

Visualizing Modeling Heuristics

An Exploratory Study

Laurie B. Waisel¹, William A. Wallace², &
Thomas R. Willemain²

¹Carnegie Mellon University

²Rensselaer Polytechnic Institute

ICIS '99

Charlotte, North Carolina

December 12, 1999

Visualization of Modeling Heuristics

■ Characteristics of the heuristics

- generally applicable
- supported by the literature

■ Heuristics to be visualized:

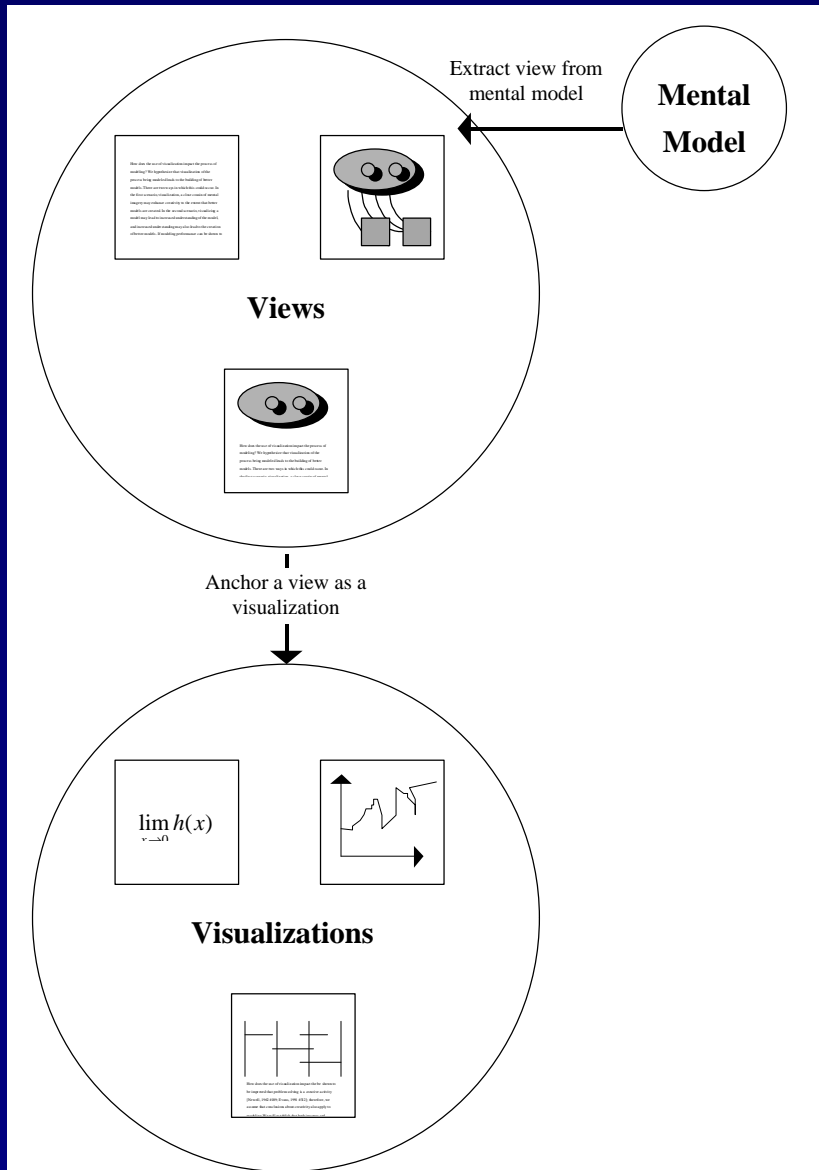
- 1) Output interesting
- 2) Input interesting
- 3) Far from known
- 4) Near high error

Definition of terms

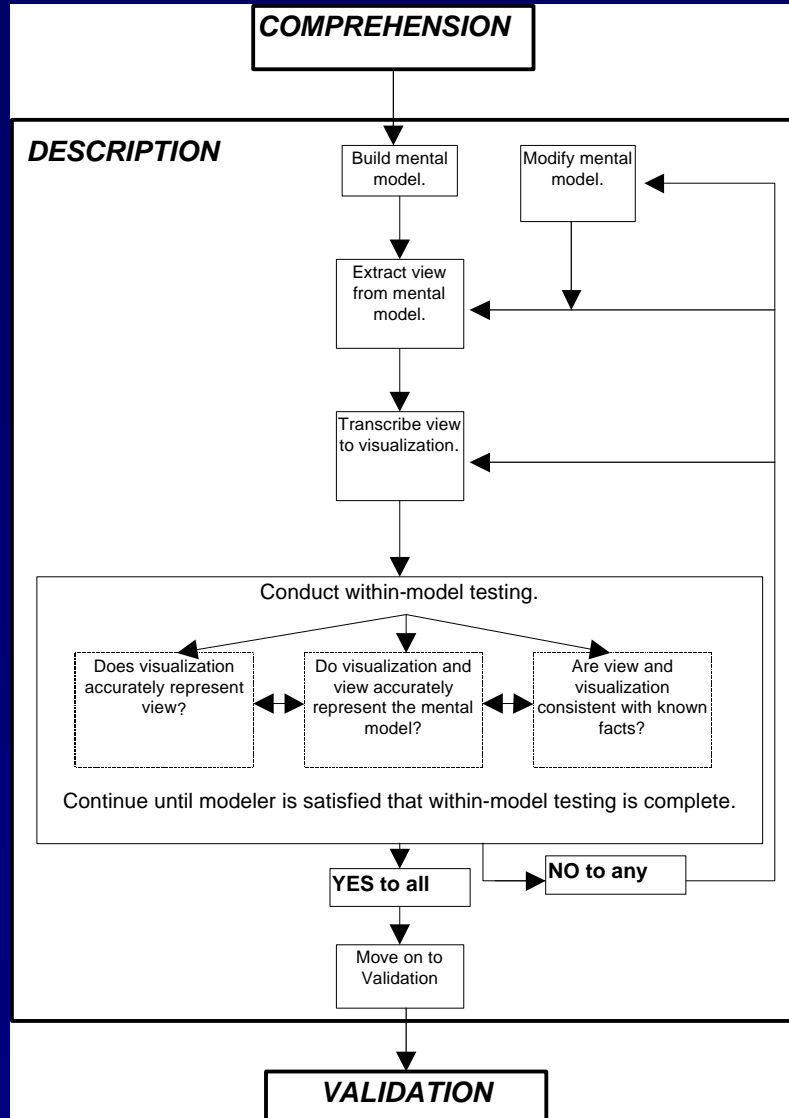
- A **model** is “any system of relations used to represent another system of relations.”
- A **visualization** is an image in the external world used to enhance insight.
- A **mental model** is a mental system of relations that has a structure similar to some other system of relations.
- A **view** is a mental image of or visual perception of a visualization.

Mental Models and Deduction

- Comprehension
 - Gathering information
- Description: within-model testing
 - Creative process of manipulating data and ideas in the effort to find a solution.
- Validation: between-model testing

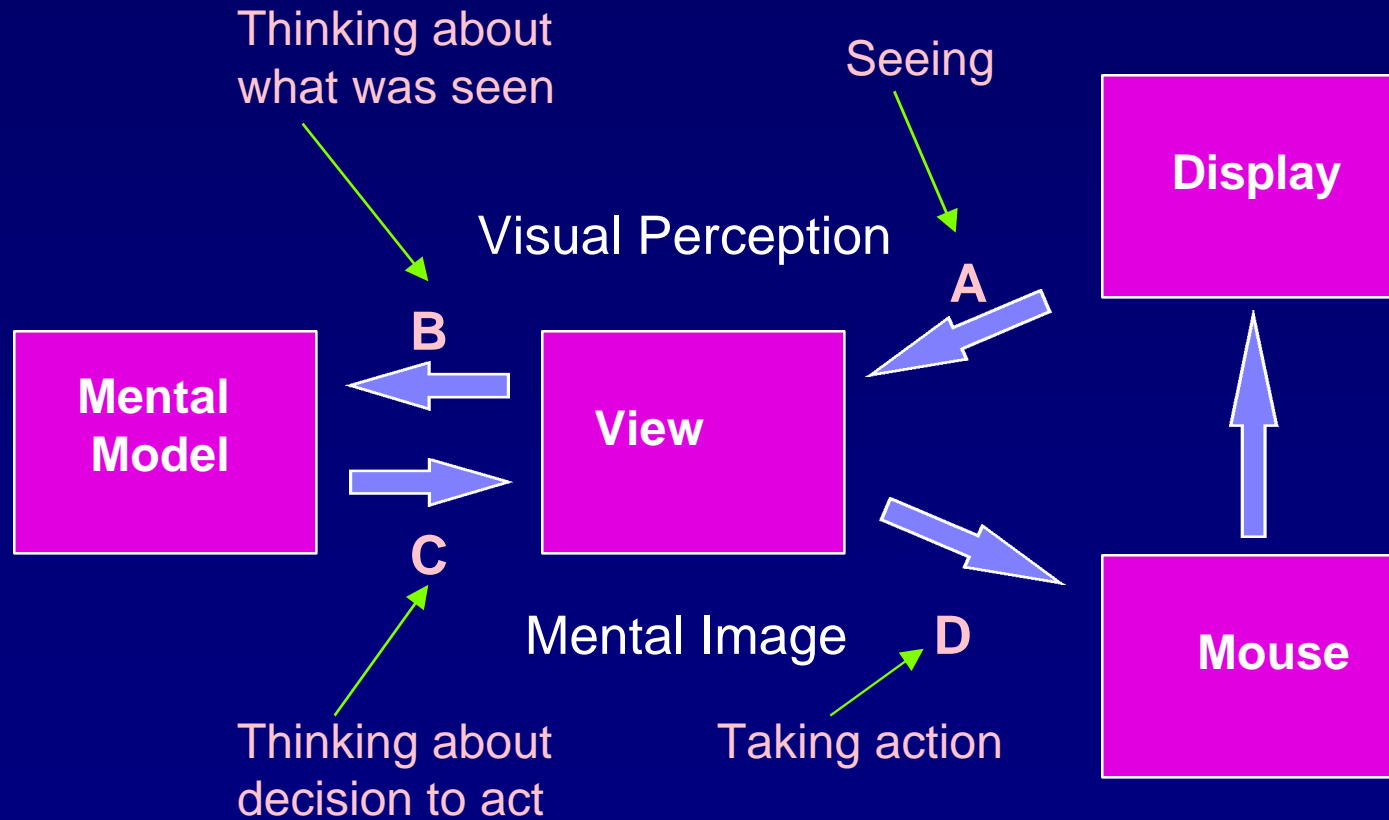


A Theoretical Framework for Modeling with Visualization



The Framework As Description Phase of Deduction

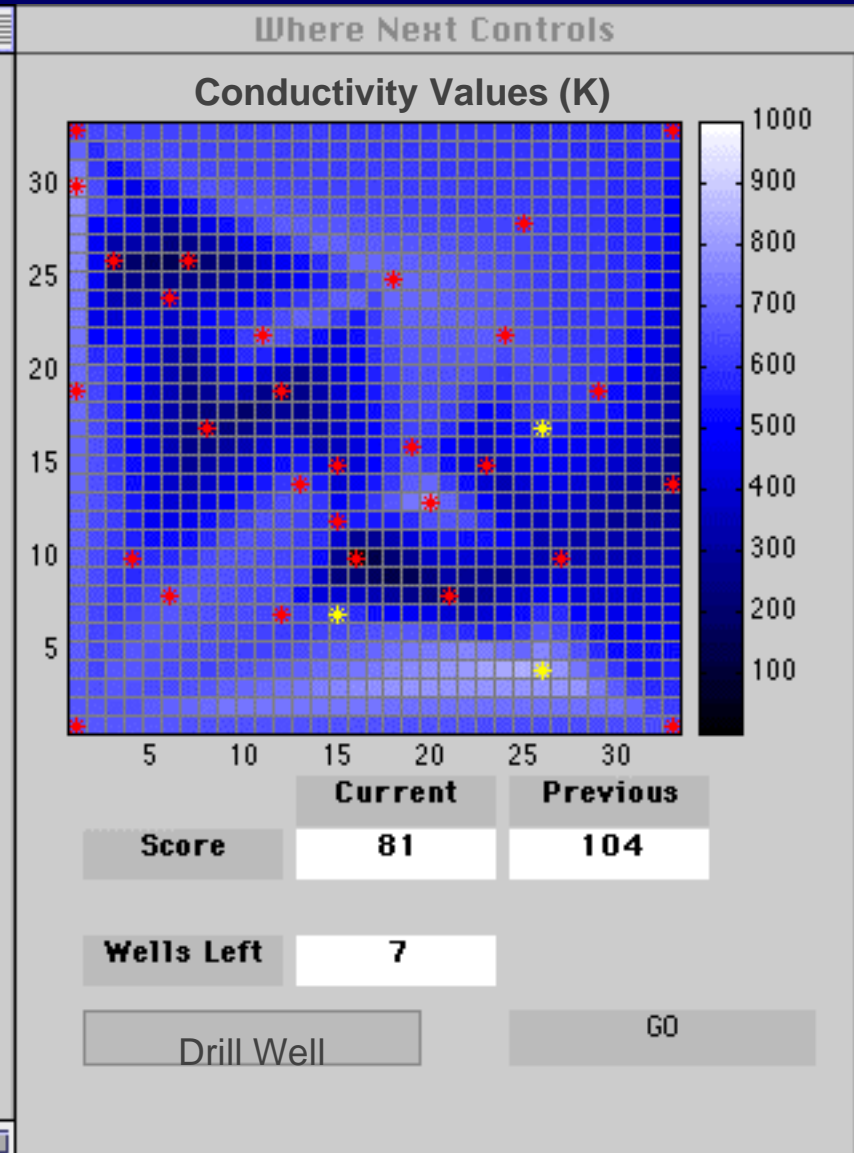
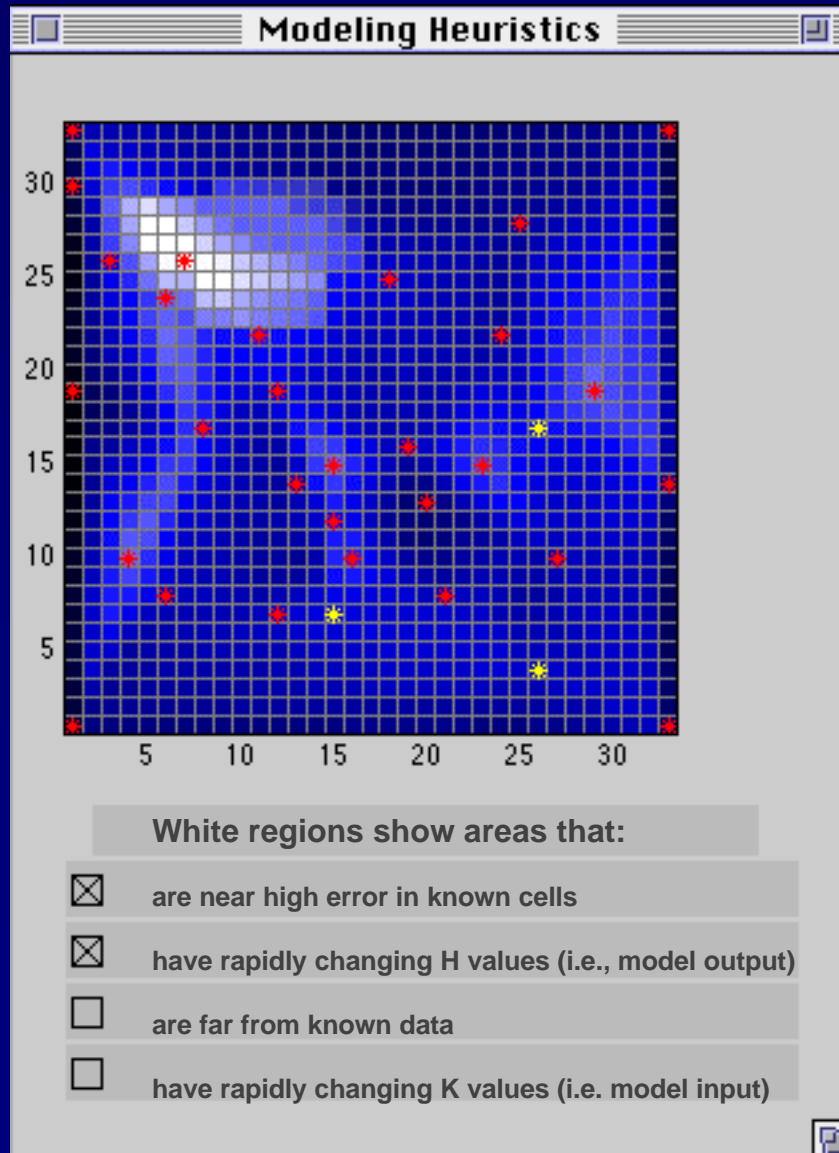
Modeling With Computer Visualization



The Experiment: Visualization of Heuristics for Groundwater Modeling

- The task: where to drill the next well
- The model:
 - Darcy's Law: $q = -K(dH/dx)$
 - K = hydraulic conductivity
 - H = hydraulic head or energy (i.e., water level)
- What the system does with the model:
 - Accepts a few sample K and H measurements
 - Interpolates to a full K data set
 - Uses model to solve for H
 - Calculates accuracy by comparing predicted H with sampled H

Modeler's Assistant



Implementation of Heuristics Visualizations

■ c_{mn} = colormap value
of the cell in row m ,
column n

■ FAR

- c_{mn} = distance to nearest
neighbor with known
value

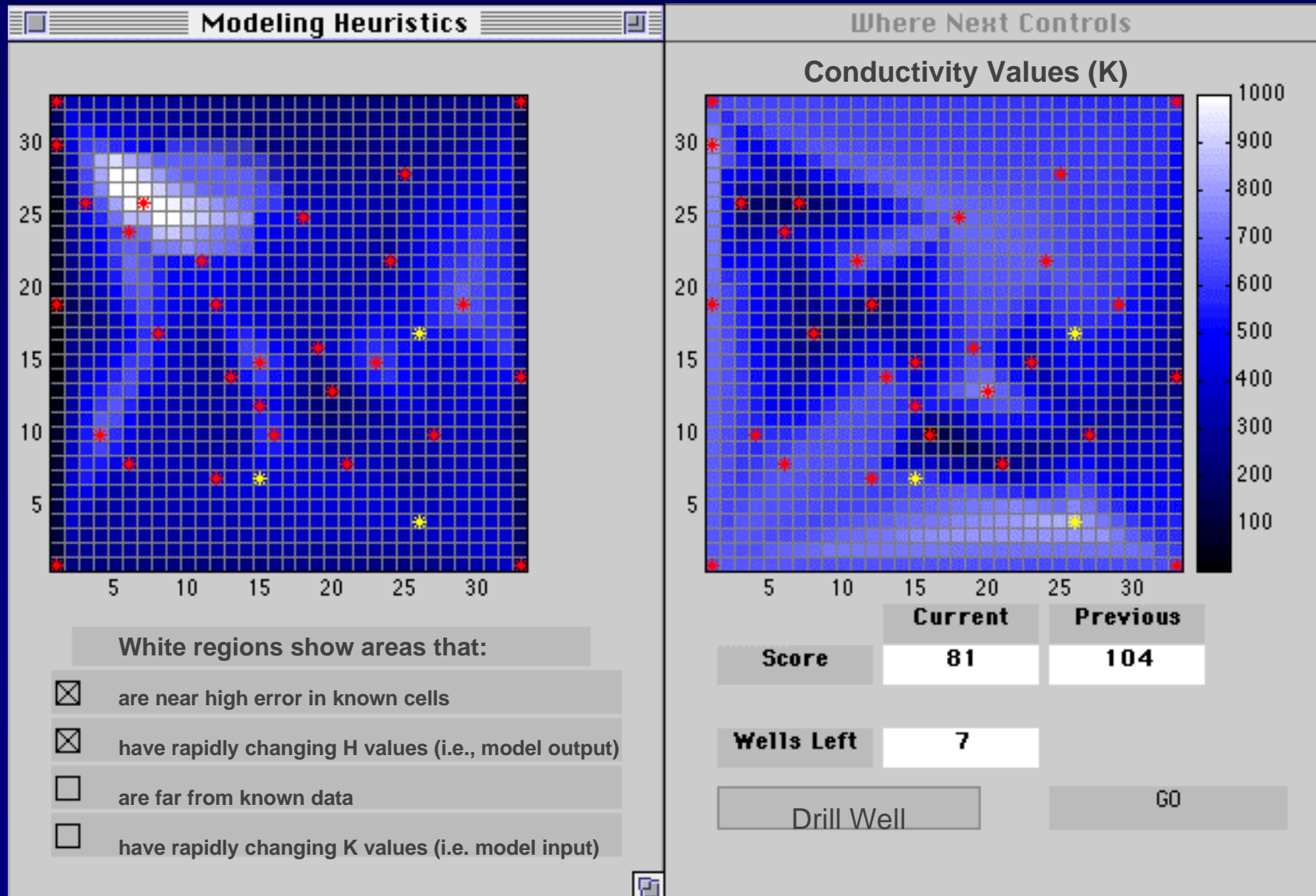
■ ERROR

- c_{mn} = interpolation of
error at known locations

■ INPUT and OUTPUT

$$c_{mn} = \sum_{s=i-1}^{i+1} \sum_{t=j-1}^{j+1} \left| \hat{x}_{ij} - \hat{x}_{st} \right|$$
$$1 \leq s \leq 33, 1 \leq t \leq 33$$

Modeler's Assistant



Experimental Design

■ Treatment factors

- Visualization level
- Skill level

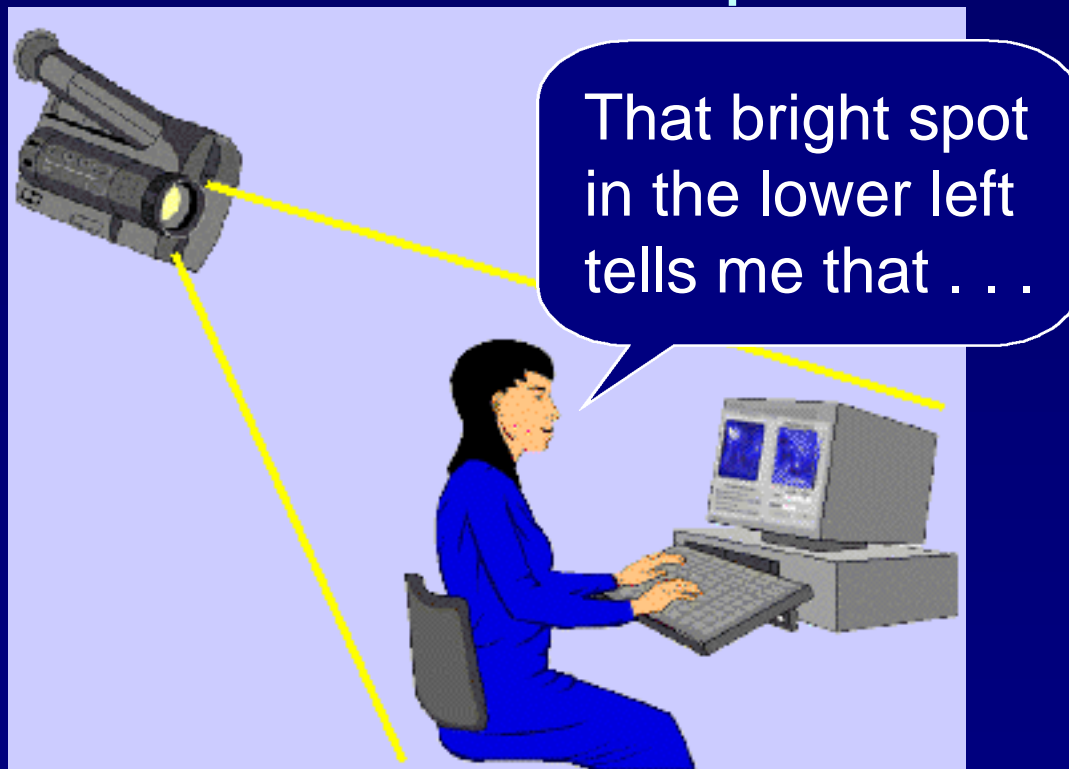
■ Blocking factors

- Modeler
- Order of heuristics text

Rule Order →	1	2	3	4
A (Novice)	L	M		
B (Novice)	H	L		
C (Novice)	M	H		
D (Novice)			L	M
E (Novice)			H	L
F (Novice)			M	H
G (Expert)	L	M		
H (Expert)	H	L		
I (Expert)	M	H		
J (Expert)			L	M
K (Expert)			H	L
L (Expert)			M	H

Experiment Execution

Think-aloud protocol



Protocol transcribed

LW: Tell me what you are thinking as you are working on the problem.
A: I see a dark patch over here that tells me that there's a problem in that corner. I think I'll drill a well there and see if I can make the score go down.
LW: What are you looking at now?
A: I'm looking at the conductivities grid. I want to pick up the outlines of that dark area.
LW: Tell me what you are thinking as you are working on the problem.
A: I see a dark patch over here that tells me that there's a problem in that corner. I think I'll drill a well there and see if I can make the score go down.
LW: What are you looking at now?
A: I'm looking at the conductivities grid. I want to pick up the outlines of that dark area.

Coded transcript

"I see a dark patch . . ." A
". . . which tells me there's a problem on the left . . ." B
"So I think I'll drill there." C
[mouse click] D

User Tracking System

- At every mouse click, the system records:
 - Which action was taken
 - Heuristic selected or unselected
 - Well location selected
 - Well drilled
 - Current system state
 - Current time and elapsed time
 - Which heuristics are currently visible
 - Current, average, and weighted average scores
 - Number of wells left to drill

Results

- Coding of protocols
 - Inter-coder reliability
- Assessing the model
 - Analysis of patterns of data flows
- Qualitative analysis of protocols
 - Transcripts and questionnaires

Coding of Protocols and Inter-coder Reliability

Verbal Protocol Text	CODE
So, see, we could use some wells in any of	
These light areas.	
So I'm going to put one where	
the permeability is high	
So let me get somewhere in	
here.	
[SELECTED WELL SITE FOR DRILLING]	D

Table
Method

Raw
Text
Method

All right.

[DE-SELECTED 'FAR FROM KNOWN' HEURISTIC]

My strategy would be to make sure I have enough data around areas that are variable so we can characterize that well.

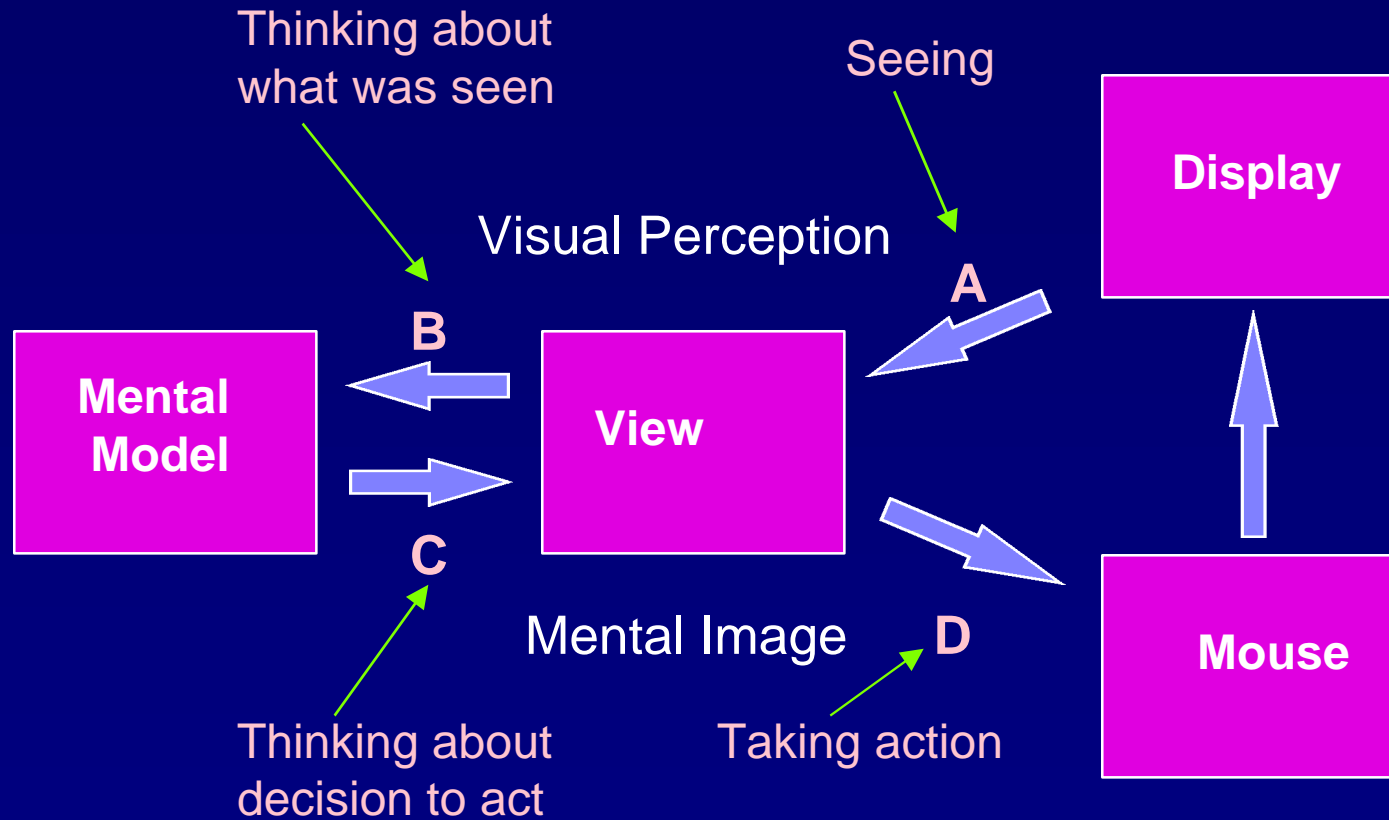
[SELECTED 'RAPIDLY CHANGING K' HEURISTIC]

[SELECTED 'RAPIDLY CHANGING HEAD' HEURISTIC]

That's what I'm going to do is put another well right up here even though it may cause me a lot of problems, because they're close together

[SELECTED CELL FOR DRILLING]

Modeling With Computer Visualization



Assessing the Model

Segment	Counts	Proportions
A	27	.029
B	160	.172
C	158	.170
AB	25	.027
BC	56	.060
AC	11	.012
ABC	24	.026
Subtotal of partial and full "ABC" segments	461	0.496
Total of all segments	932	1

Qualitative Analysis: Transcripts

Modeler	Skill Level	Preference	Comments
A	E	N>P	"Can simultaneously see the error screen and head value screen and K screen — I make mental estimates of the gradients, so don't need to rely on the gradients being calculated."
B	E	P>N	"Slightly [prefer partial heuristics visualization], seemed to have more info. Would have been nice to have four windows available simultaneously for [partial heuristics visualization]."
C	E	F>N	"Love of statistics. 1) Not just single points, more spatial, and 2) ability to combine different [heuristics]."
D	E	F>N	"Seems more user friendly being able to layer heuristics, layer is easier than visual comparison."
E	E	F>P	"Overlays of 'error' with other info was useful in pinpointing areas that needed more in-depth analysis."
F	E	F>P	"Being able to view heuristics singularly or in combination aids in the decision making process."
G	N	P>N	"I can return to the choices to see what I need and what happened [to the heuristics visualizations] after I made my decision."
H	N	P>N	"In [partial heuristics visualization] the error heuristics screen was much easier to use and understand because it showed the error for points around the well."
I	N	N>F	"Error box was more helpful. I felt like I could understand what was going on better."
J	N	F>N	"I could combine the heuristics rules, which made it easier to see where I wanted to drill a well."
K	N	F>P	"It makes it easier to compare two or more displays at a time and combined the thoughts [rather than making me] memorize it."
L	N	F>P	"It allowed you to visualize two separate ideas superimposed upon each other, eliminating the time it takes to think that way."

Qualitative Analysis: Questionnaires

■ Simultaneity

- Number of different displays visible simultaneously

■ Layering

- Displaying multiple layers of data, i.e., heuristics

■ Complexity

- Raw data, first differences, interpolated data

Conclusions

- Different people use visualization in different ways when building a model.
- Experimental participants had strong preferences regarding simultaneity, layering, and complexity.
- The theoretical framework, which incorporates visualization into the mental models paradigm, allows for reproducible measurement of modeling behavior.

Contributions

- Rich Data Set
- Modeler's Assistant
 - Hydrology, Decision Support
- Methodology for Visualizing Modeling Heuristics
 - Modeling Science
- Theoretical Framework
 - Modeling Science, Cognitive Science

Ongoing Research

- Expand the Modeler's Assistant to modeling tasks in other application domains.
- Exploring the cognitive role of visualization in exploratory data analysis