9^{th} International Conference on Computational Social Science IC^2S^2
July 17-20, 2023, Copenhagen, Denmark

Academic Support Network Reflects Doctoral Experience and Productivity

Keywords: science of science, network science, text mining, scientific careers

Extended Abstract

In recent years, well-being and mental health concerns for PhD students have been increasing. According to a recent survey by Nature on 6,300 PhD students, 36% responded that they sought help for anxiety or depression caused by their studies [1]. Another devastating fact is that doctoral students are 2.43 times more likely to have a common psychiatric disorder than the rest of the highly educated population [2]. It is therefore important to examine the journey of doctoral students not only through the lens of academic "success measures" such as publication numbers, citation counts, fellowships received *etc.* but also at their overall well-being and the quality of the environment that supports them in fulfilling their potential.

Acknowledgements contain such profound details of their authors' academic journey and environment; however, research efforts to study how they vary concerning disciplinary and demographic differences have remained limited. To fill this gap, we investigated 26,236 PhD dissertations, obtained from *ProQuest Open Access Dissertations & Theses* database, 99% of which are from the United States in the last 20 years. We systematically identified acknowledged individuals and institutions, by using a data-driven approach supported by manual inspection and found the entities in the acknowledgements such as mother, brother, advisor, colleague etc. We then enriched our dataset by adding the number of publications before and after the graduation of associated PhD students from *Dimensions.ai* API to examine the relationship between the characteristics of acknowledgements and productivity levels of PhD students.

Firstly, we uncovered the academic support network. For this task, we employed the Doc2Vec model and represented each support provider in a vector format by looking at the verbs, adverbs, nouns and adjectives that are used to acknowledge them. Then, we used these vectors to create an "academic support network" by applying the Girvan-Newman community detection algorithm. It revealed five distinct communities that support students along the way: Academic, Administration, Family, Friends & Colleagues, and Spiritual as shown in Fig.1(a).

Secondly, we compared gender and disciplines in terms of who and how they are acknowledging (Fig.2). We showed that female students, compared to their male counterparts, mention fewer people from each of these communities except for their families. However, sentiment analysis results demonstrated that they use a more positive language when acknowledging each community. Our results also suggested that the total number of people mentioned in the acknowledgements allows disciplines to be represented on an individual science - team science plane as their magnitudes change, in which case the Social Sciences & Humanities and Life& Earth Sciences are on the associated extremities, respectively.

Lastly, we applied an Inverse Gaussian regression model to describe the relationship between students' demographics, their acknowledgement characteristics and productivity levels. Results indicated that women are publishing slightly less number of papers than men. This can be explained by the fact that they are more likely to play a significant role in parenting ([3]), or the systematic undervaluation of women's involvement and their invisibility in scientific research, known as the "Matilda Effect" ([4]). This is critically important because it means that

studying the doctoral process may help us better understand the adverse conditions women face early in their academic careers. We also showed that male students who mention more people from their academic community are associated with higher levels of productivity. Our results point to the importance of academic support networks by explaining how they differ and how they influence productivity.

References

[1] Woolston, C., PhDs: the tortuous truth. Nature, 575(7782), 403–407 (2019)

[2] Levecque, K., Anseel, F., De Beuckelaer, A., Van der Heyden, J. and Gisle, L., Work organization and mental health problems in PhD students. Research policy, 46(4), 868–879 (2017)

[3] Kyvik, S., & Teigen, M., Child care, research collaboration, and gender differences in scientific productivity. Science, Technology, & Human Values, 21(1), 54-71 (1996)

[4] Rossiter, M. W., The Matthew Matilda effect in science. Social studies of science, 23(2), 325-341 (1993)

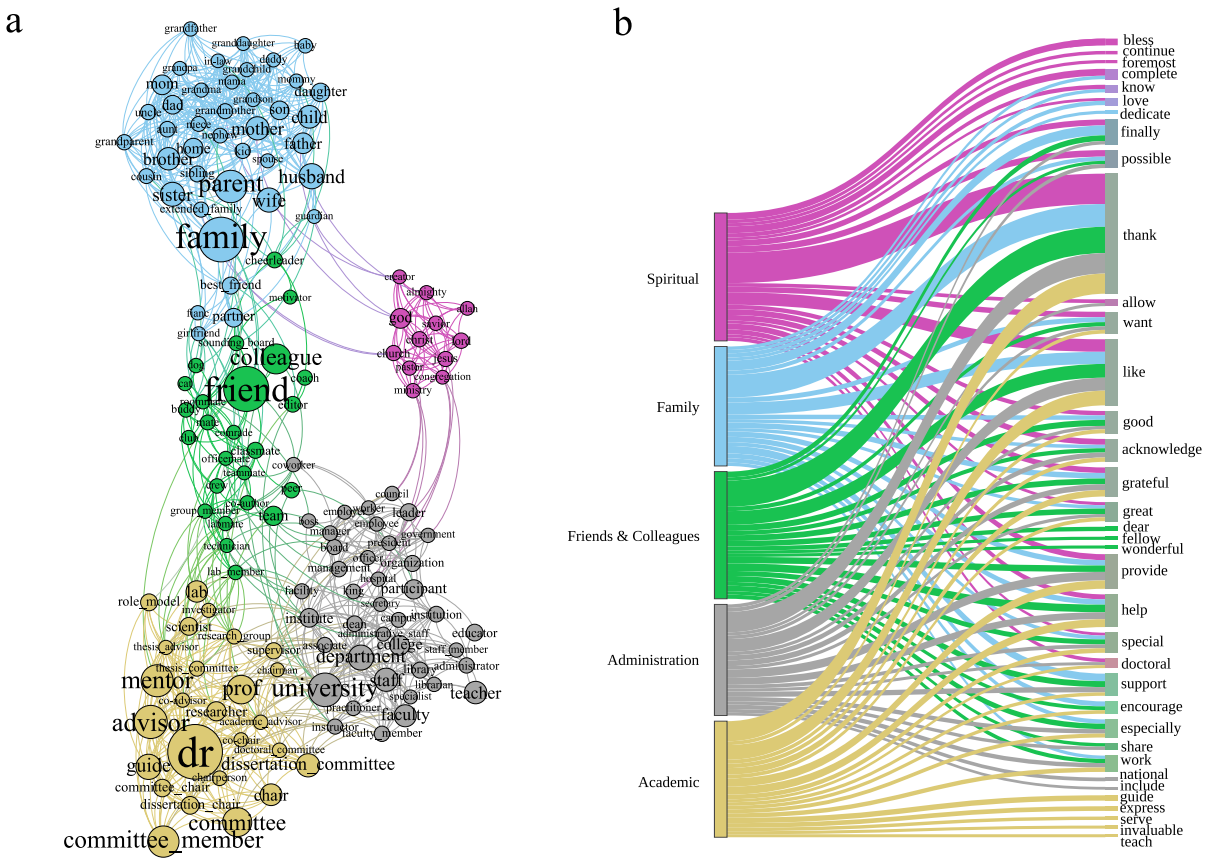

Figure 1: Support providing entities represented as nodes and their contextual similarities learned from document embeddings used as edge weights (a). These communities are acknowledged using specific words and bi-partite relation points group specific properties (b).

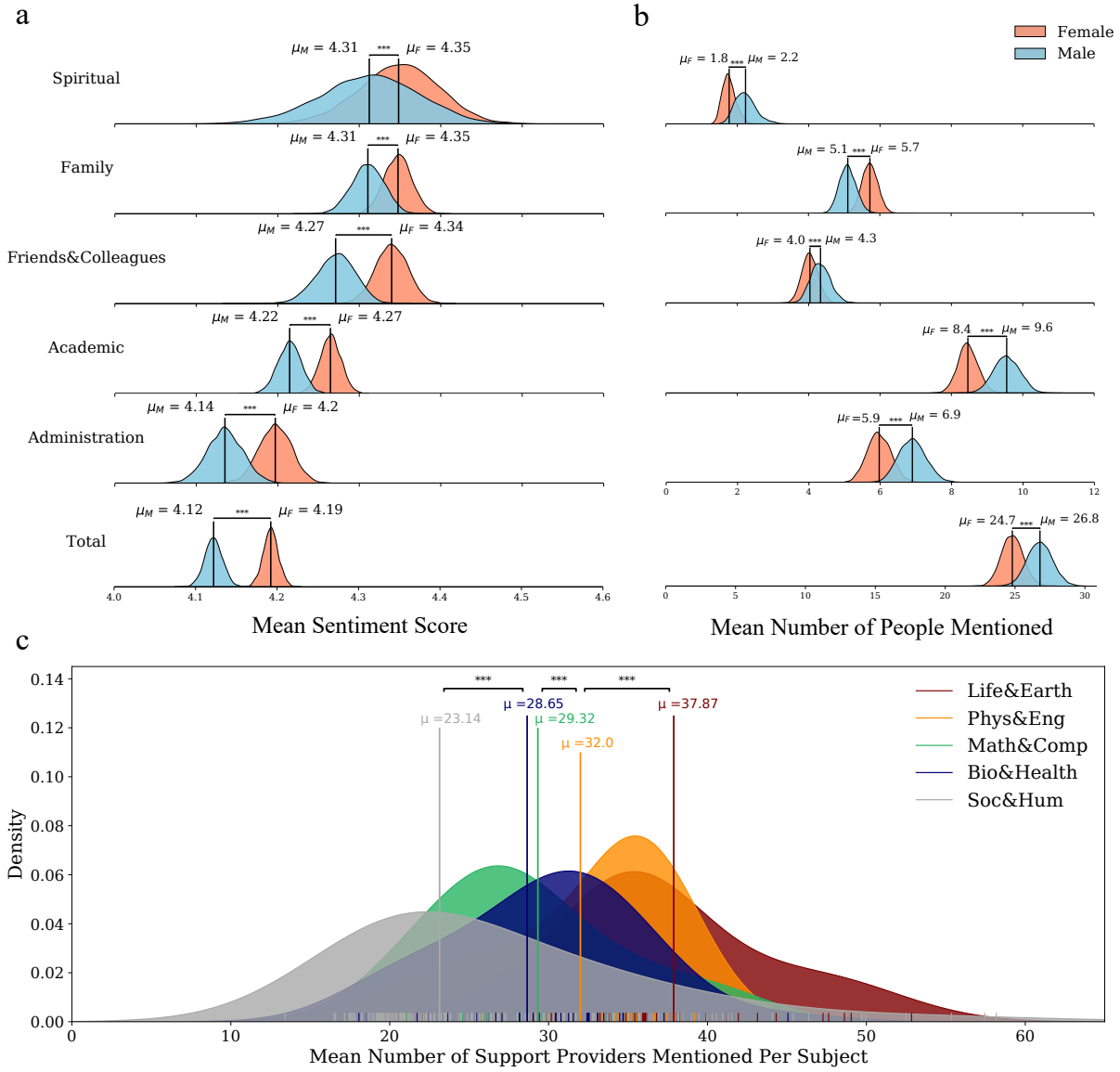

Figure 2: Sentiment scores differ when mentioning the support providers (a). Mean number of people mentioned alter across genders for different support provider categories (b).Bootstrapped mean sentiment scores differ when mentioning the support providers (a). Bootstrapped mean number of people mentioned alter across genders for different support provider categories (b). Distribution for mean number of support providers mentioned by each academic discipline subcategory (e.g. Soc&Hum consists of subcategories such as Economy and Management) illustrates the transition from individual sciences to team sciences (c).