Healthcare and Medical Devices, Vol. 130, 2024, 132–139
https://doi.org/10.54941/ahfe1004846

Words About Work Performance and the Company They Keep

Tim Arnold[1,2], Helen J. A. Fuller[1], and Ruth M. Reeves[1,3,4]

[1]Human Systems Integration, Clinical Informatics and Data Management Office, Office of Health Informatics, Veterans Health Administration, Department of Veterans Affairs Washington, DC, USA
[2]University of Michigan, College of Pharmacy Ann Arbor, MI, USA
[3]Tennessee Valley Healthcare System, Department of Veterans Affairs, Nashville, TN, USA
[4]Vanderbilt University Medical Center Department of Biomedical Informatics, Nashville, TN, USA

ABSTRACT

When approaching patient safety questions, initial perspectives and philosophies frame the approach and how work is studied. When there is an over emphasis on an individual's work performance related to safety, this perspective may distract from considering latent design and systems issues. Additionally, it is of concern among safety advocates that narrow perspectives can dominate safety conversations. Advocates for safety continue to deliver messaging that challenges traditional ways of thinking about safety. An important approach for raising awareness about these different perspectives on safety and developing communication approaches between diverse communities is learning about their respective word usage. With the goals of more effectively communicating concepts and aiding in shaping dialog specifically related to patient safety, we sought to understand how best to leverage existing language use intended to represent patient safety and adjacent concepts. It is through these learning experiences that it becomes possible to design and study how translational representations may extend or grow collective views and interprofessional approaches to patient safety.

Keywords: Computational linguistics, Natural language processing, Patient safety, Human error, Resilience

INTRODUCTION

In the study of safety and specifically in patient safety, there are many perspectives on the concept of *human error*. Safety advocates have shared their concerns about the risks of using the term, label, or construct of *human error* when referring to undesirable performance in work systems (Woods et al., 1994; Hollnagel & Amalberti, 2001; Long, 2016; Arnold et al., 2021; Le Coze, 2022). Yet, this construct is still often used when communicating about undesirable work performance and accidents. Although, there has been a call for renewed or refreshed conceptualizations of work performance and safety (Safety II), it has only slowly taken hold. Though at times it can be important to discuss how inadvertent or incorrectly chosen human actions

may contribute to unintended outcomes, focusing primarily on humans as fallible can undermine safety and can come with irreparable costs to the people directly impacted and to the wider society.

What are some approaches for reflecting on these divergent philosophies critically and collectively? The study of historical trajectory is one way for reflecting on different views. Arnold and colleagues (2021) outline an interactive interprofessional patient safety fellowship experience for facilitating group discussion into the history of patient safety. This experience walks learners through a history of patient safety. Traditional views and efforts, sometimes referred to as *Safety I*, emphasize "what went or could go wrong?" and suggest "barrier methods" for solutions. On the other hand, emerging views and efforts consistent with resiliency engineering concepts, sometimes referred to as *Safety II*, focus on and learn from 'what is going right?"

Another way to reflect on different viewpoints is by describing the language used to convey patient safety concepts and share in community conversations through presenting and publishing. To facilitate awareness of the various use patterns of patient safety and related concepts, understanding the subtleties of language could help with explaining varied word usages and with how dialog is shaped. Describing and fostering language awareness of the semantics of *human error* and words associated with Safety II such as *resiliency* or *resilience* and *resilience engineering* may help facilitate and shape dialogue across divergent communities around these words and the concepts they are intended to represent. Conversations are ongoing about the varied meanings of *human error* and *resiliency* or *resilience* and have been for some time. To bring forth greater understanding of how these distinct but related patient safety perspectives have interacted, our work focuses on understanding semantic neighbourhoods and explores words and word connotations that accompany select terms and the sentiment of the encompassing sentences to examine differences in word usage. To aid our investigation from a healthcare perspective, we extracted and examined Medline abstracts that contained the terms *human error*, *resiliency* or *resilience*, and *Safety II* or *resilience engineering*.

APPROACH

We used PubMed to extract Medline abstracts that contained the key terms listed above. Abstracts were imported into a Natural Language Processing (NLP) visualization tool called Orange (Demšar et al., 2013). Natural language pipelines within Orange were designed iteratively exploring a variety of features, methods, and visualizations. Previously, we illustrated team-based text analytics exploration of user comments including iterative participatory walkthroughs of a variety of NLP visualizations (Arnold et al., 2023). These methods included reviewing keyword frequency and significance measures, word co-occurrence network visualizations, linguistic features most important to text classifiers, corpus and concordance viewers, topic models, word clouds, and findings from sentiment analysis. The work described here expands this prior work in team-based text analytics.

To accomplish our analysis, we constructed three corpora with titles and/or abstracts that contain (1) *human error*, (2) *resiliency* or *resilience*, and (3) *Safety II* or *resilience engineering*.

We pre-processed the texts, extracted linguistic characteristics as well as co-occurring and contrastive features, developed visualizations, and subsequently reviewed features and select accompanying contexts. Pre-processing included sentence extraction, transforming to lower case, tokenization at the word level, and removal of function words. Word stemming was not performed.

Sentiment analysis was performed on sentence extracts within our working corpus using SentiArt. This was performed in Orange using the sentiment analysis widget. The SentiArt "tool uses vector space models together with theory-guided, empirically validated label lists to compute the valence of each word" (Jacobs, 2019). Then an overall sentiment for each sentence is determined and assigned.

Next, a naïve bayes text classifier was trained to distinguish between sentences that included the word *error* and sentences including *resiliency* or *resilience*. Linguistic features (e.g., words, tokens, parts of speech, semantic classes, etc.) that were strongly associated with each class as defined by the naïve bayes text classifier were collected (Table 1). Finally, to inform our investigative process, we collected uniquely co-occurring words and reviewed associated network visualizations (Arnold, 2022) to identify possible distinctions in word usage between the groups. During these activities, the team alternated between the NLP visualizations and the narrative text, in search of an ideal balance between contextual understanding and patterns among normalized representations of the narrative text.

FINDINGS

A total of 2400 titles/abstracts were extracted that included the term *human error*, 4552 titles/abstracts were extracted that included the terms *resiliency* or *resilience*, and 181 titles/abstracts were extracted in which the full text contained *Safety II* or *resilience engineering*. Because of their relative larger sizes, corpora (1) *human error* and (2) *resiliency* or *resilience* were the primary focus of the analysis. The much smaller (3) *Safety II* or *resilience engineering* corpus was used as a reference while exploring the other two corpora. Table 1 enumerates some of the significant word frequencies per corpus type.

We extracted sentences from both the (1) *human error* and (2) *resiliency* or *resilience* corpora. Sentences were isolated from the (1) *human error* corpus if they contained the word "error" and from the (2) *resiliency* or *resilience* corpus if sentences contained the words "resiliency" or "resilience." We decided to include sentences in the (1) *human error* corpus that contained the word "error" because of syntactic variation among abstracts that were included in the extract due to MESH term assignment. For example, the following sentence excerpts suggest referencing some aspect of the concept of *human error*: "the extent to which human intervention can exacerbate the problems by introducing new errors" and "the extent to which prescribing error rates are influenced by" (Hughs et al., 2010).

After sentences were extracted and filtered by keywords, sentiment analysis was performed on the two classes. Preliminary findings suggest that sentences that contain the word "error" are much less likely to be classified as positive (37%) during sentiment analysis compared to sentences that contain "resiliency" or "resilience," which are more likely to be classified as positive (85%). Figure 1 illustrates word co-occurrences with *human error* sentence extracts, while Figure 2 illustrates that of *resilience* or *resiliency* sentence extracts.

For the text classifier activity, we were most interested in the features that were important for drawing a boundary between the two classes. Sentences that contained the word "error" also included but were not limited to the following words that were important to the text classifier: *accidents*, *mistake*, *wrong, incorrect, manual,* and *automated*. Some of these words suggest the presence of a judgment statement. Whereas sentences that included the word "resiliency" or "resilience" were more likely to include words such as *family*, *communities*, *coping*, and *mindfulness* (Table 1).

Table 1. Important features identified by the text classifier to distinguish classes.

Human Error Corpus			**Resiliency / Resilience Corpus**		
Word	***Error Sentences***	***Resiliency Sentences***	***Word***	***Error Sentences***	***Resiliency Sentences***
error	4264	10	resiliency	0	6958
human	3397	90	resilience	11	3103
errors	1821	5	family	0	586
accidents	225	0	ego	0	510
prone	216	2	coping	2	446
manual	168	2	youth	0	263
consuming	124	1	women	8	209
equipment	114	2	adolescents	0	189
automated	107	1	mindfulness	0	144
wrong	97	0	communities	0	140
transfusion	76	0	personality	0	131
accident	73	0	adversity	0	112
aviation	69	0	depressive	0	111
computer	62	1	stressors	0	109
mistakes	61	0			
correct	52	0			
manually	41	0			
calculation	40	0			
incorrect	34	0			
anaesthesia	33	0			
anesthesia	33	0			
sherpa	33	0			

For the unique co-occurring network visualization activity and as expected, there were similar features important to the text classifier. We were able to focus in on additional contextual features for probing narratives. Figures 1

and 2 offer a static view of how a team may interact with these visualizations to collect information for probing the larger narratives. (Figures are included in this paper to help readers visualize how a team could interact with dynamic views and are not intended to provide deeper context around language use.)

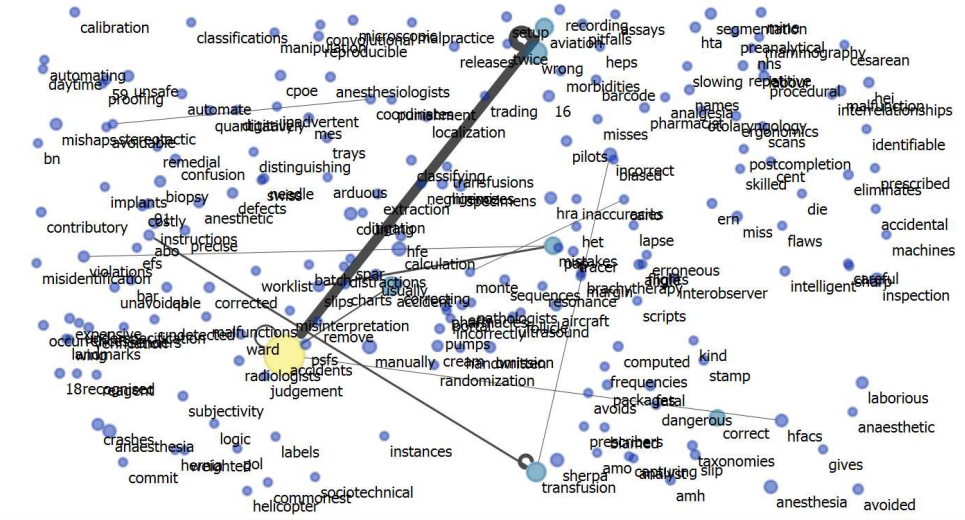

Figure 1: Unique co-occurring word network from *human error* sentence extracts.

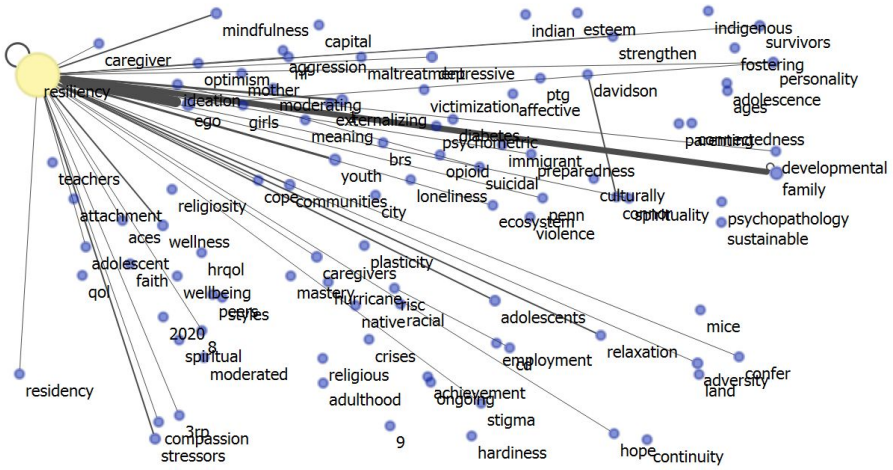

Figure 2: Unique co-occurring word network from *resiliency* or *resilience* sentence extracts.

DISCUSSION

Findings from this work suggest that sentences referring to some aspect of the concept of human error are less likely to be categorized as positive sentiment and respectively contain companion words that suggest a focus on judgment

and blame regarding people as a primary contributor to system fragility. By contrast, sentences referring to *resiliency* or *resilience* are more likely to be classified as having positive sentiment and contain companion words describing community, family, aspects of coping with adversity, and characteristics that may contribute to resilient systems designs. Additional and extensive linguistic work could be helpful for describing how people write and talk about these concepts and may help explain divergent semantics and viewpoints.

Sentences from abstracts that contained the words "resiliency" or "resilience" were diverse in description and often occurred in the context of explaining some characteristic of mental health and resiliency at the individual level. Although the use of *resiliency* or *resilience* in this context is not necessarily referring to that of *resilience engineering*, there is a possible opportunity for the safety science and human factors communities to learn from the language of healthcare. The use of the term *resiliency* or *resilience* and companion words when used to describe aspects of healthcare may refer to patients' capabilities and characteristics as community members for adapting to changing conditions, uncertainty, and adversity. Borrowing words and expressions, when sensible, from clinical and social domains for sharing perspectives on systems safety may resonate with healthcare professionals and help facilitate communication.

We would like to clarify a few points about the findings from the text classifier task. Both spellings for *anaesthesia* and *anesthesia* were present in the text, and it is only a coincidence that they have identical counts for the extracts described in Table 1. *SHERPA* is an acronym for system human error reduction and prediction approach.

Limitations to this work include challenges due to time constraints for walking through the NLP pipelines, visualizations, and narrative texts. To improve team access and engagement, we are working on the right mixture of static collections and development of spaces for shared access to interactive NLP interfaces with capabilities to annotate and identify branching team and individual interactions with the system.

Much like taking photos on a journey, a screen recorder or reporting instrument can be useful for recording NLP pipelines and visualizations and context along the way. Review teams can view, collect, and view again a variety of interactive and/or static visualizations that traverse the network of linguistic features within a given narrative text or across corpora using different techniques. Team members can explore and search for information of interest while iterating on NLP pipelines. This exploratory process can help reviewers derive a perspective on the text through a variety of vantage points, promoting an understanding of common patterns across the body of the text.

In migrating to other conceptual environments for studying safety, further exploration will require adding linguistic expressions for communicating these concepts across new communities of speakers. Expressions as used by people in these communities could help with this transition. Describing and discussing the words these communities associate with concepts about safety could help fuel critical dialogue and identify new ways of reflecting on the shared world.

CONCLUSION

Sentences referring to the concept of human error are less positive and contain companion words focused on judgment and blaming people as primary contributor to systems brittleness. By contrast, sentences referring to *resiliency* or *resilience* are more likely to be positive while containing companion words related to coping with adversity and characteristics of resilient systems designs.

ACKNOWLEDGMENT AND NOTES

We thank those working at the Veterans Health Administration. There were no relevant financial relationships nor any source of support in the forms of grants, equipment, or drugs. The authors declare no conflicts of interest. The opinions expressed in this article are those of the authors and do not necessarily represent those of the Department of Veterans Affairs or the United States government.

REFERENCES

Arnold, T. (2022, September). Bridging conversation islands to connect healthcare: Introducing unique co-occurring word networks to find distinct themes. In Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care (Vol. 11, No. 1, pp. 44–47). Sage CA: Los Angeles, CA: SAGE Publications.

Arnold, T., Fuller, H. J., Gilman, S. C., & Gunnar, W. P. (2021). How might we think about safety? Inviting deeper reasoning through elaborative inquiry. *In Advances in Human Factors and Ergonomics in Healthcare and Medical Devices: Proceedings of the AHFE 2021 Virtual Conference on Human Factors and Ergonomics in Healthcare and Medical Devices*, July 25–29, 2021, USA (pp. 58–64). Springer International Publishing.

Arnold, T., Fuller, H., Yee, S., Reeves, R., & Nazeer, S. (2023). Team-based text analytics for health information systems learning. *In: Jay Kalra and Nancy Lightner (eds) Healthcare and Medical Devices. AHFE (2023) International Conference. AHFE Open Access*, vol. 79. AHFE International, USA. http://doi.org/10.54941/ahfe1003486

Demšar, J., Curk, T., Erjavec, A., Gorup, Č., Hočevar, T., Milutinovič, M.,... & Zupan, B. (2013). Orange: data mining toolbox in Python. *The Journal of Machine Learning Research*, 14(1), 2349–2353.

Hollnagel, E., & Amalberti, R. (2001, June). The emperor's new clothes: Or whatever happened to "human error". *In Proceedings of the 4th International Workshop on Human Error, Safety and Systems Development* (pp. 1–18). Linköping University.

Hughes, C., Travaglia, J. F., & Braithwaite, J. (2010). Bad stars or guiding lights? Learning from disasters to improve patient safety. *BMJ Quality & Safety*.

Jacobs, A. M. (2019). Sentiment analysis for words and fiction characters from the perspective of computational (neuro-) poetics. *Frontiers in Robotics and AI*, 6, 53.

Le Coze, J. C. (2022). The 'new view'of human error. Origins, ambiguities, successes and critiques. *Safety Science*, 154, 105853.

Long, R. (2016). The social psychology of risk, safety and leadership maturity. *AusIMM Bulletin*, (Apr 2016), 26–28.

Tallentire, V. R., Hale, R. L., Dewhurst, N. G., & Maxwell, S. R. (2013). The contribution of prescription chart design and familiarity to prescribing error: a prospective, randomised, cross-over study. *BMJ Quality & Safety*.

Woods, D., Dekker, S., Cook, R., Johannesen, L., & Sarter, N. (1994). *Behind Human Error*. CRC Press. Available at https://apps.dtic.mil/sti/pdfs/ADA492127.pdf.