Ph.D. Dissertation Defense

Candidate:	Vahid Ashrafi
Degree:	Doctor of Philosophy
School/Department.:	School of Business / Business Administration
Date:	Monday, December 8, 2025
Time:	2:00 – 4:00 pm
Location:	Babbio 601
Title:	Toward a computational framework for understanding errors in human cognition and decision making
Chairperson:	Dr. Jordan W. Suchow, Information Systems, School of Business
Committee Members:	Dr. Gert-Jan de Vreede, Information Systems, School of Business Dr. Jeffrey V. Nickerson, Information Systems, School of Business Dr. Aron Lindberg, Information Systems, School of Business Dr. Anand Goel, Finance, School of Business Dr. Nick Byrd, Bioethics and Decision Sciences, Geisinger Commonwealth School of Medicine

Abstract

My project examines cognitive errors—including cognitive biases and logical fallacies—and their impact on decision-making in real-world contexts. It aims to map, detect, and mitigate these errors by integrating behavioral science with advanced natural language processing (NLP) and machine learning techniques.

The *first chapter* develops a computational taxonomy of 370 cognitive errors using sentence-transformer embeddings derived from bias descriptions, examples, mitigation notes, and academic abstracts. By generating high-dimensional representations and clustering them, the chapter reveals the semantic structure and relationships among these errors.

Chapter 2 shifts from theory to application by analyzing about 4 million advice-seeking scenarios as real-world examples of everyday decisions. Using embeddings and clustering, it produces a classification map of common decision types—such as financial, relational, and health-related decisions—highlighting patterns in how people seek advice. This map complements Chapter 1 by organizing the landscape of everyday decision problems.

Chapter 3 integrates insights from the first two chapters by applying a Knowledge-Base–Guided Scoring (KBGS) pipeline to construct a matrix linking cognitive errors to the scenarios curated and analyzed in Chapter 2. The system retrieves error-specific knowledge and constrains the model's evaluation through structured instructions, prompting an LLM to generate justification-based relevancy scores for each error–scenario pair. The resulting matrix supports analyses of error prevalence and frequency across scenario types, providing a scalable, knowledge-grounded mechanism for identifying the cognitive processes that systematically shape everyday judgments and decisions.

Chapter 4 moves to the individual level by introducing algorithms that detect cognitive errors in personal writing, particularly on the Reddit platform. These models generate individualized cognitive-error profiles and an empirical, theory-driven rationality measure (RQ), with practical applications for self-awareness, coaching, and decision-support systems capable of providing real-time feedback.

Chapter 5 scales the analysis to the group level, using Subreddits data to examine how cognitive errors shape collective decision-making in online communities and interest groups. Examples include confirmation bias in cryptocurrency communities and the sunk cost fallacy in relationship advice groups. The chapter offers insights for improving group dynamics and strengthening online content moderation by identifying the dominant group-level cognitive errors that arise within online communities.

Together, the chapters build a multilevel system—from taxonomy to real-world scenarios to individual and group profiling—that advances both the scientific understanding of cognitive errors and the practical tools for mitigating them. These contributions offer a scalable framework for improving decision-making across diverse personal, organizational, and societal contexts.