Stepping Closer To A Science of Interaction:
A Paradigm for Studying the Cognitive Mechanisms of Interaction

Amy Rae Fox*

University of California San Diego

ABSTRACT

Interaction is crucial for effective visualization. Yet, most of our research centers on *representations* of information, and the extent to which they support human operators performing well-specified tasks. Despite impressive progress developing tools for interaction, our understanding of how to employ these techniques is based largely on intuition and ad-hoc evaluation from disparate research communities. What happens—in the mind and body—when a user "interacts" with a visualization? In this dissertation I aim to develop a paradigm for studying interaction with representations in the same principled way we have treated visual encodings. Such a paradigm necessarily involves multiple levels of analysis, and seeks to bridge a functional gap between the study of cognition and the study visualization. Only once we understand how and why different interaction techniques are effective can we offer predictive guidelines for the design of human-information systems.

Keywords: interaction, visualization, representation, cognition, human information processing.

1 INTRODUCTION

Research on interaction with visualizations is by no means a new endeavor. After all, interaction is a necessary component of every act of representation, be it reciprocal (such as a user grasping, clicking, brushing or selecting) or uni-directional (a user perceiving). The visual analytics community in particular has an established tradition of considering interaction in visualization systems, owing largely to the emergent timespan and dialogic nature of analytic tasks [1]. Nonetheless, our understanding of *interaction* falls short of our knowledge of visual *representations*. Numerous authors have called for more a focused [2], systematic [3], unified [4] and holistic [5] treatment of interaction—theoretically bound to research in human cognition [3]—in order to provide meaningful guidelines [6] for designers of visualization systems.

This has proven to be a daunting task. Interaction seems an ineffable concept, discussed in "vague and haphazard ways" across a variety of domains [7]. Although work in the visualization community [1], [3], [8] has been successful in surveying the "landscape" of interaction techniques, there is—to date—no extant classification that connects interactions to mechanisms of cognition. One reason is the pragmatic difference in priority between basic and applied research communities. Another is the inherently confounded nature of representation and interaction in most visualization systems [9].

In this dissertation, I aim to leverage the powerful affordances of a high-level declarative visualization language—Vega-lite [10], in order to systematically evaluate human performance with different interaction techniques while maintaining maximum control over confounding factors such as tasks and visual encodings. By doing so, I aim to develop an ongoing research agenda of empirical studies, developing a more nuanced understanding of how and why and under what circumstances particular interactions effectively support desired cognitive activities.

2 RELATED WORK

In developing this research agenda I draw upon work from multiple communities. First, I build on previous efforts characterizing the design space of interaction techniques in Visualization and HCI research. Secondly, I leverage methods from the long tradition of research in the Cognitive and Learning Sciences that seek mechanistic explanations for human perception of, and learning from, visual representations.

2.1 Interaction in Visualization & HCI

While there are numerous successful efforts deconstructing visual representations into composable primitives (see [11]–[13]), complementary efforts explicating interaction have proven more exiguous. A central obstacle described by Yi et. al [2] is the reticulated nature of interactions and representations: the components are rarely mutually exclusive. Efforts at taxonomizing techniques have tended to be driven by the features of particular systems, and while highly descriptive, fall short of operating at a grammatical level. Yi et. al. make a substantial contribution to this effort by bridging a gap between low-level interaction techniques and user *intent*. A natural extension of this work, is to evaluate the extent to which the described interaction techniques effectively support the user in *realizing* these intentions. This type of investigation can help move taxonomies of interaction from purely descriptive to empirically-predictive.

The conceptual infrastructure necessary to engage in this inquiry is bolstered by Sedig & Parsons' recent contributions [3], [7]. Most notably, they present a theoretical model for examining the computational processing that unfolds between a representation and human at multiple levels of abstraction. This multi-tiered approach affords the opportunity to integrate with models of memory, attention and information processing in Cognitive Science, yielding an empirically-testable system. They further offer the most comprehensive catalogue (to date) of primitives, deemed "epistemic interaction patterns", that describe the interaction design space, "at the action-reaction level of human-information discourse" (pg. 97) [3]. Importantly, while their efforts are an attempt to catalogue the *current* scope of realizable interactions, the descriptions are extendable and composable, serving as a tool for researchers to explore and develop novel interaction techniques.

2.2 Mechanisms and Representations

Several models for this type of research exist at the boundary between cognitive science and human-computer interaction, exploring the cognitive mechanisms supporting representational *forms*. An exemplar is the recent work of Danielle Albers Szafir, who applies the methodologies from vision science to foundational questions about the visual processing of information displays [14]. The explanatory value of her work on color perception in visual displays serves not only as a source of guidelines for designers, but

* amyraefox@ucsd.edu

also to generate further, highly nuanced questions about visualization behavior. For applied researchers conducting case studies there may be a number of confounding contextual factors to explain unexpected performance. But by revealing the mechanisms behind select phenomena, this type of research can expose the precise circumstances under which we might expect to observe certain behaviors. Most simply, by seeking mechanistic explanations we improve the predictive power of applied science. Pragmatically, by developing a community of practice between researchers in basic and applied cognition, the problems encountered by visualization researchers can inspire research agendas for researchers in basic science as well.

I am similarly inspired by the role models in Learning Science exploring the mechanisms of multimedia learning. Researchers like Mary Hegarty, Tom Shipley and Rich Mayer draw inspiration from tangible problems in the classroom: How can we help students learn organic chemistry, the geologic time scale, or integrate multiple streams of information? Then through series' of systematic experimental investigations, they develop working models of *how* students perceive and construct meaning with cognitive artifacts, appealing to theories of working memory, information processing and attention. From these theoretical accounts they then derive practical guidelines for instructors and designers of learning content, and often partner with practicioners to test their validity. It is this level of integration between basic (cognitive) and applied (visualization) research I seek to develop in my own research agenda, and a type of mechanistic explanation I believe is required to develop an empirically-driven "science of interaction".

3 Preliminary Work

I am currently preparing my dissertation proposal and plan to advance to candidacy in the Winter of 2018. At this stage I have chosen a well-motivated research topic with an appropriately unreasonable scope. I will spend the next several months shaving away intractable and tangential components until a clear, operationalized question and appropriate plan of action are developed. In the sections that follow I detail my best approximation of what the resulting proposal will be. It is my hope that participation in the Doctoral Colloquium will be largely influential in refining the more fully-developed plan of work in place by Fall 2018.

3.1 Preparatory Work

The research I've completed to date—while likely not included in the dissertation—has offered both methodological preparation as well as provenance of the central role of interaction in visualization discourse. In an early project [15] visualization was my data-collection tool, as I used quantitative content analysis to systematically analyze student-generated visualizations of activity calendars in order to explore learners' conceptualizations of time. In [16] I constructed interactive visualizations as stimuli for an applied study examining the role of timeline orientation in causal reasoning. Our results demonstrated that interaction with the visualizations served to confound any effects of manipulations on the design of the visual representation. In ongoing work [17] we are exploring how students read visualizations with unconventional coordinate systems, finding that any kind of interaction—regardless of the information content—significantly improves comprehension.

While together these projects elucidated for me the power of representation in nearly all aspects of daily life—clarifying that I have indeed chosen the 'right' field of research—they also expose a disheartening gap between visualization research communities in education, psychology and computer science. If I wish to contribute to bridging this gap, my work must draw broadly from theories, models and frameworks in these domains.

Toward this end, I have spent substantial time considering which frames of reference are most relevant to the topic of interaction. I will draw largely from theories of distributed [18] and embodied [19] cognition in my generation of hypotheses for the possible mechanisms of function. I will seek to explain these mechanisms through the lens of human cognitive architecture, and in particular, the Embedded Process Theory of Working memory [20] and selective attention [21]. To describe the range of visualization tasks and associated activities that are accomplish I will draw upon work by Brehmer and Munzner [22]. I situate more broadly within the frame of Peircian semiotics, which holds that meaning is fundamentally a triadic relation between a thing in the world, its representation and an interpreter. It is this triadic relation that allows meaning to be *constructed* by the interpreter (rather than "contained" in the representation). Finally I draw upon the multi-tier EDIFICE-AP framework developed by Parsons & Sedig [3] to elaborate the "spaces" in which different "kinds" of computational activity occurs, resulting in an ongoing chain of interaction between representation and interpreter (**Figure 1**)

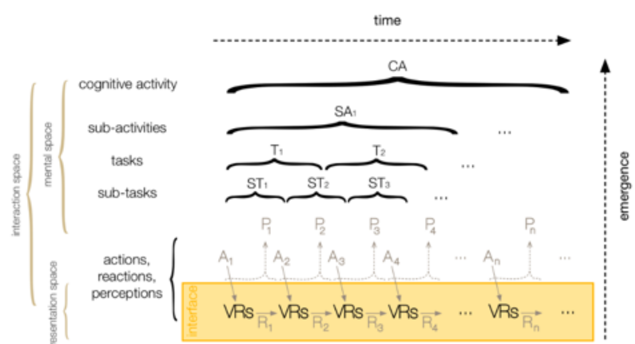

Figure 1: The Hierarchical Structure of a Complex Cognitive Activity and its Emergence over Time

Note: Figure reproduced from [3] without permission

3.2 Tool Building

To facilitate the work described in this dissertation I am presently directing[1] the development of a web-based application called *Cognitive Canvas*. Built upon the webstrates [23] framework, we intend to use the tool for prototyping interaction techniques and studying their efficacy *in situ*. To support the desired level of factor-based control for this mechanistic-research, the system must include a wide range of representational forms and be easily customized to afford different interaction behaviors. Our plan is to utilize Vega-lite [10] to develop representations and various combinations of interaction techniques. Our long-term vision is for a browser-based development framework that supports the design of representations with *dynamic* behaviors that respond to changing context governed by a *cognitively-inspired physics* [24], [25]. This system will serve as the basis for generating experimental stimuli for the work described in this research plan.

[1] I am responsible for the system architecture and design, and directly supervise a large team of research assistants in the development of the system. I am in turn supervised by my research advisor Dr. Jim Hollan.

4 PROPOSED WORK

"Given the close coupling between interaction and cognition, the science of interaction must empirically validate theories about cognitive processes on its way toward producing knowledge-construction interfaces."

— The Science of Interaction, (pg. 264)

As presently conceptualized, I envision the proposed work will proceed in three phases.

First, during a period of observation and cognitive ethnography I will embed with experts in a chosen problem domain. Through this observation, I will select a narrow set of representation(s), associated interaction techniques and "cognitive activities", which will become our objects of study. By observing the subjects in the course of their work, I will generate testable hypotheses as to how the interactions they perform support particular cognitive activities. I will also gain the cursory domain knowledge required to generate experimental stimuli.

Next, through a series of laboratory studies I will design and refine a paradigm for determining the cognitive mechanisms engaged by the selected set of interaction techniques. The laboratory studies will utilize novice learners with pre-requisite domain knowledge but lacking the tacit and procedural knowledge for performing the experts' work. Through these studies, we will systematically measure operationalized aspects of cognitive activity with different versions of the interaction techniques, in order to disambiguate how the interactions are the human cognitive architecture during information processing.

Finally, the work will culminate in a situated study evaluating the external validity of the findings from phase two outside of the laboratory. This phase will serve to further develop the model of contextual factors that intervene in interactions with visualizations.

4.1 Phase One : Observation and Ethnography

The aim of the first phase of research is to develop a cursory understanding of the chosen problem domain (e.g. a data science, physics, or neuroscience lab), and the way it is conceptualized by its practicioners. Using the theoretical framework of distributed cognition [18], the unit of analysis will encompass not only the human mind, physical representations and systems, but also bodies, situated in the context of the processes, social, temporal and material culture instantiated at the site of investigation. A necessary component of this fieldwork will be two close readings of interactions between subjects and their representational systems. First comes a description of the representing *gestures* [26] that intervene in the process of meaning-making. Such gestures serve as clues to the way a subject conceptualizes the representation with which they are interacting, as well as the emergence of cognitive processes as they construct meaning with the artifact. Secondly, attention will be paid to the extraneous tools subjects appropriate during acts of interaction. These may include material objects, such as a pencil or hand pointed at a screen, or alternatively verbal metaphors, similarly indicating sites where intervention or extension is required to elicit the desired result.

Through this field work I will select the set of representations and interaction techniques to be tested in Phase Two of the research, as well as generate hypotheses as to how the interactions may (or may not) be supporting the desired cognitive processes. Depending on the selection of problem domain, I anticipate this phase will last 2-4 months and commence in the Winter of 2018.

4.2 Phase Two : Controlled Experimental Studies

In the second phase of research I will employ a combination of existing experimental paradigms from Cognitive Psychology and Learning Science to investigate the functioning of memory, and attention, for the selected interaction techniques. Using the Cognitive Canvas platform and Vega-lite, we will develop simplified versions of the selected representations from Phase One, with experimental conditions that vary the affordances for interaction. We will engage participants with appropriate prior knowledge (for example, students who have taken an introductory course in the given topic), and systematically investigate the differences between the interaction techniques with respect to the most appropriate cognitive constructs.

What are these "cognitive constructs"? The precise design of the laboratory studies will be dependent on the interaction techniques selected for evaluation, as well as our hypotheses about how they function. For example, a hypothesis implicating the construct of cognitive load (perhaps by offloading extraneous processing onto pre-attentive perceptual processing) would necessarily involve a design with measures of attention, comprehension, and working memory. Alternatively, a hypothesis implicating a nexus of attention (perhaps a filtering interaction) would involve measuring the sequence of visual processing of the scene. The most substantial contribution of this phase will be a mapping between experimental paradigms (and associated constructs) developed in Cognitive Psychology, and the corresponding hypotheses for how certain interactions might *function*. A propitious starting point for this mapping is Sedig & Parsons' framework of epistemic interaction patterns, which is specified at a sufficiently abstract level to be mapped to corresponding cognitive constructs [3].

The present scarcity of detail regarding the design of the laboratory studies is frustrating, though intentional. Although the range of potential explanatory mechanisms is vast, I will realistically constrain the scope of investigation to a limited number of interaction techniques, which then determine the corresponding paradigms and constructs they can elucidate. I am confident that I am well-equipped for this task, on the basis of my prior experience with range of experimental methods investigating the cognitive mechanisms of graph comprehension, my expertise in developing behavioral tasks for measuring learning, comprehension, and decision making, experience designing both explicit and implicit measures, as well as my ongoing collaboration with an advisor in experimental psychology [17]. By utilizing the infrastructure described in section 3.3 I anticipate this phase of work will last approximately 12 months and commence in the Spring of 2019.

4.3 Phase Three : Situated Study

The final component of the dissertation will require returning to the site of field study, and conducting a situated case study based on the empirical findings from Phase Two. This will likely involve either: a) offering an alternative interaction for a particular representation that we hypothesize will more effectively support the desired cognitive activities, or b) evaluating an existing interaction across a range of tasks requiring "different" cognitive processing (for example, forecasting, vs. analyzing). Although the precise timeline will depend on the scope of the intervention, I anticipate this phase will last approximately 9 months and commence in the Fall of 2020.

5 QUESTIONS FOR THE COMMITTEE

In addition to the committee's wisdom on areas of prior work, relevant theories and methods, I am particularly interested in discussing the following questions.

1. Which knowledge domains might be most propitious for this kind of investigation? Given the resources at UCSD, my inclination is to work with data scientists. Here I could observe both novice and expert users. However, I am concerned that the range of representations may be somewhat narrow. I am generally interested in interactions with more traditional

'statistical graphs' as well as the larger class of *diagrammatic* forms one might find in math and physics laboratories.

2. Of the vast range of interaction techniques defined in existing taxonomies, is there a natural starting point for mechanistic investigation? Through my preliminary observation of neuroscience laboratories, I've seen both the representations and interaction techniques that are relatively advanced compared with the affordances of commercial-grade tools like Tableau. My intuition is to start with the simplest possible cases, but wish to contribute findings that are non-obvious, which may well be the case with the simplest techniques.
3. Due to the large number of participants required for the experimental studies in Phase Two, we are planning to work with novice learners in the domain of interest. If we chose to focus on data science, this opportunity may coincide with the development of the new Data Science major at UCSD. This approach raises the risk, however, of spurious effects due to differences in expertise. Our present plan to mitigate this risk is to run a small set of studies with expert users to verify that we find congruent effects. Do you have any alternative recommendations?
4. Through participation in interdisciplinary conferences (such as the Gordon Research Conference on Visualization in Science and Education) I've been impressed by both the size and enthusiasm of the "wider" (aka non-InfoVis) visualization community. From entire subdisciplines in science education devoted to studying visualizations of particular scientific constructs, to museum designers and producers of public education content, the population of visualization researchers and practicionners expands far beyond the "walls" of the VIS community. I have also been disapointed, however, by the lack of awareness of complementary research activity, shared knowledge and language across these disciplinary boundaries. As a scholar who values interdisciplinarity, what can I do to maximize the likelihood of my work being relevant to and appropriate for an audience beyond VIS?

I sincerely appreciate the time and effort of the VIS Doctoral Colloquium committee in reviewing this proposal, and welcome all feedback and suggestions for collaboration that may be offered.

6 Conclusion

It has been thirteen years since Thomas and Cook's [27] call for a "science of interaction", and deeper "understanding of the different forms of interaction and their respective benefits" (pg. 73). In the intervening years our community has made impressive progress expanding the variety and sophistication of interaction techniques, developing taxonomies, patterns and models of interactive processing. Through this dissertation I intend to apply my knowledge of human cognition, expertise with experimental methods and unabashed curiosity for information and visualization to contribute the *next* block in our joint endeavor to build a sustainable, empirically-driven science of human-information interaction.

References

[1] W. A. Pike, J. Stasko, R. Chang, and T. A. O'Connell, "The Science of Interaction," *Inf. Vis.*, vol. 8, no. 4, pp. 263–274, 2009.

[2] J. S. Yi, Y. Kang, J. T. Stasko, and J. A. Jacko, "Toward a Deeper Understanding of the Role of Interaction in Information Visualization," *IEEE Trans. Vis. Comput. Graph.*, vol. 13, no. 6, pp. 1224–1231, 2007.

[3] K. Sedig and P. Parsons, "Interaction Design for Complex Cognitive Activities with Visual Representations: A Pattern-Based Approach," *AIS Trans. Human-Computer Interact.*, vol. 5, no. 2, pp. 84–133, 2013.

[4] M. Beaudouin-lafon, "Towards Unified Principles of Interaction," in *CHItaly 2017 - Proceedings of the 12th Biannual Conference of the Italian SIGCHI Chapter*, 2017, pp. 1–2.

[5] Z. Liu and J. Stasko, "Mental models, Visual Reasoning and Interaction in Information Visualization: A Top-Down Perspective," *IEEE Trans. Vis. Comput. Graph.*, vol. 16, no. 6, pp. 999–1008, 2010.

[6] C. Tominski, *Interaction for Visualization*, vol. 3, no. 1. 2015.

[7] K. Sedig, P. Parsons, M. Dittmer, and R. Haworth, "Human-Centered Interactivity of Visualization Tools: Micro and Macro-level Considerations," in *Handbook of Human Centric Visualization*, 2014, pp. 671–691.

[8] J. Heer and B. Shneiderman, "Interactive Dynamics for Visual Analysis. A taxonomy of tools that support the fluent and flexible use of visualizations.," *ACM Queue*, vol. 10, no. 2, p. 30:30–30:55, 2012.

[9] A. Satyanarayan, "The Reactive Vega Stack : Declarative Interaction Design for Data Visualization," 2017.

[10] A. Satyanarayan, D. Moritz, K. Wongsuphasawat, and J. Heer, "Vega-Lite: A Grammar of Interactive Graphics," *IEEE Trans. Vis. Comput. Graph.*, vol. 23, no. 1, pp. 341–350, 2017.

[11] J. Bertin, *Semiology of Graphics: Diagrams, Networks, Maps*. Madison, WI: University of Wisconsin Press, 1983.

[12] Y. Engelhardt, "The Language of Graphics," University of Amsterdam, 2002.

[13] L. Wilkinson, *The Grammar of Graphics*, 2nd ed. Chicago, IL: Springer Science + Business Media, Inc., 2005.

[14] D. A. Szafir, "Modeling Color Difference for Visualization Design," *IEEE Trans. Vis. Comput. Graph.*, vol. 24, no. 1, pp. 392–401, 2018.

[15] A. R. Fox, E. de Vries, L. Lima, and S. Loker, "Exploring Representations of Student Time-Use," *Lect. Notes Comput. Sci. Diagrammatic Reason. Inference.*, 2016.

[16] A. R. Fox and M. Van Den Berg, "Representing Sequence: The Influence of Timeline Axis and Direction on Causal Reasoning in Litigation Law," in *Proceedings of the Annual Meeting of the Cognitive Science Society*, 2016.

[17] A. R. Fox and J. Hollan, "Read It This Way : Scaffolding Comprehension for Unconventional Statistical Graphs," in *Lecture Notes in Computer Science: Diagrammatic Reasoning and Inference.*, Springer International Publishing, 2018, pp. 1–17.

[18] J. Hollan, E. Hutchins, and D. Kirsh, "Distributed cognition: toward a new foundation for human-computer interaction research," *Trans. Comput. Interact.*, vol. 7, no. 2, pp. 174–196, Jun. 2000.

[19] M. Wilson, "Six views of embodied cognition.," *Psychon. Bull. Rev.*, vol. 9, no. 4, pp. 625–36, Dec. 2002.

[20] N. Cowan, "An embedded-processes model of working memory," in *Models of Working Memory*, 1999, pp. 63–101.

[21] N. Cowan, "Evolving Conceptions of Memory Storage, Selective Attention, and Their Mutual Constraints Within the Human Information-Processing System," *Psychol. Bull.*, vol. 104, no. 2, pp. 163–191, 1988.

[22] M. Brehmer and T. Munzner, "A Multi-Level Typology of Abstract Visualization Tasks," vol. 19, no. 12, pp. 2376–2385, 2013.

[23] C. N. Klokmose, J. R. Eagan, S. Baader, W. Mackay, and M. Beaudouin-Lafon, "Webstrates: Shareable Dynamic Media," *UIST '15 Proc. 28th Annu. ACM Symp. User Interface Softw. Technol.*, pp. 280–290, 2015.

[24] J. D. Hollan, "TIMELINES: Multiscale Zooming Interfaces: A Brief Personal Perspective on the Design of Cognitively Convivial Interaction," *Interactions*, vol. 18, no. 1, pp. 71–74, 2011.

[25] J. D. Hollan and A. Satyanarayan, "Designing Cognitively Convivial Physics for Dynamic Visual Information Substrates," in *Rethinking Interaction: From Instrumental Interaction to Human-Computer Partnerships. CHI 2018 workshop*.

[26] A. Becvar, J. Hollan, and E. Hutchins, "Representional gestures as cognitive artifacts for developing theories in a scientific laboratory," in *Resources, Co-Evolution and Artifacts: Theory in CSCW*, 2008, pp. 117–143.

[27] J. J. Thomas and K. A. Cook, "Illuminating the path: The research and development agenda for visual analytics," *IEEE Comput. Soc.*, p. 184, 2005.