Research Statement

Poorna Talkad Sukumar

My research in **human-computer interaction** and **information visualization** spans many application domains, from designing programming tools for mobile devices, to visualizing personal sensing data, to studying biases with visualizations. My particular interests are in **designing effective strategies to mitigate biases and improve decision-making in visualization and expanding visualization usage contexts by creating effective visualizations for mobile devices and personal data**. In my prior work, I have created new interactive visual systems for personal data exploration and specific domain problems, designed quantitative empirical experiments to examine the mitigation of implicit and cognitive biases with visualizations, and employed qualitative and mixed-methods approaches to elicit design requirements and evaluate visualization systems.

My research methodology interweaves theory, design, and evaluation. I combine my engineering and computer science background with a strong design sense to build well-engineered, effective, and clear visualization systems and use a range of research methods drawing from human–computer interaction, information visualization, and the social sciences to understand and evaluate people's use of these systems. My research focuses on emerging topics in visualization having broader impacts and I plan to publish in top human–computer interaction and visualization research venues, such as, CHI, IEEE VIS, EuroVis, and CSCW.

I am also interested in initiatives towards advancing open practices and transparency in research. I have organized a panel on transparency in qualitative research in human–computer interaction [1] and contributed to the ACM CHI conference reviewing guidelines. I believe that open practices, such as, transparent reporting of the research process and data sharing, are not only important for establishing the rigor and validity of empirical studies but can also provide invaluable guidance to students and beginners starting out in the field.

1 MITIGATING BIASES WITH VISUALIZATIONS

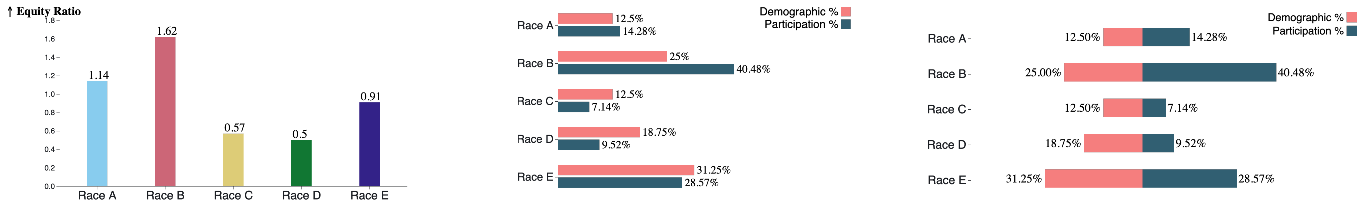
Biases associated with visualizations can be categorized into three types—perceptual biases, cognitive biases, and implicit or social biases. Perceptual biases are those where how we *see* data distorts our judgments, cognitive biases refer to flawed reasoning occurring during decision-making, and social or implicit biases refer to unconscious, ingrained attitudes and stereotypes about groups of people. I focus on studying cognitive and implicit biases in visualization (although perceptual biases are invariably at play when using visualizations and can be difficult to delineate).

1.1 Cognitive Biases

I became interested in studying cognitive biases in visualizations after encountering potential biases exhibited by reviewers when making college admissions decisions [2,3]. Holistic college admissions happens to be a domain where the decisions to be made on applications are not obvious and reviewers make subjective assessments on applications. Furthermore, reviewers typically review several applications in a short period of time while experiencing significant stress and time pressures. I identified potential reviewer biases in the process and presented numerous strategies for their mitigation. The strategies are centered around the use of visual representations of application data and interactive visual tools to ease the cognitive load of the reviewers and support note-taking and other data-exploration activities occurring during the application-review process, which currently uses text-based tools.

There are more than a hundred known types of cognitive biases and only a handful of these have been studied in visualization so far. Given the ubiquity of visualization systems for analyzing data and making decisions in numerous domains, generating knowledge on how biases can occur when using visualizations and how the biases can be alleviated can be extremely useful and contribute towards improving data-analysis and decision-making with visualizations in general. Building on my work in the admissions domain, I plan to conduct empirical studies to examine **how visualizations can inadvertently cause biased interpretations of the data, how new strategies can mitigate cognitive biases in visualizations, and how these strategies can be incorporated into visual analysis systems in different domains.**

My approach is to study specific types of bias which are discussed in the literature as being relevant to visualization, such as, confirmation bias and framing effect, and investigate each type independently. Every type of bias is associated with or is known to occur in the context of particular type of tasks. For example, confirmation bias, where people tend to favor evidence that supports their existing beliefs, is associated with hypothesis



(a) (b) (c)

Fig. 1. Sample representations for equity ratio in classroom participation data [2]—representing the equity ratio directly using a bar chart (a); representing the individual percentages in the equity ratio instead of the equity ratio by using a paired-bar chart (b) and a mirror-bar chart (c).

assessment tasks. When participants were shown visualizations of their activity data (Fig. 2) in one of our studies [5], they sought to confirm their hypotheses about their data (e.g., "I would expect myself to be probably above the cohort and all people because I'm a pretty active person") in the visualizations. I plan to find both generic and domain-specific tasks mirroring the tasks associated with each bias type, design mitigation strategies informed by theory, and use various methods, including crowdsourced studies and controlled experiments, to evaluate the mitigation strategies.

1.2 Implicit Biases

In addition, I also study implicit biases in the domain of student participation in classrooms. More specifically, my research focuses on examining **how teachers can be made aware of their implicit biases through visualizations of student participation data, disaggregated by race and gender, in their respective classrooms.** For example, the data could reveal to the teachers that students belonging a particular race were almost never called on to answer questions in their classroom. In collaboration with researchers in education, I created potential solutions for visualizing classroom participation data [4]. While these solutions are valid solutions in terms of the represented data types and supported tasks, we need to examine how teachers interpret *participatory inequities* (that is, disparities in classroom participation) using each of the visualizations and which among the solutions are most effective in leading teachers to make more equitable decisions.

For example, Fig. 1 presents a few potential representations for the *equity ratio*, a metric defined as the ratio between the participation percentage of a demographic group and the percentage of the demographic within the classroom. When teachers were shown the visualization type (a) where the equity ratio is represented directly, some of them incorrectly interpreted equity as equality and inferred that they should strive to get all the student groups to have an equity ratio of 1 and further, decrease the participation rates of groups having equity ratio > 1. The equity ratio is not intended to suggest that the ratios should be equal to 1 for all groups but is instead intended to highlight when certain marginalized student groups have a disproportionately low level of participation. The effectiveness of the alternative representations, where the individual percentages in the equity ratio are visualized instead (Fig. 1(b) and 1(c)), is yet to be examined.

I plan to conduct comparative evaluations of the proposed space of visualization solutions using measures, such as, accuracy and mental effort, as well as qualitative studies using interviews and think-aloud methods to elicit teachers' interpretations of the data shown in the different visualizations. I intend to distill findings from these studies into **design guidelines for visualizing demographic inequities also in other similar domains (e.g., selection and hiring processes) which can enable stakeholders in the domains to make more equitable decisions.**

2 EXPANDING VISUALIZATION USAGE CONTEXTS

Visualization research has primarily focused on designing for people who are familiar with using visualizations and/or data and for use on large screens. However, visualizations are increasingly being consumed in less formal contexts (e.g. on mobile phones) and by people who have limited experiences with visualizations and/or data. Expanding visualization research to make data accessible to these broader audiences and in non-traditional contexts gives rise to important and interesting research questions. For example, these wider audiences have different needs and goals, which would require the use of research methods different from those typically used in visualization research. I focus on two connected areas—mobile and personal data visualization, where the end-users are very likely to be people who are not visualization/data experts. I plan to build on my prior work in these areas to address the research challenges they pose.

2.1 Personal Data Visualizations

Personal data comprises of any data that is relevant to one's personal life, such as, health and fitness data, social media interactions, and travel diaries. Visualizations of personal data can enable people to better understand themselves, share personal insights with others, and make changes to their behaviors. Personal visualizations, however, have distinctive goals and usage characteristics; they aim to support self-reflection, are consumed in less formal contexts and often on mobile devices, and by people with different motivations, interests, and resources.

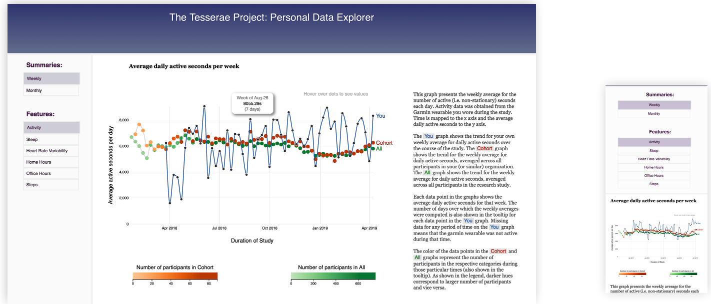


Fig. 2. Personal visualization interface presenting the various personal attributes gathered from each participant [8]. The left image presents the original layout designed for viewing on large screens. The right image presents an adapted (scrollable) design for viewing on mobile devices.

I am interested in exploring research methods that are more suited to study personal visualizations and emphasize realism of context as well as generalizability to inform the design of effective visualizations for personal data. In particular, I am interested in studying the **commonalities and differences in individual usage behaviors of personal visualizations, types of personal insights that visualizations help generate, potential barriers to effective interpretations of personal data, such as, confirmation bias and misleading baselines, and contextual information that can be embedded within visualizations to help people recall past behaviors and interpret their data.** My approach involves leveraging personal data gathered as part of large-scale, longitudinal sensing studies aimed at studying individual and collective human behavior.

I designed an interface (shown in Fig. 2) presenting interactive visualizations of various personal data attributes, including activity, sleep, and hours spent at home and office, gathered from 757 participants in a year-long study [8]. I employed both quantitative and qualitative research methods to study how participants explored visualizations of their respective data on the interface. First, I sent links to a subset of the participants (564) to their respective visualizations and logged and analyzed their interactions over a period of 18 days to characterize their collective exploratory behaviors on the interface [8]. The participants spent varying amounts of time and used a variety of devices, including mobile phones, to explore their visualizations and many of them also revisited their visualizations. Further, the participants appeared to have made more comparisons between their own data instances than with the provided baselines. Following this study, I conducted an insight-based evaluation (using a mixed-methods approach by combining think-aloud method with analysis of interactions) of the interface with a different subset of the participants [5].

Next, I plan to algorithmically combine the data (including activity, sleep, and hours spent at home) gathered from information workers during the work-from-home conditions of COVID-19 for ten weeks [6] to build visualizations of their temporal patterns or rhythms. I will conduct studies to gather user input on **how well the visualizations of users' personal rhythms align with their perceptions of their own rhythms and how these visualizations can help users coordinate their activities and improve their productivity.**

2.2 Visualizations on Mobile Devices

Data visualizations, such as those in activity-tracking apps and news articles, are increasingly being consumed on mobile devices. Much of the existing visualization design guidelines, however, are mostly meant for designing for large screens and are strongly tied to the *WIMP* (windows, icons, menus, pointer) paradigm or the interaction style of desktops. In my previous work, I studied the ways in which *programming*, also a domain with strong affinities with the WIMP paradigm, is supported on mobile devices [9]. I plan to apply relevant knowledge from my work in this domain to design visualizations for use on mobile devices. My approach is to **apply and evaluate relevant design decisions from the programming domain as well as employ suitable research methods for designing mobile-device-interactions, such as, gesture elicitation studies.**

One of the key takeaways from my previous work is that the interaction mechanisms of mobile devices are *fundamentally different* from those of conventional laptop/desktop setups and hence, designing for mobile devices should proceed from leveraging these mechanisms rather than directly adapting existing guidelines [9]. Relevant design strategies from the programming domain to overcome the drawbacks (e.g. small screen size) and leverage the interaction affordances of mobile devices include selectively hiding and displaying parts of or whole visualizations (similar to “code folding”), use of transparent overlays for displaying additional content, switching between overview and detail representations depending on the task being performed, use of in-line visualizations or “sparklines”, use of gesture and multimodal interactions, and selecting more mobile-friendly charts (e.g. radial charts over line graphs). I plan to implement and evaluate these strategies and distill the findings into visualization design guidelines for mobile devices.

3 TOOLS FOR VISUAL DATA EXPLORATION

I have designed visualization solutions and developed tools for data exploration in differing domains. I helped design *GameViews* [7], consisting of two distinct interfaces presenting key (basketball) game information—one intended to help sportswriters construct their game reports and another intended for fans to discuss the game. The paper received a Best Paper Honorable Mention award at the CHI conference.

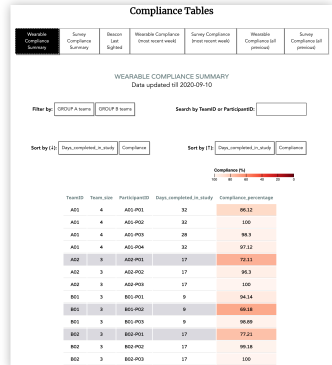


Fig. 3. Interactive tables for monitoring participant compliance in a large-scale, longitudinal study [6].

I developed an interactive visualization system for monitoring participant compliance in a large-scale, longitudinal study where data is being collected through sensors and surveys from teams of information workers [6]. The system presents data in the form of *interactive tables* (see Fig. 3) and allows personnel to check and reason about missing data from participants in the study and also supports asynchronous collaboration among the non-co-located personnel. I plan to build tools for exploring the team information collected from this study to examine the alignment of temporal rhythms among the team members and how they can inform the design of “awareness” and “availability-sharing” systems for the current work-from-home conditions.

I plan to continue to collaborate with stakeholders in different domains, especially where the use of visualizations is under-explored, to iteratively design visualization solutions to specific domain problems. My goal is to not only create suitable solutions for various domain problems but to also select problems with broader impacts and containing interesting visualization research questions that can be applied to other contexts.

REFERENCES

1. **Talkad Sukumar, P.**, Avellino, I., Remy, C., DeVito, M. A., Dillahunt, T. R., McGrenere, J., & Wilson, M. L , “Transparency in Qualitative Research: Increasing Fairness in the CHI Review Process.”, in *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-6).

2. **Talkad Sukumar, P.,** and Metoyer, R., “A Visualization Approach to Addressing Reviewer Bias in Holistic College Admissions”, in *Cognitive Biases in Visualizations* (pp. 161-175). Springer, Cham.

3. **Talkad Sukumar, P.,** Metoyer, R., & He, S., “Making a Pecan Pie: Understanding and Supporting The Holistic Review Process in Admissions,” in *Proceedings of the ACM on Human-Computer Interaction, 2(CSCW)*, 1-22.

4. **Talkad Sukumar, P.**, Reinholz, D., Shah, N., and Striegel, A., “Visualizing Participatory Inequities in Classroom Data,” in *IEEE VIS 2020 Electronic Conference Proceedings [Poster].*

5. **Talkad Sukumar, P.,** et al., “An Insight-based Evaluation of a Personal Visualization Interface.” (In Preparation)

6. **Talkad Sukumar, P.,** Breideband, T., Martinez, G., Caruso, M., Rose, S., Steputis, C., D’Mello, S., Mark, G., and Striegel, A., “Designing an Interactive Visualization System for Monitoring Participant Compliance in a Large-scale, Longitudinal Study,” in *CHI 2021 Case Studies*

7. Zhi, Q., Lin, S., **Talkad Sukumar, P.,** & Metoyer, R., “GameViews: Understanding and Supporting Data-driven Sports Storytelling,” in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1-13). **Best Paper Honorable Mention Award**

8. **Talkad Sukumar, P.,** Martinez, G.J., Grover, T., Mark, G., D’Mello, S.K., Chawla, N.V., Mattingly, S.M. and Striegel, A.D.,”Characterizing Exploratory Behaviors on a Personal Visualization Interface Using Interaction Logs,” in *EuroVis 2020 - Short Papers.*

9. **Talkad Sukumar, P.,** & Metoyer, R., “Mobile Devices in Programming Contexts: A Review of the Design Space and Processes,” in *Proceedings of the 2019 on Designing Interactive Systems Conference* (pp. 1109-1122).