VIS 2006
BALTIMORE·MARYLAND·USA

Is there Science in Visualization?

T.J. Jankun-Kelly (Mississippi State University)
Robert Kosara (UNC Charlotte)
Gordon Kindlmann (BWH, Harvard Med School)
Chris North (Virginia Tech)
Colin Ware (U. of New Hampshire)
E. Wes Bethel (Lawrence Berkeley National Laboratory)

VIS 2006

BALTIMORE·MARYLAND·USA

Is there Science in Visualization?

NO!

VIS 2006
BALTIMORE·MARYLAND·USA

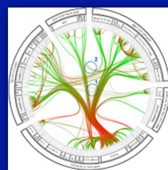
Is there Science in Visualization?

YES!

VIS 2006
BALTIMORE·MARYLAND·USA

Is there Science in Visualization?

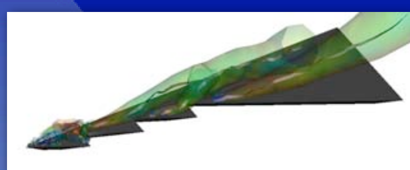
Maybe?



Information Visualization

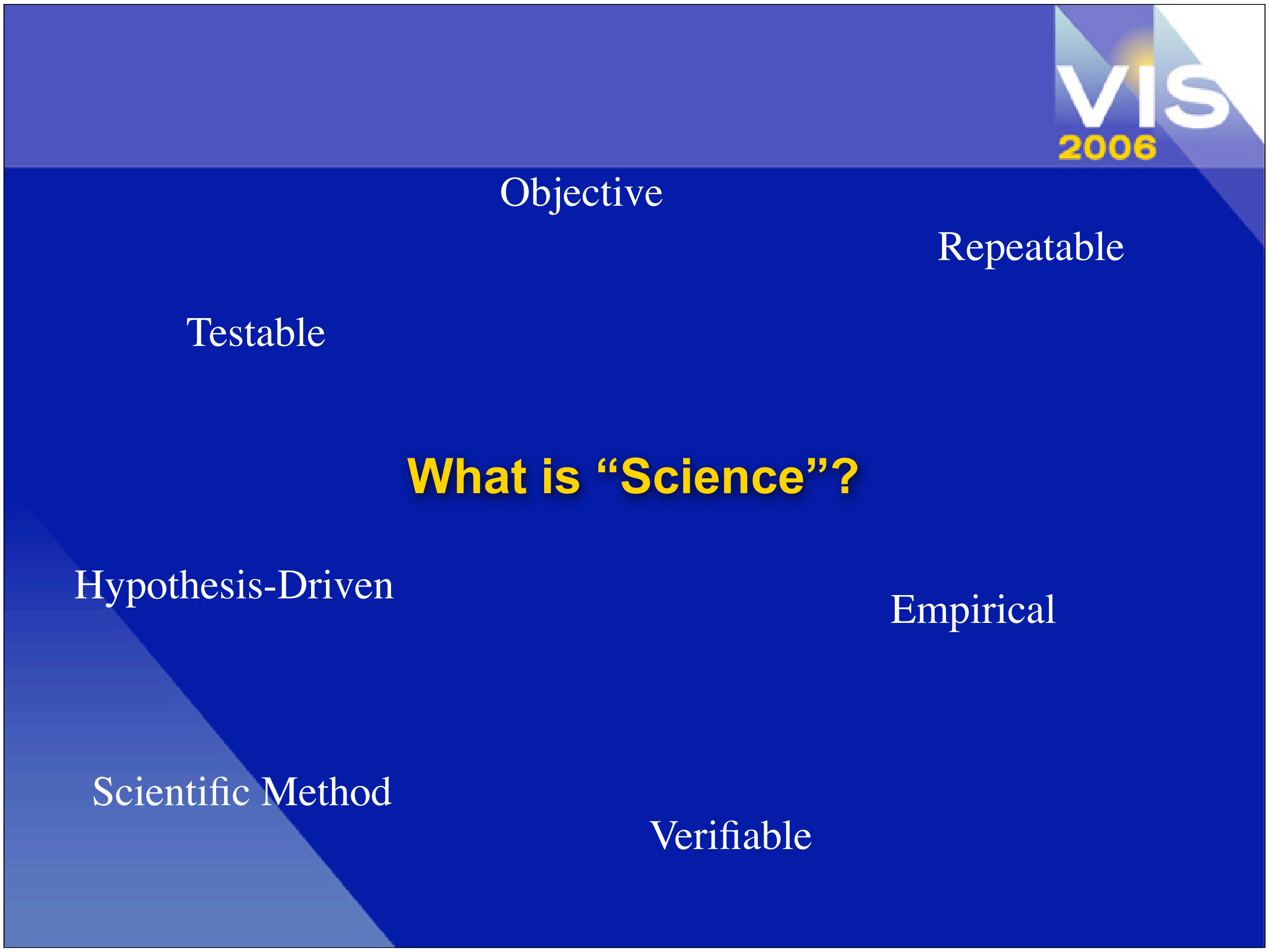
But is it “Science”?

tics



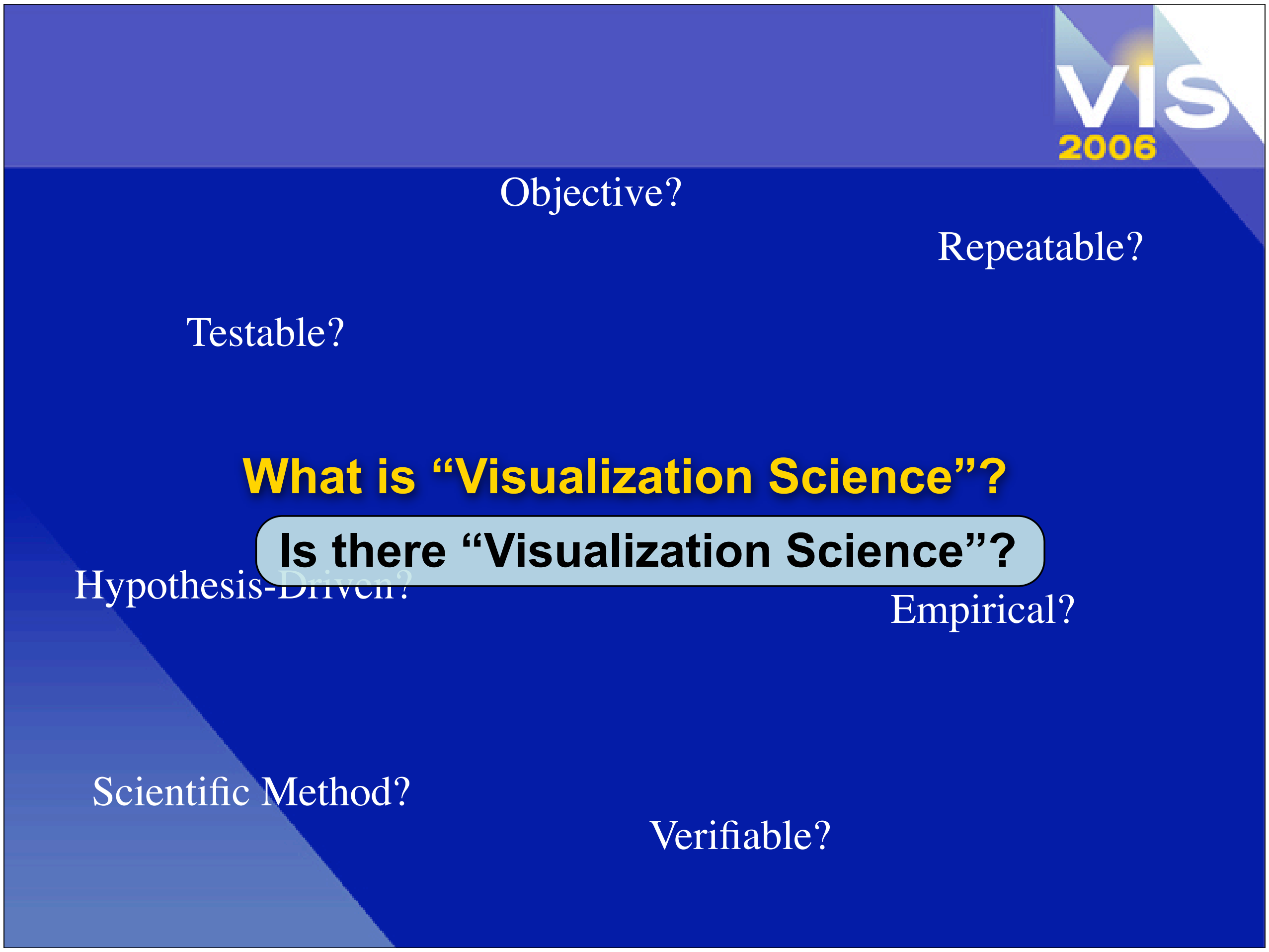
Scientific Visualization

Both call for more fundamental science



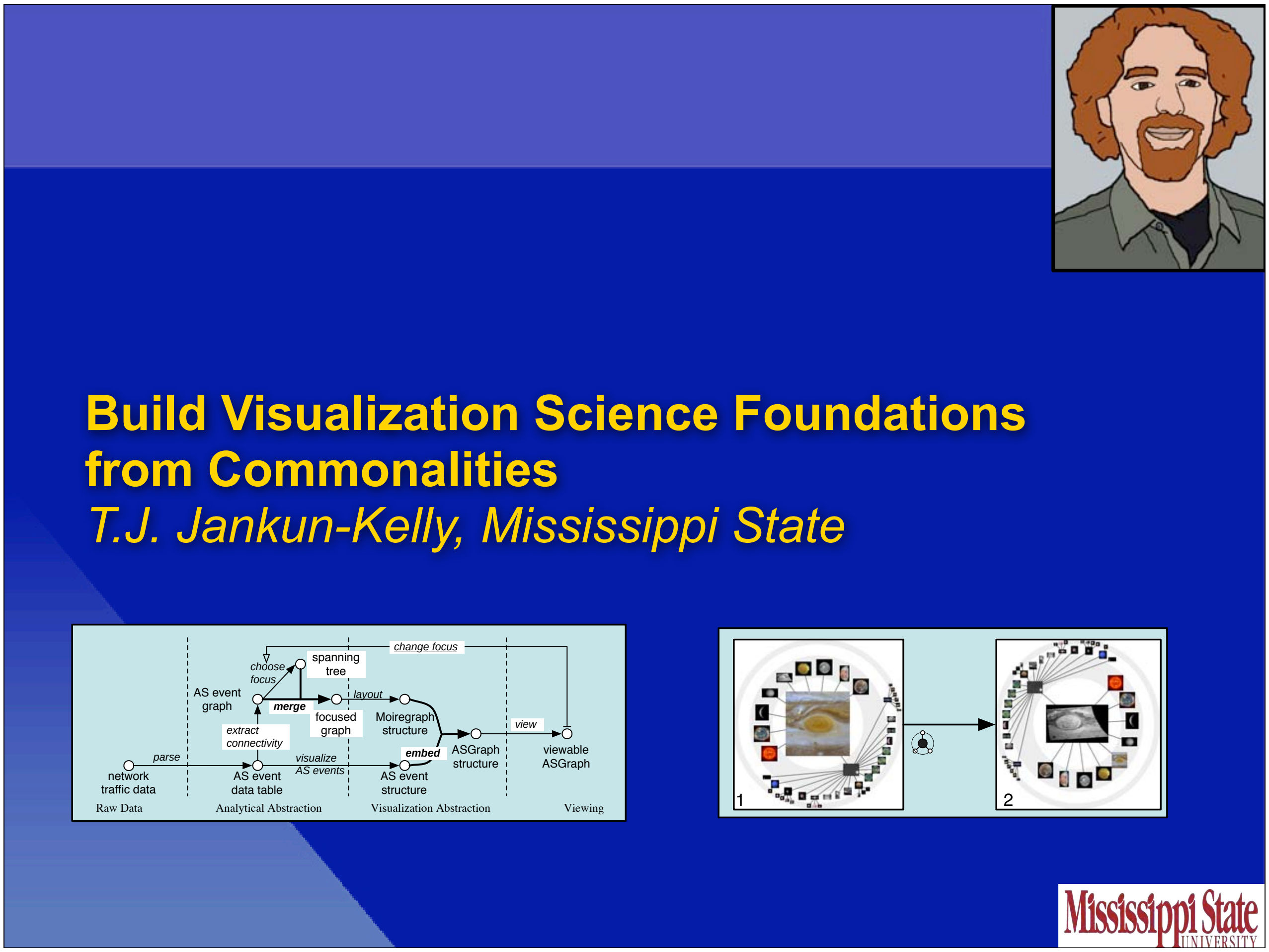
Science in the broadest sense refers to any system of knowledge attained by verifiable means.[1] In a more restricted sense, science refers to a system of acquiring knowledge based on empiricism, experimentation, and methodological naturalism, as well as to the organized body of knowledge humans have gained by such research. This article focuses on the meaning of science in the latter sense.

Scientists maintain that scientific investigation must adhere to the scientific method, a process for developing and evaluating natural explanations for observable phenomena based on empirical study and independent verification.



Science in the broadest sense refers to any system of knowledge attained by verifiable means.[1] In a more restricted sense, science refers to a system of acquiring knowledge based on empiricism, experimentation, and methodological naturalism, as well as to the organized body of knowledge humans have gained by such research. This article focuses on the meaning of science in the latter sense.

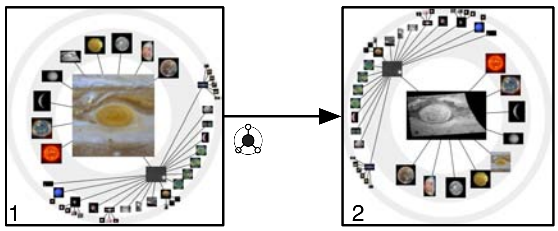
Scientists maintain that scientific investigation must adhere to the scientific method, a process for developing and evaluating natural explanations for observable phenomena based on empirical study and independent verification.



Parts of foundational models to build visualization science exists, but we must synthesize them and reward their development.

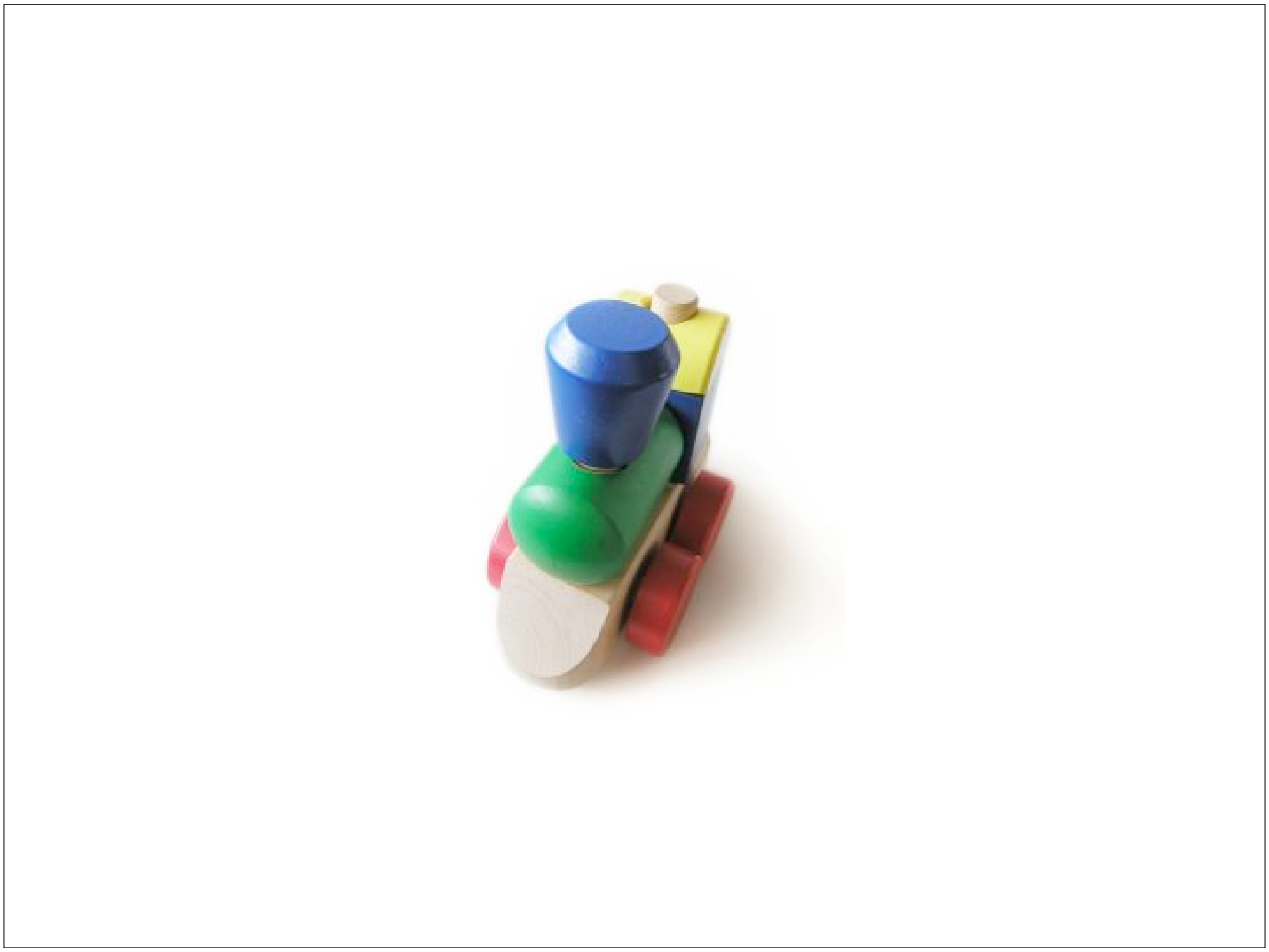
Build Visualization Science Foundations from Commonalities

T.J. Jankun-Kelly
Mississippi State University

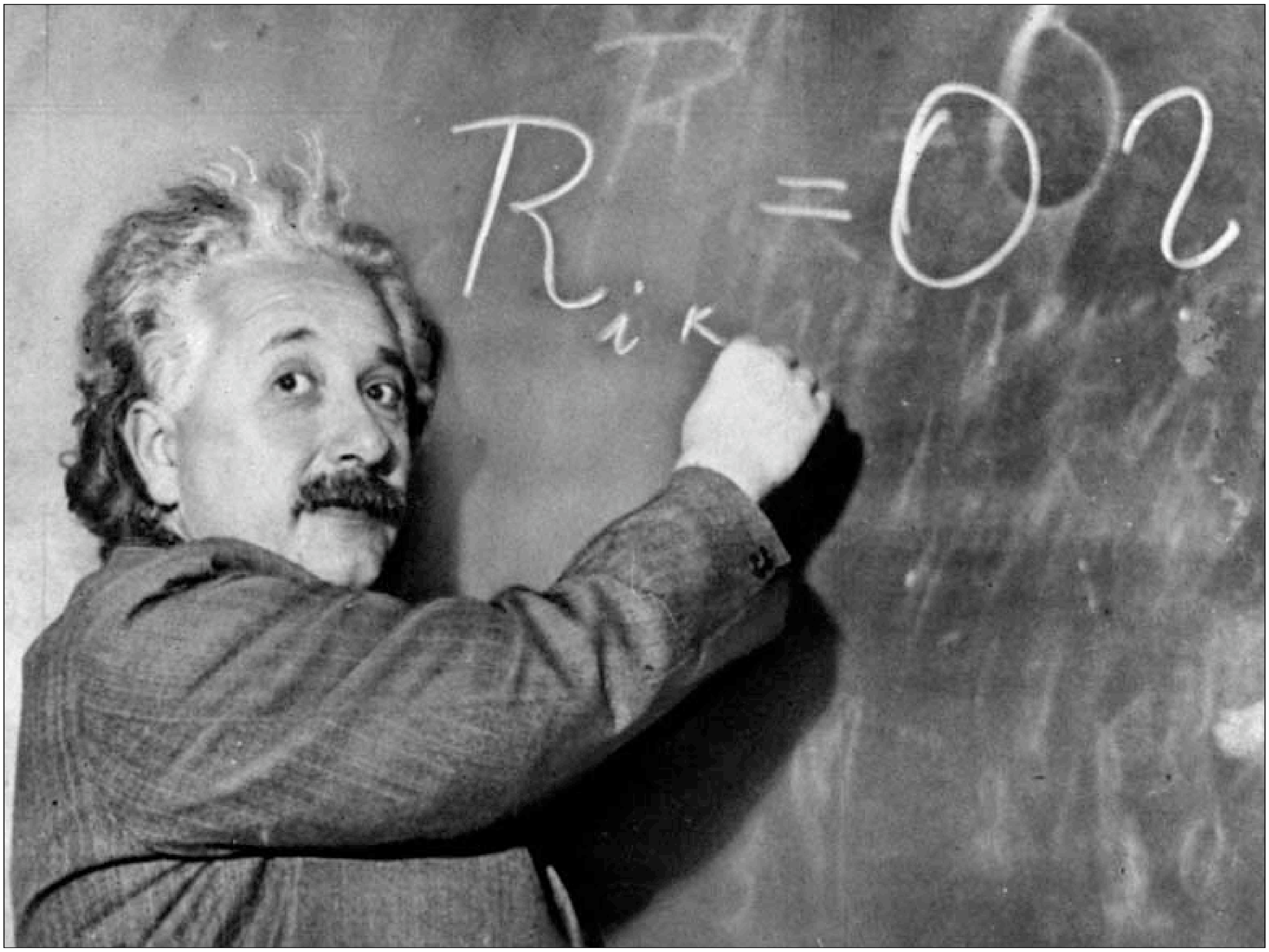


Mississippi State University

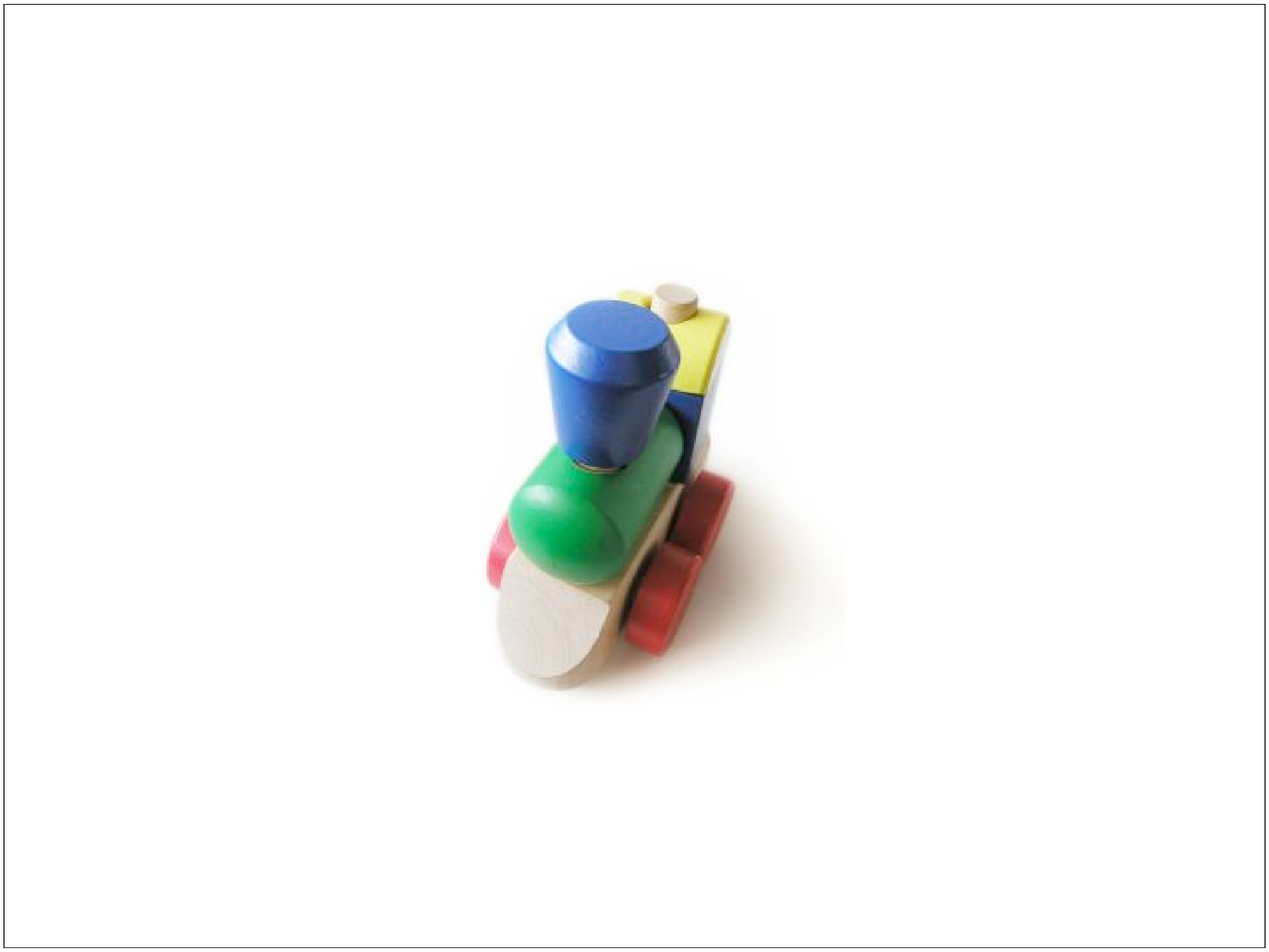
JAMES WORTH BAGLEY COLLEGE OF ENGINEERING MISSISSIPPI STATE UNIVERSITY



Today, I’m going to talk about models. Specifically, the models I think we need to make a “real” “visualization science.” It is my position that any science cannot develop without such models. But first, you need to know a bit about me.



I was trained as a physicist. I also have a minor in the history of science. So I have spent a good bit of time in the “hard” sciences doing “hard” science things. And one of the chief things you learn as a physicist is to develop a good “intuition” about how physical behavior works---i.e., to intuit physical models of reality.



Because of this training, I tend to think of things in a very “model-centric” point of view. Just like CS is built on abstractions, natural sciences are built on abstractions as well. But these abstractions are models---**testable** representations of reality. These models provide a consistent description of an aspect of reality, and also provide a context for the field to work with that chunk of reality.

But are visualization models scientific?

Natural scientific models tend to be **empirical** in a nature---i.e., they are formed by observation of the real world and tested by hypothesis, experimental design, and empirical validation. An empirical model describes the world as we see it (or how our tools see it) now and should be in the future; and thus needs to be able to predict subsequent behavior and be amenable to correction from new observations.

But this is not the case in visualization. Our models are **constructive** in nature. We still have hypothesis. We can still test them. We can empirically validate them; some of them more easily than others. All those running time tests you see in a SciVis paper is a classic case of a constructive proof of a constructively derived model. But constructive models are not all scientific. Consider string theory---a mathematical, constructed model that describes physical behavior, but several think cannot be tested and thus hinders science. So we must be careful about what we call “scientific.”

The textbook definition of a scientific model is something that “describes reality.” But our “visualization” models are not things which exist a priori in reality. We create them. They are not something we **see** or **measure** but something we **create** or **theorize** about. Is this really science?

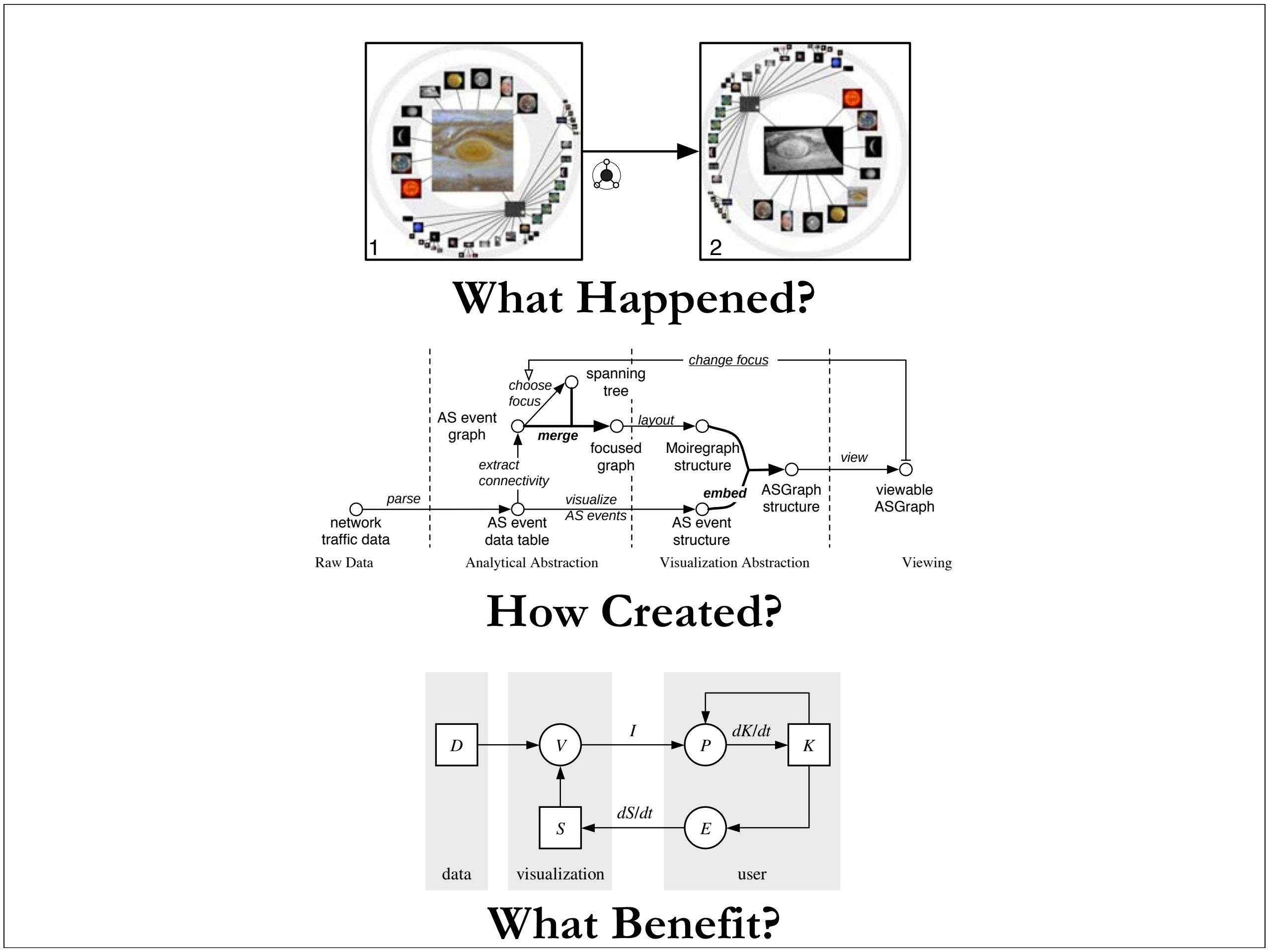
I claim these models are “scientific” if they can be treated by scientific mechanisms. I.e., they are created, tested, and validated via the scientific method. So while these models may not have a priori natural existence and cannot be “observed” per se, if they describe behavior---even if it is computational---in a way that is observable, testable, and refutable, we should be set.

But what are these models for visualization? And are they scientific?

Natural sciences are built from “foundational” models. These are models that are the building block of the science. Newton's Laws and Light-Particle duality in Physics. Geonomic Theory in Microboiology.

Foundational models provide an ontology that defines the science’s language, and they provide a scope marking the boundaries of what that the science is about. And, most importantly to me, they provide a description of the overall activity that goes on in that science. These models are **general** to the entire science but specific enough to accurately **encapsulate** the science through possible descriptions and predictions of phenomena.

So what are these in visualization?



I propose that there are three major models that visualization science needs to delineate its scope, provide a framework for the field, and actually describe and drive what is going on:

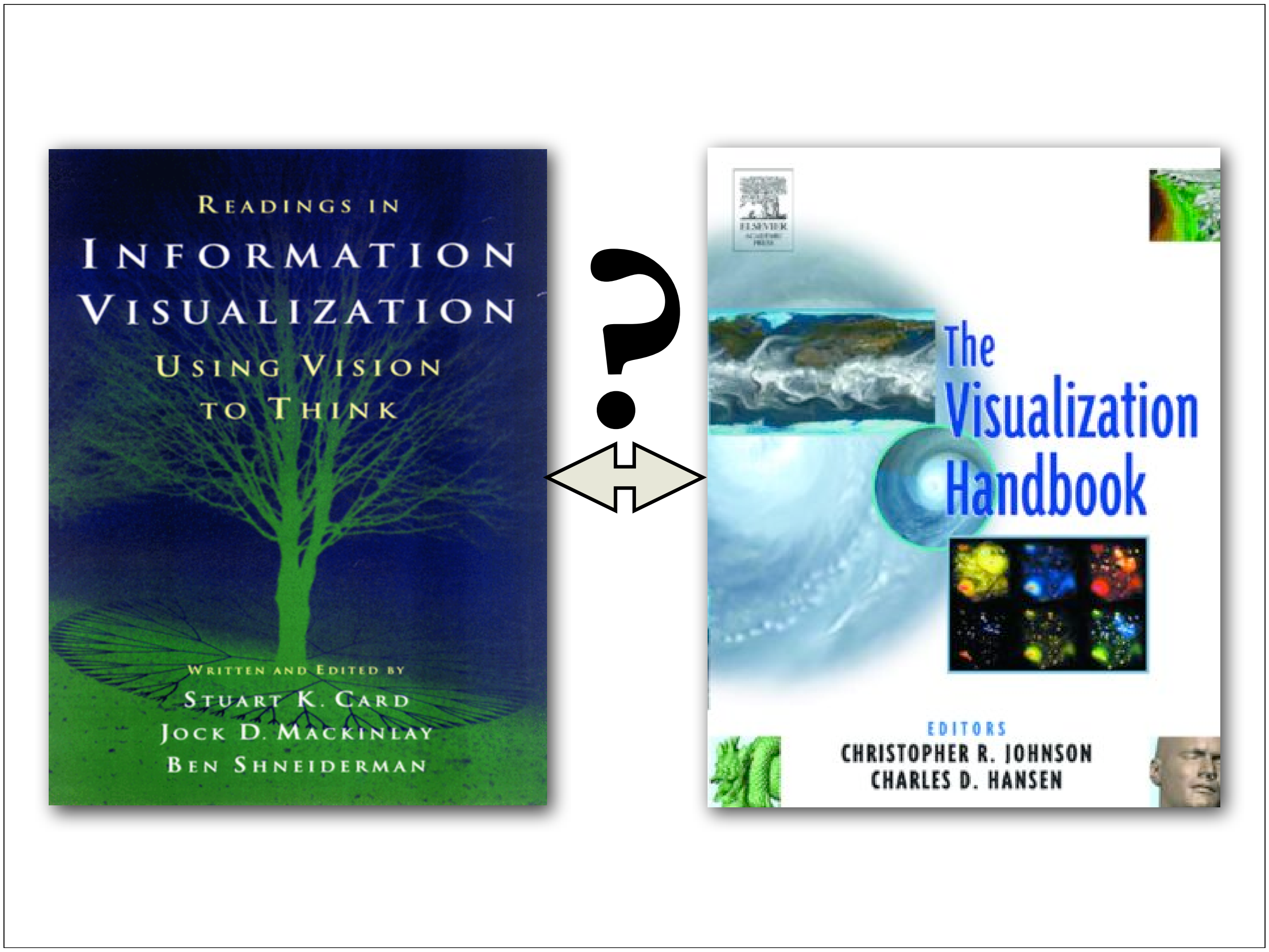
- **Exploration Models** They tell us what a user actually does during a visualization. This models is what people do with visualization.
- **Visual Transform Models** They describe how a visualization was actually created. This models how we create visualizations.
- **Visualization Design Models** These models predict what benefit a visualization will provide based upon perceptual, cognitive, economic, and other principles.

These models provide a foundation to describe other activities in visualization. They provide a **rigorous** basis for the science of visualization. Every aspect of visualization is encapsulated in these three types of models. To name a few examples:

- Transfer function design is an aspect of visualization design models. The model would predict what t.f.'s are better in what situations based upon grounded principles.
- Data caching for very large data streaming to GPU-based renderers is an applied aspect of visualization transforms models. The transform model would suggest the trade-offs from previous approaches.
- The appropriate visual mappings for ordinal and nominal data in multivariate visualization such as parallel coordinates is a visualization

There is good news. Elements of these models **exist today!** They are scattered over the visualization literature, and in some case, in the literature of related fields. We just need to find and remember them.

But there several problems. And these are potential deal breakers

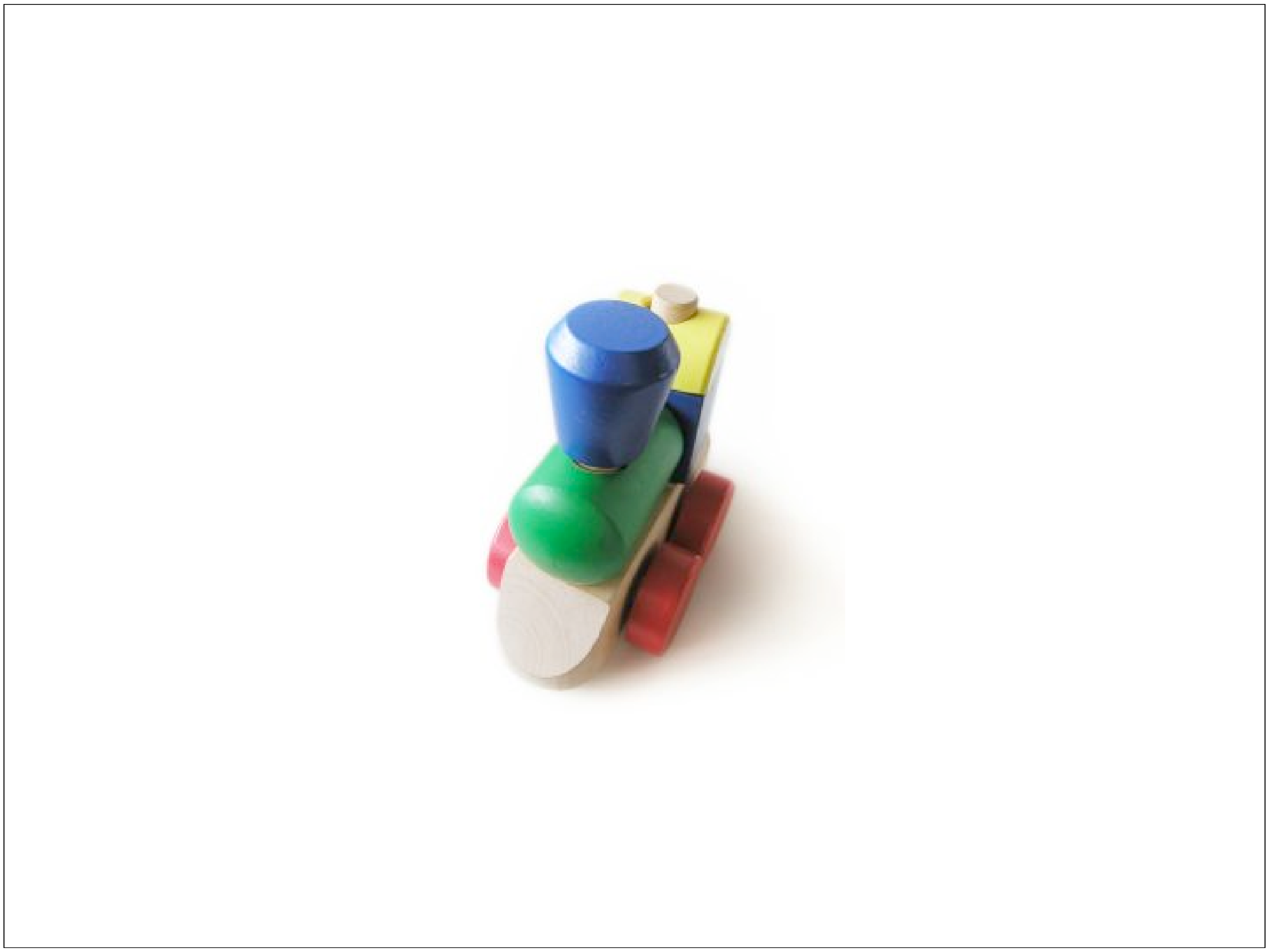


The first problem is while elements of these exist today, they are not ready yet. These proto-models **are not generalizable.** In my mind, visualization is a single science. There are specializations of it, to be sure, but it is a single science. And there is no reason why we need a data flow and visualization lattice model for scientific visualization transforms and a data state and visualization relational language models for information visualization when they do not even describe all the possible forms of visualizations that exist today or could exist in the future.

The second problem is that of the partial, incomplete models that we do have, **most have not been "scientifically" validated**. In some sense, many of these models have some validation---they describe the problem there were originally designed for, don't they? But is this really a rigorous validation? Is there a formal method to test these? Quantum physics and special relativity make predictions that can be tested. Does the P-Set Model of Visualization Exploration provide a similar facility? A model is not rigorous if it does not posses such capabilities, and it is not scientific either.

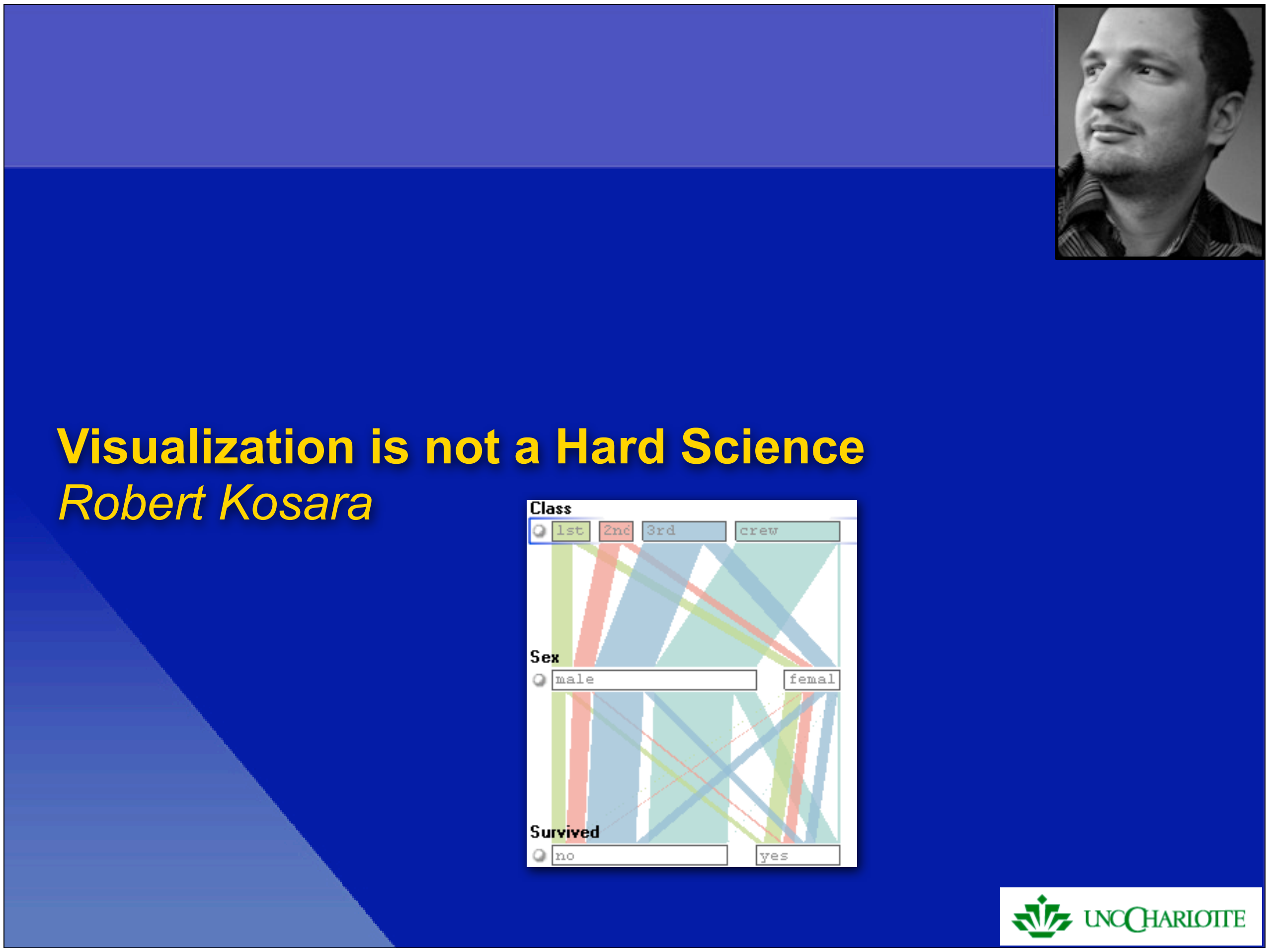
Of course, part of this difficulty in validation comes because it is not easy or well known how to validate many of these things. How do you measure “insight?” How does predict the value of a visualization design? How does one emprically validate the success of a visualization foundational model? New methods of research and standards of evaluations must be designed and rigorously vetted to establish our “visualization science.”

The final deal breaker, and perhaps the largest, is that there is little reward for such fundamental work. Visualization---scientific visualization in particular, though information visualization also has bought of this---is a very practically minded field. Formal “navel gazing” is not rewarded in visualization. Research papers and research funding has traditionally gone to algorithms and case studies. While groundbreaking modeling work has passed muster in the community occasionally, these are often passed as part of a larger more “practical” application at the same time---The P-Set Model is a means to efficiently share visualizations, the Data State and Data Flow Models are efficient means to share implementations of operations. These are not model formulations in and of themselves. If we do not reward researchers for doing such “fundamental” work, it will never be done.



In summary, I think formal models of visualization are needed to bring together and drive the field forward in a rigorous, scientifically robust manner. My personal opinion is that without them, we will miss the fundamental findings of visualization and instead focus on iterative refinement of the latest, and greatest millionth visualization algorithm. A research program focused on the development, refinement, and validation of such fundamental models can drive visualization forward for the next decade and beyond---each model has the potential to give us new ways of thinking about visualization that spark new improvements that require new models. And that is the way knowledge---and science---is done.

Thank you.



Visualization is not a hard science. We need to look beyond the fields we are currently considering to understand what we are doing.

Visualization is not a Hard Science

Robert Kosara

UNC CHARLOTTE

How many of you are researchers?

How many of you are scientists?

What is Wrong with Vis?

- No foundational theory
  - Visual representation
  - Perception
  - Interaction
- Ad hoc processes
- Focus on producing new things
- Very little learning and digestion

Visualization:

Making graphical representations to help people communicate information.

— Holly Rushmeyer

Stealing Others' Science

- Mathematics
  - Geometric modeling
  - Statistics
- Computer Graphics
  - Is that a science?
- Psychology
  - Perception
  - Cognition

Standing on the shoulders of giants

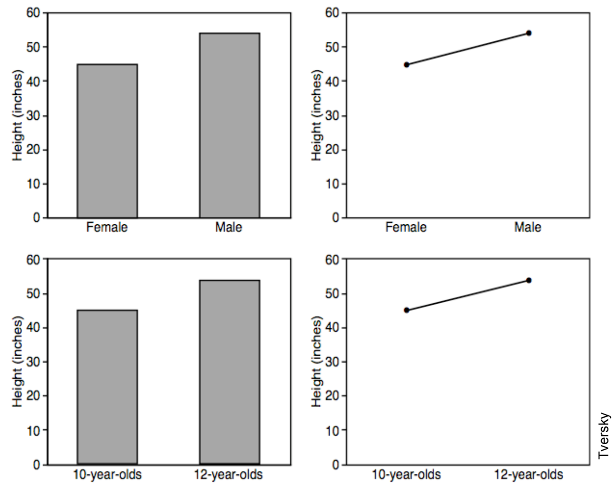
Pick the right giants!

A Broader Perspective

- Communication
- Design
- Aesthetics
- Illustration
- Representation
- Visual rhetoric
- Story-telling

Soft topics!

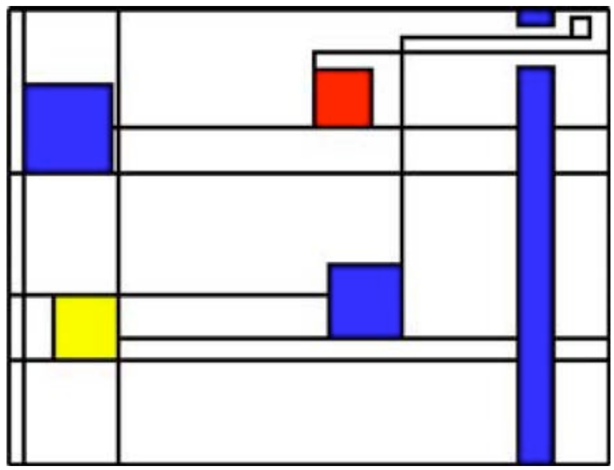
Visual Representation



Robert Kosara

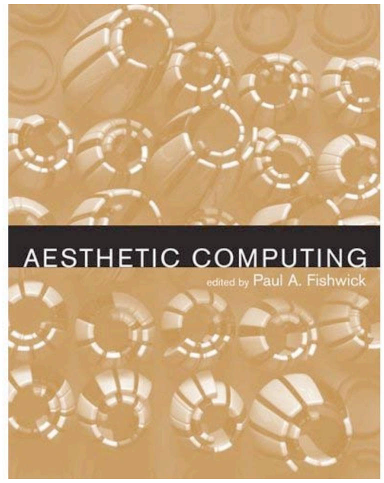
Visualization is not a Hard Science

Representation



Aesthetics

- Beyond pretty pictures
- Perception
- Meaning
- Representation
- Framework
  - Vis vs. Art



Propose a way of using this …

Visualization Criticism

- Apply the VisCrit
  - from your classes
  - to your research
- Criticize
  - Explain
  - Understand
  - Improve
- Develop a Language

Robert Kosara Visualization is not a Hard Science

There is disconnect between the teaching and research communities.

Used by lots of people, I got it from David Laidlaw

A Language for Visualization

- Describe
  - the visualization,
  - not the graphics
- Understand the interaction
  - What does it mean?
  - Not: Where do I click?
- Publish visualization
  - Not pretty pictures

Art theory is art criticism!

Languages are important, see Jeffrey Heer's talk at InfoVis: design patterns are a language

Christopher Alexander, A Pattern Language

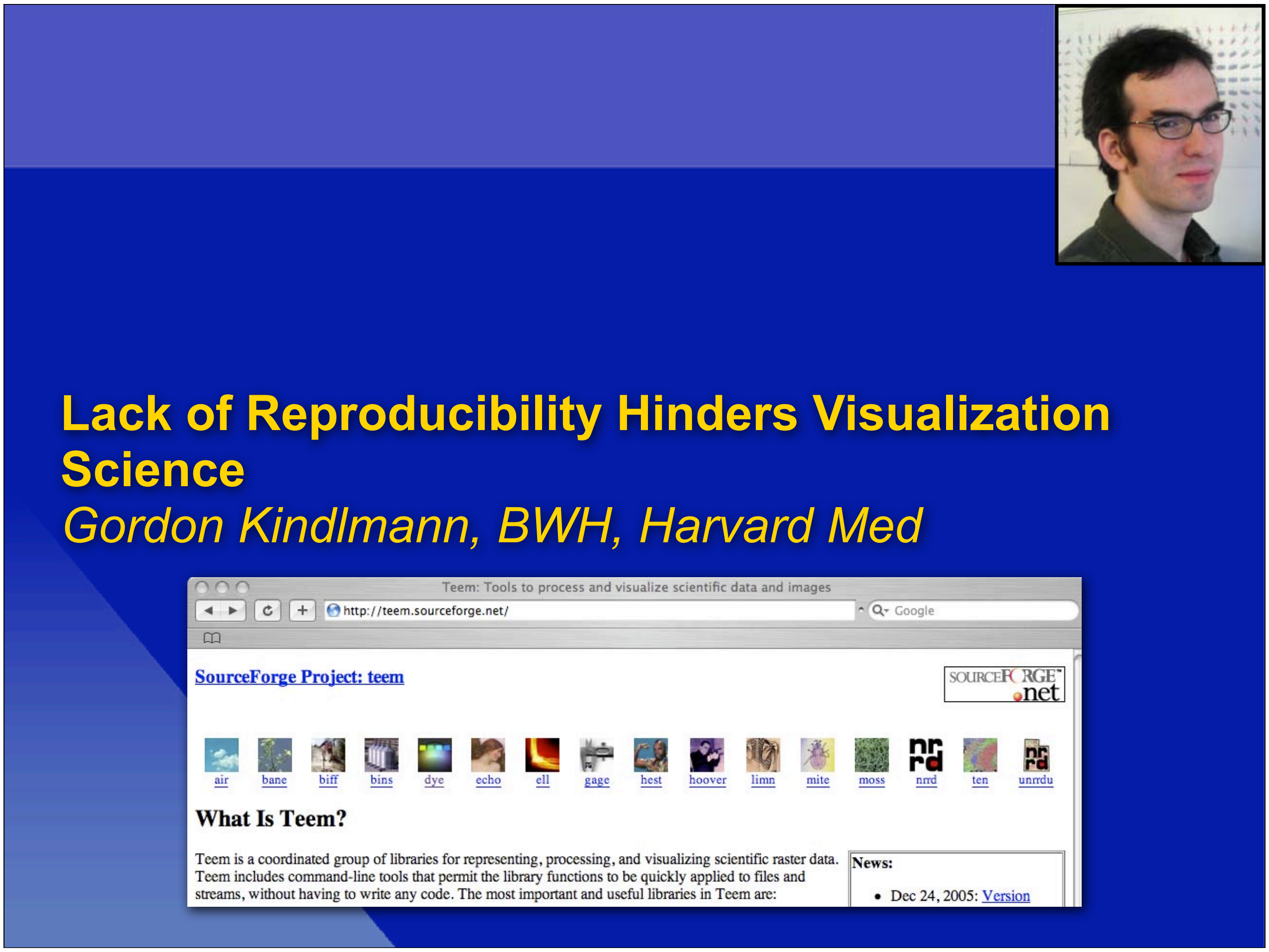
Too hung up on technical details

Rocket science

EagerEyes.org

Take a Step Back

- How is Visualization related to …
  - Computer Graphics?
  - Information Graphics?
  - Design?
  - User Interfaces?
  - Communication?
  - Illustration?
  - Art?



While the lack of reproducibility in visualization papers hinders Visualization Science, recent developments in software development and scientific publishing may inspire solutions.

Visualization is not Scientific without Reproducibility

Gordon Kindlmann

Laboratory of Mathematics in Imaging
Department of Radiology
Brigham & Women's Hospital
Harvard Medical School

Bullet Points

- There is science, but we can do better
- Whether or not visualization is a science unto itself, scientific impact is bounded by its reproducibility
- Reproducing previous results becomes more daunting as research become more sophisticated, bigger integrations of simpler methods
- Steps for improving reproducibility can be inspired by modern software development, and publishing (open source, and open science)

Reproducibility is non-negotiable

Supposing visualization is in support of science:

Visualization reproducibility enables enhancements and experimentation, and tools for visual debugging

Can the reader re-implement the method?

Does he/she have to (can get published without)?

Will it generate the same results (are the parameter settings the same)?

http://www.sci.utah.edu/~vgc/vistrails/

Point of comparison

Lab equipment (lenses, reagents, etc) are essential commodities for science

Lorensen's "Death of Visualization" ('04): the field of research suffers if vis is merely a commodity (e.g. volume rendering), but we should go further:

Compilers meet standards

OpenGL meets standards

Marching Cubes? LIC? Flow field topology?

How is Vis 2006 doing?

Of 63 Vis papers, I found 5 that made reference to available implementations

What can we do to increase the incentives for this?

Insight Journal

http://www.insight-journal.org/

Mechanism for vetting new code for ITK

Very high bar for reproducibility

- Code has to be multiplatform (e.g. CMake)
- Includes tests to verify correct behavior
- Includes code to generate figures

We can scoff at this as too restrictive for researchers

But is your research for the community, or for you?

Example of 3rd party evaluation

"Have you done a user study?"

"Ah, well, its future work."

Who is going to do that work, and how?

Why not outsource the evaluation?

Van Wijk's Vis '05 "Value of Visualization": we can increase its value by enabling others to evaluate it

Example: Laidlaw et al. "Quantitative Comparative Evaluation Of 2D Vector Field Visualization Methods" used Turk & Banks "Image Guided Streamline Placement".

Why you should release your code

The people who would benefit most from seeing your code up close and person is probably **the most** forgiving about the short-cuts, hacks, and lack of flexibility, etc. that you're uncomfortable with.

http://www.opensource.org

http://teem.sf.net used in Blaas Vis '05 "Fast and Reproducible Fiber Bundle Selection in DTI Visualization"

Public Library of Science

http://www.plos.org

New model of electronic publishing

Open Access: source data available, additional electronic resources

Not the author’s responsibility to maintain on their web page

Reproducibility is non-negotiable, part 2

If visualization is a science unto itself (strong idea):

With accurate and complete models of data and visualization, we could predict the success of a vis method in a novel context (e.g. other panelists).

Then the visualization result is scientific by definition only if it is reproducible.

Problems

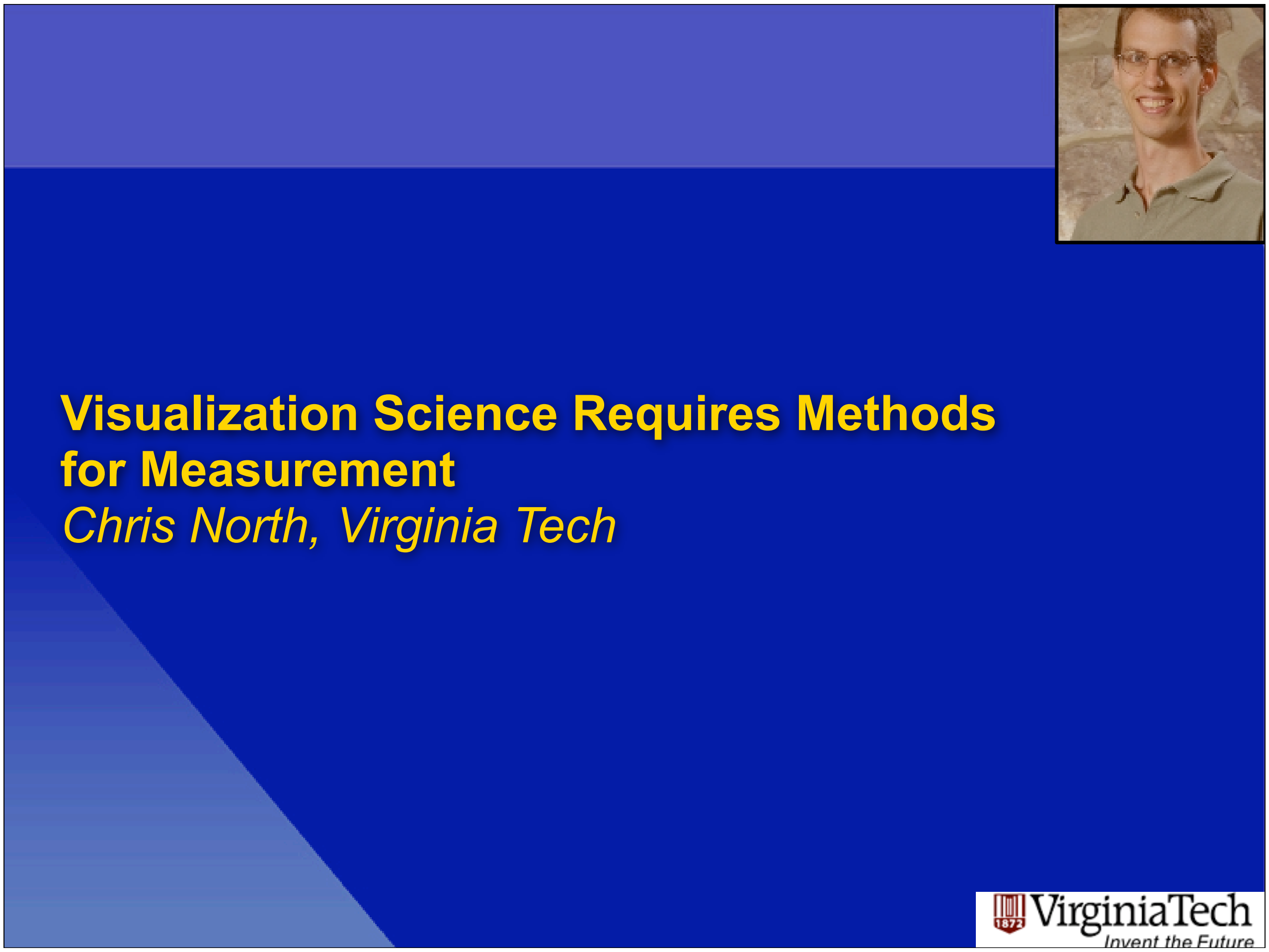
Failure to test and evaluate is a credibility problem for the community (Peter J. Denning '05 "Is computer science science?")

http://cs.gmu.edu/cne/pjd/PUBS/CACMcols/cacmApr05.pdf

Ability to test/evaluate is a software distribution problem

Incentive to test/evaluate is a community problem (what counts as a publication?)

The scientific power of visualization, and the science of visualization, will be enriched and amplified through reproducibility.



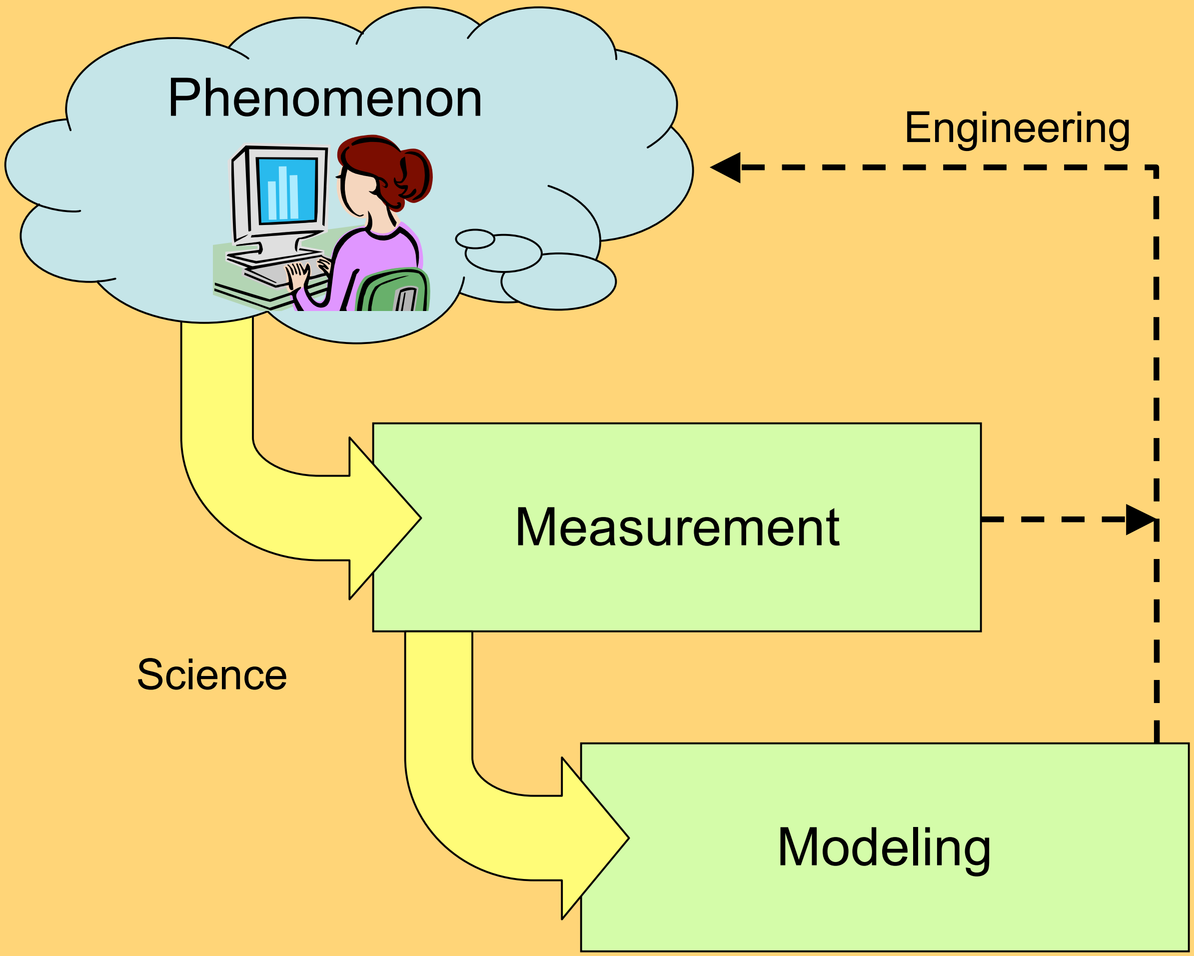
We must determine what we need to measure before we can effectively study visualization; this will lead to new means of measurement appropriate for visualization science.

Visualization Science requires Methods for Measurement

Chris North, Virginia Tech

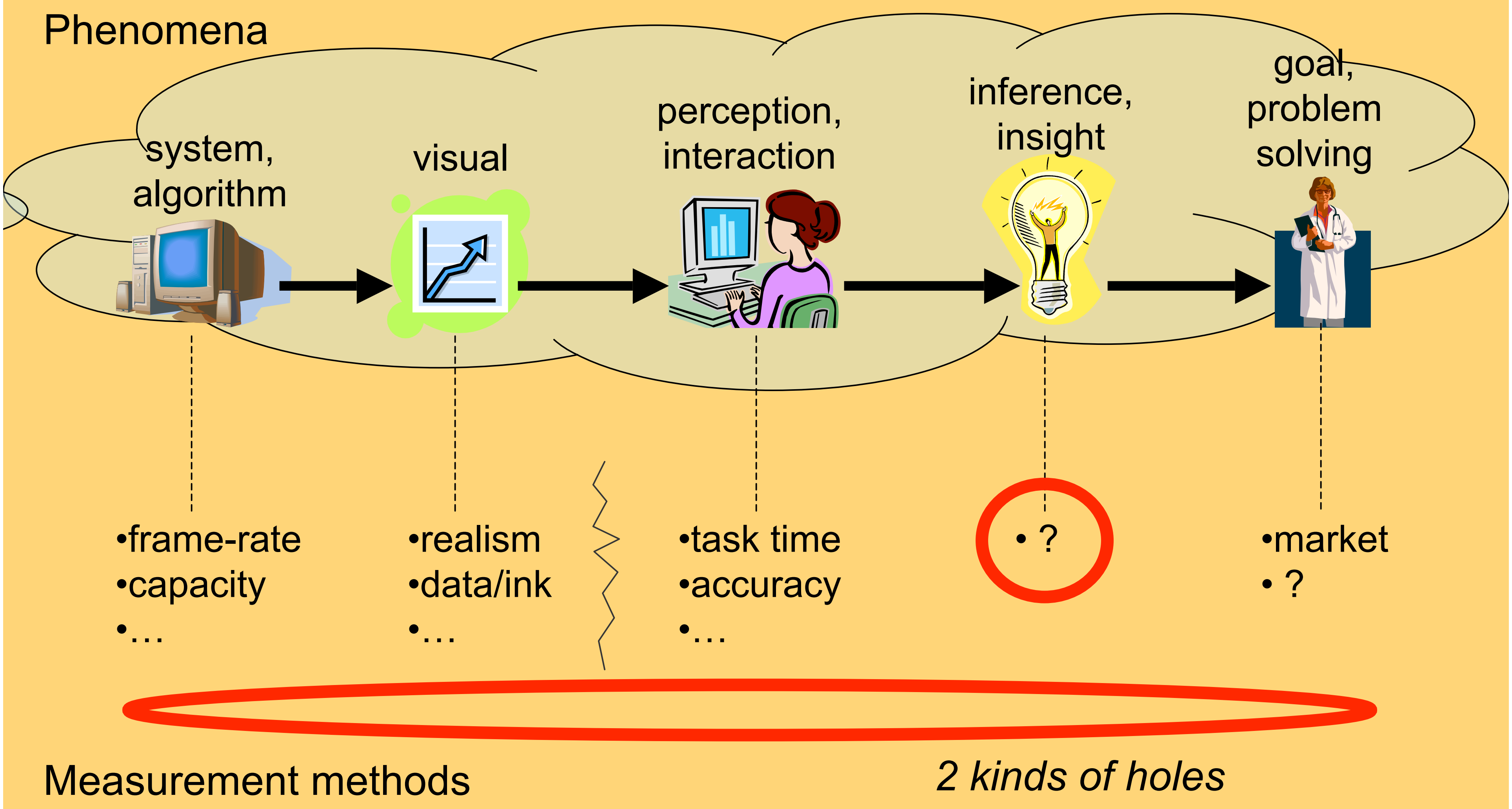
*<Insert disclaimer here about **gross generalizations**>*

Is there Science in Visualization?



…analogy to biology

Measurement in the Science of Vis.



WANTED!

New instruments and methods for measurement

Dead
Or
Alive

REWARD

$000.003

Example: Insight-based Evaluation

- Problem: current measurements focus on low-level task performance and accuracy
  - What about Insight?
- Idea: treat tasks as dependent variable
  - What do users learn from this Visualization?
  - Realistic scenario, open-ended, think aloud
  - Insight coding
  - Information-rich results
- Short-term vs long-term

More examples: BELIV'06

- "BEyond time and errors: novel evaLuation methods for Information Visualization"
- A workshop of the **Advanced Visual Interfaces** (**AVI 2006**) International Working Conference, May 2006
- Organized by Bertini, Plaisant, Santucci

(caution to HCI folks…)

Science of Visualization

Science of HCI

Conclusions

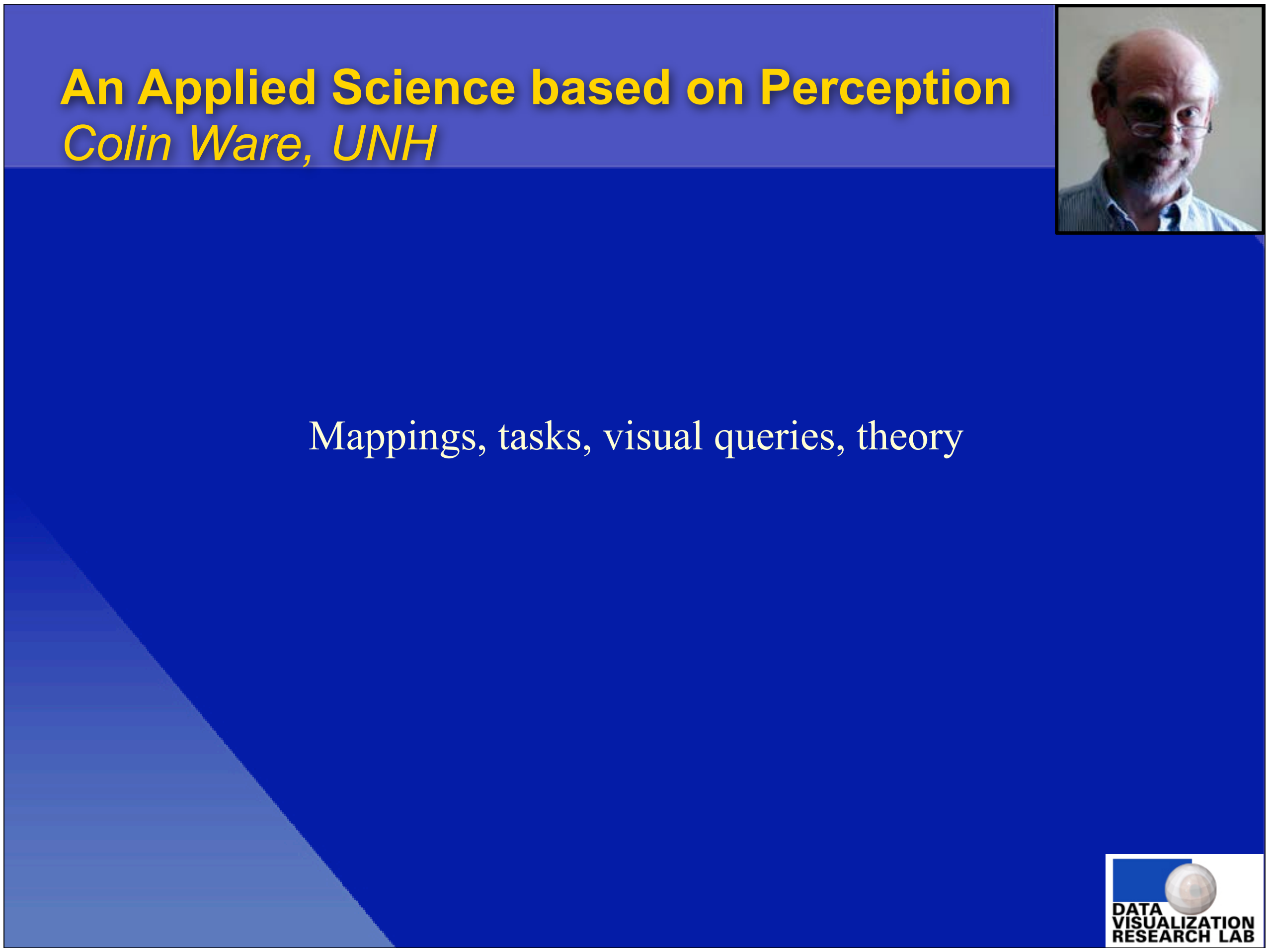
1. Is there science in visualization?

   Yes! (if we want it)

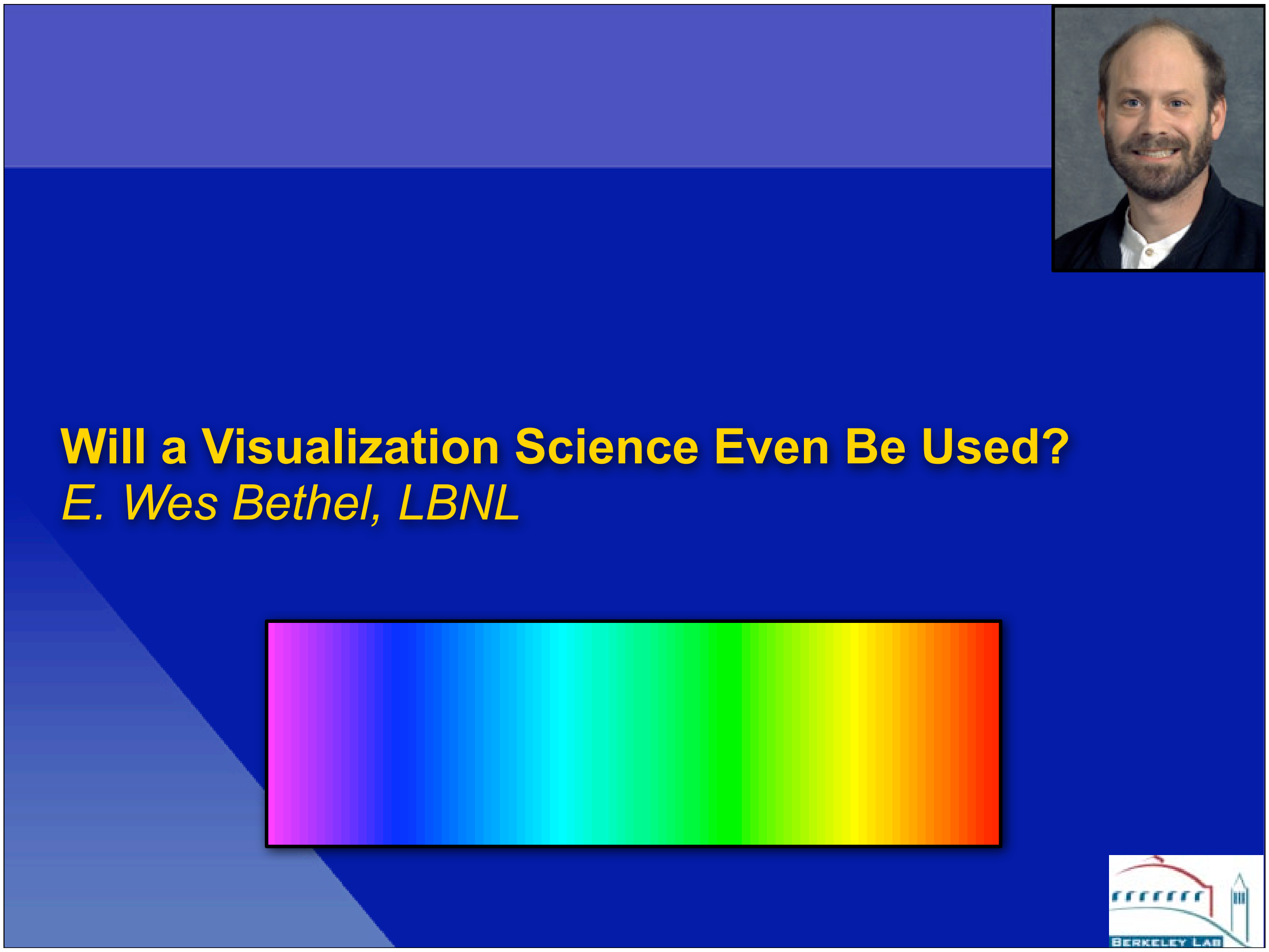
2. To get it, we must:
   a) Invent new measurement instruments and methods

      → start new Vis track on methods

   b) Decide who will do the measuring?
      (left as an exercise for the reader)



A rigorous applied discipline of vision science can build on existing perception theory



To be effective, a 'science of visualization' needs to put into practice, accepted as part of our culture, attuned to the needs of our customers, and gracefully accommodate Brooks' 80/20 rule.

Is there Science in Visualization?

VIS 2006
BALTIMORE·MARYLAND·USA

Wes Bethel – Lawrence Berkeley National Laboratory

Summary:

- *What good is science if we don't use it?*
- *Practice of science needs to be a part of our culture.*
- *What problems do Vis Science solve?*
- *Brooks 80/20 rule.*

Disclaimer

- Opinions are like noses – everyone has one.
- My biases:
  - I work with domain scientists all the time.
  - I am constantly in a position of having to defend and justify visualization to customers and funders.
  - One foot in research, another in development, another in production deployment.
  - Constant optimization of competing needs.
- Your biases are probably different.

Motivation

- Our stakeholders (and funders) tend to think vis is about "pretty pictures."
  - One of our field's strengths – stunning visuals – is also our Achilles Heel.
- "Pretty pictures" are good for many purposes, but tend to not be so good for "hard science."
  - A physics user: "I don't need help from visualization; I don't need that fancy 3D stuff, my 1D/2D histograms work just fine."
- Our CS cousins tend to have a much better developed scientific basis than visualization.
  - The term "visualization" coined in 1987 report.
  - Visualization is a nexus of hard and soft science.
  - It is often difficult to impose scientific rigor on elusive processes and ideas.

What Good is Science if We Don't Use It?

- We know certain things are bad, yet we do them without thinking about them very much.
  - Science of Applying visualization: Hue ramp colormaps.
    - A meme phenomenon?
  - Science of implementing visualization: Sequential scan/search.
    - In a hurry to whip out a quick prototype with intent to upgrade to something better at a "later time?"
  - "Bad scientific visualization example"
  - Arguably weak comparative analysis example.
- Some Useful "Science"
  - Cognitive: 3D motion parallax increases comprehension of depth relationships by 100%.
  - Computational complexity: Span space indexing increases isocontouring algorithms from O(N) to O(sqrt(N) + k).
- Applying science
  - Computational complexity in Petascale visualization.

Science Needs to Be Part of Our Culture

- Published work needs more scientific rigor.
  - Airtight experimental methodology.
  - Quantifiable results.
  - My new vis algorithm makes a nice picture, but …
- In one session at this year's meeting, there are six papers.
  - Three have no results: no experimental methodology, no quantifiable results, etc.
  - Two present results but don't have an experimental methodology.
  - Only one has an experimental methodology and quantifiable results.
- Proposals should describe projects that have well-grounded scientific methods.
  - Alternative is "research sandbox" with no clearly defined metrics.
- Scientific methods need to be "clearly defined."
  - E.g., "user study."

What Problems does Vis Science Solve?

- Impact on stakeholders.
  - Without a clear statement/idea of stakeholder impact, we have a hard time “selling” vis to funders and collaborators.
  - Examples:
    - Find X more quickly than before.
    - New ability to understand relationships in complex data.
- Traction on some elusive problems:
  - Cognitive: Is one vector field vis technique better than another? By what measure?
  - Impact: clear measure of contribution.
- Increased credibility; increased understanding of value of vis.

Brooks 80/20 Rule

- Pursuit of perfection is a noble objective, but getting that last 20% may not be practical in a reasonable amount of time.
- Don’t short-change the scientific methodology and rigor.
  - It is always good to have “Future Work.”
- How to get there?
  - Clarity of purpose.
- The “flip side”
  - Need for rapid discovery.
  - Build new techniques using scientifically solid components and ideas.

Specific Suggestions

- Teach scientific method, instill it as part of our culture.
  - No experiment & methodology = no degree.
- Enforce its practice as part of the publication review process.
  - No experimental methodology, no quantifiable results = no pub.
- Write proposals that have a solid basis for scientific methodology.
  - No hypothesis, no experimental methodology = no funding.
  - Will make it easier to review proposals and obtain funding.

On a Positive Note

- Think about the history of science.
  - You stand on the shoulders of giants: Pythagorus, Newton, Curie, Seaborg, etc.
  - Who is your role model?
  - The desire for knowledge and unanswered questions should keep you awake at night.

Summary

VIS 2006

-  We need foundational models
-  Visualization science needs a different method
-  Reproducibility needs to drive visualization science
-  Visualization science requires new measurements
-  Visualization is the science of visual thinking
-  No visualization science without practical grounding

Is there Science in Visualization?

Questions? Comments? Reactions?