunday), we concatenate all tweets posted on that day within a given period. Second, we identify the time (in seconds) at

which these tweets were posted in the day. Third, we calculate the interquartile range (IQR) of tweet frequency. Once we have determined borders, we define metrics to calculate permeability and flexibility, given these are the two key characteristics of borders (Clark 2000). To determine temporal permeability, we calculate the ratio of the overlap of the IQR for work and life and the length of the IQR:

$$\text{Permeability}(d,t) = \frac{Q3_{min}(d,t) - Q2_{max}(d,t)}{Q3_{max}(d,t) - Q2_{min}(d,t)} * 100 \tag{1}$$

where $Q2_{min}(d,t)$ and $Q2_{max}(d,t)$ are the minimum and maximum of the second quartile and $Q3_{min}(d,t)$ and $Q3_{max}(d,t)$ are the minimum and maximum of the third quartile on a given day d for a given period t. We define temporal flexibility using two complementary metrics: the distance of the IQR and the shift in the IQR for work and life tweets over time. The flexibility-distance (F-dist) and flexibility-shift (F-shift) are computed as follows:

$$\text{F-dist}(i,d,t) = Q3(i,d,t) - Q2(i,d,t) \tag{2}$$

$$\text{F-shift}(i,d,t) = [|Q2(i,d,t) - Q2(i,d,t+1)|, |Q3(i,d,t) - Q3(i,d,t+1)|] \tag{3}$$

for $i \in \{work, life\}$ for a given day d in a given period t. Prior literature has shown a relationship exists between work-life borders and well-being (Sahay and Wei 2021; Carvalho et al. 2021). Therefore, measuring well-being is an important step to complete the operationalization of border theory. We adapt the measure of subjective well-being proposed by Zhou et al. (2020) to compute the daily subjective well-being score for all tweets. We apply the above methodology to study the language of 447,668 tweets posted by 95 teachers in the UK from the beginning of 2019 to the end of 2022, before, during, and after COVID-19 lockdown periods.
Results. The key finding from our work is that border flexibility increases in line with well-being during the second COVID-19 lockdown for all UK teachers considered in our case study. As work-life borders become more flexible (changes in the IQR of tweet frequency, as displayed in Figures 1 and 2), well-being shifts towards more positive values (displayed in Figures 3 and 4). This however is not the case with border permeability, which falls as well-being increases during the second COVID-19 lockdown.

References

Ashforth, B. E.; Kreiner, G. E.; and Fugate, M. 2000. All in a day's work: Boundaries and micro role transitions. *Academy of Management Review*, 25(3): 472–491.

Carvalho, V. S.; Santos, A.; Ribeiro, M. T.; and Chambel, M. J. 2021. Please, do not interrupt me: work–family balance and segmentation behavior as mediators of boundary violations and teleworkers' burnout and flourishing. *Sustainability*, 13(13): 7339.

Clark, S. C. 2000. Work/family border theory: A new theory of work/family balance. *Human Relations*, 53(6): 747–770.

Grootendorst, M. 2022. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv preprint arXiv: 2203.05794*.

Hobfoll, S. E. 1989. Conservation of resources: A new attempt at conceptualizing stress. *American Psychologist*, 44(3): 513.

Sahay, S.; and Wei, W. 2021. Work-family balance and managing spillover effects communicatively during COVID-19: Nurses' perspectives. *Health Communication*, 1–10.

Zhou, X.; Jin, S.; and Zafarani, R. 2020. Sentiment paradoxes in social networks: Why your friends are more positive than you? In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 14, 798–807.
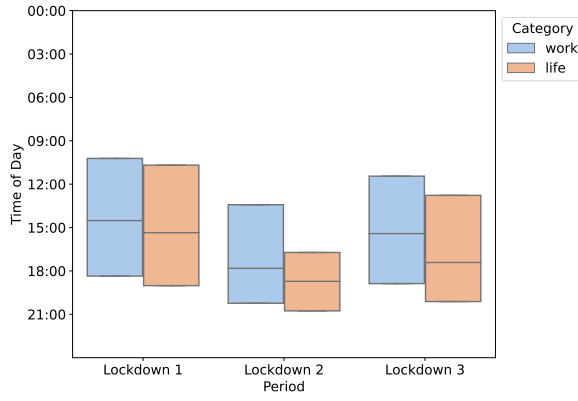
Figure 1: IQR of tweet frequency for Scottish teachers on weekdays, for each COVID-19 lockdown period.
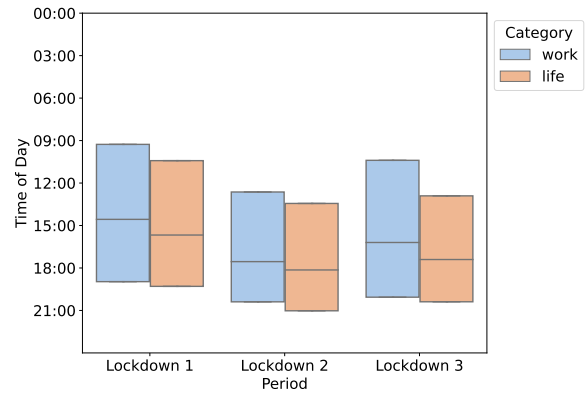

Figure 2: IQR of tweet frequency for Irish teachers on weekdays, for each COVID-19 lockdown period.
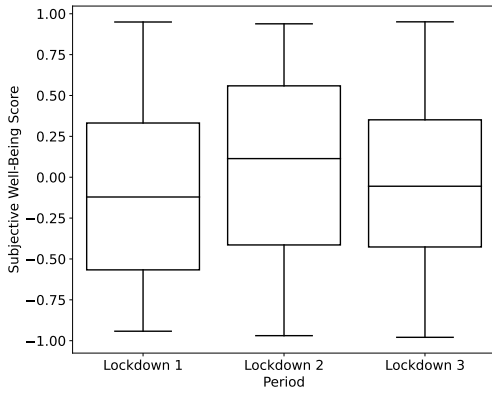

Figure 3: Subjective well-being of Scottish teachers on weekdays, for each COVID-19 lockdown period.
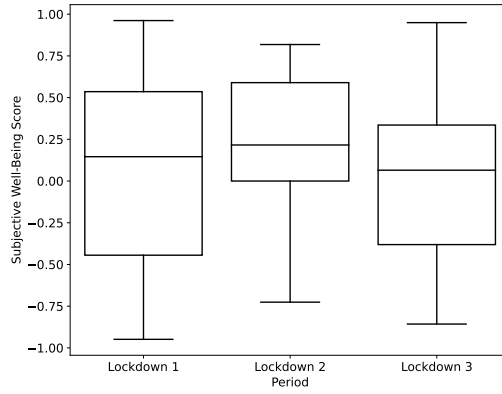

Figure 4: Subjective well-being of Irish teachers on weekdays, for each COVID-19 lockdown period.