Erudit Report 01

Measuring Burnout, Engagement, Turnover

1 March 2022

Overview

In this report, we review the quality of Erudit's data analysis in three sections:

1. Evaluation of the psychological measurements for burnout, engagement, and turnover used by Erudit's psychologists to train its Artificial Intelligence network
2. Assessment of data analysis by Erudit's psychologists based on said psychological measurements
3. Comparison of traditional psychological methodologies and Erudit AI

This report will show that the results produced by Erudit's psychologists are consistent with literature, which serves as a strong foundation for achieving high performance in machine-learning technology.

Introduction

Managing behaviors and feelings in the workplace brings tremendous value to the team, yet it comes with its own set of challenges (Eisman E., 2016). Among the problems that a manager may find is how to best measure work-related emotions that may result in undesirable behaviors, such as the burnout or turnover risk, which may lead to a absenteeism and a drop in productivity.

Traditional methodologies, such as interviews and questionnaires, do not enter into an employee's everyday routine and may not accurately portray true sentiment. They often lack an employee's self-introspection and are easily subject to lies. It is here where Artificial Intelligence sets a new standard, as it offers solutions to monitor the wellbeing of employees and managers of a company within their daily work routines and conversations.

Erudit AI has tremendous respect for data and believes that transparency can only improve our software. We seek to ensure that Erudit's machine learning algorithms contain scientific rigor and high quality to best measure human behavior through communications.

In this report we provide the background on our psychological measurements and dimensions, review evidence of how accurate and reliable our data is, as well as compare our software versus traditional methodologies.

Traditional Psychological Measurement

Erudit counts on a certified team of psychologists to ensure the scientific precision of our Artificial Intelligence models. In this section, we discuss the traditional means of measuring the three psychological domains of major relevance to companies: the risk of burnout, the level of engagement towards the company, and the risk of turnover.

Burnout is a syndrome associated with prolonged, overwhelming exposure to work-related stressors, and thus, is associated with depression, anxiety, and exhaustion. It is characterized by a negative emotional response generated in the psychological, emotional, and even physical sphere of an individual. Our team is able to measure the degree of employee burnout risk based on **Maslach Burnout Inventory** (Maslach, C. et al., 1997), which is a widely used, and validated method to measure burnout in the workplace.

When it comes to **Turnover**, we measure the voluntary act of an employee abandoning the company, which is mainly related to job dissatisfaction due to the discrepancy between the employee's expectations and the actual experience received. Our psychologists are able to measure this using the **Turnover Intention Scale** (Bothma, C. et al., 2013), which is a common instrument to measure it in the workplace, having been validated worldwide.

Lastly, the **Engagement** dimension is an employee's affinity for an organization and its goals and values, which is associated with personal commitment towards the company and a psychological state of wellbeing. The engagement level is being measured based on the **Utrecht Work Engagement Scale** (Schaufeli, W et al., 2003), which is the most widely accepted, globally validated method of measuring engagement in organizations.

These three psychological measurements are traditional ways of measuring wellbeing, used by psychologists when assessing employees. It is important to determine whether our AI can produce results consistent with traditional methods, as observed in patterns and correlations.

Data analysis

In order to assess the relevance of the burnout risk, turnover risk and engagement level dimensions, we performed a data analysis by asking trained psychologists to rate the presence and intensity of these dimensions in random messages. The messages were extracted from social media (reddit, twitter) in English and 100 of them were randomly sampled. See for instance different kinds of messages with varying intensities:

- Low intensity:
  "*I think this is a fair wage for where I live and the work I do. I have always been told that hospice is a good "first social work job" and I've found that to be true. I would like to make more eventually but for right now I'm comfortable with this salary.'"*"

- Moderate intensity:
  "Is there a back story of other stuff, or this just started happening? It sounds like you're dissociating."

- High intensity:
  "THEN WHY WOULD YOU HIRE ME KNOWING I CAN'T WORK THAT DAY?! This is just... baffling! I'm so mad! I've been mad all dang day!"

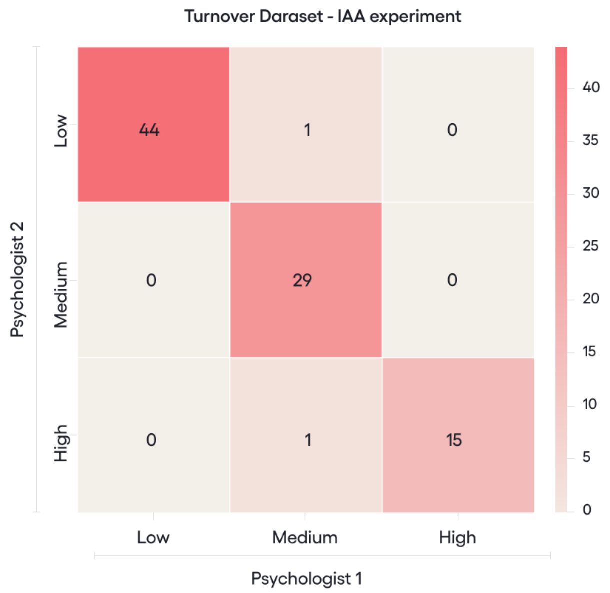
For each message, we asked two psychologists to label, in parallel, both the intensity of the speaker's engagement level (on the scale Low, Medium, High), and whether the message provided any clues about turnover and burnout risk (again, on the scale Low, Medium, High). The psychologists were also allowed to skip if the message was too noisy and uninterpretable.

The results show two interesting findings. First, we computed the Cohen's kappa which measures the level of agreement between the two psychologists. It is a value between 0 and 1. If the value is 1, it means that both psychologists are in total agreement and they rate the same level for the same messages; if the value is 0 it means that they never agree. The goal is to check whether each dimension can be reliably interpreted without ambiguity from the messages. The results (Table 1) show that there is a high level of agreement between the two psychologists for each dimension, meaning that each dimension can be reliably analyzed from the messages alone.

Table 1

	Kappa
Burnout	0.87
Engagement	0.95
Turnover	0.96

Figure 1



We further explored the confusion matrices for each dimension, which allows us to see at a glance on what level the psychologists are in agreement (see, for example, the turnover in Fig. 1). From this chart, we can notice that the psychologists are mostly in agreement, except in two cases: in one message, Psychologist1 rated Medium, while Psychologist2 rated Low, and for another, Psychologist1 rated Medium while Psychologist2 rated High.

The **second** most important finding is obtained after computing Pearson's correlation between the different dimensions. For two given dimensions, we can check whether the different levels are correlated. It is a value between -1 and 1. If it is 1, it means that the two dimensions are exactly the same and that the psychologists are actually rating the same phenomena. If it is -1, it means that the two dimensions are diametrically opposed, for instance each time the turnover is High, the engagement is Low. If it is 0, it means that the two dimensions are completely unrelated. We restrict ourselves to the messages for which both the psychologists agree. We show the correlation in Table 2. along with the p-value, which measures the probability of having these results if the metrics are unrelated. The lower the p-value, the stronger the results. We also provide the number of messages (N) that could be compared between the dimensions.

Table 2

	Correlation	p-value	N
Burnout vs Turnover	0.802	2.04e-16	68
Burnout vs Engagement	-0.778	1.01e-14	67
Turnover vs Engagement	-0.712	4.92e-12	70

The results show that the dimensions are highly correlated or anti-correlated. Burnout and turnover show the highest amount of correlation meaning it is very likely that there exists a relationship between the two dimensions as rated. The very low p-value strengthens the likelihood that this is the case. We also found out that engagement is anti-correlated with burnout or turnover, meaning that when engagement is low, there is a much higher chance for turnover or burnout happening.

Taking this together, **we can confidently say that our results are consistent with literature that focuses on the relationship between engagement and burnout** (see for instance Wu et al. 2020). However, note that these other literature references rely on traditional methodologies (questionnaires, interviews…), which are costly and difficult to set-up, among others. On the contrary, our software relies on the analysis of text messages and natural language processing which can be fruitful to understand wellbeing in companies without the issues the aforementioned method presents. Despite both approaches having benefits, Erudit AI is adjusted to modern times as it adapts to current day necessities such as daily management of our workers. This scientific report highlights that we are one step closer to ensuring data quality - an important step towards achieving high performance in machine-learning technology. All in all, these achievements set the standards for future linguistic improvements and research (such as model validation) in our software.

Comparison

In studying the consistencies, we also noted differences in using traditional methodologies as compared to using artificial intelligence, which we outlined in Table 3 below.

Table 3

	Erudit	**Traditional methodologies**
Speed	Daily update	Yearly or quarterly update
Time cost	No time cost	At least 1 psychologist must interview employees (an average of 45 min per person), correct the psychological questionnaires and interpret them translating to weeks or months until evaluation results are ready to be presented.

	Erudit	Traditional methodologies
Cost	Subscription cost	Personnel cost for at least 1 employee qualified to conduct psychological assessments
Framework requirements	- Use communication channels (Slack, Microsoft Office, Gmail…)	- Qualified evaluation personnel - Psychological instruments - Results interpretation and report
Bias of sources	- Difficulty with unclear messages - Difficulty in evaluating when there is lack of communication	- Difficulty of information transparency (tell lies, pressure to maintain a certain department level, lack of self-awareness or introspection…). - Psychologists own bias towards employees (discrimination, pre-judging, tiredness, identifying with employees struggles…). - Difficulty to access work-from-home employees
Types of measurements	Artificial Intelligence neural network (internal linguistic model)	Self-assessments (questionnaires, surveys), individual interviews, group interviews
Measures	Risk of turnover and burnout, and engagement level	Burnout, personality, cognitive abilities, job satisfaction…
Validity	Psychologists must be trained in the specific AI models, and must go under a self-agreement test to ensure data quality. They must be certified in Psychology	Psychologists must be trained in the specific instruments being used, and must be certified to conduct evaluation and assessments with people. They must be certified in Psychology
(Worad L., 2019; Tay, L. et al., 2022)		

It is safe to conclude that a high level of accuracy in psychological measurement and diagnosis can be achieved by both traditional methods and through artificial intelligence. The stark difference lies in the amount of time it would take to see the results. The time factor is even greater for companies with a large number of employees. Time would also affect cost, as traditional methods rely on personnel.

In conclusion, for executives that are interested in seeing accurate and timely psychological measurements of their entire organization, artificial intelligence provides a quicker, potentially more cost effective, better solution.

References

- Bothma, C. F., & Roodt, G. (2013). The validation of the turnover intention scale. *SA Journal of Human Resource Management*, *11*(1), 1-12.
- Eisman E. (2016) Expanding the Practice Spectrum - APA. State Leadership Conference
- Maslach, C., Jackson, S. E., & Leiter, M. P. (1997). *Maslach burnout inventory*. Scarecrow Education.
- Schaufeli, W. B., Bakker, A. B., & Salanova, M. (2003). Utrecht work engagement scale-9. *Educational and Psychological Measurement*.
- Tay, L., Woo, S. E., Hickman, L., Booth, B. M., & D'Mello, S. (2022). A Conceptual Framework for Investigating and Mitigating Machine-Learning Measurement Bias (MLMB) in Psychological Assessment. *Advances in Methods and Practices in Psychological Science*, *5*(1), 25152459211061337.
- Worad L. (2019, December 17). How Often Should Employees Have Performance Reviews? *Lawpath*.
- Wu, F., Ren, Z., Wang, Q., He, M., Xiong, W., Ma, G., ... & Zhang, X. (2021). The relationship between job stress and job burnout: the mediating effects of perceived social support and job satisfaction. *Psychology, health & medicine*, *26*(2), 204-211.