Linguistic Complexity Reduces the Perceived Validity of Moral Arguments

Keywords: collective decision making, natural language processing, moral issues, linguistic complexity, large-scale experiments

Extended Abstract

The belief that using more complex and lengthier words will make a text seem better written (and the author more intelligent) is highly prevalent among college students [1]. However, previous literature has proven this to be a myth: writing simpler and clearer provides the best results [1, 2]. Processing fluency, or the subjective experience of ease with which people process information, reliably influences people's judgments across a broad range of social dimensions. In particular, linguistic fluency is positively associated with intelligence [1], solidarity [3], valuation [4], familiarity [5], persuasiveness [6], trustworthiness [7], and truthfulness [8]. Previous studies on this phenomenon have primarily focused on how people process written texts at the individual level. However, in many real-world situations, we find ourselves interacting with others, and trying to solve problems that may be complex in nature, through argumentation and deliberation. In this work, we ask whether linguistic complexity reduces the perceived validity of arguments about controversial moral issues, and whether it negatively impacts the probability of people reaching consensus on those issues. To address this overreaching aim, we focused on the effect of word length as a previously established standard index of linguistic complexity [9, 10]. Across four studies, performed in different contexts, we show that the use of longer words is a driver of poorer argumentation in:

- An open asynchronous online discussion forum (Study 1);
- A large-scale behavioral experiment eliciting changes of mind (Study 2);
- A behavioral study where participants deliberate in online chatrooms (Study 3);
- A randomized controlled experiment where we manipulate word length (Study 4).

In Study 1, we looked for real-world evidence for the hypothesized effect. To do so, we scrapped a website on the online platform Reddit called "Change My View", where individuals post an opinion that may be flawed in an effort to understand other perspectives through the exchange of arguments (Figure 1A). We downloaded 21,839 posts and observed that arguments with lengthier words were rated more negatively than arguments with shorter words (Figure 1B).

In Study 2, we aimed to replicate the effect of word length in a more controlled setup. In a large-scale behavioral study (N=10,548) performed online, we tested whether arguments that induce changes of mind are perceived as less valid if they use longer words. Participants first rated their agreement with a moral statement (e.g. "Sex work should be legal"), and then they were presented with two arguments that were randomly selected to be either in favor or against the initial position (Figure 1C). We asked participants to rate the validity of each argument and gave them the opportunity to revise their initial position. Results show that the selected arguments induced changes of mind and that word length negatively correlated with the validity ratings (Figure 1D), replicating the effect observed in Study 1.

In Study 3, N=768 participants were invited to deliberate about moral issues in a three-stage behavioral study (Figure 1E). In the first stage, participants individually rated their agreement on six highly polarizing moral statements (on a scale of 0 to 10). In stage 2, they deliberated

with two other individuals through online text chats under the instruction to try to reach consensus. They had five minutes for discussing each moral statement, and received four of the six statements from stage 1. The third stage consisted of a second opportunity to provide individual ratings for the six moral statements. We observed that the probability of reaching consensus, controlling for the initial disagreement in the group, was reduced by the use of longer words (Figure 1F).

Finally, in Study 4 (N=600), we directly tested the hypothesis that word length causally reduces perceived argument validity, while also examining the role of perceived disfluency as a potential mechanism (pre-registered at https://aspredicted.org/VTR_86P). We created a set of stimuli consisting in moral arguments with identical content but different word length based on the use of different synonyms of the same words. For each statement, participants were randomly assigned to read two arguments in favor and two against a given moral statement, either in their shorter or longer versions. After each argument, they were asked to rate their perceived validity and disfluency (argument presentation order was randomized). We found that word length significantly increased the perceived disfluency, and that this variable negatively impacted on perceived validity (Figure 1H).

In conclusion, this work provides converging evidence across four studies that using long words (a proxy of linguistic complexity and disfluency) makes moral arguments seem less valid. In short: these results suggest that brevity and simplicity are key drivers for good argumentation and deliberation on controversial moral issues.

References

[1] Oppenheimer, D. M. (2006). Consequences of erudite vernacular utilized irrespective of necessity: Problems with using long words needlessly. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 20(2), 139-156.
[2] Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and social psychology review*, 13(3), 219-235.
[3] Dragojevic, M., & Giles, H. (2016). I don't like you because you're hard to understand: The role of processing fluency in the language attitudes process. *Human Communication Research*, *42*(3), 396-420.
[4] Alter, A. L., & Oppenheimer, D. M. (2006). Predicting short-term stock fluctuations by using processing fluency. *Proceedings of the National Academy of Sciences*, *103*, 9369-9372.
[5] Whittlesea, B.W.A., & Williams, L. D. (1998). Why do strangers feel familiar, but friends don't? The unexpected basis of feelings of familiarity. *Acta Psychologica*, 98, 141-166.
[6] Lowrey, T. M. (1998). The effects of syntactic complexity on advertising persuasiveness. *Journal of consumer psychology*, 7(2), 187-206.
[7] Alter, A. L., & Oppenheimer, D. M. (2009). Suppressing secrecy through metacognitive ease: Cognitive fluency encourages self-disclosure. *Psychological science*, 20(11), 1414-1420.
[8] McGlone, M. S., & Tofighbakhsh, J. (2000). Birds of a feather flock conjointly (?): Rhyme as reason in aphorisms. *Psychological Science*, *11*, 424-428.
[9] Lewis, M. L., & Frank, M. C. (2016). The length of words reflects their conceptual complexity. *Cognition*, 153, 182-195.
[10] Fenk-Oczlon, G., & Pilz, J. (2021). Linguistic complexity: relationships between phoneme inventory size, syllable complexity, word and clause length, and population size. *Frontiers in Communication*, 6, 626032.

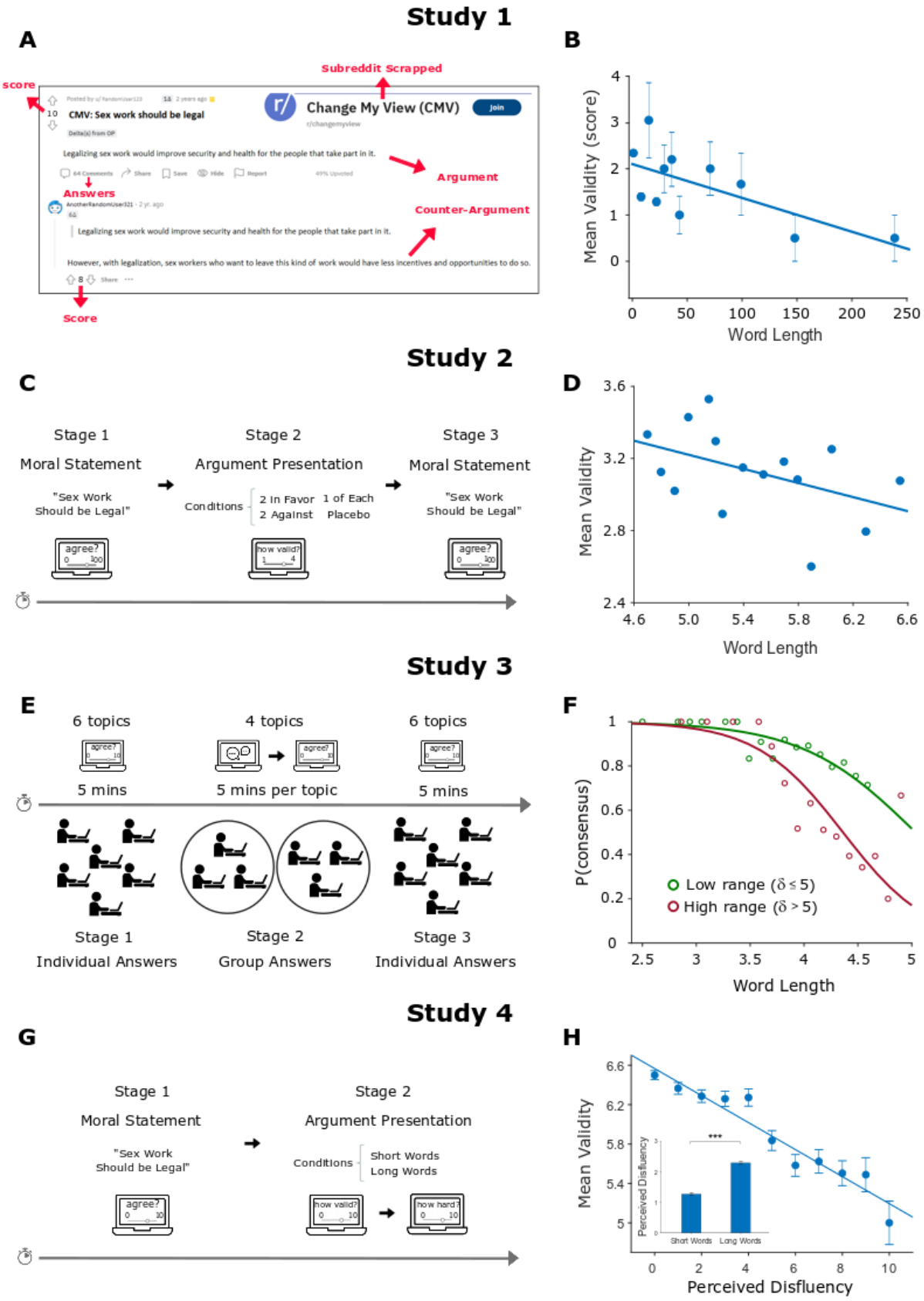

Figure 1. A. Study 1: Reddit scrapping. We extracted score and word length for each post on the Change My View website. B. Mean validity (score) of the scrapped posts as a function of word length. C. Study 2: we presented moral statements and asked participants to rate their

agreement. Then, each participant received two arguments from one of four conditions, and rated the validity of each argument. Finally, they again rated their agreement with the moral statements. D. Mean validity of the presented arguments as a function of word length. Error-bars are smaller than the dot size. E. Study 3: participants individually rated their agreement with six moral statements. Then, they were prompted to try to treach consensus on four of these statements with two other participants in a text chat. Finally, they again indiviudally rated their greement with the six moral statements. F. Probability of reaching consensus as a function of word length of the written messages. We separate groups with an initially low range of opinions (δ = max opinion – min opinion in the group) from groups with an initially high range of opinions (green: low range, brown: high range). G. Study 4: participants rated their agreement on moral statements, and then received two arguments in favor and two arguments against each statement (the order of presentation was randomized). Arguments could be in either their shorter or longer forms. They rated the validity and disfluency (reading difficulty) of each argument. H. Mean validity of arguments as a function of perceived disfluency. We also show that perceived disfluency is directly affected by word length.