HCI: THINK ALOUD AND HEURISTIC

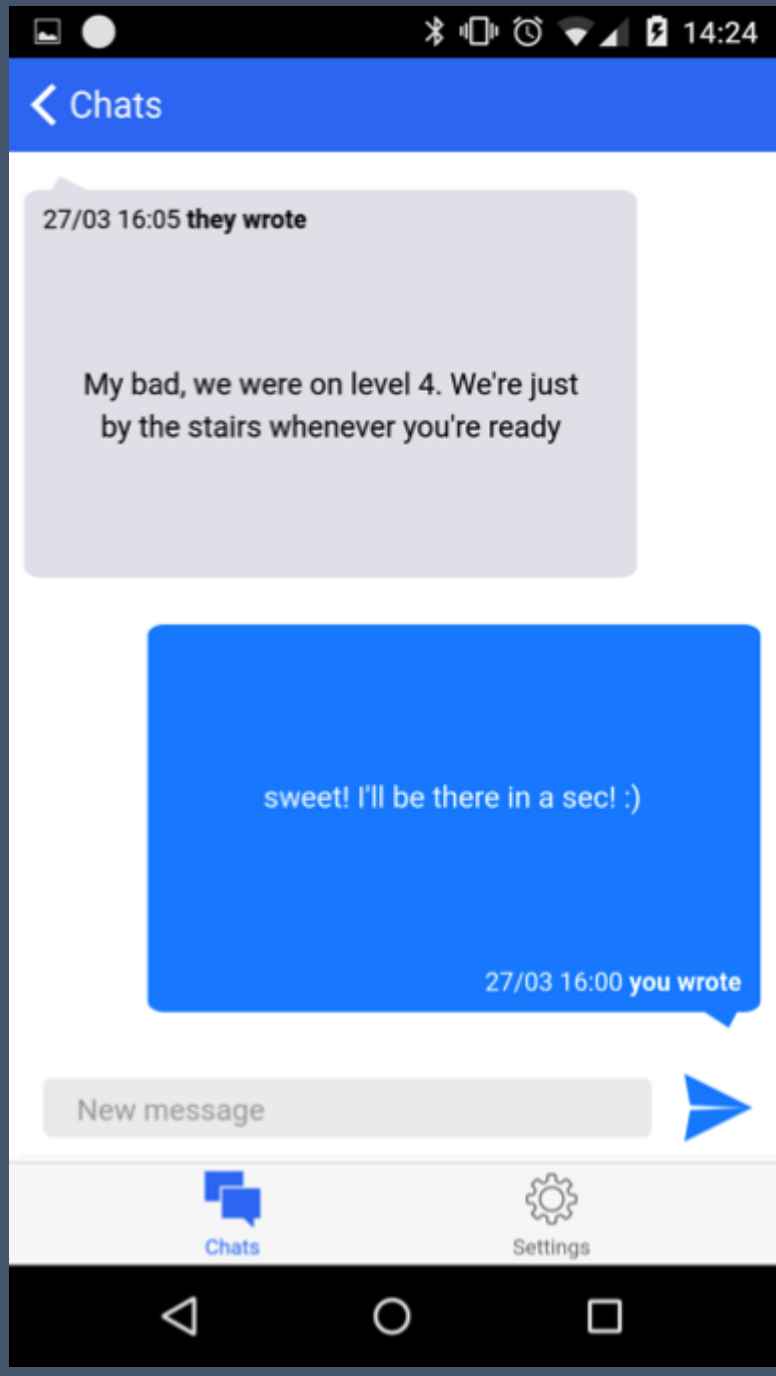
Dr Kami Vaniea

First, the news…

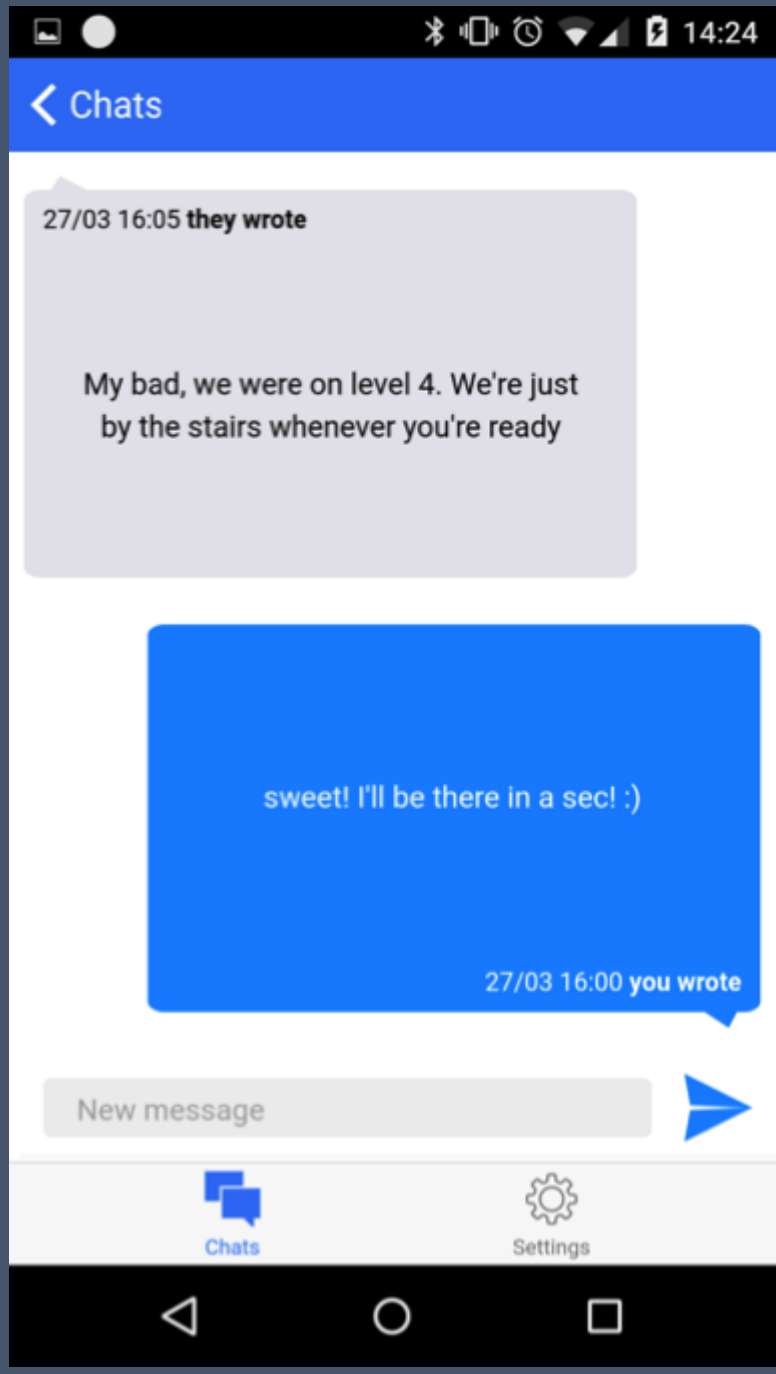
- Forget-me-not: History-less mobile messaging
- Mattias Rost, Christos Kitsos, Alexander Morgan, Martin Podlubny, Pietro Romeo, Edoardo Russo, and Matthew Chalmers. 2016. Forget-me-not: History-less Mobile Messaging. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (CHI '16). ACM, New York, NY, USA, 1904-1908. DOI: http://dx.doi.org/10.1145/2858036.2858347

What if a mobile conversation was more like a verbal one?

The Forget-me-not app only shows you the last thing you and they said, no other history.



“We ran a trial with 10 participants for 2 weeks. The participants were recruited from two groups with existing social ties.”



Tutorials

- There are now more people signed up for this class than fit in our tutorials….
  - 125 students according to EUCLID
  - 120 according to the Doodle poll
  - 116 according to Learn
  - And we have space for a max of 120 in the tutorials …

- I will be announcing a new tutorial session as soon as the ITO tells me when and where it is, you are welcome to switch to the new tutorial

Coursework 1 Q&A

- Q&A drop-in session for CW1
- 2-4pm today (Thursday)
- Informatics Forum 1.15

- I will be posting answers to any common sounding questions onto the coursework page

Coursework 1 Groups

- Please sign your group up on Learn
- 30 of you currently have no group on Learn
- 7 groups have 2 or fewer members

- I have turned on all the random group features of Learn so each group should have access to the ability to share documents, a wiki, a forum, and the email of the other group members

Usability testing & research

Usability testing

- Improve products
- Few participants
- Results inform design
- Usually not completely replicable
- Conditions controlled as much as possible
- Procedure planned
- Results reported to developers

Experiments for research

- Discover knowledge
- Many participants
- Results validated statistically
- Must be replicable
- Strongly controlled conditions
- Experimental design
- Scientific report to scientific community

Think Aloud

Think aloud

- Basic idea: Have a participant use the interface and speak aloud while they do so
- Think aloud is a very versatile, can be long or short, detailed or minimal, planned or ad-hoc
- Pros
  - Get a sense of what the user is trying to do and why they click on some things
  - Very detailed information
  - Testing with 5 users will find the majority of major issues
- Cons
  - Small sample sizes
  - Talking aloud changes how long a user spends on tasks so this method cannot be combined with timing

Think aloud

- Think aloud sessions are typically scripted, that is, you write down everything you will say in advance
- Everything you say to the participant will change their behavior so you have to be very careful
- Typical session
  1. Tell the participant what the session will involve including things like how long it will be and what kind of data recording you will be doing
  2. Train them in thinking aloud
  3. Ask them to accomplish several tasks which have been previously written down, reading aloud each task before starting it
  4. End by thanking them and offering to answer any questions they may have

Usability testing the iPad

- 7 participants with 3+ months experience with iPhones
- Signed an informed consent form explaining:
  - what the participant would be asked to do;
  - the length of time needed for the study;
  - the compensation that would be offered for participating;
  - participants' right to withdraw from the study at any time;
  - a promise that the person's identity would not be disclosed; and
  - an agreement that the data collected would be confidential and would be available to only the evaluators
- Then they were asked to explore the iPad
- Next they were asked to perform randomly assigned specified tasks

Examples of the tasks

App or website	Task
iBook	Download a free copy of *Alice's Adventures in Wonderland* and read through the first few pages.
Craigslist	Find some free mulch for your garden.
eBay	You want to buy a new iPad on eBay. Find one that you could buy from a reputable seller.
Time Magazine	Browse through the magazine and find the best pictures of the week.
Epicurious	You want to make an apple pie for tonight. Find a recipe and see what you need to buy in order to prepare it.
Kayak	You are planning a trip to Death Valley in May this year. Find a hotel located in the park or close to the park.

Table 14.1 Examples of some of the tests used in the iPad evaluation (adapted from Budiu and Nielsen, 2010).

Source: Copyright Nielsen Norman Group, from report available at http://www.nngroup.com/reports/.

Think aloud training

“In this observation, we are interested in what you think about as you perform the tasks we are asking you to do. In order to do this, I am going to ask you to think aloud as you work on the task. What I mean by “think aloud” is that I want you to tell me everything you are thinking from the first time you see the statement of the task until you finish the task. I would like you to talk aloud constantly from the time I give you the task until you have completed it. I don’t want you to try to plan out what you say or try to explain to me what you are saying. Just act as if you were alone, speaking to yourself. It is most important that you keep talking. If you are silent for any long period of time, I will ask you to talk. Do you understand what I want you to do?”

Observe-pair-share

- I am going to do a live think-aloud with myself as the subject
- You need to record:
  - Tasks and subtasks I engage in
  - Any critical issues I have
  - Any unexpected behaviors

Observe-pair-share

- I am going to do a live think-aloud with myself as the subject
- You need to record:
  - Tasks and subtasks I engage in
  - Any critical issues I have
  - Any unexpected behaviors I engage in
- I need:
  - A website
  - A task

Pair-share

- Share your observations with your neighbor
- If you could change how this website is designed, what would you change?

Think aloud Analysis

- Task analysis
- Critical incident analysis

Think aloud sounds easy but…

- Wording is important. Everything you say to the participant will have an impact and change their behavior
  - Read from a script, or better yet, memorize the script
  - Some researchers hire actors to run think aloud sessions because they can memorize scripts
- It is VERY important that you not talk to the participant during the tasks (this is surprisingly hard to do)
- The only things you can say off-script are
- Train the participant in thinking aloud, this takes time, but it is very important to do

Think aloud sounds easy but…

- Never call the protocol “think aloud” in front of a participant. It is always “talk aloud” or “speak aloud”.
  - The word “think” implies “explain” which will cause participants to start explaining what they are doing to you. This is bad, because when they start explaining they stop behaving normally, because they are thinking about it more.
- Immediately stop the protocol, or step in if the participant becomes distressed
- Help the participant if they ever spend more than 3 minutes on a sub-task

Usability Inspections

Inspections

- Several kinds.
- Experts use their knowledge of users & technology to review software usability.
- Expert critiques can be formal or informal.
- Heuristic evaluation is a review guided by a set of heuristics.
- Walkthroughs involve stepping through a pre-planned scenario noting potential problems.

Heuristic evaluation

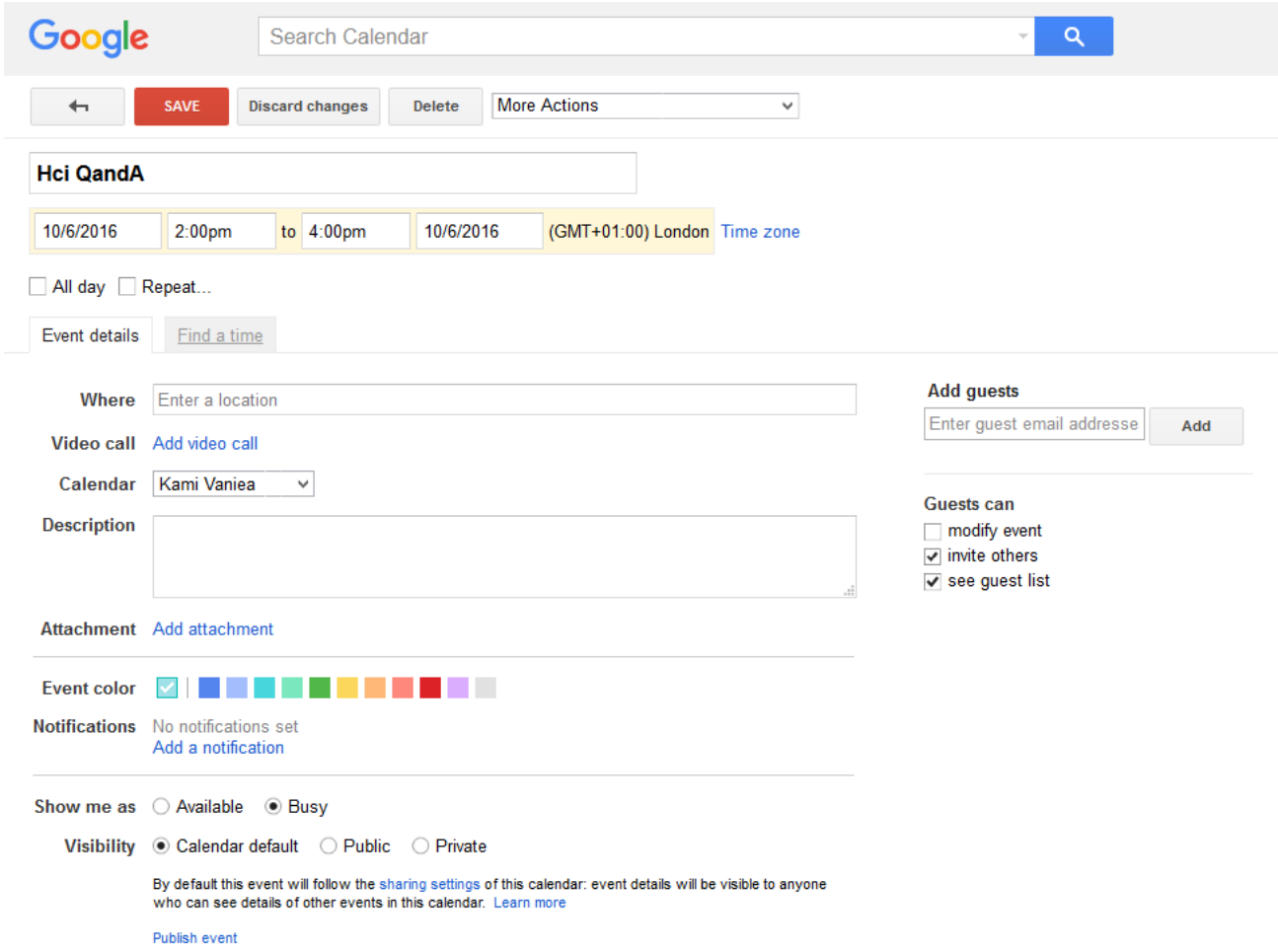
- Developed by Jacob Nielsen in the early 1990s.
- Based on heuristics distilled from an empirical analysis of 249 usability problems.
- These heuristics have been revised for current technology by Nielsen and others for:
  - mobile devices,
  - wearables,
  - virtual worlds, etc.
- Design guidelines form a basis for developing heuristics.

Revised version (2014) of Nielsen's original heuristics

- Visibility of system status.
- Match between system and real world.
- User control and freedom.
- Consistency and standards.
- Error prevention.
- Recognition rather than recall.
- Flexibility and efficiency of use.
- Aesthetic and minimalist design.
- Help users recognize, diagnose, recover from errors.
- Help and documentation.

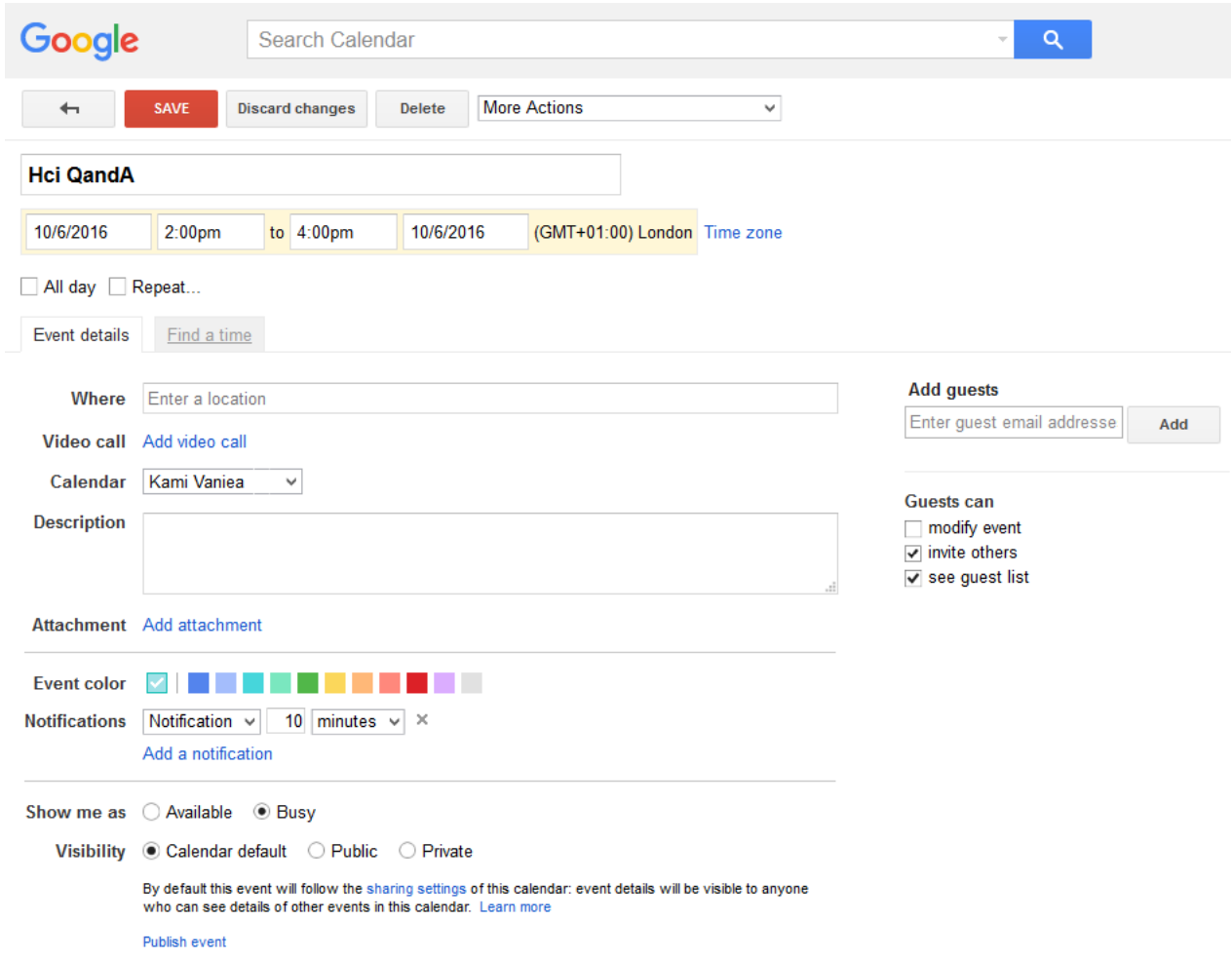
Visibility of system status

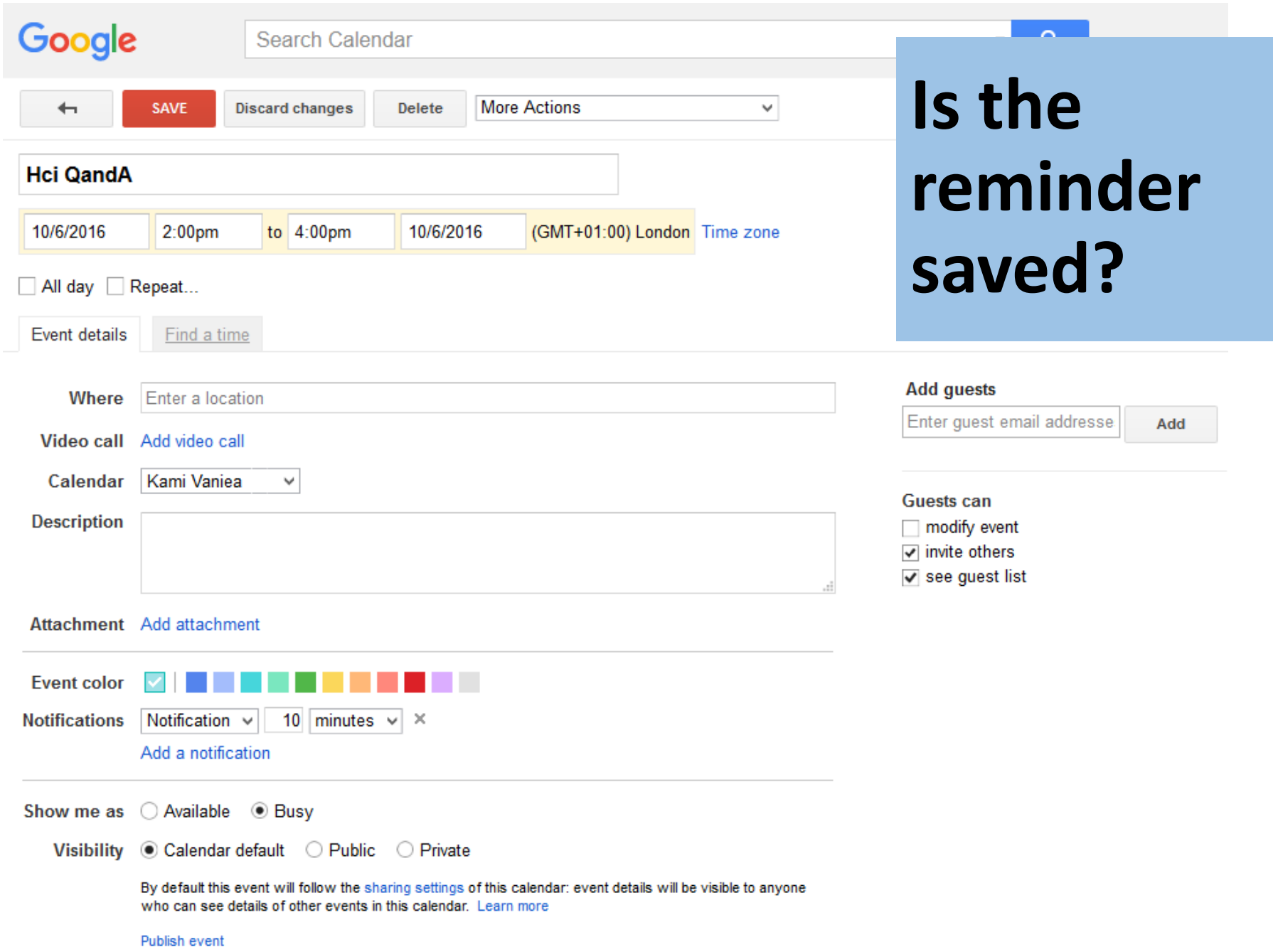
Me adding the Q&A session to my Google calendar



Visibility of system status

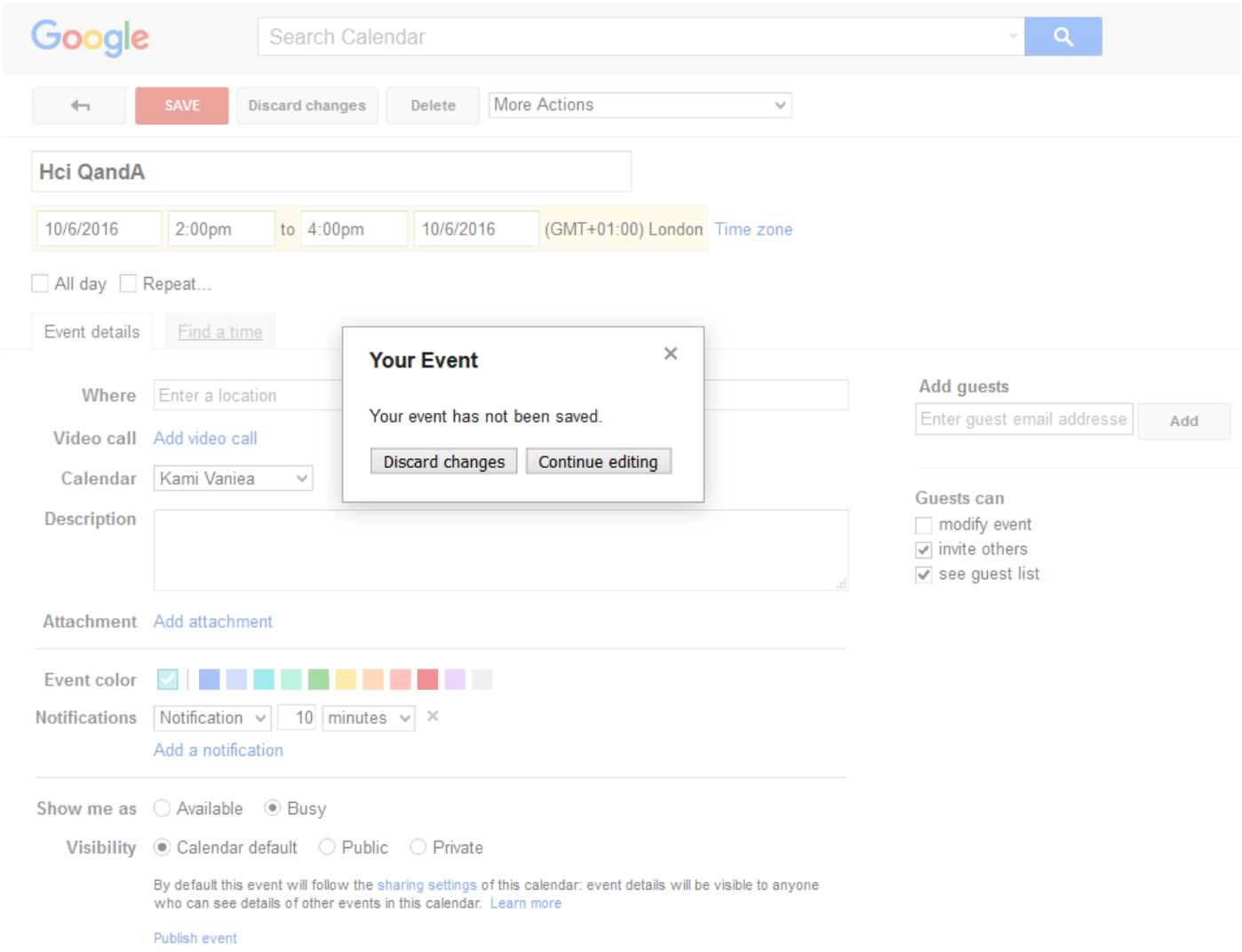
Better add a reminder or I might forget to go




Google
Search Calendar
SAVE
Discard changes
Delete
More Actions
Hci QandA
10/6/2016
2:00pm
to
4:00pm
10/6/2016
(GMT+01:00) London
Time zone
All day
Repeat...
Event details
Find a time
Where
Enter a location
Video call
Add video call
Calendar
Kami Vaniea
Description
Attachment
Add attachment
Event color
Notifications
Notification
10
minutes
Add a notification
Show me as
Available
Busy
Visibility
Calendar default
Public
Private
By default this event will follow the sharing settings of this calendar: event details will be visible to anyone who can see details of other events in this calendar. Learn more
Publish event
Add guests
Enter guest email addresse
Add
Guests can
modify event
invite others
see guest list
Is the reminder saved?

Visibility of system status

I click the back button without clicking “save” and get a warning

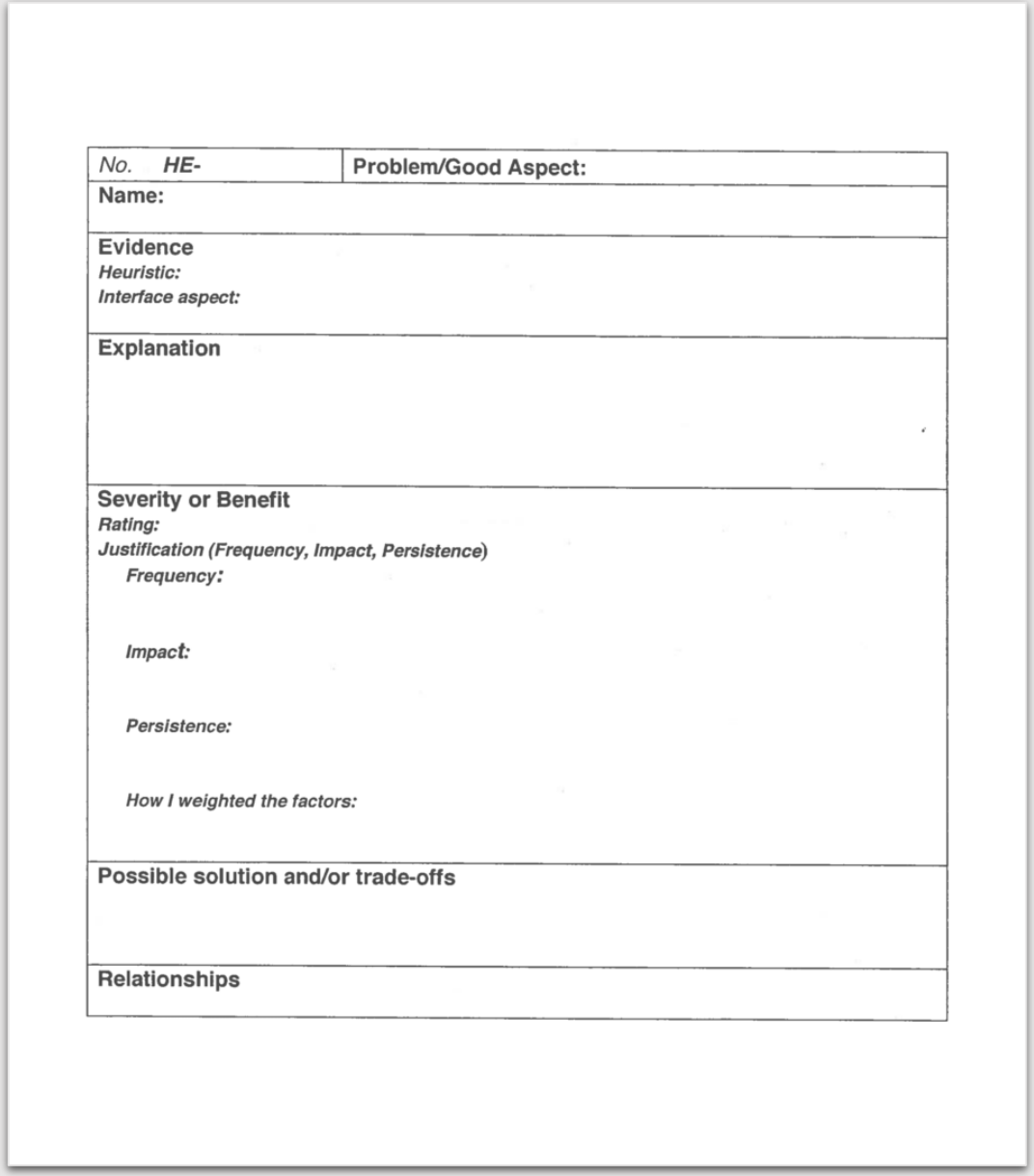


Think-pair-share

- Does Google Calendar violate the heuristic of "visibility of system status" for this task?

Usability Aspect Reports (UAR)

- Similar to a bug report, but for usability issues
- Can be a good or bad feature
- Should link to a heuristic



No. ***HE-***	**Problem/Good Aspect:**
Name:	
Evidence ***Heuristic:*** ***Interface aspect:***	
Explanation	
Severity or Benefit ***Rating:*** ***Justification (Frequency, Impact, Persistence)*** ***Frequency:*** ***Impact:*** ***Persistence:*** ***How I weighted the factors:***	
Possible solution and/or trade-offs	
Relationships	

3 stages for doing heuristic evaluation

- Briefing session to tell experts what to do.
- Evaluation period of 1-2 hours in which:
  - Each expert works separately;
  - Take one pass to get a feel for the product;
  - Take a second pass to focus on specific features.
- Debriefing session in which experts work together to prioritize problems.

No. of evaluators & problems

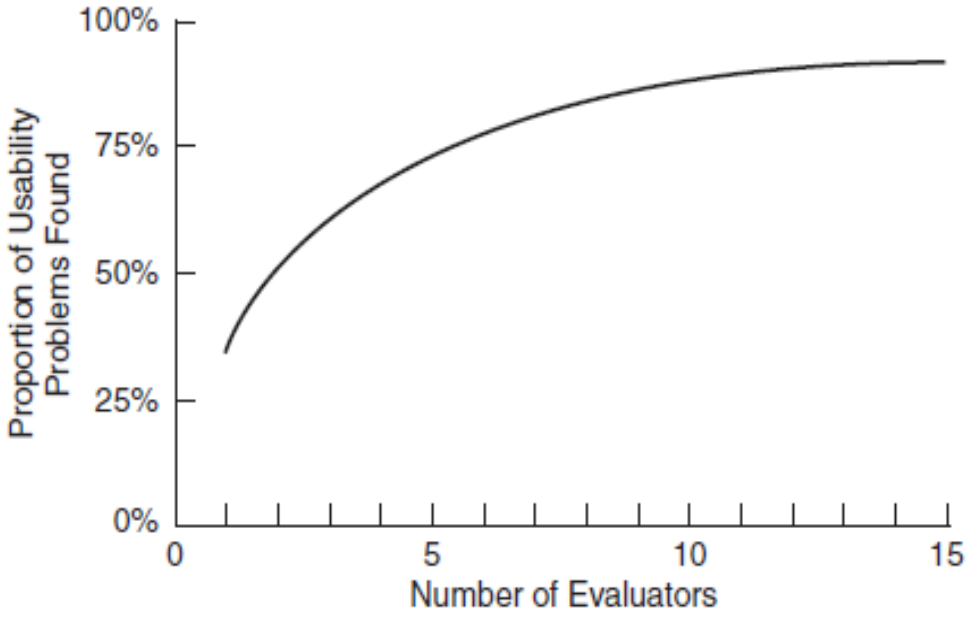


Figure 15.1 Curve showing the proportion of usability problems in an interface found by heuristic evaluation using various numbers of evaluators. The curve represents the average of six case studies of heuristic evaluation

Source: Usability Inspection Methods, J. Nielson & R.L. Mack ©1994. Reproduced with permission of John Wiley & Sons Inc.

Number of evaluators

- Nielsen suggests that on average 5 evaluators identify 75-80% of usability problems.
- Cockton and Woolrych (2001) point out that the number of users needed to find 75-80% of usability problems depends on the context and nature of the problems.

Heuristics for websites focus on key criteria (Budd, 2007)

- Clarity
- Minimize unnecessary complexity & cognitive load
- Provide users with context
- Promote positive & pleasurable user experience

Advantages and problems

- Few ethical & practical issues to consider because users not involved.
- Can be difficult & expensive to find experts.
- Best experts have knowledge of application domain & users.
- Biggest problems:
  - Important problems may get missed;
  - Many trivial problems are often identified;
  - Experts have biases.

Questions?