Jake Sosine, M.S., M.B.A., BCBA

Data Analyst

**Dr. Tiffany Kodak**

Department of Psychology

Cramer Hall, 317

604 N. 16th St.

Milwaukee, WI 53233

Dear, Dr. Kodak:

I am writing to submit the manuscript titled, “Identifying Trends in the Behavior Analytic Literature via Natural Language Processing” to *The Analysis of Verbal Behavior*. The purpose of this study was twofold – both of which we believe will be of interest to the readership of *The Analysis of Verbal Behavior*. First, humans created an average of 154 GB of data each day in 2020 and are projected to generate 463 exabytes (4.63x1011 GB) of data each day in 2025. Much of this data is verbal behavior and, thus, presents new opportunities for behavior analysts interested in the description, prediction, and control of verbal behavior across laboratory and nonlaboratory settings. However, handling these data efficiently requires the use of technologies to measure and analyze verbal behavior at scale. The first purpose of this proof-of-concept study was to introduce the readers of *The Analysis of Verbal Behavior* to Natural Language Processing (NLP). NLP is a technique from computer science that allows the user to turn verbal behavior into numbers for subsequent quantitative analysis in a replicable way. Here, we show how readers can use NLP techniques to collect, clean, and analyze large amounts of verbal behavior. In turn, NLP might open the door for behavior analysts to ask and answer questions about verbal behavior that were not previously possible.

Second, the texts used for the analysis were all published articles up through 2019 for the journals: *Behavior Analysis in Practice*, *Journal of Applied Behavior Analysis*, *Journal of the Experimental Analysis of Behavior*, *The Behavior Analysts*, *The Analysis of Verbal Behavior*, and one control journal – *Health Service Research*. Using NLP, we found several patterns such as changing article lengths over time, changing prevalence of research topics, and drift in the language used between applied and basic areas of behavior analysis. To our knowledge, this kind of a large-scale analysis on the history of scientific writing in behavior analysis has never been conducted. Importantly, the resulting dataset and availability of NLP tools make replication and extension of this work comparatively easy for readers of *The Analysis of Verbal Behavior*.

We hope that you enjoy this unique analysis and find potential value in these tools for behavior analysts researching verbal behavior. I confirm that this manuscript has not been published elsewhere and is not under consideration by another journal. All authors agree with its submission to this journal*.*

Best,

Jake Sosine

Data Analyst, Behavioral Health Center of Excellence

Researcher, Behavioral Data Science Research Lab, Endicott College

E: jacobsosine@gmail.com