Research Statement – V.N. Vimal Rao

My research investigates the psychology of statistics. That is, what does it mean to do statistics or think statistically, and how does one learn to do this? I draw on theories and methods from the cognitive and learning sciences to investigate statistical thinking, and consider what new types of evidence we might observe and how this affects our theories of statistical cognition.

Current Research on P-value Cognition To this end, one of my major projects since 2020, and something that will be my main focus for the next several years, is the investigation of the effects of a century of categorical interpretation of *p*-values on individuals' cognition. In 2019, the American Statistical Association gave a warning not to categorize any statistical measure. Yet, to cognitive psychologists, categorization is fundamental to cognition. This '*p*-value project' began with a study (PV1) applying theories of numerical cognition and methods from the study of categorical perception to model graduate students' perceived differences between pairs of *p*-values (e.g., '$p = .047$' and '$p = .053$'). We found a boundary effect at .05 consistent with the theory of categorical perception, and antithesis to statisticians' recommendations to no longer dichotomize *p*-values. My paper, co-authored with Dr. Sashank Varma and Dr. Jeffrey Bye, won the 2021 Disciplinary Diversity and Integration Award from the Cognitive Science Society. We have replicated and extended the finding (PV2), and have a manuscript we will submit to the journal Psychological Science.

We are currently working on two follow-up studies. The first (ES1) is to examine the way graduate students categorize Cohen's *d* effect sizes – we will present these results at the 2022 International Conference on Teaching Statistics (subsequently published in its proceedings) – and how Cohen's *d* effect size estimates are labeled in published manuscripts. This extension to a different statistical measure, one for which benchmarks are widely known rather than categorical boundaries, will support the development of a theory of statistically situated numerical cognition.

The second follow-up study (PV3) is to investigate whether boundary effects extend beyond the initial information processing, and to what extent they are correlated with self-regulation/inhibition abilities. Our initial studies (PV1 and PV2) only measured prepotent responses within the first second or two of exposure to a stimulus. Interestingly, in a post-task survey in PV2, participants indicated that they were trying to follow consistent arithmetic heuristics when selecting responses, even though their response profiles indicate otherwise. Thus, PV3 examines statistical cognition on a longer time scale to distinguish whether graduate students are able to self-regulate out of prepotent categorical interpretations of *p*-values, or whether the categorical boundary we identified affects longer-term cognition than the 1-2 second interval we initially investigated. These findings will help us understand how graduate students process, store, and then recall information about *p*-values (i.e., their numerical value and/or their category membership as 'statistically significant') in a more ecologically-valid time-span.

Future Research on P-value Cognition Over the next several years, I plan to pursue two further extensions of this project as we build a theory of statistically situated numerical cognition. The first extension (PV4) will investigate the effect of categorical learning and training activities on students' thinking about *p*-values. Current recommendations for breaking out of the '$p < .05$' world focus on conceptual instruction (e.g., reasoning about *p*-values as a 'continuum of evidence'). However, reasoning and thinking, as advanced cognitive processes, depend on information processing, encoding, and retrieval. Therefore, if initial processing and encoding of *p*-values favors categorical information, so too will subsequent reasoning. In order to ameliorate the categorical effects on these initial cognitive processes that underlie statistical thinking and reasoning, we then must target information processing, encoding, and retrieval with specific training. I propose to develop and compare three training methods in terms of their effects on students' *p*-value cognition: (1) categorical training, (2) self-regulation/inhibition training, (3) conceptual training.

The second future extension (PV5) will focus on the development of research methods to study contextually situated numerical cognition. In PV1 and PV2, we used tasks known as 'discrimination tasks' often used in the study of categorical perception of physical stimuli such as speech sounds and color. However, adapting these tasks to numeric stimuli required unique considerations, and to some degree leaves open questions as to the exact mechanism underlying the effects we have documented. Therefore, we plan to investigate, develop, and test a variety of tasks, based on methods in categorical perception as well as methods in numerical cognition, in order to obtain a reliable measure of categorical effects in the reasoning about numerical stimuli, and to better understand the cognitive mechanisms underlying statistical cognition.

Current Research in Statistical Cognition Methods Beyond the *p*-value project, I also conduct methodological research more broadly in the field of statistical cognition and statistics education. For example, I am in the process of writing a manuscript detailing a scoping review of uses of the Comprehensive Assessment of Outcomes in Statistics (CAOS), and will submit the paper to the Statistics Education Research Journal. We use validity theory as a framework to evaluate how CAOS is utilized in statistics education research. The study purpose was to critique the scope of inferences made from CAOS and multiple choice assessments, and to educate the statistics education community in validity theory.

My dissertation also has a methodological focus, inspired by the learning sciences. I examined graduate students' thinking in statistical tests, with specific focus on their thinking about null models, in a multi-modal multiple descriptive case study. The study method includes participatory concept mapping, task-based interviews, video and audio recordings, eye tracking, and video-cued interviews. While the primary focus is to describe graduate students' thinking to inform teaching and curricular development, I am also using the data to author a paper in which we compare and contrast the differences in the traces of students' thinking captured by each modality and task. Our paper will serve as a novel contribution to the study of metacognition, and help researchers to choose from a variety of novel tasks and data sources, including video.

Broader Scholarship The research I have described above is only a small part of the broader scholarship that I engage in. My projects often allow me to generate pedagogical papers, instructional materials, and other resources for instructors and students of statistics. For example, I worked on a project to develop a framework for a social justice oriented data science curriculum. I adapted the framework to create activities focused on demonstrating the potential of statistical analysis for social action, activities that focus on marginalized groups and social issues, and workshops for instructors to use these resources. I am in the process of preparing a manuscript about this project to be submitted to the journal Teaching Statistics. Similarly, I served as a consulting statistician on a project investigating instructional sequences for rational numbers. The data collected had significant ceiling effects, and through this example, I have worked on a pedagogical paper to help school psychologists detect when a ceiling or floor is sufficiently robust to significantly bias results obtained from an ANOVA-based analysis of group comparison, and provide instructions on what to do in such cases.

My career research goal is to develop a social cognitive theory of statistics in the hopes that this work and its applications can empower individuals in the study and practice in statistics, both the formal and informal quantification of uncertainty, and its subsequent use in decision making.