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. This research guides how I practice statistics and the way I train my students to practice statistics.

Statistically Situated Numerical Information Processing My nain project is the investigation of the categorical interpretation of *p*-values on individuals' cognition. After a century of '$p < .05$' as a categorical boundary for "statistical significance", in 2019 the American Statistical Association gave a warning not to categorize any statistical measure. Yet, to cognitive psychologists, categorization is fundamental to cognition. Has training and practice in categorical thinking distorted the way graduate students perceive decimals in a statistical context?

This 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 '$p = .053$'). In an adapted AX discrimination task, 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. In a follow-up study (PV2), the effect was replicated with a larger sample and larger number of stimuli. PV2 also extended the study to a different task (the ABX discrimination task) and included a control group, establishing that the observed boundary effect was one above and beyond the effect expected by standard models of numerical cognition.

A study of the way graduate students categorize Cohen's *d* effect sizes (ES1) extended the project beyond *p*-values. ES1 measured graduate students' category boundaries implicitly with a boundary identification task and explicitly with a survey. Participants consistently drew boundaries between common benchmark values, as predicted by psychological theories of categorization. An extension (ES2) focused on how Cohen's *d* effect size estimates are labeled in published manuscripts. These extensions to a different statistical measure, one for which benchmarks are widely known rather than categorical boundaries, supports the development of a general theory of statistically situated numerical cognition.

Statistically Situated Numerical Cognition (PV3) extends the project beyond the initial information processing and into encoding and retrieval and the role of working memory. 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 that were seen in PV1 and PV2, or whether categorical information permeates throughout encoding and retrieval. These findings will help us understand how graduate students process information about *p*-values (i.e., their numerical value and/or their category membership as 'statistically significant') in a more ecologically-valid time-span.

Training and Methods in Statistical Cognition I will pursue two further extensions of this project as I 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 extension (PV5) develops research methods to study contextually situated numerical cognition. PV1 and PV2 used '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 leaves open questions as to the exact mechanism underlying the documented effects. Therefore, I plan to investigate, develop, and test a variety of tasks, based on methods in categorical perception as well as methods in numerical cognition, to obtain a reliable measure of categorical effects in the reasoning about numerical stimuli, and to better understand the cognitive mechanisms underlying statistical cognition.

Methods Research 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 conducting a scoping review of uses of the Comprehensive Assessment of Outcomes in Statistics (CAOS). I use validity theory as a framework to evaluate how CAOS is utilized in statistics education research. The study purpose is 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 investigating the differences in the traces of students' thinking captured by each modality and task. This 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 and gaze recordings.

Broader Scholarship I prioritize improvement of statistical methodology through pedagogical papers, instructional materials, and other resources for instructors and students of statistics. 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 how to use censored regression models in these cases.

Similarly, I worked on a project to develop a framework for a social justice oriented data science curriculum. I then 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. This work is the basis of a grant proposal I am working on to continue to develop resources and provide professional development to teachers interested in integrating social justice initiatives in their classroom.

Research Values My research focuses on a set of questions about statistical cognition, and I utilize a variety of research methods in order to answer these questions. This interdisciplinary diversity is not possible without collaborations built on a foundation of both social and academic relationships. As both a statistician and a psychologist, I am uniquely positioned to serve as a bridge between fields. I often serve as a quantitative methodological consultant for psychologists, and a qualitative methodological consultant for statisticians. By bringing my connections in each of these two fields together, and building new connections, I am able to develop a generative environment for research and scholarship in the psychology of statistics.

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.