Is There Science in Visualization?

Panel Organizer:
T. J. Jankun-Kelly, Mississippi State University

Panelists:
T. J. Jankun-Kelly, Mississippi State University
Robert Kosara, University of North Carolina at Charlotte
Gordon Kindlmann, Brigham and Women's Hospital, Harvard Medical School
Chris North, Virginia Tech
Colin Ware, University of New Hampshire
E. Wes Bethel, Lawrence Berkeley National Laboratory

INTRODUCTION

The field of visualization is at a crossroads. Advances in computer graphics technology and computing power have enabled the development of visualization techniques that have had a positive impact on medicine, computational science, bioinformatics, and finance. However, this focus on transitional efforts has not sufficiently addressed the basic science needed to create universal, validated principles on which to ground future visualization efforts. Without such principles, visualization risks becoming a niche or service field concerned only with iterative refinement of new and existing methods. Rigorous study of the *science* behind visualization systems could enrich existing efforts and suggest novel applications based upon predictive theories. With several recent calls to investigate visualization principles [1–6], now is the time to consider what is a science of visualization.

The purpose of the panel is to (1) assess whether a science of visualization is necessary and (2) discuss what is needed for such a science. We seek to discover the tools we need in order to examine why and how visualizations work. The panelists present different approaches to visualization science, from foundational theories to issues of practicality. The goal of this panel is to spark discussion about the need for a science of visualization and the real world barriers to its acceptance and adoption.

Keywords: science of visualization, foundational models, aesthetics, reproducibility, open science, evaluation, visual thinking, cognitive science, practical issues

POSITION STATEMENTS

T.J. Jankun-Kelly
Build Visualization Science Foundations from Commonalities

Every science is built from a core set of abstractions that define what that science is about. These foundational models not only provide a universal referent, but can be built upon to form predictive theories and measurable findings. Any rigorous science of visualization must be constructed from such foundational models. But what is needed for such a science of visualization?

To understand visualization—to move beyond the "implement, visualize, iterate" cycle—we need to break it down into its constituent parts: *How* the visualization was created, *what happened* during the visualization, and *what benefit* did the user get from the visualization/what did they want to do with the visualization. These core questions of visualization suggest three foundational models: A model to describe the visualization method (a *visualization transform model*), a model to describe the use of the visualization (a *visualization exploration model*), and a model to predict/measure the success of a given visualization method or session (a *visualization design model*). Similar models have been developed in HCI, in cognitive psychology, in other fields; such models would give visualization a rigorous foundation, instigate new directions of research, and provide a basis from which to build curricula for training.

Luckily for us, elements of these models exist, scattered throughout visualization literature. Unfortunately, these efforts have not been significantly developed or rigorously tested for their validity and universality. Many of these disparate efforts focus on specific problem domains, and do not all generalize to visualization—scientific and information—as a whole. Why is this? Visualization, as a community, rewards new methods and systems, not "navel gazing" and searching for fundamental truths. While methods and systems are important—especially to our users—a scientific field lives or dies based upon how it can describe and predict reality. We do not need the 1001st isosurfacing paper; we need testable, validated theories that predict when an isosurface is useful—and when alternate approaches are better.

It is time to look back at the commonalities amongst the empirically-based results of the last 20 years to build a rigorous, predictable science of visualization.

Robert Kosara
Visualization is not a Hard Science

When looking at a science of visualization, there are some obvious candidates to take ideas and concepts from: computer graphics, statistics, and perceptual psychology.

But in order to develop a fundamentally new scientific basis for what we are doing, we have to take a step back to look beyond the current state of the art, and rethink the field from the beginning. Why did visualization start in computer science in the first place? What are we really doing, and why are we doing it? In addition, we should question the idea of a science of visualization right from the beginning. This is not to say that we do not need such a science (we certainly do), but that it will likely have to be very different from many other sciences in order to be useful.

The most basic description of visualization is that we produce images from data. In other words, we depict data, and we do that to communicate information. This is not at all unlike what artists, illustrators and designers do (and have done for a long time). Like artists—and unlike traditional scientists–we build artifacts. Unlike artists, we design fairly general tools rather than specific pieces for communicating just one idea for one purpose; and we also need to produce things that are useful.

A question that emerges quite often in visualization—but is usually ignored—is that of aesthetics. After all, the purpose of visualization is insight, not (pretty) pictures. But is an aesthetic image not also a visually effective one? What is the connection between aesthetics and utility in visualization? How does aesthetics apply to visualization? What is aesthetics, anyway?

In addition to depiction, visualization almost always involves interaction. How can communication research inform the design of visualization systems, and how can it help to provide meaningful interaction? How can understanding visual representation make it easier to understand visualization, even if the images are completely abstract? How can we enhance the usefulness of a visualization by making it tell a story?

In addition to all that, we are living in a highly visual world: television, video games, commercials, billboards, PowerPoint, etc. What if visualization was just an extension of this? In the context of postmodernist thinking, it certainly is: we are dealing with visual means of communication (the domination of visual culture is a defining property of postmodernism), as well as new ways of communication and the use of signs. As such, we can connect with many fascinating current developments, and should be shaping and enhancing fields like visual studies, media theory, and visual rhetoric.

Do we need a science of visualization? We absolutely do. But that science will not be a hard science, and and not just a single one, either: we need many different sciences. We need a semiotics of visualization, an aesthetics of visualization, a philosophy of visualization, a communication science of visualization etc. And we need to connect to the other sciences that our work is linked to, to learn—but also to enrich them with new ideas and questions.

Gordon Kindlmann
Lack of Reproducibility Hinders Visualization Science

Reproducibility is one of the principles of the scientific method: A result is meaningful if other researchers can conduct the same experiment and arrive at the same result. To date, however, most visualization research is not easily reproduced, in both a specific and a general sense. Given a visualization paper, an independent researcher would likely have a hard time exactly recreating the final images from the same underlying data. This is due to the typical unavailability of the author's implementation (or any reference implementation) for comparison and for detailed algorithmic understanding, as well as the absence of complete parameter settings information. The difficulty of recreating specific visualization results contributes to a more general problem with the scope and power of visualization research: the reported success of a particular method in a particular problem domain (such as medical imaging) may be hard to reproduce in other fields of research (say, computational fluid dynamics), due to differences in data representation, exploratory process, user expertise, or research goals.

I argue that visualization must be reproducible if it is to be scientific, whether visualization is merely a supporting technology or a scientific field unto itself, and that we can make visualization more reproducible by following the example of recent developments in other fields. Considering visualization as a supporting technology for some overlying scientific or engineering application domain, visualization reproducibility would simplify the task of reproducing and extending results in that domain, as well as offering a standard tool for visually debugging and explaining its methods. In this context, visualization reproducibility is an essential prerequisite to making visualization a commodity, and an every-day ingredient in modern science. In contrast to lab supplies sold by various vendors that meet established technical specifications (e.g. lenses, pipettes, reagents), visualization solutions tend to be more customized and home-brewed. The problem becomes more acute as visualization techniques increasingly build upon each other, and the cost of reimplementation becomes a greater and greater portion of the work involved in evaluating and applying the research of others.

The problem here is connected to how visualization research is disseminated, and solutions may be inspired by recent advances in software development and scientific publishing. Making our software Open Source would at very least permit detailed understanding of how published algorithms are actually implemented. Ideally, Open Source visualization research would also foster a common pool of code that multiple researchers extend and build upon, complementing tools like the Visualization Toolkit (www.vtk.org) for more established methods. The recent shift away from traditional scientific publishing models towards Open Access (exemplified by the Public Library of Science journals; www.plos.org) is based on principle that the fruits of publicly funded research should be a publicly accessible resource. Thus, Open Access publications are freely electronically available without subscription fees, but more pertinent for visualization research, data and supplementary materials (such as parameter settings) can be published alongside the text of the paper, to facilitate the evaluation and exploration of the work by the community. A more radical model is the Insight Journal (www.insight-journal.org) for image processing, which requires that the software compile on multiple platforms and exactly reproduce the submitted results.

The goal of reproducibility becomes more challenging when visualization is considered a scientific domain of inquiry unto itself, more general than any single application area. In that case, a visualization result is (by definition) scientific only if it is reproducible, in the strong sense of being able to generalize the method to different research areas that have similar opportunities for insight enabled by visual analysis. In this context, reproducibility becomes a telling indicator of the maturity of the field of visualization. I believe this more general kind of visualization reproducibility is currently hampered by a range of factors that differentiate the various research areas in which visualization is used, from what may be entirely superficial (such as differences in data structures and formats, terminologies, user interfaces, or image aesthetics), to aspects that are perhaps more fundamental (potential influence of visualization on the research process, or cultural acceptance of new technologies). These differences highlight the importance of developing basic conceptual models for visualization methods, and for how visualization enables insight, a theme that my fellow panelists address. Better visualization models could help us predict and describe exactly how and where visualization successes can be extended and applied to other areas of inquiry. The scientific power of visualization, and the science of visualization, will be enriched and amplified through reproducibility.

Chris North
Visualization Science Requires Methods for Measurement

Two primary components of science are measurement and modeling. Scientists must observe and measure phenomena in the world around them in order to create theories that model those phenomena. Models can then be further refined and validated through additional measurement. The models then serve as foundational principles for engineers to construct new tools that exploit the modeled phenomena, and the measurement methods enable them to hone the tools.

A science of visualization should establish measurement methods and models that address the fundamental purpose and goals of visualization. In visualization science, what should be modeled? What should be measured and how do we measure it? In current visualization research, measurement methods focus primarily on human performance time and accuracy on benchmark tasks. This leads to straightforward mechanical 'rules of thumb' about low-level task performance. But these methods make it difficult to advance models for visualization science that address the higher level purposes of visualization.

For example, one potential claim is that the purpose of visualization is insight. In that case, visualization scientists will require methods for measuring insight, in support of establishing higher level models. But what is insight? How can it be measured? How can the process of insight acquisition be observed? How can we map the visualization affordances to insight gained? Thus, new methods will be needed for measuring the interactive processes that take place during visualization usage, and measuring the effects of visualizations on human users.

To enable major advances to occur in visualization science, visualization scientists must first make a significant research investment into the development of such new and innovative measurement methods. Then, armed with a toolbox of a variety of useful methods for visualization measurement, visualization scientists can begin to measure and model visualization phenomena.

Colin Ware
Visualization Science is the Science of Visual Thinking

Visualization is about creating visual thinking tools. For our purposes visual thinking can be thought of as a form of distributed cognition where some computational processes occur in the visual pattern finding processes of the human brain and other computational processes occur in the algorithms that map data to a visual representation. Although visualization is usually treated as an algorithmic problem, at a fundamental level it is also a perceptual and cognitive problem. This implies that we should understand how visual perception works in order to create the optimal visual representations for this pattern finding to occur. But the question here is. Do we need a special new science of perception relating to data visualization, or can we simply take advantage of the vigorous science of perception that already exists?

There are lots of reasons to use the methods of science in empirical evaluation of visualizations but this is not enough to make visualization a science. By usual definition visualization we must own a body of theory that makes testable predictions.

On the perception side, I believe that this is possible. What makes perception of visualizations different from perception of the everyday world partly has to do which what we put on the screen in many cases there may be no equivalents in nature of the patterns we can produce. Another difference derives from the cognitive tasks that are executed when we think about data using a visualization. These tasks are special because they deal with the analysis and communication of information.

Nevertheless, the basic neural machinery is used for both everyday world perception and visualization, which means that the perceptual part of visualization theory has to be closely linked to perceptual theory. This requires a huge commitment because vision research is highly specialized and developing at a frenetic pace. But ignoring existing research means doing work that is amateurish and irrelevant.

Classical sciences (e.g. physics, chemistry and biology) aim to describe and explain the natural world. Visualization is one of the emerging disciplines of the cognitive sciences (psychology, cognitive neuroscience, artificial intelligence, augmented intelligence. human-computer interaction). These aim not only to describe and explain cognitive processes but to enhance and develop them.

Wes Bethel
Will a Visualization Science Even Be Used?

Many in the "hard sciences" view Computer Science as a "Johnny-come-lately" and lacking rigor in terms of scientific methodology; the field of visualization is often viewed by outsiders as being even "more soft" than pure CS. As a result, there is a "credibility gap" between a segment of our customer population—scientific researchers—and us. They make statements like "3D visualization won't help us understand petascale datasets" and "How is it that these funny goggles will help me better understand how a particle accelerator works?" Armed with better-grounded and proven theories, we will be better equipped to answer such fundamental questions and to achieve more significant impact on our customers' problems.

As all in our community know, visualization as a discipline cuts across many different fields—computer science, computational science, art, cognitive psychology, speech and communication, sociology, as well as corners of our customer's disciplines. One of the many challenges in "implementing scientific rigor" in our field is the sheer breadth of the problem space.

My fellow panelists present excellent, well-considered and cogent arguments germane to the need for scientific rigor in our field—I am in complete agreement with them in that regard. There is no question in my mind that more scientific rigor would be beneficial to our community, our field and our customers. In practical terms, there are several angles we may wish to consider:

- There is a plethora of hard evidence showing the hue ramp color table is the worst possible one to use. Given such evidence, why do we keep using it? In other words, what good is "visualization science" if we don't put it into practice?
- Related to the above: What impetus or reward is there for putting scientific rigor into practice? The next generation of the world's most popular "information visualization" application has chosen to ignore input from our community favoring 3D-pie charts with drop shadows over more useful concepts like intelligent layout and colorization schemes.
- What impact will scientific rigor have on users? Interestingly, all major visualization reports on my bookshelf (including those from 1987, 1998 and 2006) tend to emphasize the research needed in new and better visualization algorithms and systems; there is very little emphasis upon the need for scientific rigor. Our funders base their programs on these documents—we must convey to our funders the value of scientific rigor in terms of impact to the community and our customer base.
- We must sensitive to striking harmony between the tension created between the opposing forces of "absolutely correct" versus "close enough to get the job done." More than one major project has failed due to the obstinate desire for absolute perfection.

Biographical Sketches

T.J. Jankun-Kelly

T.J. Jankun-Kelly is an assistant professor of computer science and engineering within the James Worth Bagley College of Engineering, Mississippi State University. His research areas are at the intersection of scientific and information visualization. His goal is to make visualization techniques and systems more effective by improving interaction methods and visualization utilization. Towards this end, he focuses on visualization interfaces, visualization modeling, and applications such as volume, graph, and security visualization. T.J. has a Master's and Ph.D. (Computer Science) from the University of California, Davis and a B.S. (Physics/Computer Science) from Harvey Mudd College. He is a member of the ACM, SIGGRAPH, IEEE, and the IEEE Computer Society and was a founding Contest Co-Chair for IEEE Visualization during 2004–2006.

Robert Kosara

Robert Kosara is an Assistant Professor of Computer Science at the University of North Carolina at Charlotte (UNCC). He received his M.Sc. and Ph.D. degrees in Computer Science from Vienna University of Technology in Austria. His interest in information visualization comes from his enjoyment of all things visual, which is why he believes that a good visualization researcher has to be as competent in visual representation and communication as in the technical side of the field.

Gordon Kindlmann

Gordon Kindlmann is a post-doctoral research fellow in the Laboratory of Mathematics in Imaging, in the Department of Radiology at Brigham and Women's Hospital, Harvard Medical School. His current research interests include diffusion tensor imaging and analysis, direct volume rendering, and color science. He received his PhD in computer science from University of Utah in 2004, and his MS (computer graphics) and BA (mathematics) from Cornell University in 1999 and 1995, respectively.

Chris North

Chris North received the PhD degree from the University of Maryland, College Park. He is an assistant professor of computer science at Virginia Polytechnic Institute and State University, is head of the Laboratory for Information Visualization and Evaluation, and a member of the Center for Human-Computer Interaction. His current research interests are in the HCI aspects of data visualization, including interaction techniques for large high-resolution displays, evaluation methods, and multiple-view strategies. In applied work, he collaborates with faculty in bioinformatics, network security, and construction engineering.

Colin Ware

Colin Ware's research is focused on applying our understanding of human perception to information display. He has published more than 100 articles on this subject and a recent book: *Information Visualization: Perception for Design*. He is currently working on a new book on Visual Thinking. In addition to theory-based research Ware likes to build useful visualization systems. Fledermaus, a GIS visualization package originally by him and his students, is now the leading 3D visualization system used in oceanography. Ware is a professor of Computer Science and Director of the Data Visualization Research Laboratory at the University of New Hampshire. He has degrees in both experimental psychology (PhD, Toronto) and Computer Science (M.Math, Waterloo).

Wes Bethel

Wes Bethel is head of the Scientific Visualization Group at Lawrence Berkeley National Laboratory. Their mission is applied research and development of high performance, remote and distributed visualization and analytics technologies aimed at solving some of the world's most challenging computational and experimental science data understanding problems.

REFERENCES

[1] C. R. Johnson, R. Moorehead, T. Munzner, H. Pfister, P. Rheingans, and T. S. Yoo, editors. *NIH-NSF Visualization Research Challenges Report*. IEEE Press, Los Alamitos, CA, USA, 1st edition, 2006. http://tab.computer.org/vgtc/vrc/index.html.

[2] Chris Johnson. Top scientific visualization research problems. *IEEE Computer Graphics and Applications*, 24(4):13–17, 2004.

[3] Bill Lorensen. On the death of visualization. In *Position Papers of the NIH/NSF Proceedings of the Fall 2004 Workshop on Visualization Challenges*, 2004. http://visual.nlm.nih.gov/evc/meetings/vrc2004/position_papers/lorensen.pdf.

[4] Theresa-Marie Rhyne, Bill Hibbard, Chris Johnson, Chaomei Chen, and Steve Eick. Can we determine the top unresolved problems of visualization? In *Proceedings of IEEE Visualization 2004*, pages 563–566. IEEE Computer Society, 2004.

[5] James J. Thomas and Kristin A. Cook, editors. *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. IEEE Computer Society Press, 2005.

[6] Jarke J. van Wijk. The value of visualization. In *Proceedings of IEEE Visualization 2005*, pages 79–86. IEEE Computer Society, 2005.