

# Abstract Poster presentation EARLI SIG1+4

## Unravelling language use of narrative comments in ePortfolios: a text analysis approach

Sofie Van Ostaeyen, Orphée De Clercq, Mieke Embo, Tammy Schellens, Martin Valcke

### Abstract

Narrative comments reported in ePortfolios allow to ground competency assessment and development during workplace learning in healthcare education. However, not all narrative comments are considered effective. The present study is a first step in exploring whether automatic text analysis could support the authors of narrative comments. Therefore, the aim of this study was to determine whether high-quality narrative comments can be characterised by certain language use. First, 2,348 narrative comments retrieved from ePortfolios of 149 Flemish healthcare students were manually labelled to determine their quality. Subsequently, these comments were analysed using the Linguistic Inquiry and Word Count (LIWC) tool. The results reveal that word count is the single lexical dimension which can be associated with quality differences. The LIWC dictionary categories did not vary across low-, moderate- or high-quality comments. This suggests potential shortcomings in the currently available LIWC dictionary categories. More specialized dictionary categories might be required.

### Extended summary

#### **Theoretical background**

The ongoing move to competency-based education has challenged assessment and evaluation strategies within healthcare education. While assessments focus largely on numeric scores, more attention is being paid to the potential of narrative comments (Ginsburg et al., 2017). Narrative comments - often written down in ePortfolios - have been recognised as useful and valid data sources to assess and direct student performance during workplace learning. In addition to evidence pointing to the more reliable nature of narratives, research also emphasizes how narratives seem easier to interpret for students, stimulate reflection, and are considered more effective to provide constructive feedback; as compared to numeric scales (Ginsburg et al., 2017; Marcotte et al., 2019). Nevertheless, healthcare students report narrative comments being nonspecific, generic, vague and ineffective (Branfield Day et al., 2020; Shaughness et al., 2017). This suggests that the authors of those narratives would also benefit from receiving feedback in order to improve. However, analysing the quality of narrative comments and providing feedback to those authors is time-consuming. Recent technological advances – in particular related to automatic text analysis - might be helpful to meet these challenges. These analytics do not interpret the content of narrative comments but analyse the literal text in terms of language use. Therefore, the aim of this study was to identify whether high-quality narrative comments reflect a certain language use.

#### **Method**

A two-stage study was set up. In the first stage, 2,348 narrative comments retrieved from ePortfolios of 149 Flemish (Belgium) healthcare students (specialist medicine, general practice, midwifery, speech therapy and occupational therapy) were manually labelled in the annotation platform

INCEPTION (Klie et al., 2018), according to four quality criteria (performance, judgment, elaboration and improvement). In this way, narrative comments were labelled as of low (meeting none or one criterion), moderate (meeting two or three criteria) or high quality (meeting all four criteria). To ensure reliability, the codebook used was tested by three researchers and a subset of the comments (n=100) was double coded by two researchers. In the second stage, all narrative comments were analysed using the Linguistic Inquiry and Word Count (LIWC) tool (Pennebaker et al., 2015). This software quantifies language use along multiple lexical dimensions. These dimensions can be summations (e.g. word count) or percentages of words that match available LIWC dictionary categories such as part of speech (e.g. nouns, negations) and word categories (e.g. cognitive processes, time orientations). Consequently, an LIWC analysis results in a categorization of all words used following the available lexical dimensions, along with their summations or relative percentages.

## Results

After the first study stage, 29% of the comments were labelled as of low quality, 56% as of moderate quality, and 15% as of high quality. In the second study stage, the three quality sets of comments were analysed by applying the Dutch lexicon of the LIWC (van Wissen & Boot, 2017). This resulted in an overview of the summations and relative percentages per lexical dimension for all narrative comments. To study potential differences due to variations in quality level, the average per lexical dimension across all comments at each quality level was calculated. When looking at the average word count across the three quality levels, we observed a clear difference. Low-quality comments counted on average 20 words, which was three times less as compared to moderate-quality comments (n=59) and more than eight times less compared to high-quality comments (n=165). When comparing the highest and lowest average percentages per dictionary category, only minor differences could be identified. The part-of-speech and word categories with percentages higher than 10% did not differ when comparing low-, moderate- and high-quality comments. These categories included dictionary words, total function words, words with more than six letters, punctuation marks, prepositions, common verbs, relativity, total pronouns, cognitive processes, present focus, adverbs, and social processes. Similarly, no differences could be observed when comparing the five dictionary categories with the lowest relative percentages, being anger, religion, death, swear, and non-fluencies. Only small differences in relative percentages could be identified.

## Theoretical and educational significance

The LIWC analysis helped quantifying language use in narrative comments that vary in feedback quality. Our results showed that word count was the only lexical dimension that differed across the quality levels, which does not really offer insights into language use. Differences were not present when comparing the currently available LIWC dictionary categories while looking at low-, moderate- or high-quality comments. This suggests potential shortcomings in the currently available dictionary categories. More specialized dictionary categories might be needed to identify the unique language use of high-quality narrative comments.

## References

- Branfield Day, L., Miles, A., Ginsburg, S., & Melvin, L. (2020). Resident Perceptions of Assessment and Feedback in Competency-Based Medical Education: A Focus Group Study of One Internal Medicine Residency Program. *Academic Medicine, 95*(11), 1712–1717. <https://doi.org/10.1097/ACM.0000000000003315>
- Ginsburg, S., van der Vleuten, C. P. M., Eva, K. W., & Lingard, L. (2017). Cracking the code: residents' interpretations of written assessment comments. *Medical Education, 51*(4), 401–410.

<https://doi.org/10.1111/medu.13158>

Klie, J.-C., Bugert, M., Boullosa, B., de Castilho, R. E., & Gurevych, I. (2018). The INCEpTION Platform: Machine-Assisted and Knowledge-Oriented Interactive Annotation. In *Proceedings of the International Conference on Computational Linguistics: System Demonstrations*, 5–9.

Marcotte, L., Egan, R., Soleas, E., Dalgarno, N. J., Norris, M., & Smith, C. A. (2019). Assessing the quality of feedback to general internal medicine residents in a competency-based environment. *Canadian Medical Education Journal*, *10*(4), e32–e47. <https://doi.org/10.36834/cmej.57323>

Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2014). *The Development and Psychometric Properties of LIWC2015*. University of Texas at Austin

Shaughness, G., Georgoff, P. E., Sandhu, G., Leininger, L., Nikolian, V. C., Reddy, R., & Hughes, D. T. (2017). Assessment of clinical feedback given to medical students via an electronic feedback system. *Journal of Surgical Research*, *218*, 174–179. <https://doi.org/10.1016/j.jss.2017.05.055>

Van Wissen, L., & Boot, P. (2017, September). An electronic translation of the LIWC Dictionary into Dutch. In *Electronic lexicography in the 21st century: Proceedings of eLex 2017 conference*, 703–715