Can LLMs model meaning in restorying interventions?

Cody Kommers
The Alan Turing Institute, London, United Kingdom

William Cunningham
University of Toronto, Toronto, Ontario, Canada

Abstract

While cognitive science has made great progress in modeling a range of psychological phenomenon, the processes underlying how people interpret the meaning of their own experiences has mostly resisted formalization. In this paper, we explore a method for using large language models (LLMs) to simulate the effects of this interpretive process. We compare our LLM-based simulations to extant data on restorying interventions— which show that people are more likely to endorse their life stories as meaningful after being prompted to reflect on how their experiences fit into the narrative structure of a hero's journey. Across three simulations, we show that (1) LLMs are capable of modeling the effects from these restorying interventions, (2) they are sensitive to signals from restorying interventions other than the hero's journey, and (3) this pattern of results is broadly—though not entirely—consistent across several different LLMs. Ultimately, these simulations point towards how LLM-based computational models might generate novel predictions about the effects of restorying interventions on meaning in human participants.

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*