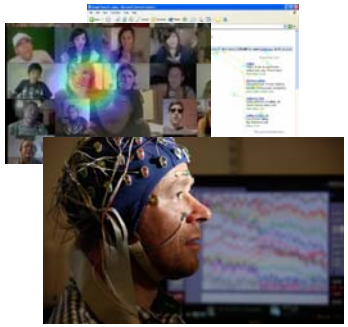
case studies in design informatics

Lecture 2: When Smart Things Go Wrong



design informatics

NRlabs neuropolitics research

Robin Hill

Institute for Language, Cognition & Computation,
School of Informatics
and
Neuropolitics Research Lab,
Politics and I.R.
www.robin.org.uk
r.l.hill@ed.ac.uk

Today's objective

- Awareness of errors.
- Holistic approach to design.
- Data overload (multimodal = multiple streams of data that have to be synchronised and combined).
  - Different sensors, biological and non-biological, relay information at different rates and with different latencies (e.g. sound and light, both externally and internally).
- System resilience is a combination of human and technology reliability.
- Things go wrong for many different reasons.

It's only a glitch.

Poor Human-Robot Interaction

- ED-209: You now have 15 seconds to comply. [Human compliance undetected by robot.]
- ED-209: You are in direct violation of Penal Code 1.13, Section 9. [Uninformative referring expression for human.]
- ED-209: You have 5 seconds to comply. [Human actions unrecognised.]
- Human: Help me! [Human speech unrecognised.]
- ED-209: Four... three... two... one... I am now authorized to use physical force!
- [Fatal robot action begins…]
- Other human: Can you pull the plug on this thing? [Insufficient safety backup – "kill switch".]
- Other human: He didn't hear it! [System failure.]
- Tech leader: I'm sure it's only a glitch. A temporary setback. [Not a good D.I. attitude.]

Multimodal

- Normal communication is multimodal.
- For embodied agents (humans or robots) it is inefficient, if not simply wrong, to assume
  - independent language system/module;
  - independent vision system/module;
  - independent motor control system/module during interaction.
  - Research specialisms can encourage this view. Design Informatics view should be **holistic**.

Errors

- **Simple failure**: something breaks or stops functioning correctly, e.g. motor burns out.
- **Accidental**: wrong choice or decision; unintentional, e.g. hit the wrong button.
- **Ambiguity and Confusion**: miscommunication or misunderstanding.
- Problem: can be slow to detect or totally undetected.
  - Known errors/bugs and unknown errors/bugs.
- Potential learning opportunity when encountered!

Error analytics

Four canonical errors in science (Mayo, 1996)

1. mistaking chance effects or spurious correlations for genuine correlations or regularities;
2. mistakes about the quantity or value of a parameter;
3. mistakes about a causal factor; and
4. mistakes about experimental assumptions.

Mayo, Deborah. 1996. *Error and the Growth of Experimental Knowledge*. Chicago: University of Chicago Press.

Human Factors Analysis and Classification System (HFACS)

- Initially devised to understand the role of human error in aviation accidents.
- Attempt to model human error in a complex technological environment.

Wiegmann, D. A., & Shappell, S. A. (2001). Human error analysis of commercial aviation accidents: Application of the human factors analysis and classification system (HFACS). [Article]. *Aviation Space and Environmental Medicine, 72(11), 1006-1016.*

Woods, David D., Johannesen, Leila J., Cook, Richard I., Sarter, Nadine B., & Dayton Univ Research Inst, O. H. (1994). *Behind Human Error: Cognitive Systems, Computers and Hindsight. Ft. Belvoir: Defense Technical Information Center.*

Skalle, P., Aamodt, A., & Laumann, K. (2014). Integrating human related errors with technical errors to determine causes behind offshore accidents. *Safety Science, 63, 179-190. doi: 10.1016/j.ssci.2013.11.009.*

Ambiguity

- Temporary – high level of uncertainty. Resolved quickly by obtaining more data. "The answer becomes clear."
- Global – high level of uncertainty. Correct choice or option may never be known.
- Often probabilistic solutions.

Who likes broccoli?

Instantaneous binary world?

- Maybe… leaning towards… perhaps… certain?
- Speed of response.
- Social response.
- Interested in the "why" ("show your working").
- Not just the end result but the process or steps getting there.

Black or white; right or wrong

- Different paths lead to the same destination.
- Not always a single correct (or best) solution.
  - Speed-accuracy trade-off.
  - Long and easy; or fast and complicated?

A major issue during interaction

- Generating referring expressions: how do people refer to objects and distinguish between multiple objects?
- The "my problem" or "somebody else's problem" is not trivial in communicating reference.
- A lot depends on whether the speaker considers the listener's view/knowledge of the world

  Audience design a.k.a. my problem
- or whether the speaker assumes their world knowledge is the "standard model"

  Egocentric a.k.a. somebody else's problem.

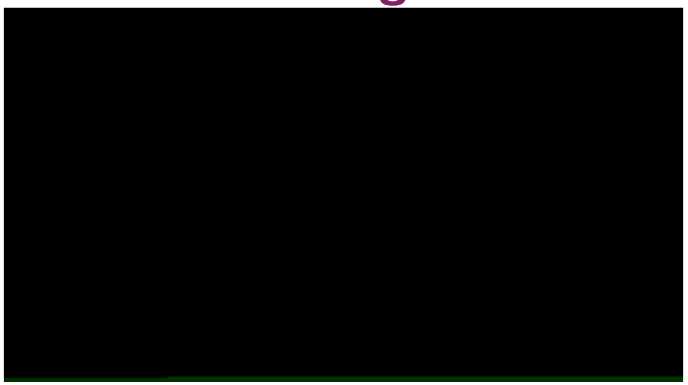
Prediction & intentionality

- Knowing the goal (or even the context or general discourse) makes life easier and reduces errors.
- Interpreting referring expressions

"I'm going to get the red square / this one / it."

reduces the need to simply be reactive to a situation or even computationally anticipatory.

Counting task



Divided attention

- Limited computational resources (still true for humans and machines).
- Competing data channels.
- Weighting or priority of data evidence (task dependent)?
- Multitasking is really a myth: typically switch between tasks (selective attention).

Pokémon Go on the go

Ayers, J. W., Leas, E. C., Dredze, M., Allem, J., Grabowski, J. G., & Hill, L. (2016). Pokémon go—a new distraction for drivers and pedestrians. *JAMA Internal Medicine doi: 10.1001/jamainternmed.2016.6274.*[Published online September 16, 2016.]

Optimal conditions can vary

Recognition system based on face and voice matching.

- Face recognition suggests Person X;
- Voice recognition suggests Person Y.
- What to do with conflicting information?
  - For this particular system face recognition is usually highly reliable, so trust it more.
  - Perhaps conditions mean you should weight or bias the evidence in favour of voice? Bright lights or night time.

Visual clutter

- Information overload! Too much presented at once.
- Again, issue of limited capacity / resources.
- Human phenomenon, but also a problem for automatic object identification.

Miniukovich, A., & De Angeli, A. (2014). Quantification of interface visual complexity. *Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces, Avi 2014, 153-160. doi: 10.1145/2598153.2598173.*

McPhee, L. C., Scialfa, C. T., Dennis, W. M., Ho, G., & Caird, J. K. (2004). Age differences in visual search for traffic signs during a simulated conversation. *Human Factors, 46(4), 674-685. doi:* 10.1518/hfes.46.4.674.56817

Rosenholtz, R., Li, Y. Z., & Nakano, L. (2007). Measuring visual clutter. *Journal of Vision, 7(2) doi: 10.1167/7.2.17.*

Obfuscation

Person walks behind wall (obfuscated). Still exists? Walks out again – is it the same person? Multi-object tracking makes it even harder.

Samsung safety truck

Equipped with a front-end wireless camera and four external monitors on the back, it can live stream oncoming traffic, hopefully removing the need for risky overtaking manoeuvres. [Need to trust the live feed.]

Spot the difference

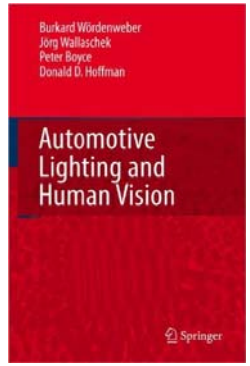
Spot the difference (2)

[Aside / comment]

- Donald Hoffman (Prof. of Cog Sci), who provided the previous images, also co-wrote a highly specific design manual:

"The safety of vehicle traffic depends on how well automotive lighting supports the visual perception of the driver. This book explains the fundamentals of visual perception, like e.g. physiology of eye and brain, as well as those of automotive lighting technology, like e.g. design of headlamps and signal lights. It is an interdisciplinary approach to a rapidly evolving field of science and technology written by a team of authors who are experts in their fields."

- Example of good design fusion.



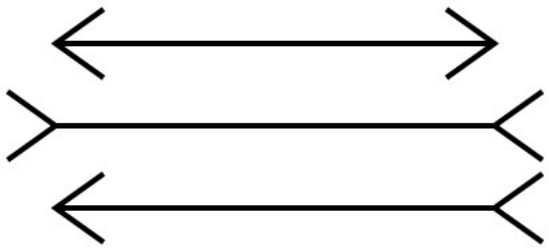
Filtering

- Inattentional blindness (Gorilla example): failure to notice something that is fully obvious but your attention is somewhere else (distracted).
- Change blindness: failure to notice a difference between what is there now and what was there a moment ago.
- Often miss things – attention not so great.
- Encode what is relevant and ignore / discard the rest.
- "Cocktail party effect": audio salience filtering [next week].

Illusions

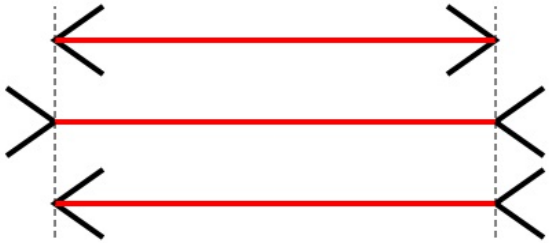
- Sometimes senses fail because of the architecture of the system (cognitive, physiological; hardware, software).
- These are usually impossible to prevent – they are an artefact of the system.
- Conjurers (magic tricks) often exploit this in combination with divided attention (distraction).

Muller Lyer illusion



Which horizontal line is bigger?

Muller Lyer illusion



Assumptions - Quirkology

Camera: 2D vision.

Problem for automated systems as well.

Tesla driver dies in first fatal crash while using autopilot mode

The autopilot sensors on the Model S failed to distinguish a white tractor-trailer crossing the highway against a bright sky

Joshua Brown, the first person to die in a self-driving car accident. Photograph: Facebook

The first known death caused by a self-driving car was disclosed by Tesla Motors on Thursday, a development that is sure to cause consumers to second-guess the trust they put in the booming autonomous vehicle industry.

The Guardian, 1 July 2016

McGurk Effect

Limitations of knowledge and knowledge limitations

- "Being aware of our limitations can help us adapt and compensate for them, allowing us to do things that prevent the really negative consequences that can happen due to failures of awareness."

  Daniel Simons, University of Illinois, Urbana-Champaign.

Chabris, C. F., & Simons, D. J. (2010). The Invisible Gorilla (And Other Ways Our Intuitions Deceive Us). New York, NY: Crown. (see www.theinvisiblegorilla.com for details).

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. Perception, 28, 1059-1074.

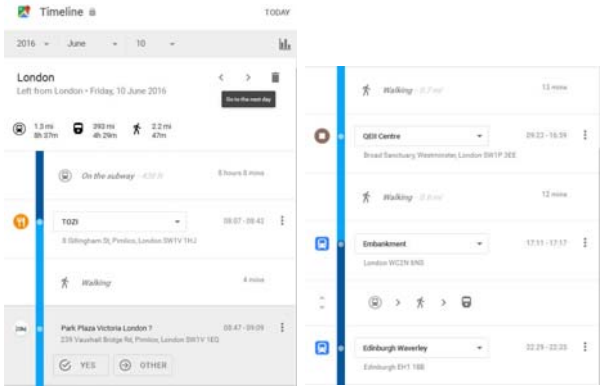
Simons, D. J., & Jensen, M. S. (2009). The effects of individual differences and task difficulty on inattentional blindness. Psychonomic Bulletin & Review, 16(2), 398-403.

Simons, D. J., & Rensink, R. A. (2005). Change blindness: Past, present, and future. Trends in Cognitive Sciences, 9(1), 16-20.

Data quality

- Links back to error analytics.
- Can you trust your data?
- No data – really none; or missing / lost / uncaptured?
  - Google tracking. Stayed at home or just left my phone at home?
- Reliable signal?
  - Google tracking. Sometimes it thinks I sleep at my neighbour's house.
- "noisy" data
  - Controlled lab conditions or uncontrolled field test.
  - Face-to-face interviews vs online surveys.

Conference trip



SpaceBook GPS

- GPS signal can be very accurate but that relies on clear line of sight with multiple satellites.
  - Not too bad for cars in the middle of roads.
  - Not too good for pedestrians walking next to tall buildings that shield the signal.
- Errors in signal resulted in people appearing to teleport (apparate) or often detected on the wrong side of the road.

SpaceBook ASR

Automatic speech recognition system defeated by:

- Rain on umbrella (held just above head and therefore microphone).
- Wind blowing against microphone.
- Bagpipes.
- Passing buses.
- Intermittent drops in mobile phone signal.

Lab data gathering (good quality)

Not so good out of the lab…

Family life and impersonation

Inferential statistics

- How reliable and complete is the data you are using?
- Enough evidence to make strong conclusions?
- Errors: failure to reject something that wasn't there (false positive); failure to detect something that was (false negative) [Type I & Type II Errors].
- Often controversial assumptions made about data, distributions, statistical power, sample size, random or unbiased sampling, levels of significance, etc.
- But not today! ☺