

## 2.2. Experimental Setup

To generate website landing page content for the company we collaborated with, the participants of the three groups of human content writers described in the IT service industry application study in the article had access to all tools and the environment that the company uses in its common SEO content production workflow. The instructions provided to the study participants are presented in Table W2.2.1. The participants were offered incentives for content production. The incentive for groups 1 (novices) and 2 (quasi-experts) was 15 € per produced content and credit for a marketing course. The incentive for group 3 (SEO experts) was 40 € per produced content.

Content production took place within the same week and in the same geographic location so that all participants had the same state of search engine results as a basis, which we controlled for via daily crawls. We controlled for writers' differences in terms of educational and other background covariates we considered relevant for the writing task. To this end we conducted a series of Kruskal Wallis group comparison tests. We find that the human content writing groups did not differ in their education ( $\chi^2(3)=.60$ ,  $\eta^2=.01$ ,  $p=.745$ ) or writing skills (for which the SEO experts scored a bit higher;  $\chi^2(3)=5.89$ ,  $\eta^2=.12$ ,  $p=.053$ ); we also could not detect any significant differences in terms of time invested conducting research on the target keyword / topic ( $\chi^2(3)=.28$ ,  $\eta^2=.00$ ,  $p=.868$ ) and content writing ( $\chi^2(3)=3.76$ ,  $\eta^2=.08$ ,  $p=.153$ ).

Table W2.2.1: Participants' Survey Instructions for Content Writing

Content Writing Group	Instructions <sup>1</sup>
Novices	<p>[Short introduction stating the goal of this study, strict anonymization, the incentive and a contact person for questions.]</p> <p>Imagine you are a marketing employee in an IT service company.</p> <p>Your manager approaches you to <b>write a Google search engine optimized (SEO) text for a single site on the website of your IT company</b>, that elaborates on a specific service. You should <b>write the text in a way that it ranks well in Google</b>. That means, it should preferably appear on page 1 in the Google search results.</p> <ul style="list-style-type: none"> <li>• The text should be written for the keyword / search term / topic: <b>“IT maintenance”</b> (i.e., for IT maintenance provided as a service by your company to firms).</li> <li>• It should be written for ranking well in <b>Google in [Country blinded], set to English language</b> (please use the link below).</li> <li>• For ranked <b>example sites see</b>:  <a href="https://www.google.com/search?num=100&amp;hl=en&amp;q=it+maintenance">https://www.google.com/search?num=100&amp;hl=en&amp;q=it+maintenance</a></li> <li>• It should be <b>original, unique content</b>, invented by you (i.e., NO copies).</li> <li>• It should be written <b>in English language</b>.</li> <li>• It should contain <b>around 700 to 800 words</b> (ca. 2 A4 pages).</li> </ul> <p><b><u>Your text:</u></b> (Please write your text in the following text field.)</p>
Quasi Experts	<p>[Short introduction stating the goal of this study, strict anonymization, the incentive and a contact person for questions.]</p> <p>Imagine you are a marketing employee in an IT service company.</p> <p>Your manager approaches you to <b>write a Google search engine optimized (SEO) text for a single site on the website of your IT company</b>, that elaborates on a specific service. You should <b>write the text in a way that it ranks well in Google</b>. That means, it should preferably appear on page 1 in the Google search results.</p> <ul style="list-style-type: none"> <li>• The text should be written for the keyword / search term / topic: <b>“IT maintenance”</b> (i.e., for IT maintenance provided as a service by your company to firms).</li> <li>• It should be written for ranking well in <b>Google in [Country blinded], set to English language</b> (please use the link below).</li> </ul>

- For ranked **example sites** see: <https://www.google.com/search?num=100&hl=en&q=it+maintenance>
- It should be **original, unique content**, invented by you (i.e., NO copies).
- It should be written in **English language**.
- It should contain **around 700 to 800 words** (ca. 2 A4 pages).

### How to write a SEO optimized text?

- **Integrate the main keyword** (“IT maintenance”) **or parts of it most often compared to the other words in your text.**
- **Write about subtopics / content** that you can find on the top ranked websites for the main keyword.
- **Align the word distribution of your text** with the word distribution of the top ranked websites for the main keyword (i.e., **put the right words with the right frequencies into your text**).
- For the **word distribution analyses** use: <https://wordcounter.net/> (Please be aware that the tool doesn't count common stopwords like "it".)
- **Prevent keyword stuffing** (i.e., [don't integrate keywords overly often and in an unnatural way into your text](#)).
- Try to give your text a **good readability and structure**.

**Your text:** (Please write your text in the following text field.)

---

Real SEO  
Experts

[Short introduction stating the goal of this study, strict anonymization, the incentive and a contact person for questions.]

Imagine you are a marketing employee in an IT service company.

Your manager approaches you to **write a Google search engine optimized (SEO) text for a single site on the website of your IT company**, that elaborates on a specific service. You should **write the text in a way that it ranks well in Google**. That means, it should preferably appear on page 1 in the Google search results.

- The text should be written for the keyword / search term / topic: **“IT maintenance”** (i.e., for IT maintenance provided as a service by your company to firms).
- It should be written for ranking well in **Google in [Country blinded], set to English language** (please use the link below).
- For ranked **example sites** see: <https://www.google.com/search?num=100&hl=en&q=it+maintenance>
- It should be **original, unique content**, invented by you (i.e., NO copies).

- It should be written **in English language**.
- It should contain **around 700 to 800 words** (ca. 2 A4 pages).

**Your text:** (Please write your text in the following text field.)

---

<sup>1</sup>Keywords and links were adapted in each survey.

Descriptive statistics on the content length and changes are provided in Table W2.2.2. As these statistics show, the produced content is of about equal length across the experimental groups and human revisers changed about 9% of the machine-made content in the revision process. We report evaluations of the performance of each of the experimental groups with regards to search engine rankings and the quality score in the next section of this document.

Table W2.2.2: Descriptives for Content Lengths and Revision

Dimension	Groups	Descriptives			
		Median	(IQR)	Min	Max
<b>Produced content length (in words)</b>	Revised machine	807	(67)	632	899
	Real SEO Experts	729	(84)	578	771
	Quasi Experts	694	(69.5)	498	749
	Novices	711	(48.5)	377	966
<b>Content change (raw vs. revised)<sup>1</sup></b>	Change in %	9.04	(3.77)	3.31	21.45
	Change in words	74.00	(36.50)	27.00	154.00

---

<sup>1</sup>This includes every possible change between the raw machine and revised machine output like added words, deleted words, and words with at least one changed letter (including changed letter capitalization).

## Appendix References

- Baayen RH, Shafaei-Bajestan E (2019) Analyzing linguistic data: A practical introduction to statistics. Package ‘languageR’. Version 1.5.0. CRAN. Accessed May 20, 2019, <https://cran.r-project.org/web/packages/languageR/languageR.pdf>
- Benoit K, Watanabe K, Wang H, Nulty P, Obeng A, Müller S, Matsuo A, (2018) “quanteda: An R package for the quantitative analysis of textual data.” *Journal of Open Source Software*. 3(30). <https://doi.org/10.21105/joss.00774>
- Berger J, Sherman G, Ungar L (2020b) TextAnalyzer. Accessed November 11, 2020, <http://textanalyzer.org>
- Bronnenberg BJ, Kim JB, Mela CF (2016) Zooming in on choice: How do consumers search for cameras online? *Marketing Science*. 35(5):693-712.
- Danaher PJ, Mullarkey GW, Essegai S (2006) Factors affecting website visit duration: A cross-domain analysis. *Journal of Marketing Research*. 43(2):182-194.
- Edelman B, Zhenyu L (2016) Design of search engine services: Channel interdependence in search engine results. *Journal of Marketing Research*. 53(6):881-900.
- Flanigan, AJ, Metzger, MJ (2007) The role of site features, user attributes, and information verification behaviors on the perceived credibility of web-based information. *New Media & Society*. 9(2):319-342. <https://doi.org/10.1177/1461444807075015>
- Jerath K, Ma L, Park YH (2014) Consumer click behavior at a search engine: The role of keyword popularity. *Journal of Marketing Research*. 51(4):480-486.
- Kamoen N, Holleman B, Bergh H (2013) Positive, negative, and bipolar questions: The effect of question polarity on ratings of text readability. *Survey Research Methods*. 7(3):181-189.
- Liu J, Toubia O (2018) A semantic approach for estimating consumer content preferences from online search queries. *Marketing Science*. 37(6):930-952.
- Maechler M, Rousseeuw P, Croux C, Todorov V, Ruckstuhl A, Salibian-Barrera M, Verbeke T, Koller M, Conceicao ELT, Palma MA (2020) Basic robust statistics. Package ‘robustbase’. Version 0.93-6. CRAN. Accessed May 20, 2020, <https://cran.r-project.org/web/packages/robustbase/robustbase.pdf>
- Pennebaker JW, Booth RJ, Boyd RL, Francis ME (2015) Linguistic inquiry and word count: LIWC2015. Austin, TX: Pennebaker Conglomerates. Accessed November 1, 2020, [www.LIWC.net](http://www.LIWC.net).

Pitler E, Nenkova A (2008) Revisiting Readability: A unified framework for predicting text quality. *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing*. 186-195.

Radford A, Narasimhan K, Salimans T, Sutskever I (2018) Improving language understanding by generative pre-training. OpenAI.

Roberts C (2010) Correlations among variables in message and messenger credibility scales. *American Behavioral Scientist*. 54(1):43-56.

Rocklage MD, Rucker DD, Nordgren LF (2018) Persuasion, emotion and language: the intent to persuade transforms language via emotionality. *Psychological Science*. 29(5):749-760.

Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser L, Polosukhin I (2017) Attention is all you need. *31st Conference on Neural Information Processing Systems (NIPS 2017)*. 1-15.