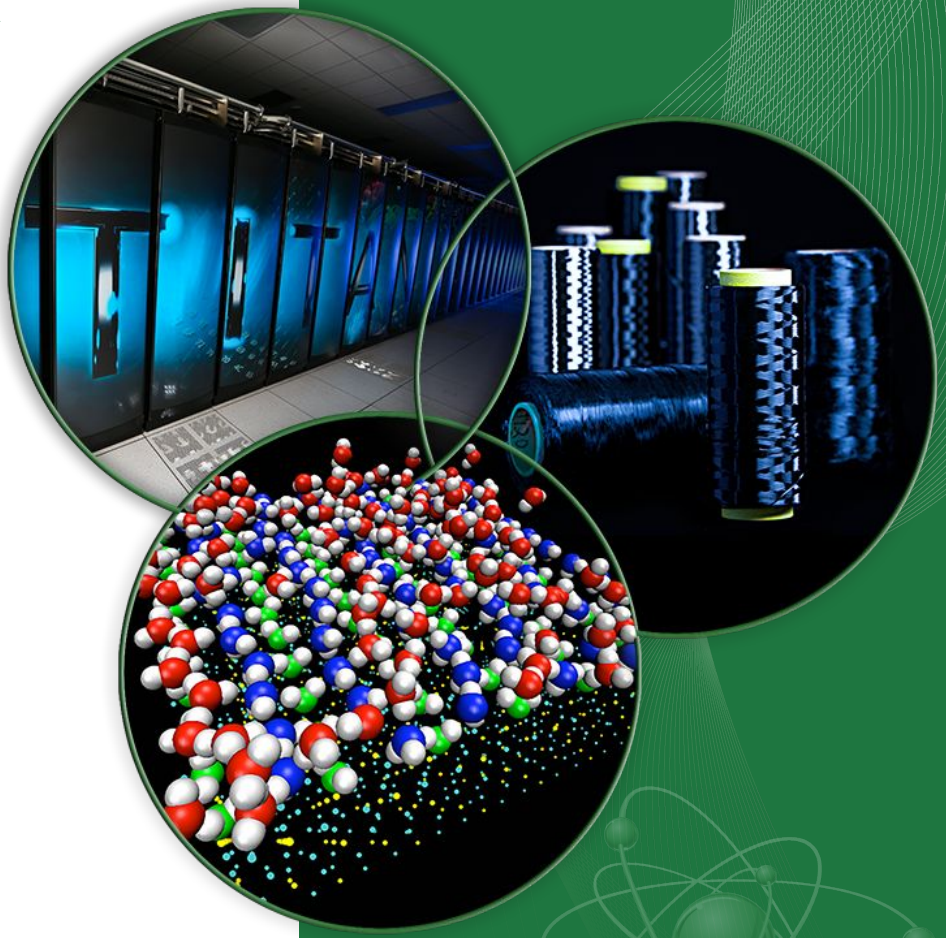
User Driven: Improving End-User Quality Through Usability Testing



Taylor Patterson
Research Associate
Oak Ridge National Laboratory
pattersontc@ornl.gov
@TCPatt

Jay Jay Billings *Research Staff, ICE Project Lead*
Oak Ridge National Laboratory
billingsjj@ornl.gov
@jayjaybillings
+jayjaybillings

EclipseCon Europe 2015
Project Quality Day
Ludwigsburg, Germany 2015-11-04

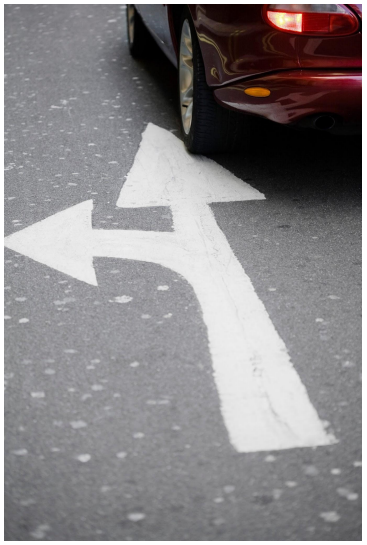
ORNL is managed by UT-Battelle
for the US Department of Energy

OAK RIDGE National Laboratory

Outline

- Motivation
- Product Tested
- Design and Planning
- Execution
- Outcomes

Outline



- Motivation
- Product Tested
- Design and Planning
  - Steps
  - Case Study
  - Tools
- Execution
- Outcomes

Motivation

- Project maturity
- Make better software
- Capture perspective of new users
- Take a unique approach
- Share our experience with others

Product Tested

Eclipse Integrated Computational Environment (ICE)

Design and Planning

Formulate a research question:

How does X affect Y (under condition(s) Z)?

Design and Planning

Our research questions:

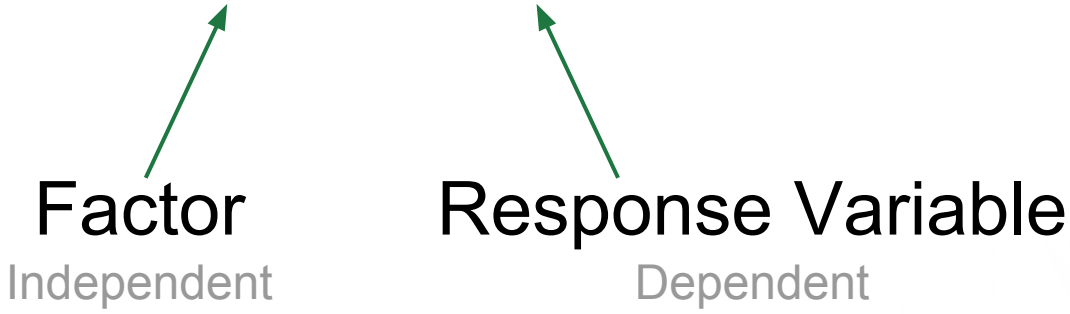
How does the use of Eclipse ICE affect user efficiency and success in completing a series of scientific simulation tasks?

How will users evaluate the usability of Eclipse ICE compared to alternative tools when completing a series of scientific simulation tasks?

Design and Planning

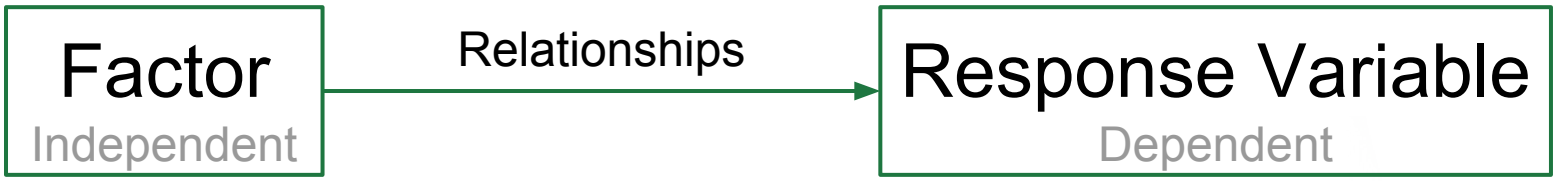
Identify variables:

How does X affect Y (under condition(s) Z)?



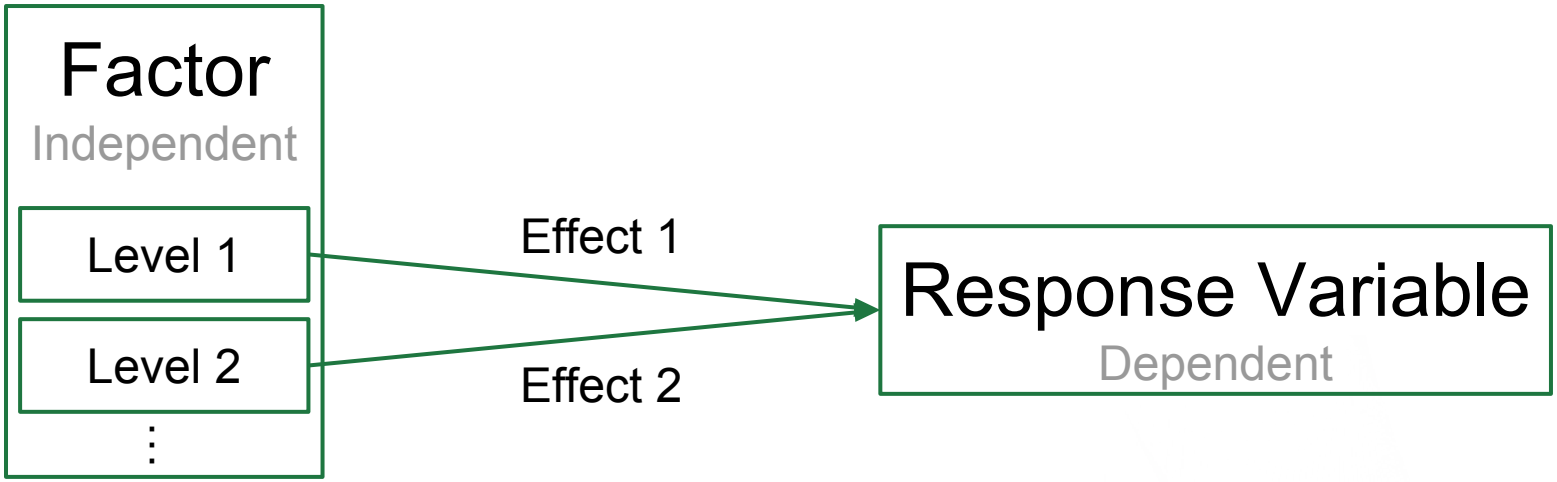
Design and Planning

Identify variables:



Design and Planning

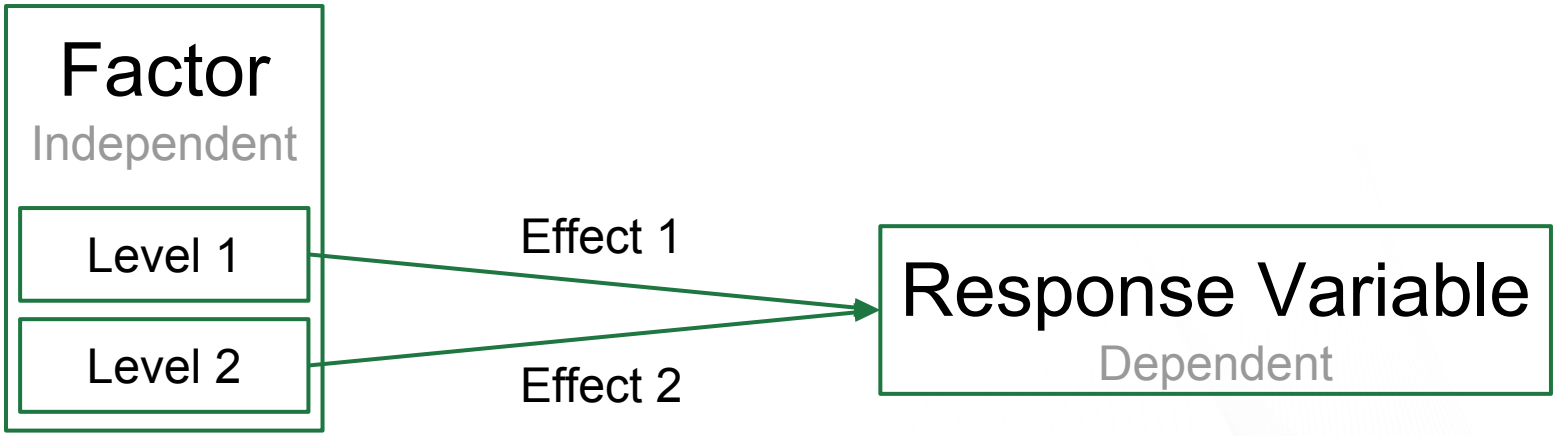
Identify variables:



Design and Planning

Our variables:

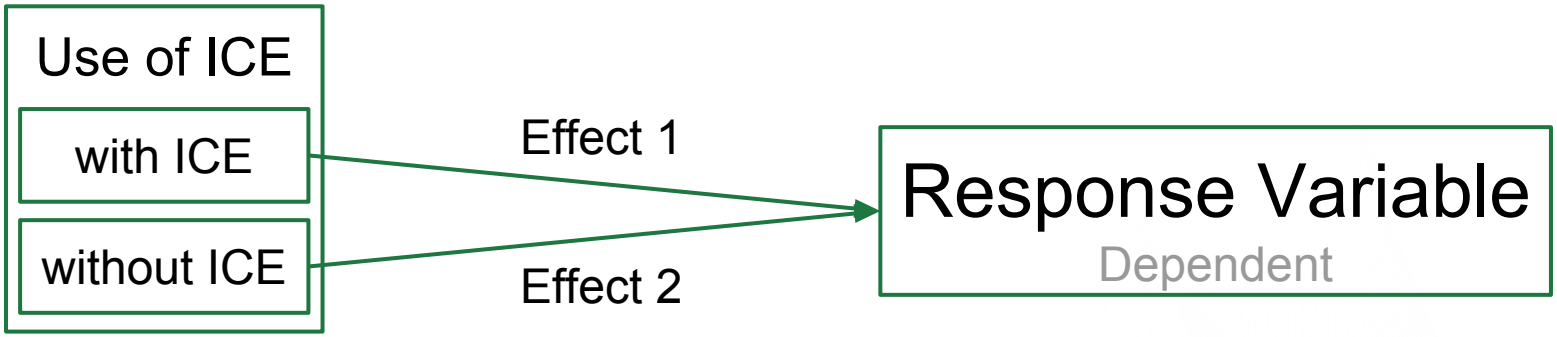
How does the use of Eclipse ICE affect user efficiency and success in completing a series of scientific simulation tasks?



Design and Planning

Our variables:

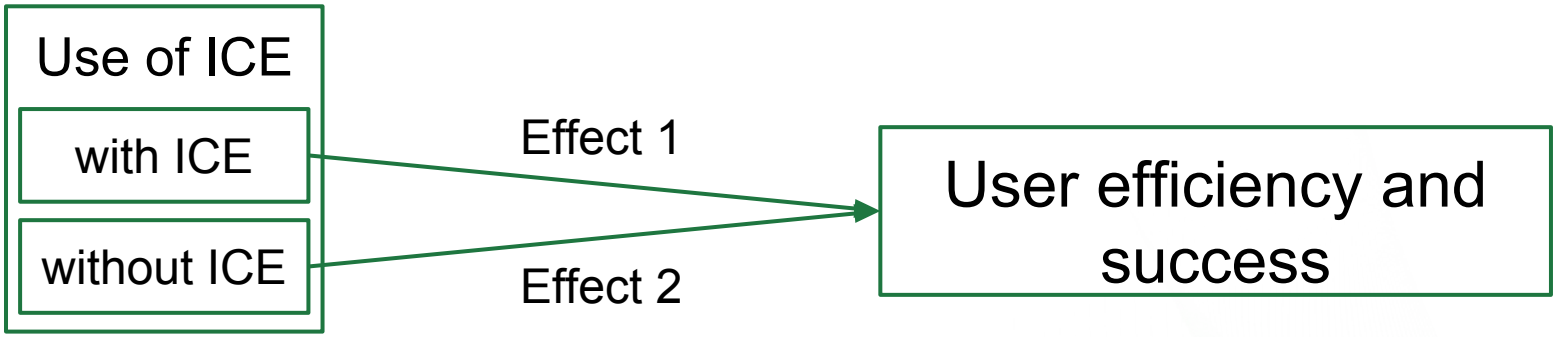
How does the use of Eclipse ICE affect user efficiency and success in completing a series of scientific simulation tasks?



Design and Planning

Our variables:

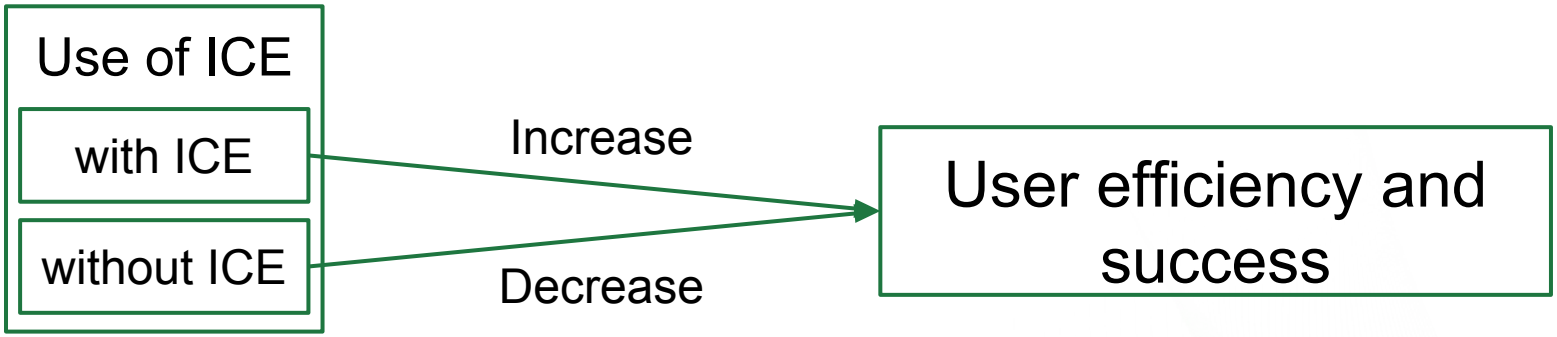
How does the use of Eclipse ICE affect user efficiency and success in completing a series of scientific simulation tasks?



Design and Planning

Our variables:

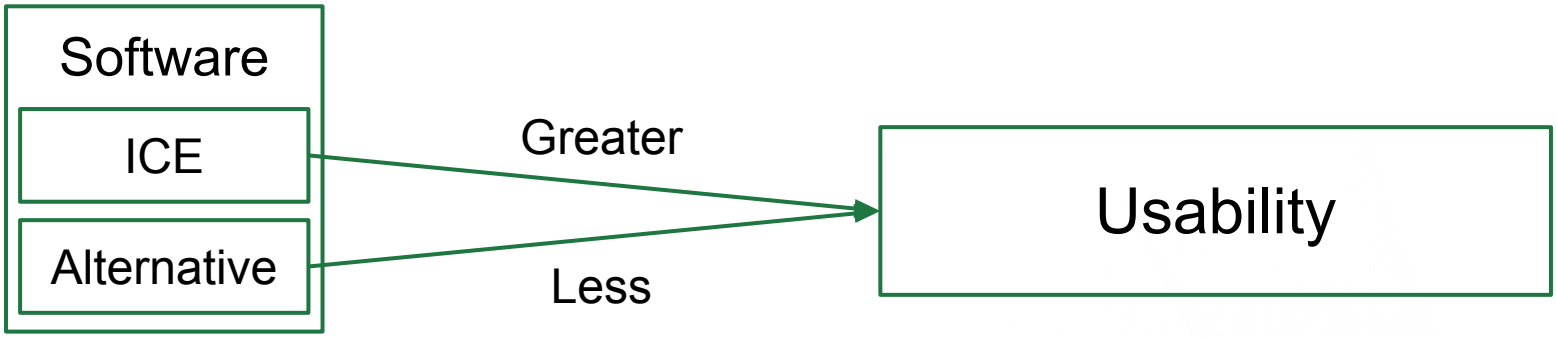
How does the use of Eclipse ICE affect user efficiency and success in completing a series of scientific simulation tasks?



Design and Planning

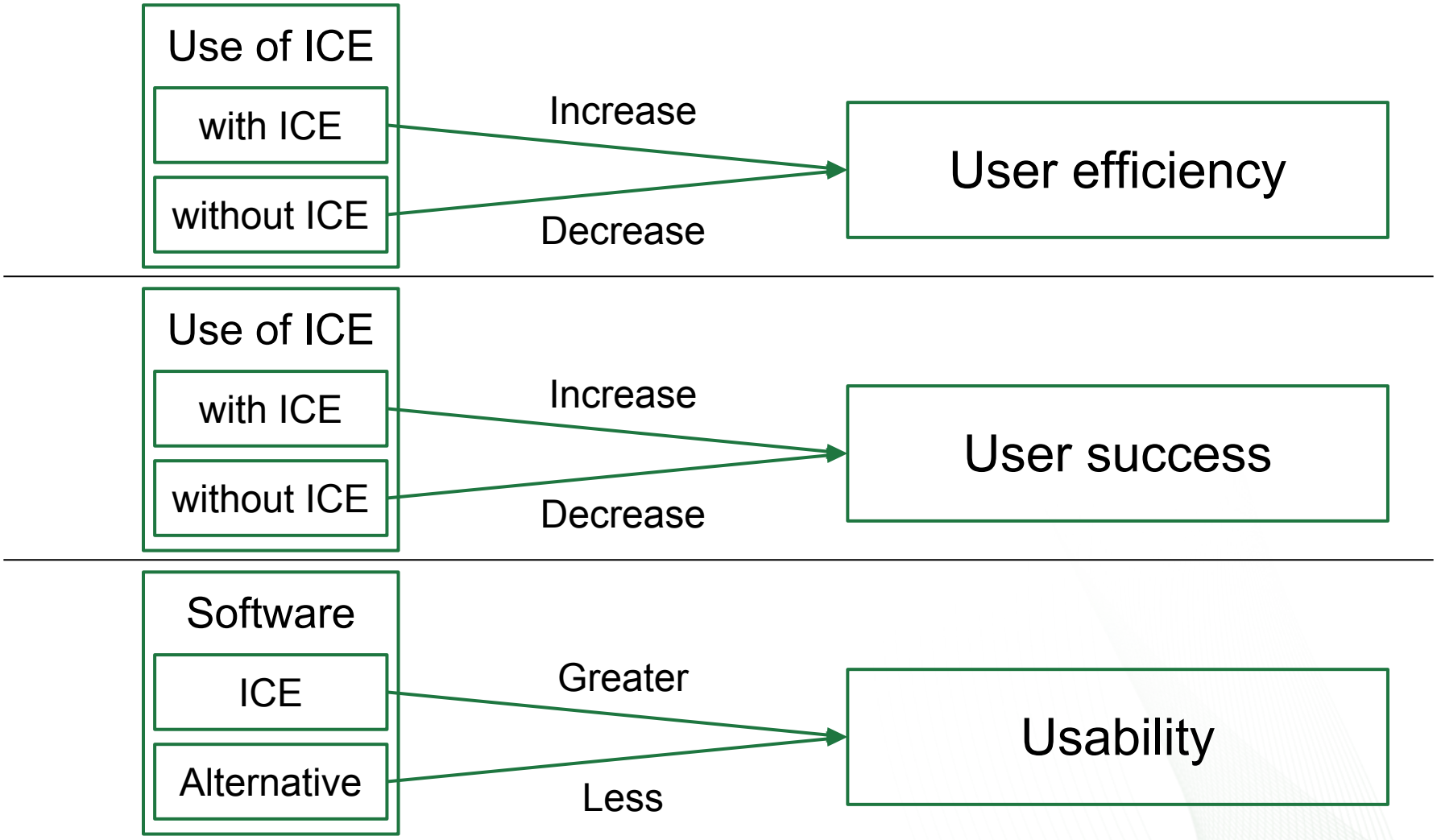
Our variables:

How will users evaluate the usability of Eclipse ICE compared to alternative tools when completing a series of scientific simulation tasks?



Design and Planning

Our variables:



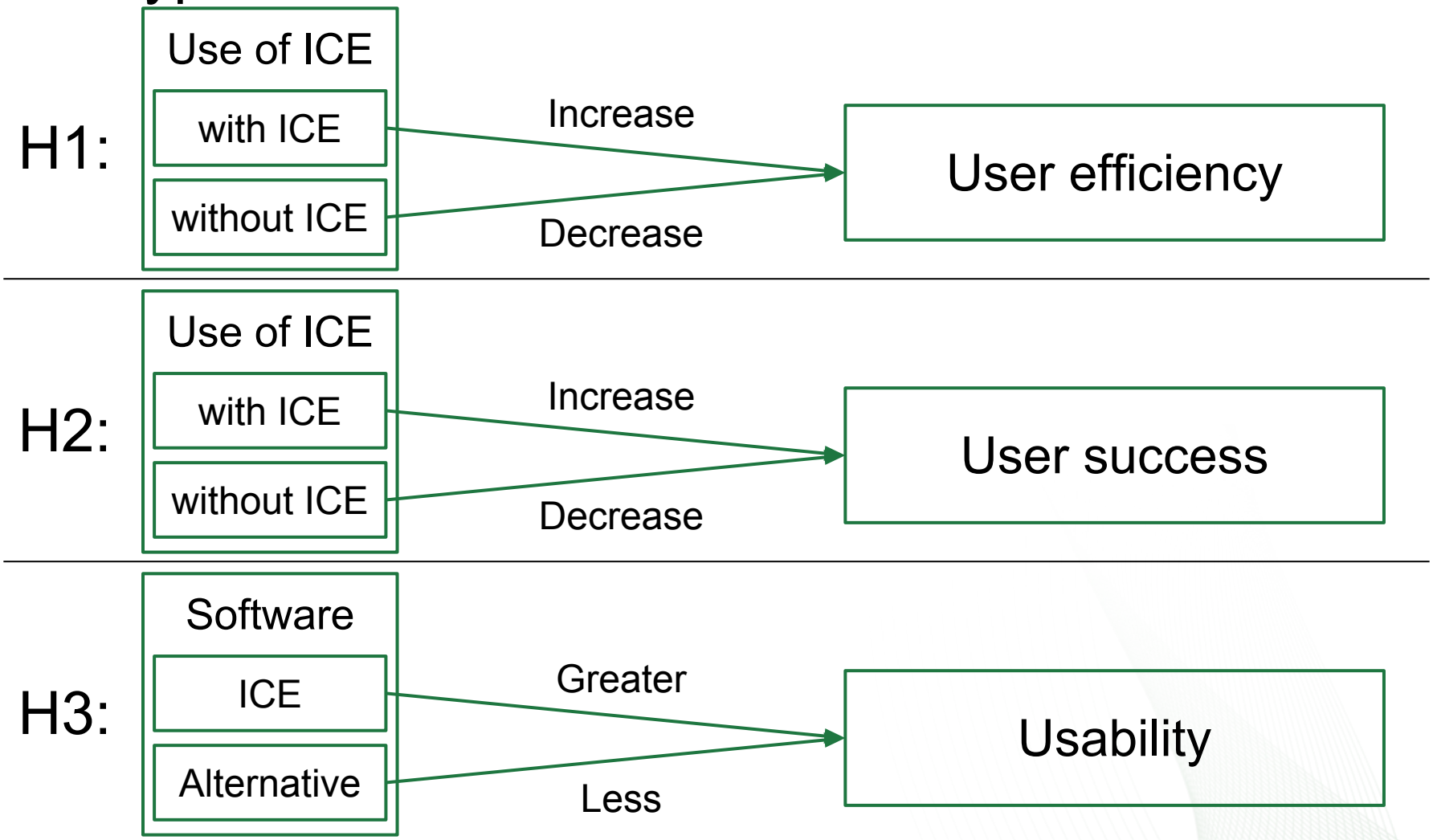
Design and Planning

Generate hypotheses:

X affects Y in W ways (under condition(s) Z).

Design and Planning

Our hypotheses:



Design and Planning

Our hypotheses:

H1: Users of Eclipse ICE will be more efficient in completing each milestone in a series of simulation tasks as compared to users of the collection of alternative tools.

H2: Users of Eclipse ICE will be more successful in completing each milestone in a series of simulation tasks as compared to users of the collection of alternative tools.

H3: Users will find Eclipse ICE more usable than alternative tools after completing a series of simulation tasks.

Design and Planning

Determine experimental design:

- Complexity
- Exposure

Design and Planning

Determine experimental design:

Complexity

Simple	Factorial
- Vary one factor at a time - No study of interactions - Possible misleading results	- Multiple factors varied - Requires more participants - Intricate statistical modeling

Design and Planning

Determine experimental design:

Exposure

Between-participants	Within-participants
- One combination of factors and levels - Eliminate transfer effects - Limit demand characteristics	- All factors and levels - Reduce variability across participants - Greater statistical power

Mixed-model
- Draw aspects of both designs - Maximize power, minimize bias

Design and Planning

Our experimental design:

Simple, between-participants design

2 groups: ICE and alternative tools

Design and Planning

Select target participants:

Representative of the population
to which results will be generalized

Rule of thumb: Minimum 8 participants per group

Design and Planning

Our target participants:

All users, exclude no one

Seek a total of 60 participants

Design, Planning, and Execution

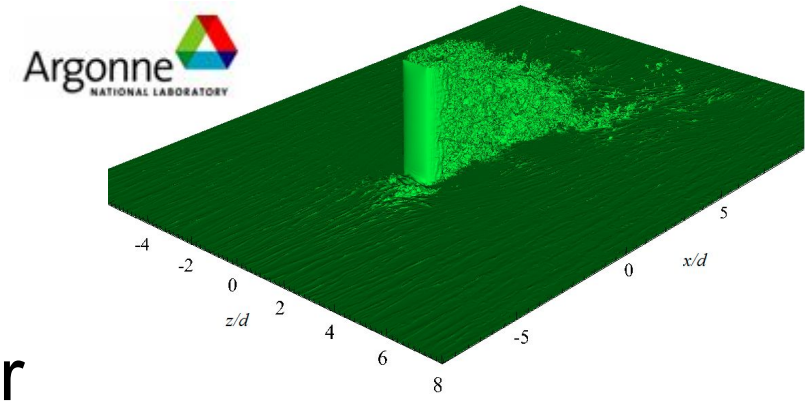
Define experimental task and procedure:

- High level of detail for publication
- Reasonable context to test hypotheses
- Accurately represent a real world scenario
- How motivation is built-in
- Ensure consistency in researcher actions

Design, Planning, and Execution

Our task:

- Nek5000
- Standard desktop computer
- Equivalent series of tasks
- No explicit motivation provided
- Run on Amazon Elastic Cloud



Credit: Nek5000 Project, http://nek5000.mcs.anl.gov/

Design, Planning, and Execution

Our procedure:

- Informed consent
- Script of verbal instructions
- Equivalent user manuals
- Morae Suite for desktop and participant recording

Design, Planning, and Execution

Establish measures:

How do we capture the ways in which different factors and levels affect response variables?

Objective, Subjective, and Behavioral measures

Design, Planning, and Execution

Our measures:

Factor: Use of ICE

Response variable:

Efficiency	→	Task completion time
Success	→	Accurate completion & expert intervention
Usability	→	Questionnaire responses System Usability Scale (SUS), NASA Task Load Index (NASA-TLX)

Outcomes

40 minutes of video per participant
- Identify difficult processes
- Modify UI and help text

Comparison to alternative tools
- Statistical analysis with R, StatET, Architect

Gather new user feedback

Validate known concerns and expose new bugs

Thank You

Questions?

Author Email: pattersontc@ornl.gov, billingsjj@ornl.gov

Acknowledgements: AJ Ierulli, April Lewis, Ryan McMahan, Eric Ragan, Teresa Robison, Rachel Volentine

Credit: Content inspired by Bilge Mutlu, University of Wisconsin-Madison


eclipsecon Europe
Ludwigsburg, Germany, 3 - 5 November 2015
Evaluate the sessions at www.eclipsecon.org
+1
0
-1