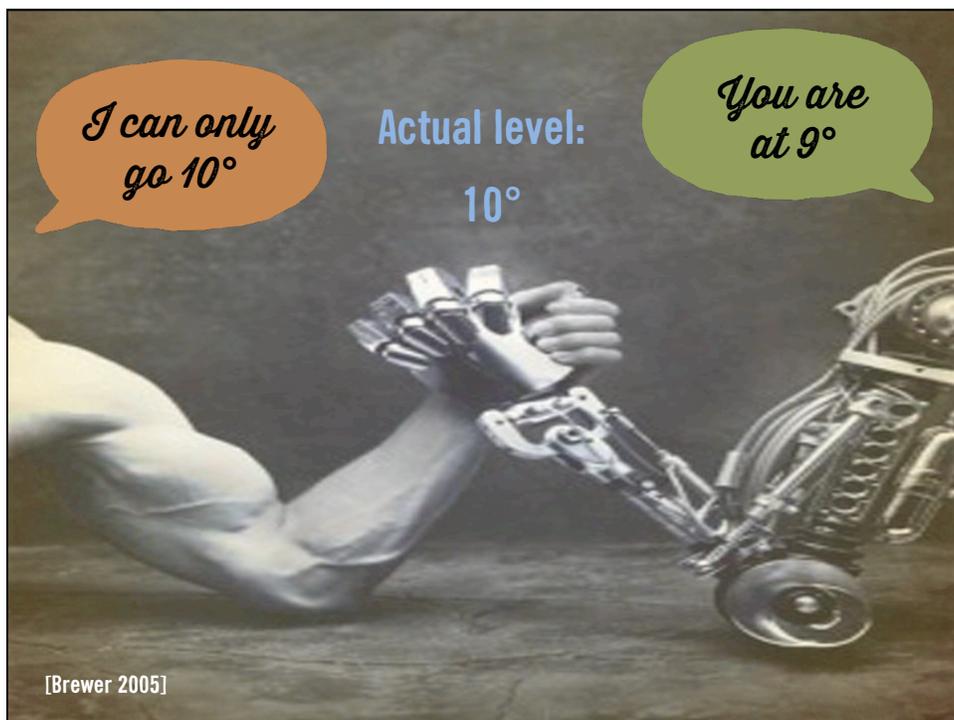


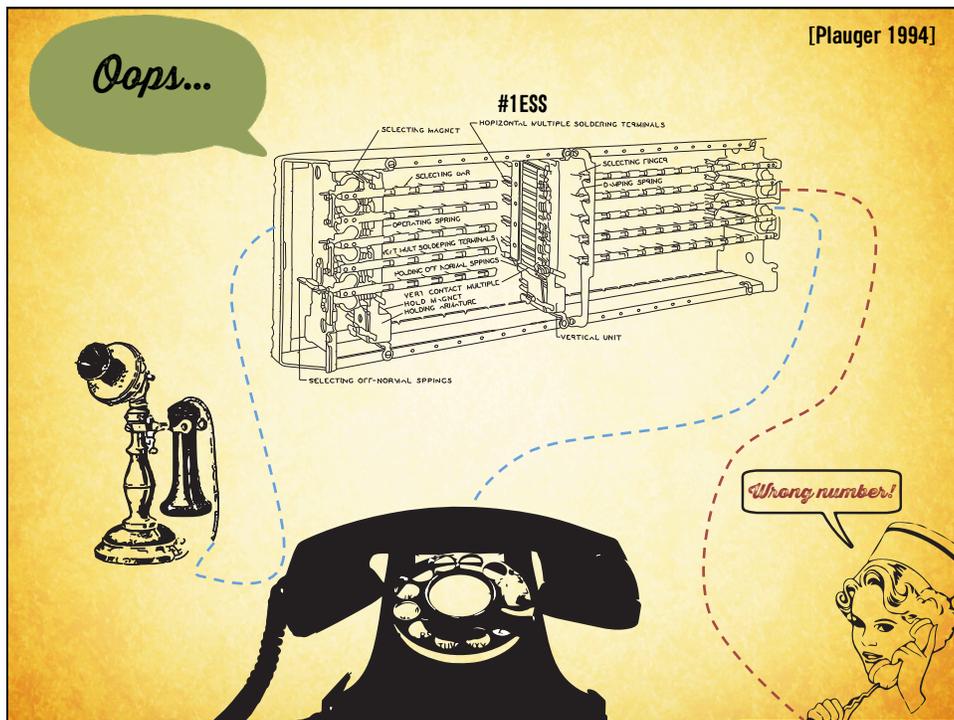
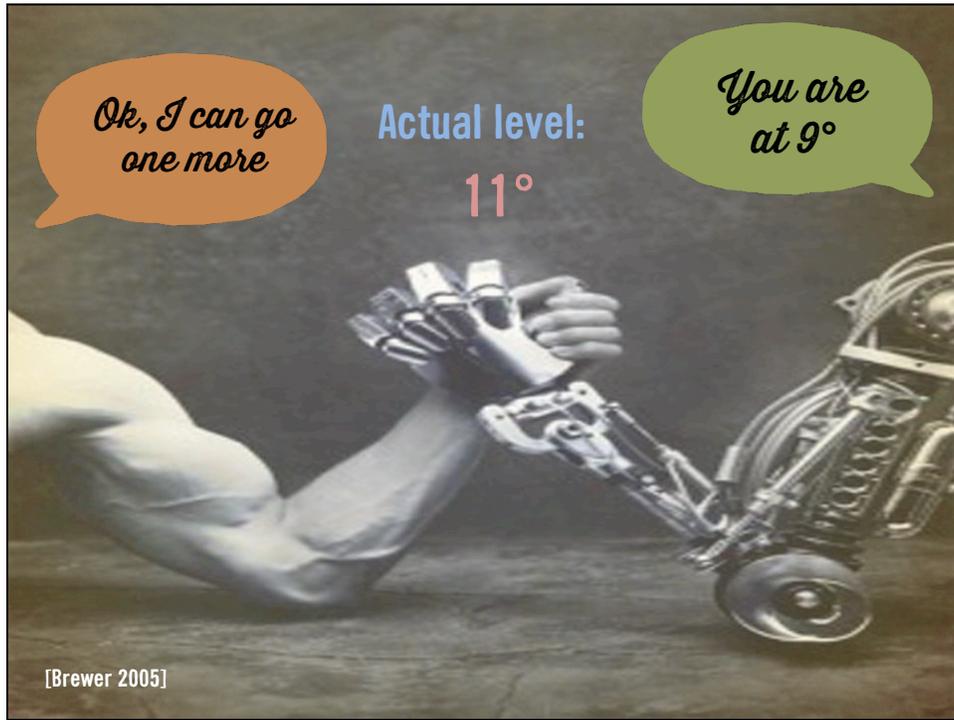
Benevolent Deception in Human Computer Interaction

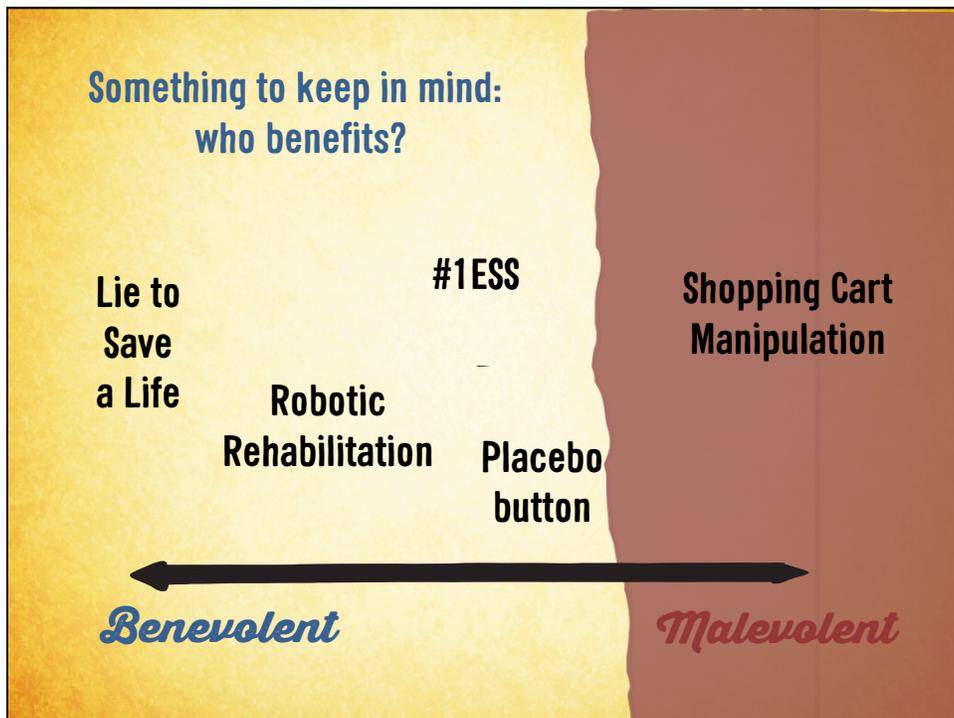
Eytan Adar, Desney Tan, Jaime Teevan

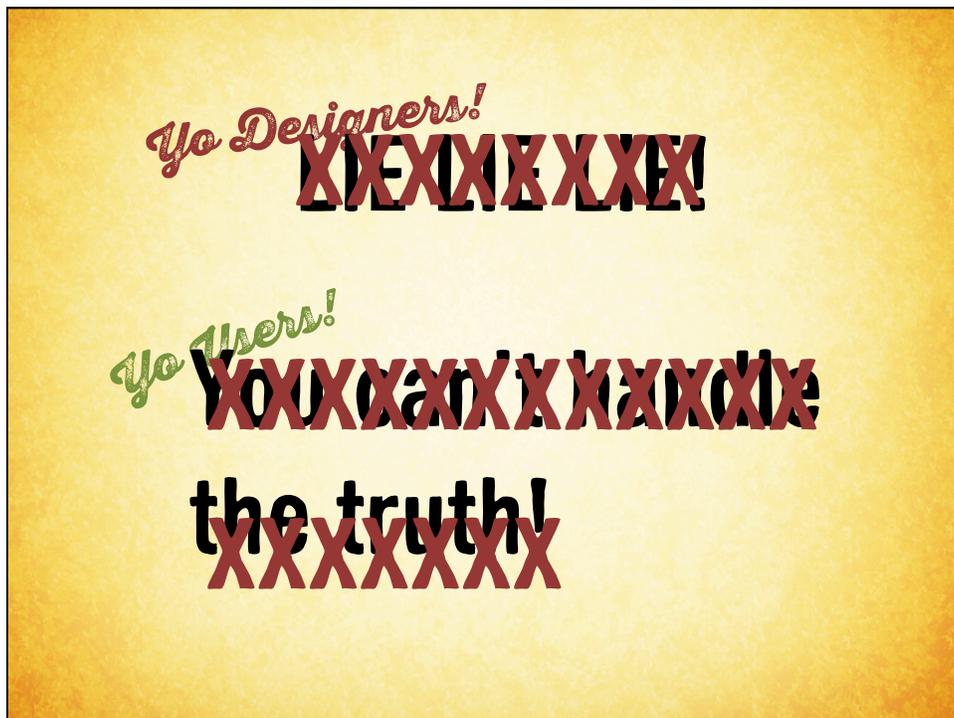
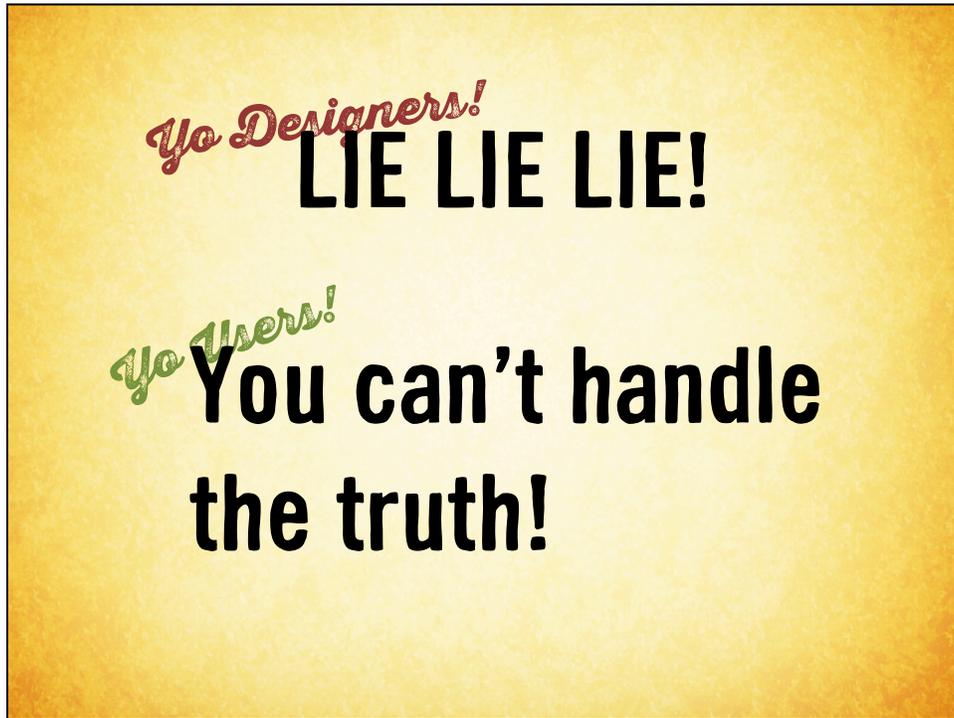


The slide features a textured yellow background with a dark green brushstroke across the top. The title and authors' names are written in a black, cursive font. At the bottom, there are two logos: the 'school of information' logo with a stylized 'i' and 'M' below it, and the 'Microsoft Research' logo.



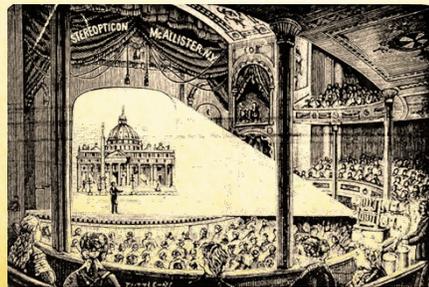






*Benevolent deceptions
are ubiquitous*

We talk “around” deception



[Tognazzini 1993; Laurel 1993]

*Benevolent deceptions
are ubiquitous*

We talk “around” deception

*Makes it extremely hard to
design benevolent deception
properly / effectively / ethically*

Metaphors for deception

Criminology

Means
Motive
Opportunity

False Advertisements

For a better start in life
start **COLA** earlier!

How soon is too soon?

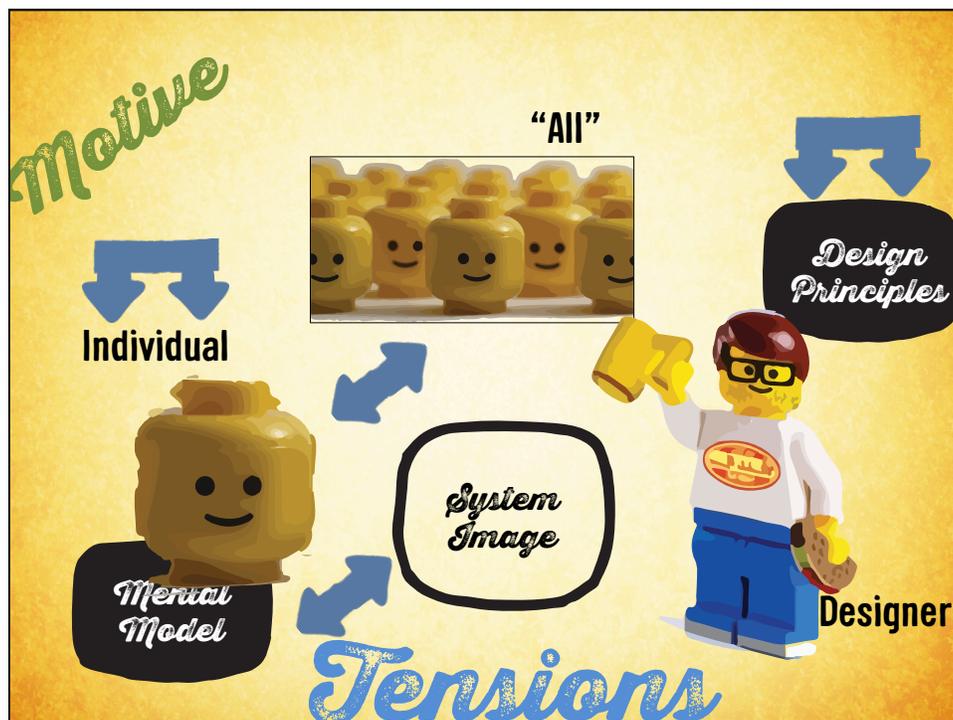
Not soon enough. Laboratory tests over the last few years have proven that babies who start drinking soda during that early formative period have a much higher chance of gaining acceptance and "fitting in" during those awkward pre-teen and teen years. So, do yourself a favor. Do your child a favor. Start them on a strict regimen of sodas and other sugary carbonated beverages right now, for a lifetime of guaranteed happiness.

The Soda Pop Board of America
1515 W. Hart Ave. - Chicago, ILL.

A Working Definition

An explicit or implicit claim, omission of information, or system action, mediated by user perception attention, comprehension, prior knowledge, beliefs, or other cognitive activity, creates a belief about a system or one of its attributes, which is demonstrably false or unsubstantiated as true, where it is likely that the belief will affect behavior, of a substantial percentage of users.

Modified from [Richards 1990]





Motive

Person vs Self

High level goals
inconsistent with
behavior

Individual Safety

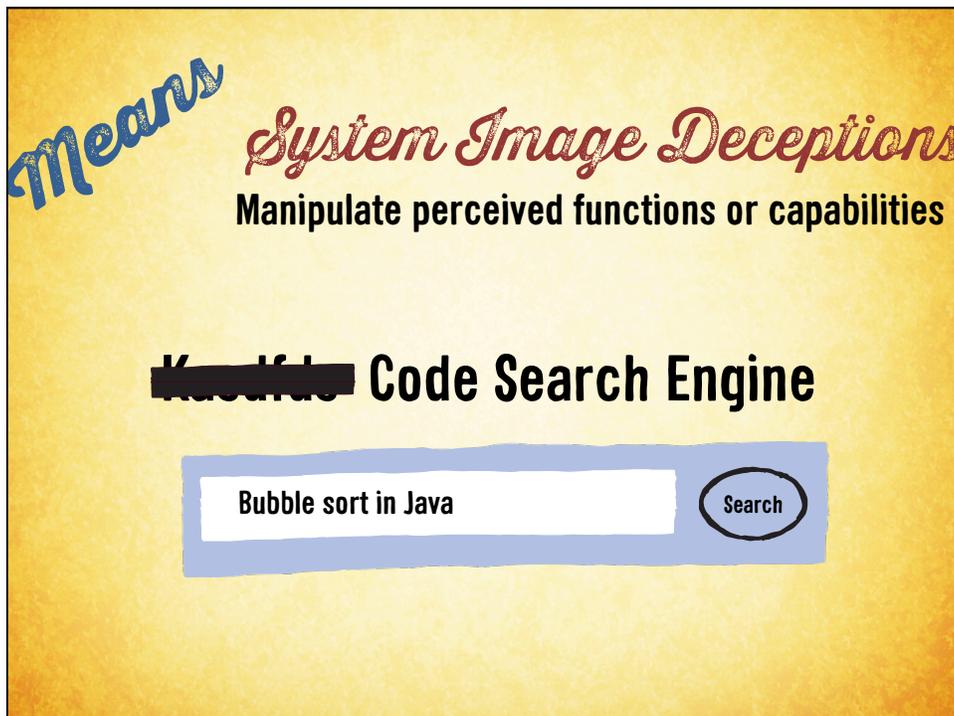
[de Quetteville 2008]

Motive

Design vs Design

Do not fail silently

Fail gracefully



Means

System Image Deceptions
Manipulate perceived functions or capabilities

~~Kasoft~~ Code Search Engine

Java

Here's some junk

Means

System Image Deceptions
Manipulate perceived functions or capabilities

~~Kasoft~~ Code Search Engine

Java

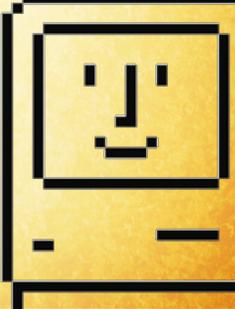
Here's a canned response

Means **Psych Deceptions**

Just Noticeable Differences (JND),
Aesthetics,
Change blindness,
Social processes

I am a human being

Speech recognition system



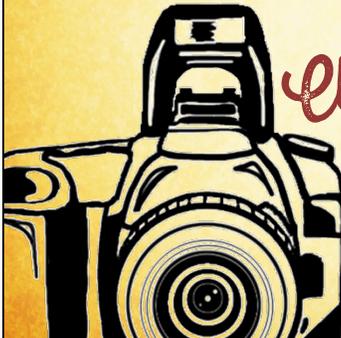
Phone Slave, Grunt [Schmandt 1991]

Means **Mental Model Deceptions**

Metaphors, Skeuomorphs, and more

Not all metaphors are deceptions
Not all metaphors change behavior

click!



Means **Mental Model Deceptions**
Metaphors, Skeuomorphs, and more

*I am an not a
police officer
I am an actor playing
a police officer*

“ [while] not believ[ing] in the role being real, [the user was ‘tricked’] into believing the meta-role is real” – [Klessen 2003]



Motive
Tension *Means*
Tension Removal
Opportunity
When is the user prepared

Opportunity

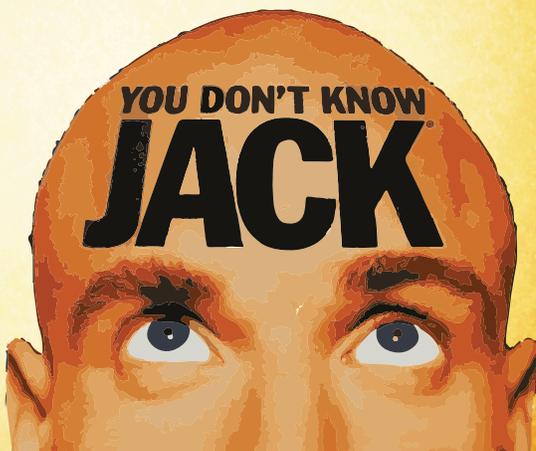
**User wants to be deceived
("willing suspension of disbelief")**

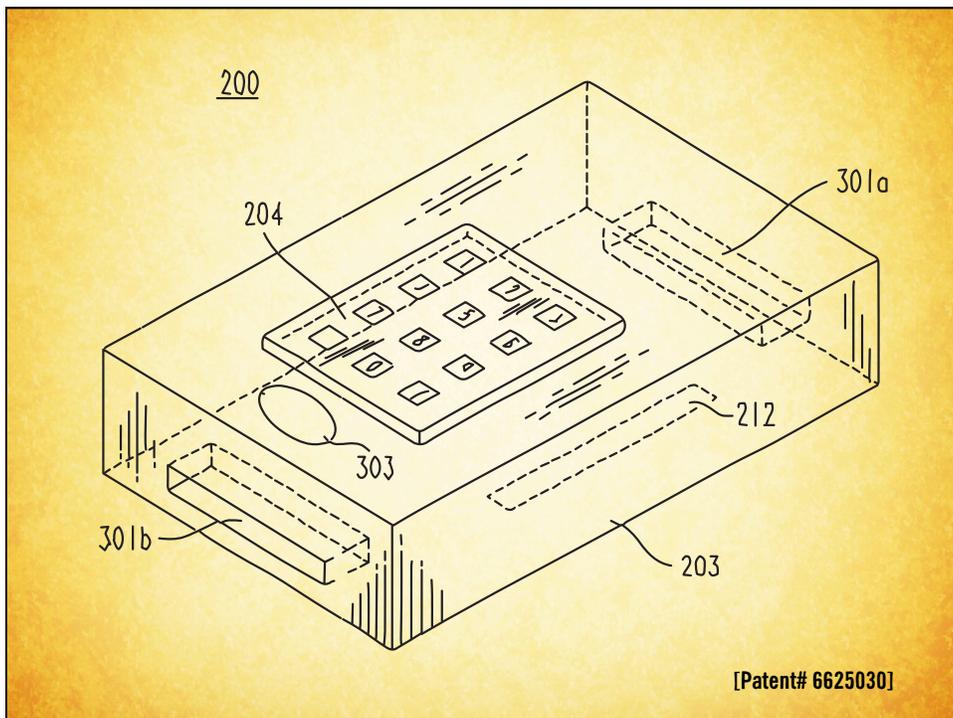
**User will not be able to distinguish truth from
deception**

Opportunity

**The user will spend a few
minutes trying to identify
the limits of the AI, get
bored, and then go back to
enjoying the game**

Paraphrased from [Gottlieb 1997]

A stylized, high-contrast illustration of a man's face, focusing on his eyes and forehead. The text "YOU DON'T KNOW JACK" is written in large, bold, black letters across his forehead. The man has a surprised or intense expression, with wide eyes and a slightly open mouth. The background is a textured, yellowish-gold color.



Because battery module [is] in a position that many less-technologically sophisticated users might expect to see a microphone, [the] module **advantageously** comprises **microphone façade** which is not a microphone at all, but is merely a cosmetic device (e.g., a series of holes, a screen, etc.) to deceive the user into thinking that the microphone is indeed where they expect it to be.

[The] ... façade provides [a] **comfortable feeling** that [the device] is constructed as they perceive it should be,

and this might make a prospective purchaser of wireless terminal assembly **more likely to purchase** it than if microphone façade were absent.

[Patent# 6625030]

Should we?

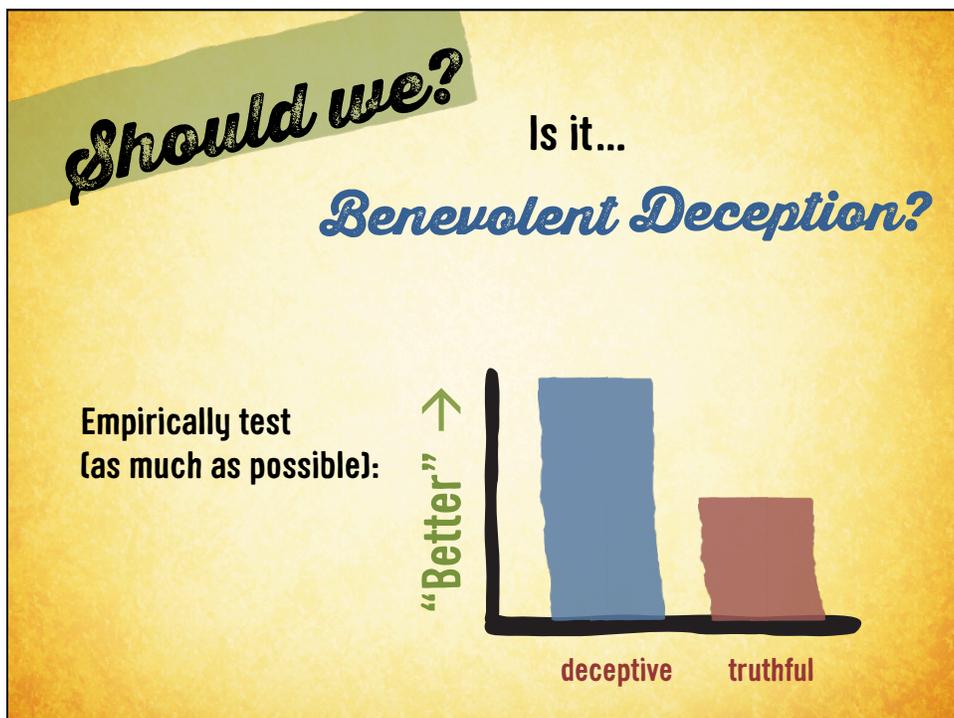
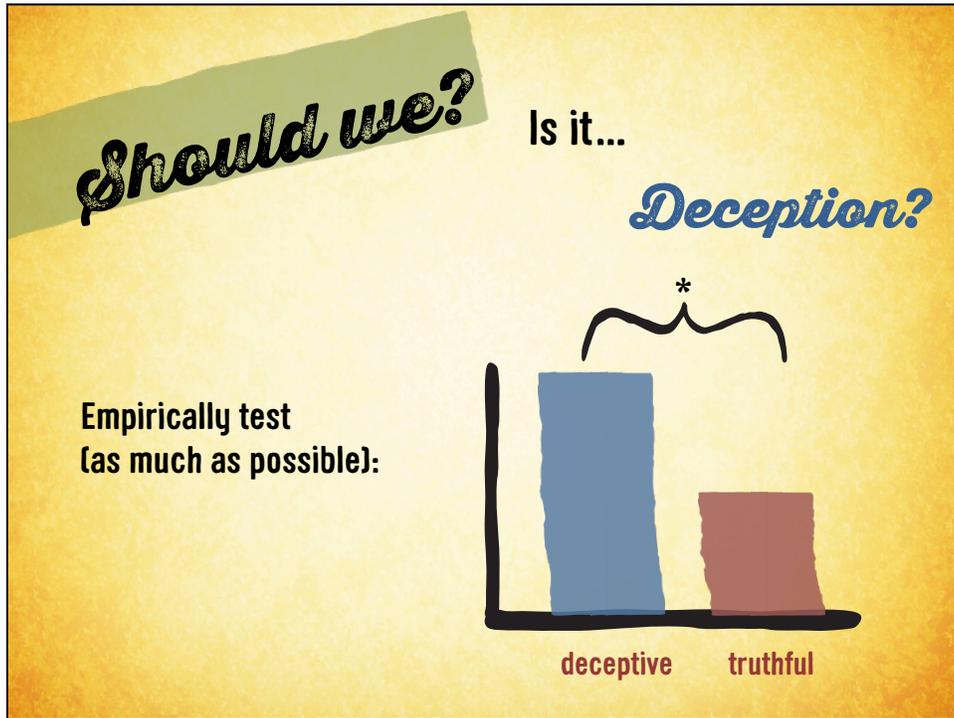
Is it...

Deception?

Empirically test
(as much as possible):

perception reality

d



Should we?

➔ **Risk**

- ⇨ commitment to the deception
- ⇨ consequences of getting caught
 - ⇨ benevolence as gateway drug
 - ⇨ ignore alternatives

Should we?

Is it...
Ethical Benevolent Deception?

Rationalization?

Category	Relative Value (Better)
deceptive	High
truthful	Low

Bok (1999)

Should we? Is it...
Ethical Benevolent Deception?

Rationalization → Justification

Bok (1999)

Principle of Veracity

Deception as last resort

Did we really try all reasonable,
truthful alternatives?

Bok (1999)

Publicity

What happens if the truth emerges?

Are we/users still satisfied with the deception?

Bok (1999)

Should we?

Is it...
Ethical Benevolent Deception?

← Preference (even when exposed)

Empirically test (as much as possible):

A bar chart with a vertical axis labeled 'Better' with an upward arrow and a horizontal axis with two categories: 'deceptive' and 'truthful'. The 'deceptive' bar is blue and significantly taller than the 'truthful' bar, which is brown. An arrow points to the top of the 'deceptive' bar with the text 'Preference (even when exposed)'.

Option	Preference Level
deceptive	High
truthful	Low

A research plan for deception

- 1) Designer studies
- 2) User studies
- 3) Improving definitions/metrics
- 4) Modeling truth
- 5) Ethics debate

Summary...

**Benevolent deception is ubiquitous
(we don't even know we're doing it)**

Not helpful to ignore

Practice → Design

Some progress towards understanding

Many research directions

Thanks!!

Mark Ackerman, Sara Adar, Yaw Anokwa, Michael Bernstein, Batya Friedman, Travis Kriplean, Yang Li, David Notkin, and Dan Weld

(and many others who gave us great examples)

 school of information
Microsoft Research

Eytan Adar
**University of Michigan
eadar@umich.edu**

Desney Tan
**Microsoft Research
desney@microsoft.com**

Jaime Teevan
**Microsoft Research
teevan@microsoft.com**

Thanks!!

