



<http://www.brainleadership.com/tag/brain/>

# Think Aloud User Testing

Jon Kolko  
Professor, Austin Center for Design

**ac4d**

# Think Aloud Testing

Evaluating the usability of your work by encouraging a user to think out loud as they use your product or service.

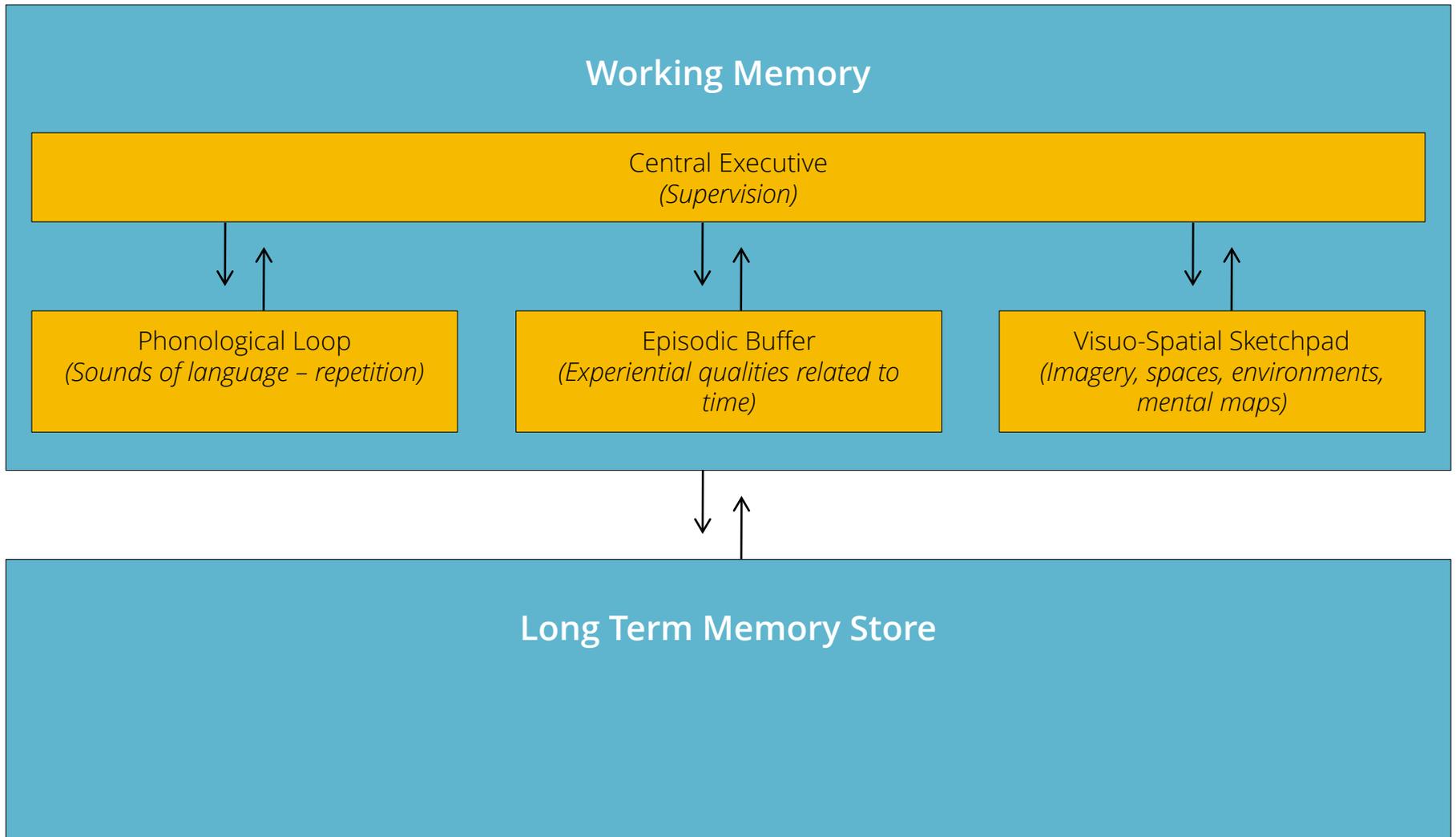
# Think Aloud Testing

Evaluating the usability of your work by encouraging a user to think out loud as they use your product or service.

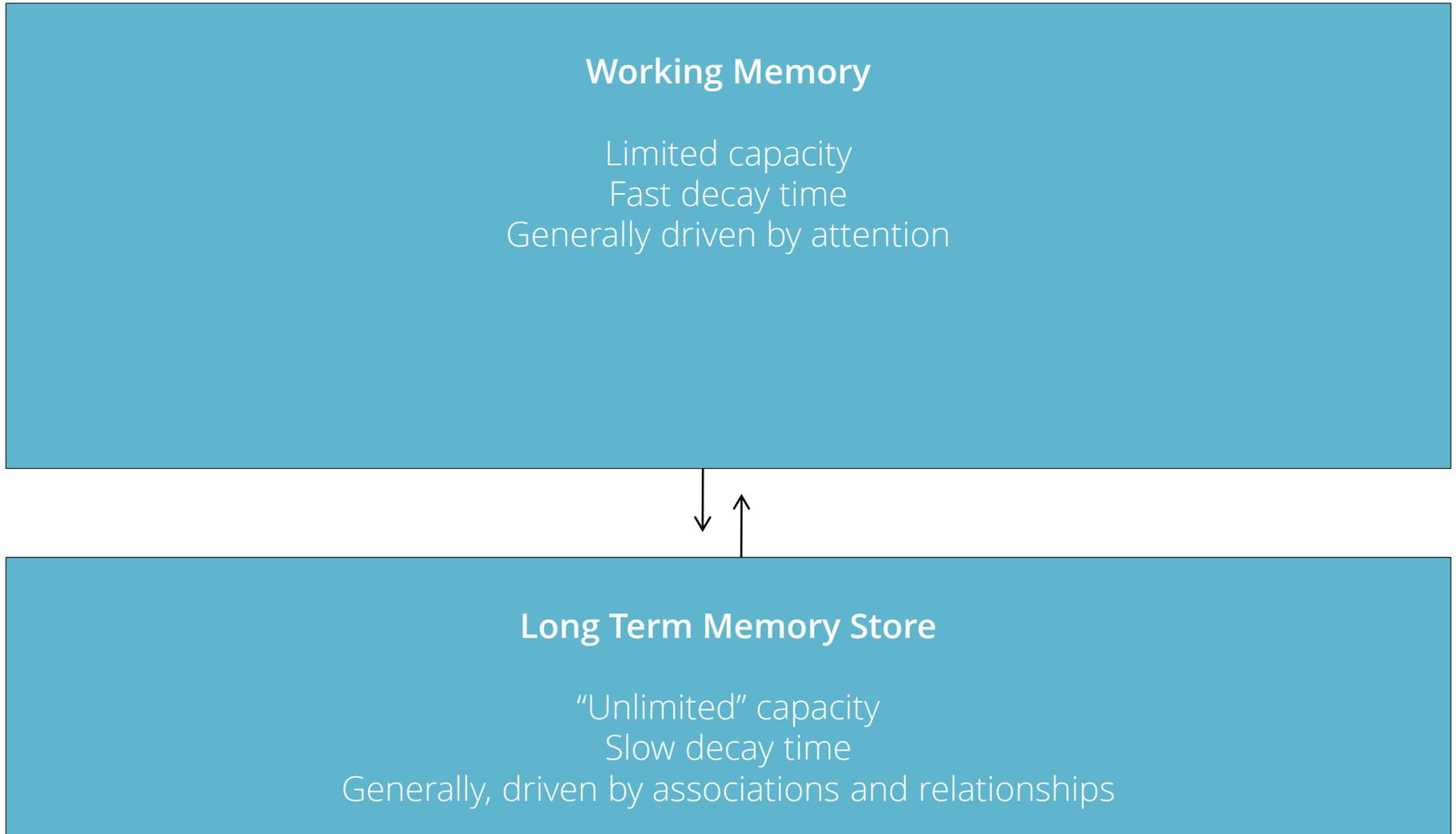
## THINK ALOUD TESTING/

1. Is evaluative, in that it responds to an existing artifact
2. Identifies defects that impact comprehension
3. Is considered a “quick and dirty” method, as compared to a formal experiment
4. Requires interacting with real people

# Background: Understanding Memory



# Background: Understanding Memory



# Background: Understanding Memory

## Working Memory

Limited capacity  
Fast decay time

Generally, driven by associations and relationships

Parameter	Mean	Range
Eye movement time	230 ms	70-700 ms
Decay half-life of visual image storage	200 ms	90-1000 ms
Visual Capacity	17 letters	7-17 letters
Decay half-life of auditory storage	1500 ms	90-3500 ms
Auditory Capacity	5 letters	4.4-6.2 letters
Perceptual processor cycle time	100 ms	50-200 ms
Cognitive processor cycle time	70 ms	25-170 ms
Motor processor cycle time	70 ms	30-100 ms
Effective working memory capacity	7 chunks	5-9 chunks
Pure working memory capacity	3 chunks	2.5-4.2 chunks
Decay half-life of working memory	7 sec	5-226 sec
Decay half-life of 1 chunk working memory	73 sec	73-226 sec
Decay half-life of 3 chunks working memory	7 sec	5-34 sec

## Long T

“Ur  
S

Generally, driven by associations and relationships

# Background: Allen Newell & Herb Simon

- Research focus on artificial intelligence, decision making, bounded rationality.
- 1972: wanted to understand how people solve problems; needed to understand the sequence of thoughts people experience as they work through a problem.
- Ultimate goal was to simulate human problem solving with Artificial Intelligence.



# Background: Why it Works

- Developed experiments where one set of participants “verbalized their thoughts as they went about a task” and the other set did not.
- Found that there is no affect on thought sequences, as long as there is no introspection:
- People can successfully verbalize what they are doing without changing the outcome of a task.

# Background: Why it Works

## Working Memory

Limited capacity  
Fast decay time  
Generally driven by attention

We can verbalize the contents of working memory.



## Long Term Memory Store

“Unlimited” capacity  
Slow decay time  
Generally, driven by associations and relationships

# Background: Why it Works

- When the information is not linguistic (ie, when it is visual), this will slow down the task, but will not alter the validity of the task.
- If there is introspection (“why am I doing this?” or “what am I thinking about?”), the outcome changes.
- This is formally called the Think Aloud Protocol (a Protocol is a standard procedure for regulating data transmission – in this case, the tasks a person is completing): a way of understanding what someone is doing, as they do it.
- Or, put another way, this is a way of understanding the contents of working memory.

“Please keep talking.”

# Planning a Think-Aloud User Testing Session

## 1. Develop a prototype.

You'll need *all of the screens that a user will see* completed, but they don't need to be working – you can do Think Aloud testing with paper.

# Planning a Think-Aloud User Testing Session

## 2. Develop tasks that represent typical user goals.

Our scenarios are a great place to start.

There must be a complete path through the interface to solve these goals (don't give the user impossible tasks, it's mean).

Print these tasks, one to a page.

# Planning a Think-Aloud User Testing Session

## 3. Schedule sessions with users that match the target audience

2-8 users, one per session

# Planning a Think-Aloud User Testing Session

## 4. Organize yourself

Video camera, batteries, audio camera, tapes, pens, etc.

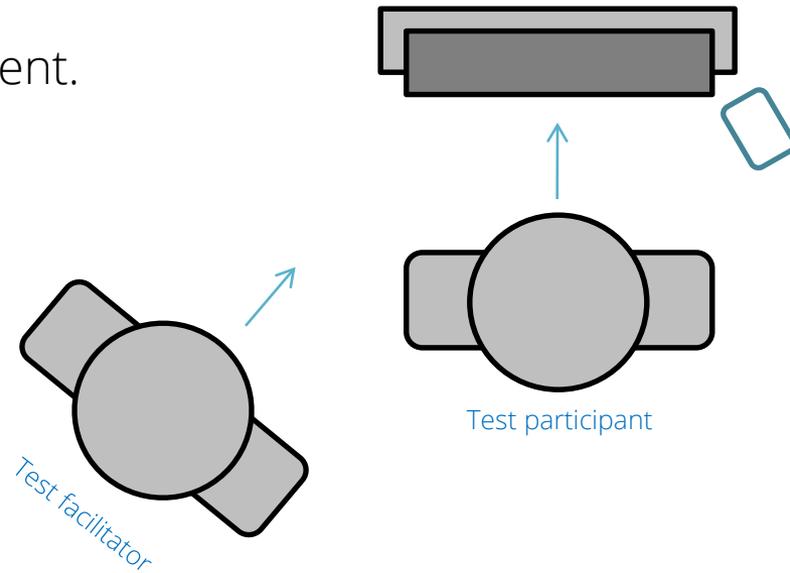
# Running a Think-Aloud User Testing Session

## 5. Get set up.

Prepare the physical arrangement.

Get written consent to tape.

Don't forget to start taping!



# Running a Think-Aloud User Testing Session

## 6. Explain the rules.

Explain to the user:

- who you are & what you are doing
- that you are testing your interface, and not testing them
- that they can quit at any time
- that you won't be able to help them
- that you require them to continue talking, and you will remind them to "please keep talking" if they fall silent
- To simply verbalize what it is they are doing, as they are doing it

Verify that the user understands the tasks (have them read the tasks aloud too, and ask if there are any questions)

# Running a Think-Aloud User Testing Session

## 7. Run the session.

Take good notes!

If the user falls silent for more than three seconds, prompt them “please keep talking”

Do not help the user complete a task (if the user asks for help, explain that you cannot help, and prompt them to try what they think is correct)

Don't defend your designs! This is not a critique of your design skills; don't even mention that they are your designs.

Be thoughtful – these are real people!

# Running a Think-Aloud User Testing Session

## 7. Run the session, continued.

While the session is running, do not say things like:

Please explain what you are doing

Note any design problems you see

Tell us if you have any suggestions

Why are you doing what you are doing

By asking questions like this, we call into play Mediated Processes, such as cognitive processes, which can disrupt the information state and alter the data. Specifically – we lose track of what was in working memory, and replace it with something else.

# After a Think-Aloud User Testing Session

## 8. Identify Critical Incidents

Determine the critical incidents that occurred:

*To be critical, an incident must occur in a situation where the purpose or intent of the act seems fairly clear to the observer... [they are] extreme behavior, either outstandingly effective or ineffective with respect to attaining the general aims of the activity.*

Flanagan, (1954), Psychological Bulletin, 51 (4), 327-358.

# After a Think-Aloud User Testing Session

## 8. Identify Critical Incidents, continued

Some (not all) criteria for identifying a Bad Critical Incident:

The user articulates a goal and cannot succeed in attaining that goal within two minutes

The user articulates a goal, tries several things and explicitly gives up

The user articulated a goal and has to try three or more things before finding a solution

The user does not succeed in the task

The user expresses surprise

The user expresses some negative sentiment, either about the interface or about their own skills

The user makes a design decision

# After a Think-Aloud User Testing Session

## 9. Calculate a SUS score per user

Strongly Disagree      Strongly Agree

	1	2	3	4	5
I think that I would like to use this system frequently.					
I found the system unnecessarily complex.					
I thought the system was easy to use.					
I think that I would need the support of a technical person to be able to use this system.					
I found the various functions in this system were well integrated.					
I thought there was too much inconsistency in this system.					
I would imagine that most people would learn to use this system very quickly.					
I found the system very cumbersome to use.					
I felt very confident using the system.					
I needed to learn a lot of things before I could get going with this system.					

# After a Think-Aloud User Testing Session

## 9. Calculate a SUS score per user

	Strongly Disagree					Strongly Agree	
	1	2	3	4	5		
I think that I would like to use this system frequently.					good		
I found the system unnecessarily complex.	good						
I thought the system was easy to use.					good		
I think that I would need the support of a technical person to be able to use this system.	good						
I found the various functions in this system were well integrated.					good		
I thought there was too much inconsistency in this system.	good						
I would imagine that most people would learn to use this system very quickly.					good		
I found the system very cumbersome to use.	good						
I felt very confident using the system.					good		
I needed to learn a lot of things before I could get going with this system.	good						

# After a Think-Aloud User Testing Session

## 10. Calculate an average SUS score for all users.

1. For odd items: subtract one from the user response.
2. For even-numbered items: subtract the user responses from 5.
3. This scales all values from 0 to 4 (with four being the most positive response).
4. Add up the converted responses for each user and multiply that total by 2.5. This converts the range of possible values from 0 to 100 instead of from 0 to 40.
5. Average across all users.

# Presenting Think-Aloud Testing Findings

You might be presenting Usability findings...

- To propose design decisions
- To defend or rationalize design decisions
- To educate a naïve development or business team
- To evangelize to a skeptical development or business team

# Presenting Think-Aloud Testing Findings

Generally, your audience wants to know...

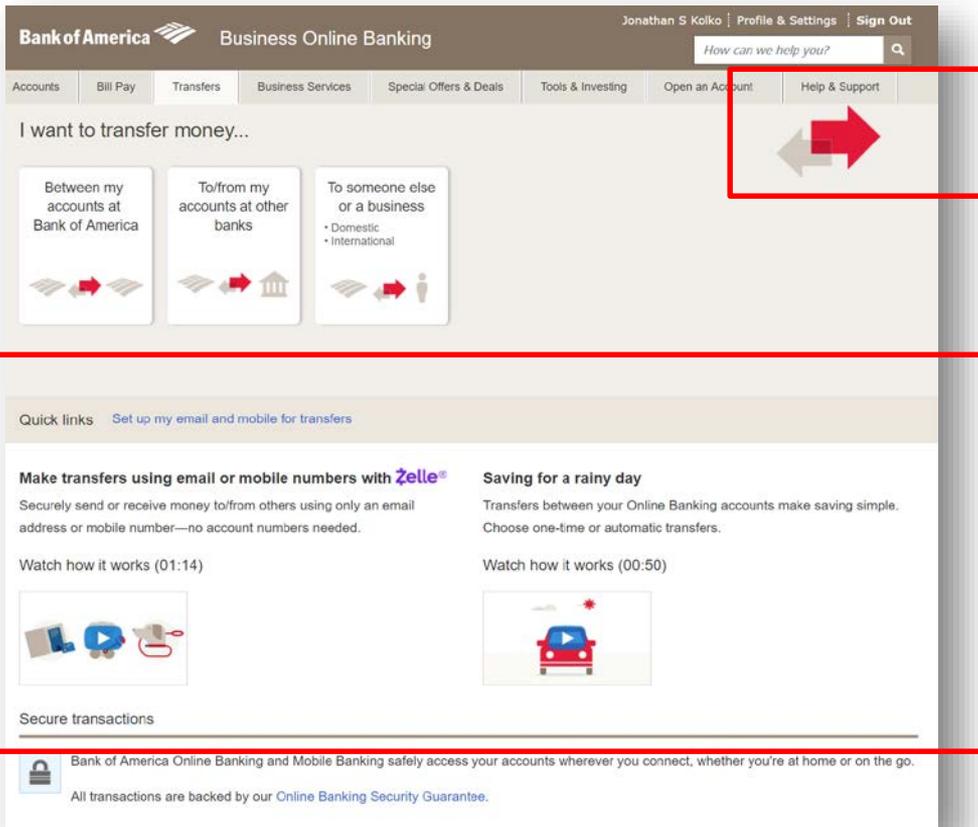
- What is the problem?
- How was the problem identified – was your methodology sound?
- How are you sure that it's a problem – what evidence do you have?
- How important is it to fix the problem, relative to other problems?
- What are the technical and business implications of solving (or not solving) the problem?

# Presenting Think-Aloud Testing Findings

- What method was used?
- Explain how think aloud works
- Explain why think aloud works
- Explain how many people you spoke with, and why such a small sample is acceptable

# Presenting Think-Aloud Testing Findings

- What are the top 3 problems?
- Show a screen shot
- Circle the problem area in a big red box
- Include an actual quote from a user
- Propose a recommendation fix
- Include meta-details (screen unique identifier, screen name, etc)



#### 4.1 Payment Transfer

Main problems found:  
Transfer – Unclear / non-functional visual elements

*"I'm clicking on this arrow to start the transfer, but nothing's happening. I'll refresh the page, but I don't understand why it isn't working..."*

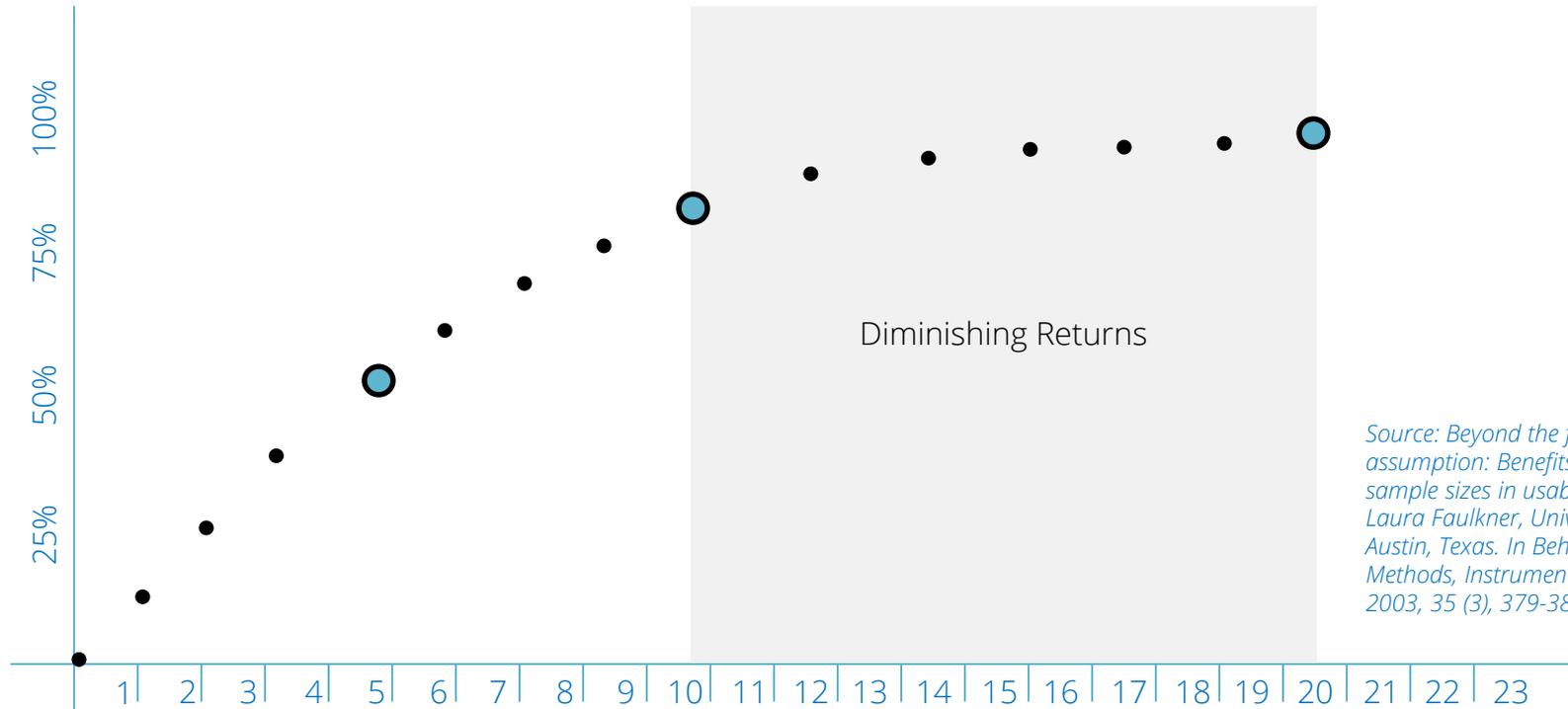
- Aesthetic elements that are decorative are distracting from actual functionality
- The distinction between elements on the screen are confusing
- There are too many choices, at the expense of the "right" choice

Recommendation:

- Remove distracting information that serves no purpose
- Remove advertisements for alternative services
- Use traditional button forms to indicate what areas are clickable

# Think Aloud – How Many Users?

Lowest Percentage of Usability Problems Found



Diminishing Returns

Source: *Beyond the five-user assumption: Benefits of increased sample sizes in usability testing.* By Laura Faulkner, University of Texas, Austin, Texas. In *Behavior Research Methods, Instruments, & Computers*. 2003, 35 (3), 379-383

# Think Aloud Testing – Recap

Evaluating the usability of your work by encouraging a user to think out loud as they use your product or service.

## THINK ALOUD TESTING/

1. Is evaluative, in that it responds to an existing artifact
2. Identifies defects that impact comprehension
3. Is considered a “quick and dirty” method, as compared to a formal experiment
4. Requires interacting with real people

## RUNNING A THINK ALOUD TEST/

1. Develop a prototype.
2. Develop tasks that represent typical user goals.
3. Schedule sessions with users that match the target audience
4. Organize yourself
5. Get set up.
6. Explain the rules.
7. Run the session.
8. Identify Critical Incidents.
9. Calculate a SUS score per user.
10. Calculate an average SUS score for all users.

Let's try it..

# ac4d

Jon Kolko  
Director, Austin Center for Design  
[jkolko@ac4d.com](mailto:jkolko@ac4d.com)