

UNRAVELLING LANGUAGE USE OF NARRATIVE COMMENTS IN EPORTFOLIOS: A TEXT ANALYSIS APPROACH

Sofie Van Ostaeyen (PhD student)

Prof. dr. Orphée De Clercq, dr. Mieke Embo, Prof. dr. Tammy Schellens, Prof. dr. Martin Valcke



BACKGROUND

- Ongoing move to **competency-based education in healthcare education** → challenged **assessment strategies**
- Traditional focus on numeric scores BUT **narrative comments** have a lot of potential!

Feedback door supervisor *Feedback by supervisor*

Stabiele hand, poliep vlot kunnen verwijderen, bij wissel instrumenten hou je de hysteroscoop stabiel. Nog wat meer oefenen met Betocchi introductie.
Stable hand, able to remove a polyp easily, keep the hysteroscope stable when changing instruments. Some more practice with Betocchi introduction.

- = **Useful & valid data** source during workplace learning
- As compared to numeric scales^{1,2}

- more **reliable**
- easier to **interpret**
- stimulate **reflection**
- more effective to provide **constructive feedback**

PROBLEM STATEMENT

Healthcare students: narrative comments = **nonspecific, generic, vague and ineffective**^{3,4}

- Authors need feedback in order to improve
- Manual analysis = **time-consuming**

Automatic text analysis to the rescue

RESEARCH AIM

To **identify** whether **high-quality narrative comments** reflect a certain language use

METHOD

Stage 1: Quality of narratives

- 2,348** narrative comments
- EPortfolios of **149 Flemish (Belgium) healthcare students**
 - Specialist medicine, general practice, midwifery, speech therapy and occupational therapy
- Collected in June and July 2021
- Manually labelled** in annotation platform INCEption⁵, according to **four quality criteria** for effective feedback⁶:
 - performance (behavioural)
 - judgment (positive – negative)
 - elaboration
 - improvement

1 Behavioural Positive Elaboration
Bloedafname gaat goed , goeie voorbereiding , steriel gewerkt , vlot aangeprikt .
2 Improvement Improvement
Haal inderdaad eerst je buisje weg Venapunctie bij baby verder oefenen .
3 Behavioural Positive
Handeling zit wel goed !

- 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

REFERENCES

- 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>
- Marcotte, L., Egan, R., Soteas, 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/cmaj.57323>
- 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.00000000000003315>
- 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>
- Klie, J.-C., Bugert, M., Boulosa, 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.
- Van De Ridder, J. M. M., Stokking, K.M., McGaghie, W.C., & ten Cate, O.T.J. (2008). What is feedback in clinical education? *Medical Education*, 42(2), 189–197. <https://doi.org/10.1111/j.1365-2923.2007.02973.x>
- Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2014). *The Development and Psychometric Properties of LIWC2015*. University of Texas at Austin
- 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

Stage 2: LIWC tool

- Analysis using the **Linguistic Inquiry and Word Count (LIWC) tool**⁷
- Quantifies language use** along multiple lexical dimensions: 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)
- Results in categorization of all **words** used following the available **lexical dimensions**, along with their **summations or relative percentages**

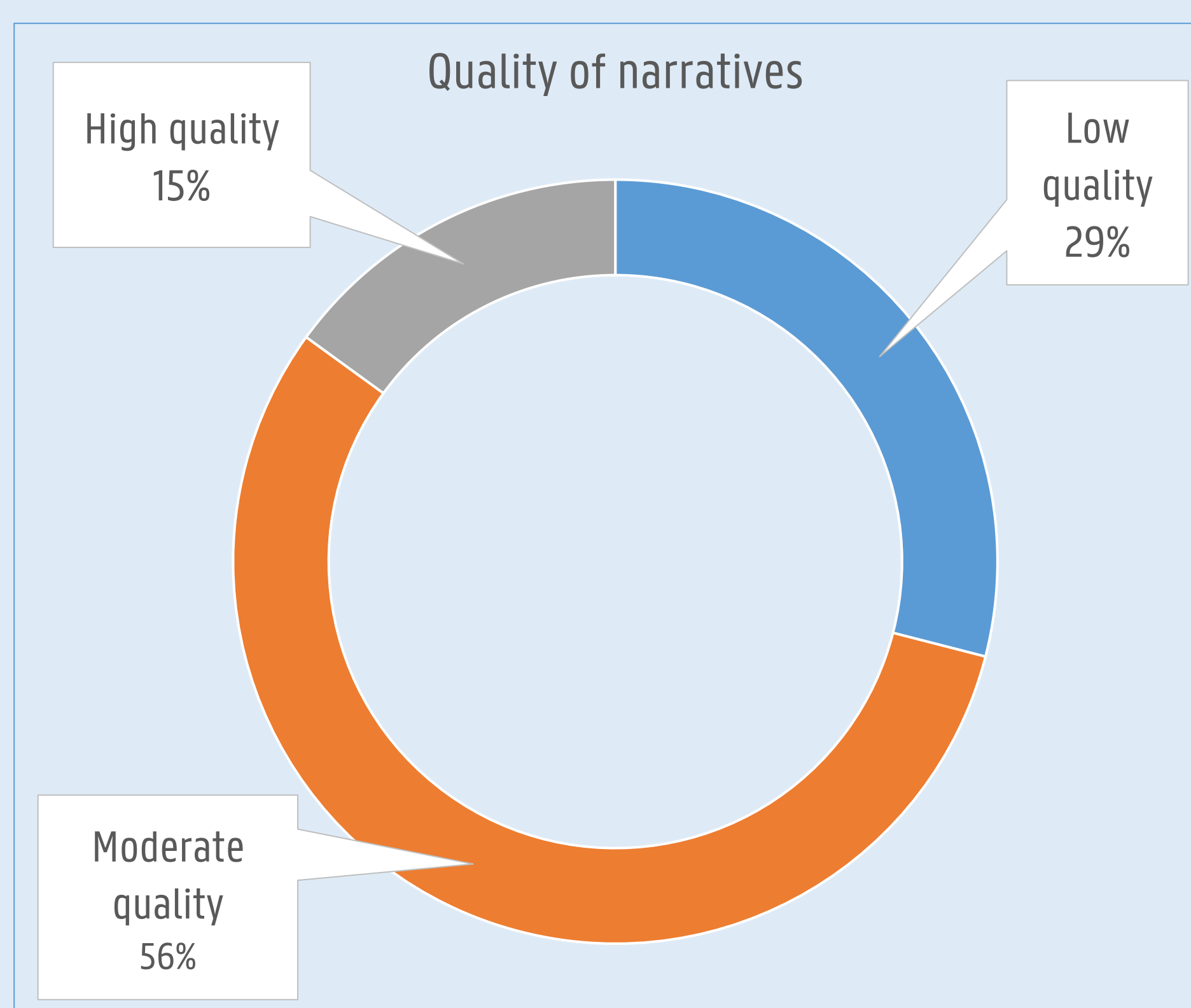
	A	B	C	D	E	F	G	H	I
1 Source (A)	Source (B)	WC	WPS	Sixltr	Dic	functi	pronoun	ppron	
2 aso_1	mooie pre:	31	15,50	45,16	77,42	48,39	12,90	6,45	
3 aso_10	oefening b	3	3,00	33,33	66,67	0,00	0,00	0,00	
4 aso_100	Meeste sta	104	11,56	11,54	93,27	66,35	23,08	11,54	
5 aso_101	schitterenc	1	1,00	100,00	0,00	0,00	0,00	0,00	

RESULTS

Stage 1: Quality of narratives

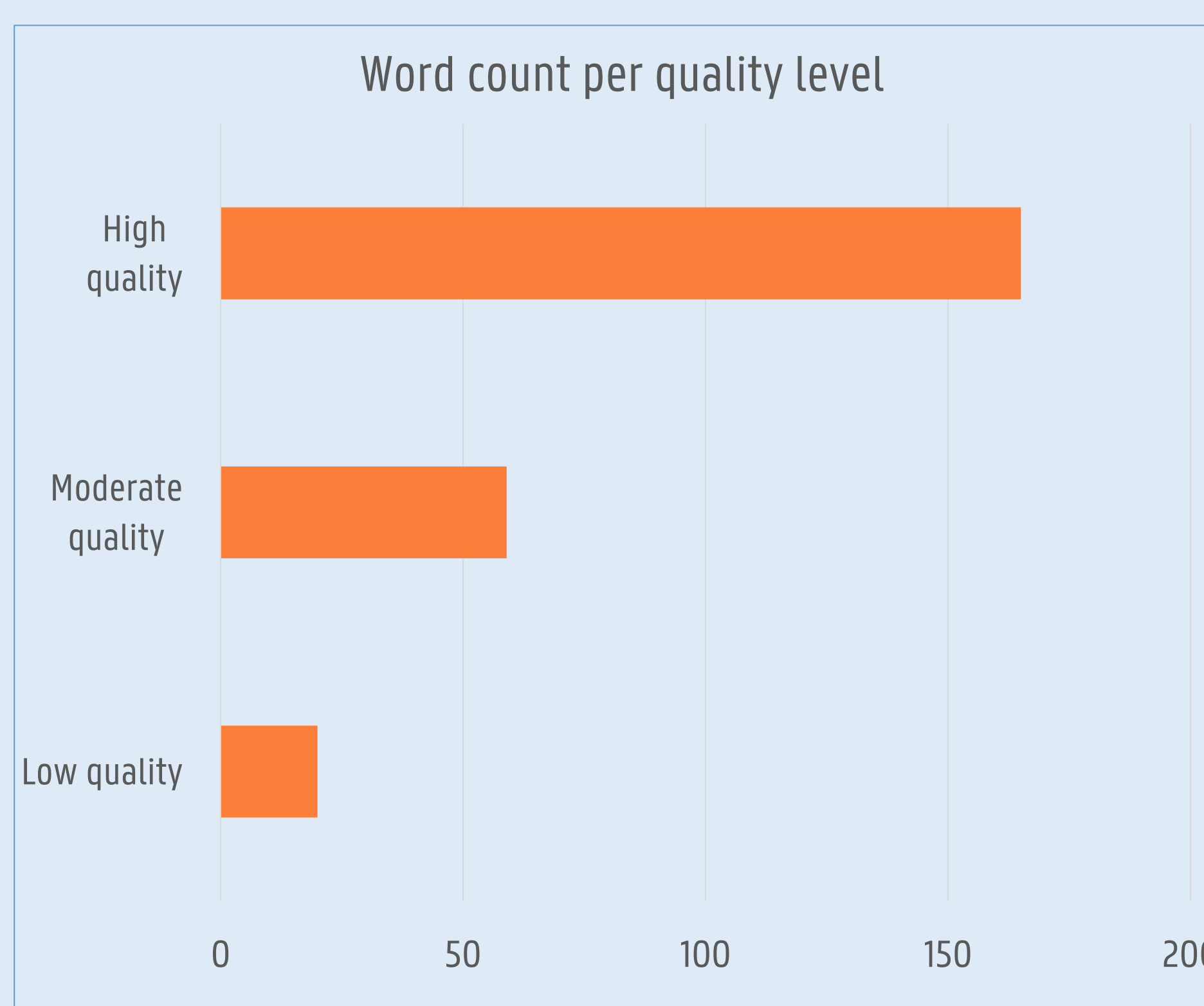
Three quality levels:

- low quality: meeting none or one criterion
- moderate quality: meeting two or three criteria
- high quality: meeting all four criteria



Stage 2: LIWC tool

Word count



Stage 2: LIWC tool

Dictionary categories

- Only **minor differences** could be identified
- The **part-of-speech and word categories** with percentages higher than 10% **did not differ**

Category	Low quality	Moderate quality	High quality
Dictionary words	76,47614	77,01622	78,43966
Function words	54,22316	55,76214	57,77691
Words with more than six letters	26,22642	24,80595	22,59647
Punctuation marks	16,78358	17,22540	15,70120
Prepositions	16,16213	14,68921	14,34848
Common verbs	16,04643	17,25017	16,93648
Relativity	14,31966	13,68276	13,78498
Total pronouns	14,28366	15,29711	16,57453
Cognitive processes	14,18955	12,49447	13,07047
Present focus	12,73013	12,48258	12,46520
Adverbs	11,16240	11,51115	12,02162
Social processes	10,23764	12,53200	12,48964

- No differences** when comparing the five dictionary categories with the **lowest relative percentages**
- Only **small differences in relative percentages** could be identified

Category	Low quality	Moderate quality	High quality
Anger	0,05057	0,07148	0,05228
Religion	0,01445	0,01299	0,00696
Death	0,00719	0,00650	0,01220
Swear	0	0,00909	0,00348
Non-fluencies	0	0,00259	0,00523

CONCLUSION

Quality of narratives

- Results confirm the **critical quality** of narratives in healthcare education^{3,4}

LIWC tool

- Word count** was the only lexical dimension that differed across the quality levels
- Differences were not present** when comparing the currently available LIWC dictionary categories while looking at low-, moderate-or high-quality comments

→ **Language use** in high quality narratives **did not vary** considerably compared to narratives of lower quality, based on currently available dictionary categories

More specialized dictionary categories might be needed?

Contact

sofie.vanostaeyen@ugent.be
<https://www.sbo-scaffold.com/en/team/van-ostaeyen-sofie>

@sboscaffold

@OstaeyenSofie

linkedin.com/in/sofie-van-Ostaeyen

