Privacy Decision Making in the Internet of Things

Bart P. Knijnenburg
@usabart
www.usabart.nl

Hello, I'm Bart
(with Disco)

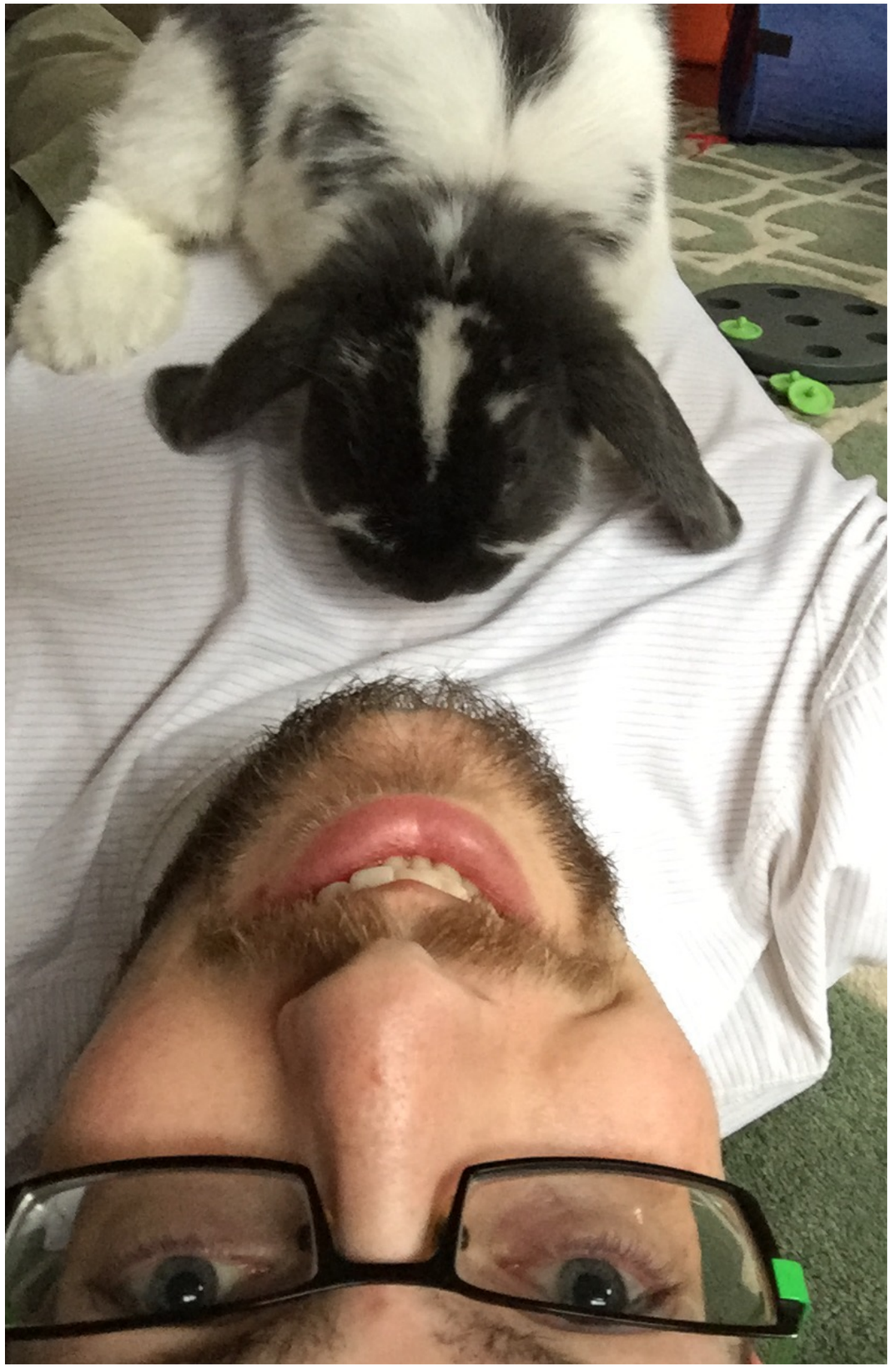
bartk@clemson.edu
www.usabart.nl
@usabart

Clemson University (Asst. Prof.)

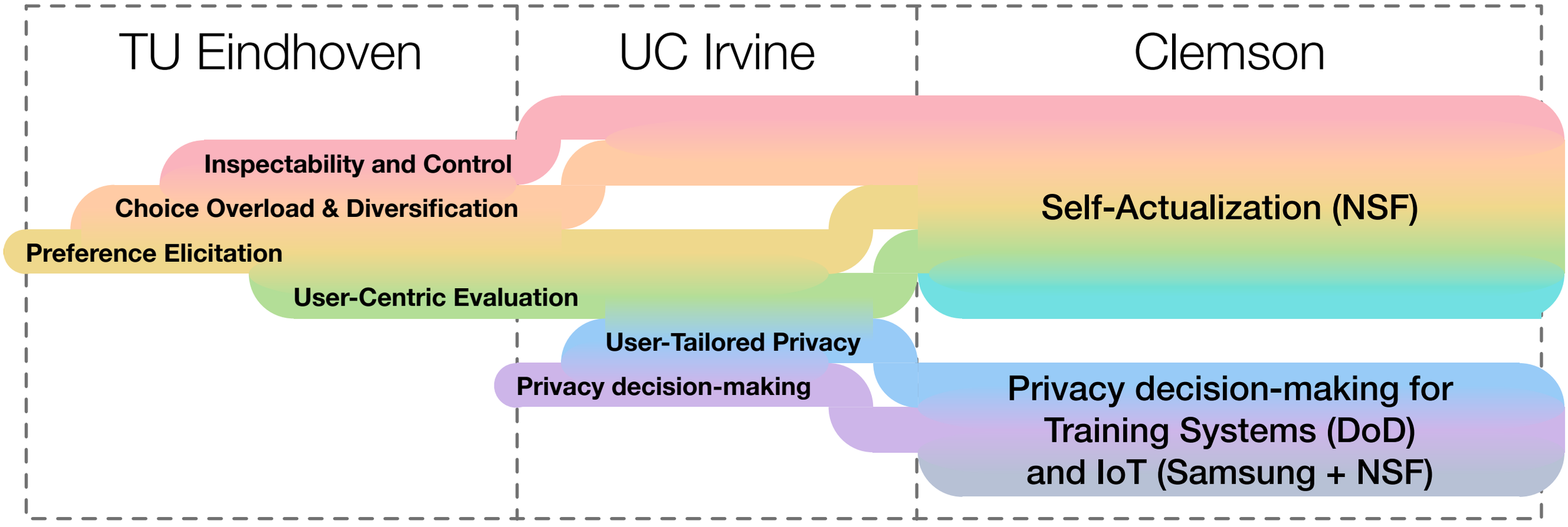
UC Irvine (PhD)

Carnegie Mellon University (M)

TU Eindhoven (BS + MS)



Past, Present, Future



Privacy is everywhere

BABY FIRST SALE

The one stop for all your baby's firsts.

$18 set
Save $5

Baby bear overalls and top set.
New baby-12 months.

20% OFF
kids' and babies' clothing, underwear and sleepwear (new baby-7)
See page 2 for exclusions.
Sorry, no rainchecks.

Also available in spot, striped, ladybird or dino.

$6
Save $3

50% OFF

$124
Save $125

The Hipod Milan convertible car seat.
- Features high side wings for superior side impact protection
- Converts from rear to forward facing.

Suitable from birth-4 years old.

No deposit lay-by
on kids' and babies' clothing (new baby-7) and nursery until 30 January. Not available online.

33% OFF

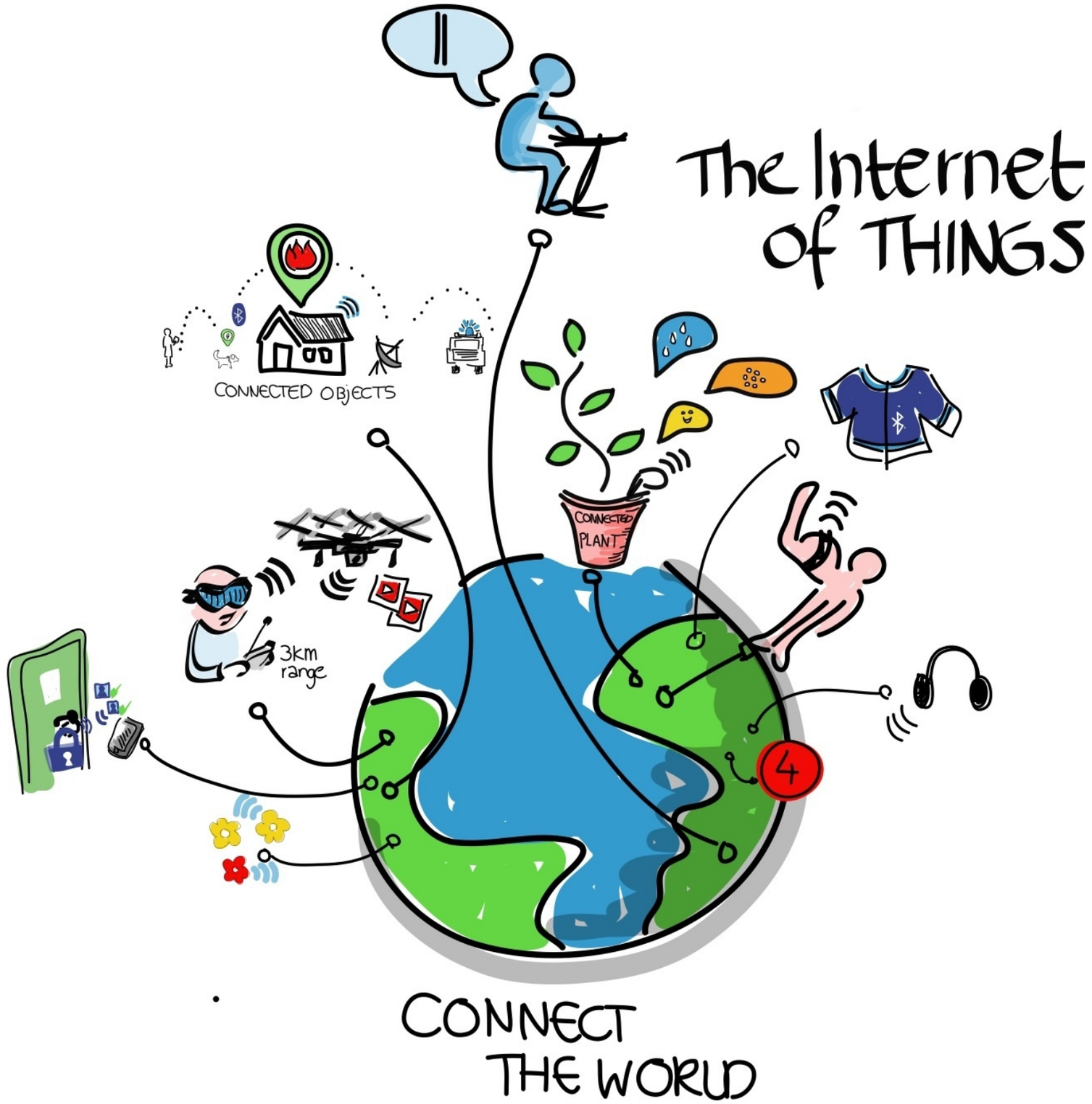
Babies' print coverall. Cupcake.
New baby-18 months.
- Snap-stud opening for easy dressing.

A huge range of kids' and babies' clothing (new baby-7) and nursery offers **ON SALE FOR 2 WEEKS** until Wednesday 30 January 2013.

NATE SILVER ON WHAT OBAMA SHOULD DO NEXT, P. 44
DON'T MOCK THE ARTISANAL PICKLE MAKERS, P. 14
A NANNY'S VIEW OF THE WORLD, P. 47
MANAGEMENT TIPS FROM 'DOWNTON ABBEY,' P. 52
GREECE CONFRONTS ITS SPARTAN FUTURE, P. 38
'It's the not doing it that's sexy.' Nina Arianda, P. 12

The New York Times Magazine
February 19, 2012

How your shopping habits reveal even the most personal information. By Charles Duhigg


The Internet of THINGS
CONNECTED OBJECTS
CONNECTED PLANT
3km range
4
CONNECT THE WORLD

Motivation

How can we **help users**

to **balance** the **benefits** and **risks** of information disclosure

in a **user-friendly** manner,

so that they can make **good** privacy **decisions**?

Notice and control, and privacy nudges

and why they don't work...

Notice and control

Privacy Calculus: People weigh the risks and benefits of disclosure

Prerequisites of the privacy calculus are:

- being able to **control** the decision;
- having adequate **information** about the decision.

Notice and control **empower** users to regulate their privacy at the desired level.

Why this doesn't work

Transparency paradox:

Simple privacy notices **aren't useful**, but detailed notices are **too complex**.

Control paradox:

Consumers claim to want full control over their data, but they **eschew the hassle** of actually exploiting this control!

Ironic effect of transparency:

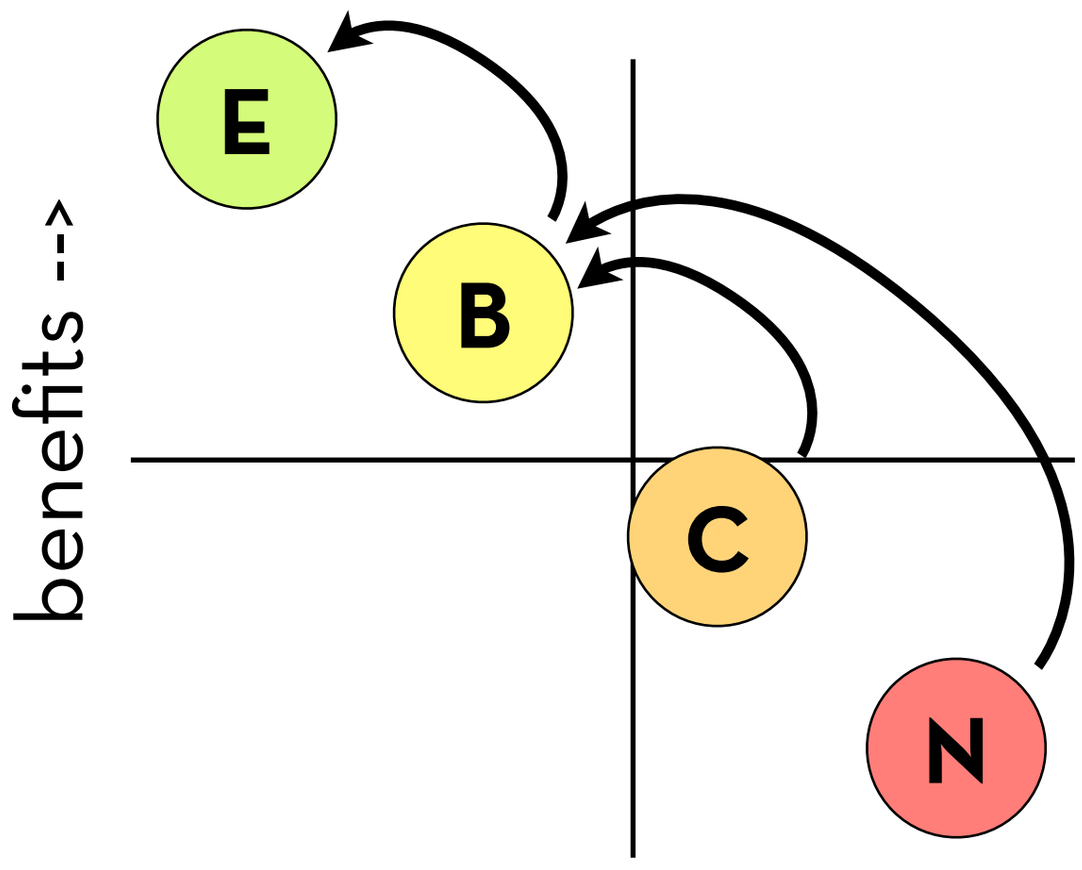
Informing users about privacy practices can make users **more wary** about their privacy.

Why this doesn't work

Decision biases:

Framing, defaults and decision context can influence disclosure.

Knijnenburg et al.,
CHI 2013



Expected:
Some will choose Exact instead of Block
Unexpected:
Sharing increases across the board!

Alternative: privacy nudges

Subtle yet **persuasive cues** that makes people more likely to decide in one direction or the other.

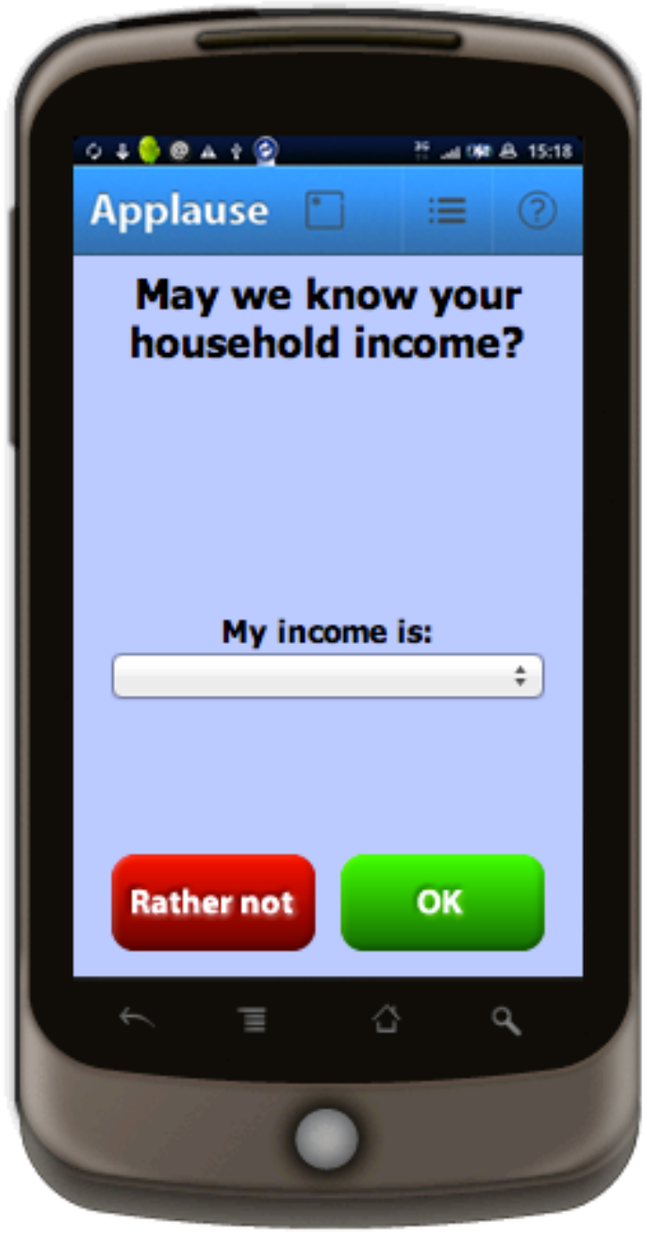
Examples of nudges:

- **Justification:** a succinct reason to disclose or not disclose a certain piece of information.
- **Default:** make the best action the easiest to perform.

Testing justifications

5 justification types

- None
- Useful for you
- Number of others
- Useful for others
- Explanation

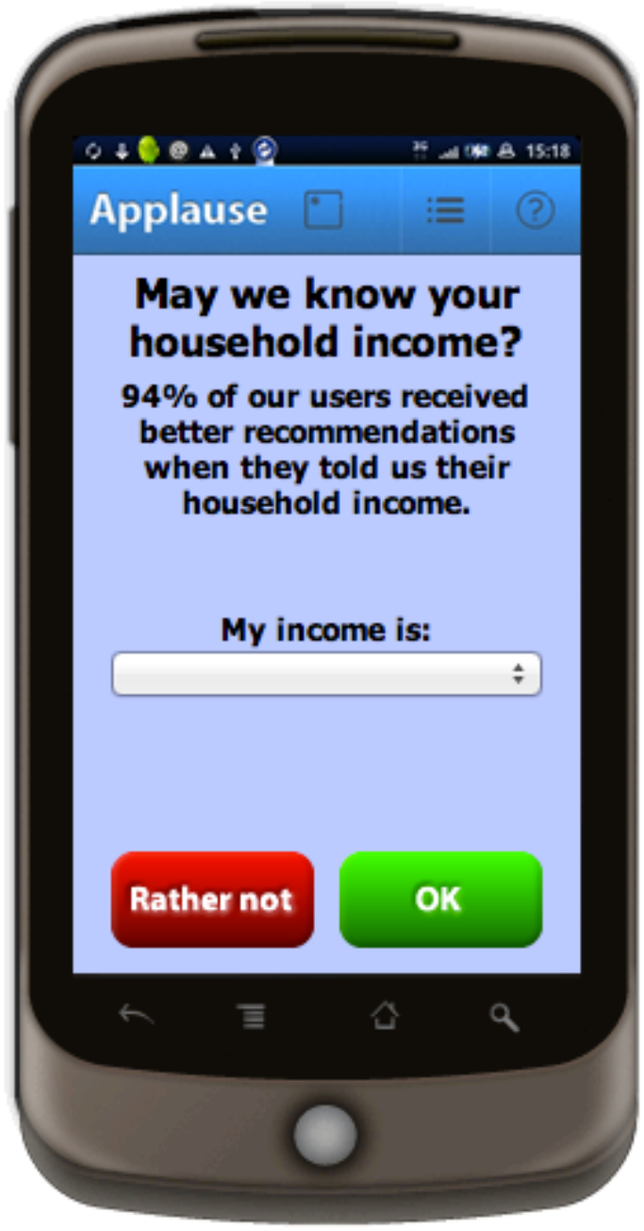


Knijnenburg and Kobsa, TiiS 2013

Testing justifications

5 justification types

- None
- Useful for you
- Number of others
- Useful for others
- Explanation

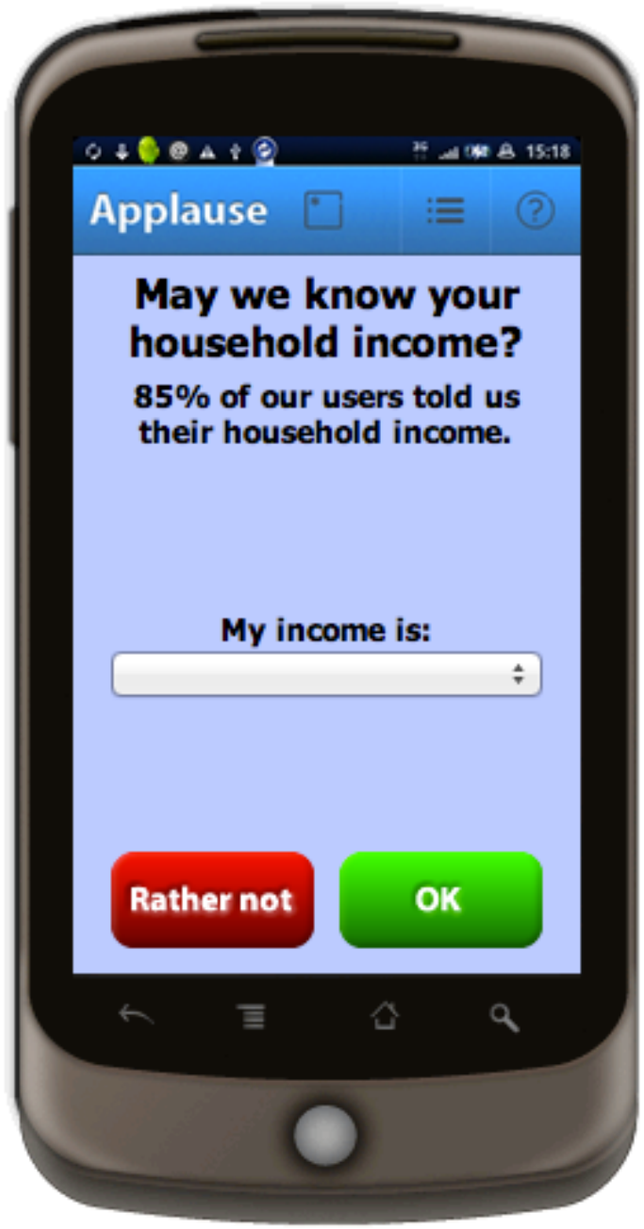


Knijnenburg and Kobsa, TiiS 2013

Testing justifications

5 justification types

- None
- Useful for you
- Number of others
- Useful for others
- Explanation

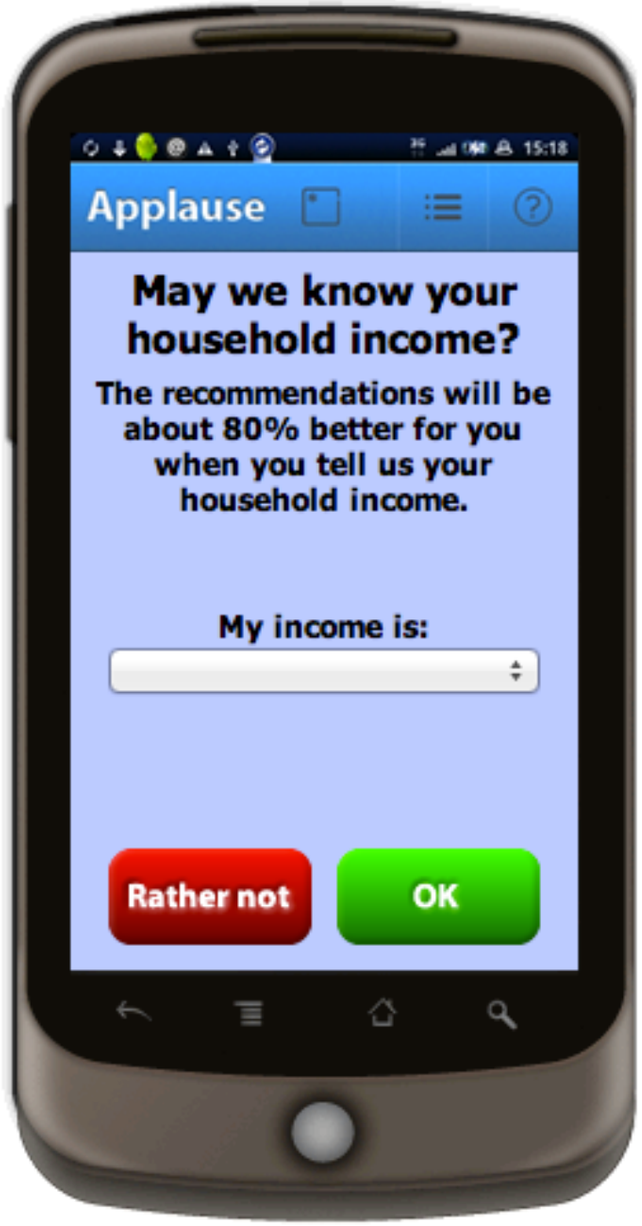


Knijnenburg and Kobsa, TiiS 2013

Testing justifications

5 justification types

- None
- Useful for you
- Number of others
- Useful for others
- Explanation

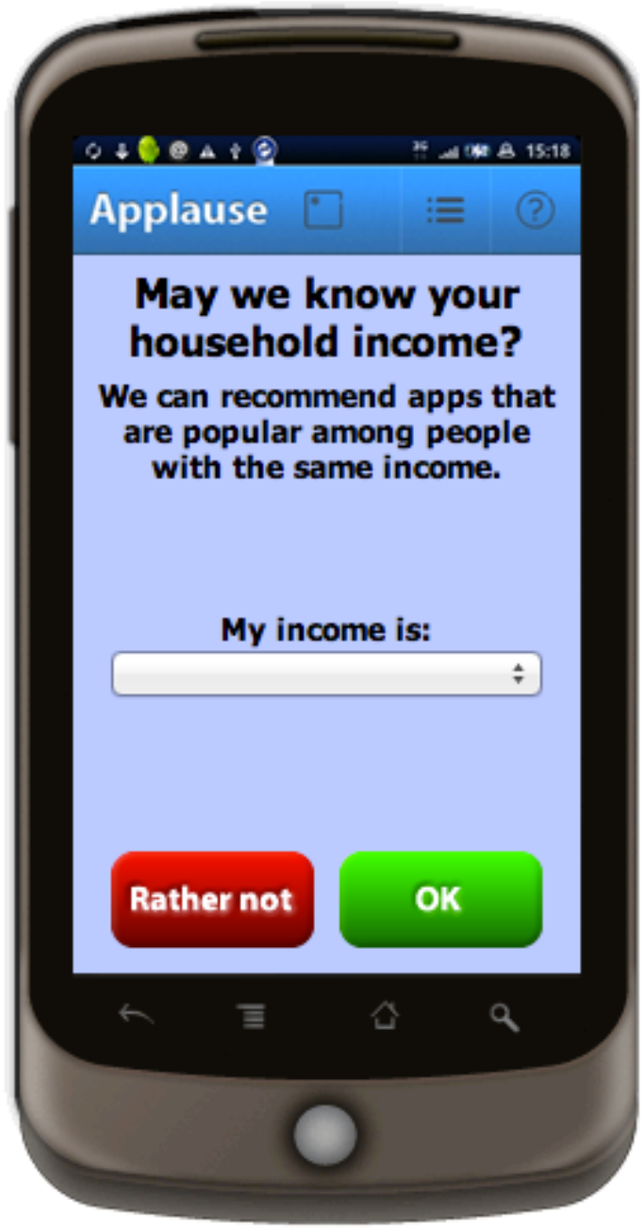


Knijnenburg and Kobsa, TiiS 2013

Testing justifications

5 justification types

- None
- Useful for you
- Number of others
- Useful for others
- Explanation



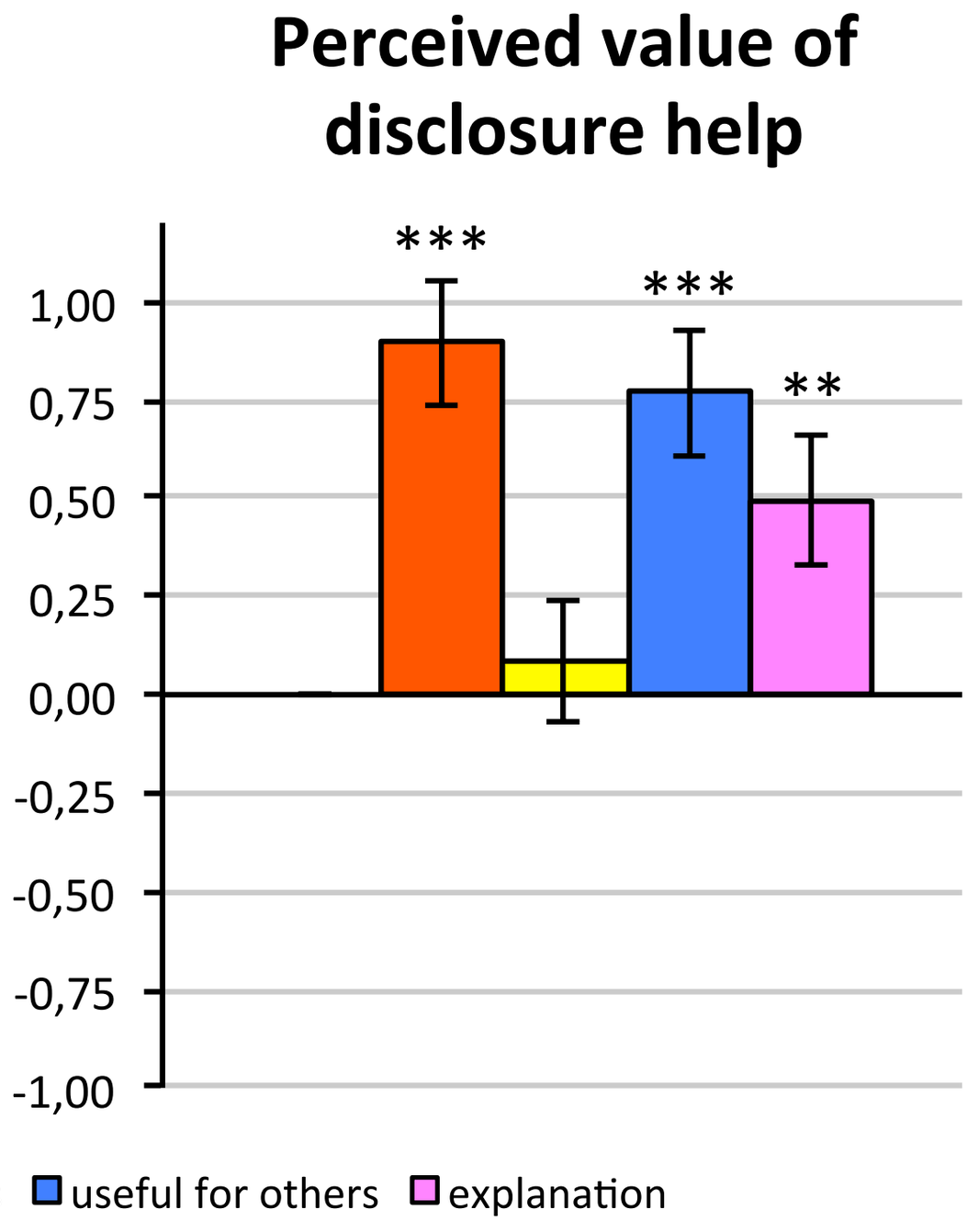
Knijnenburg and Kobsa, TiiS 2013

Results

Perceived value of disclosure help:

3 items, e.g. "The system helped me to make a tradeoff between privacy and usefulness"

Higher for all except "number of others"

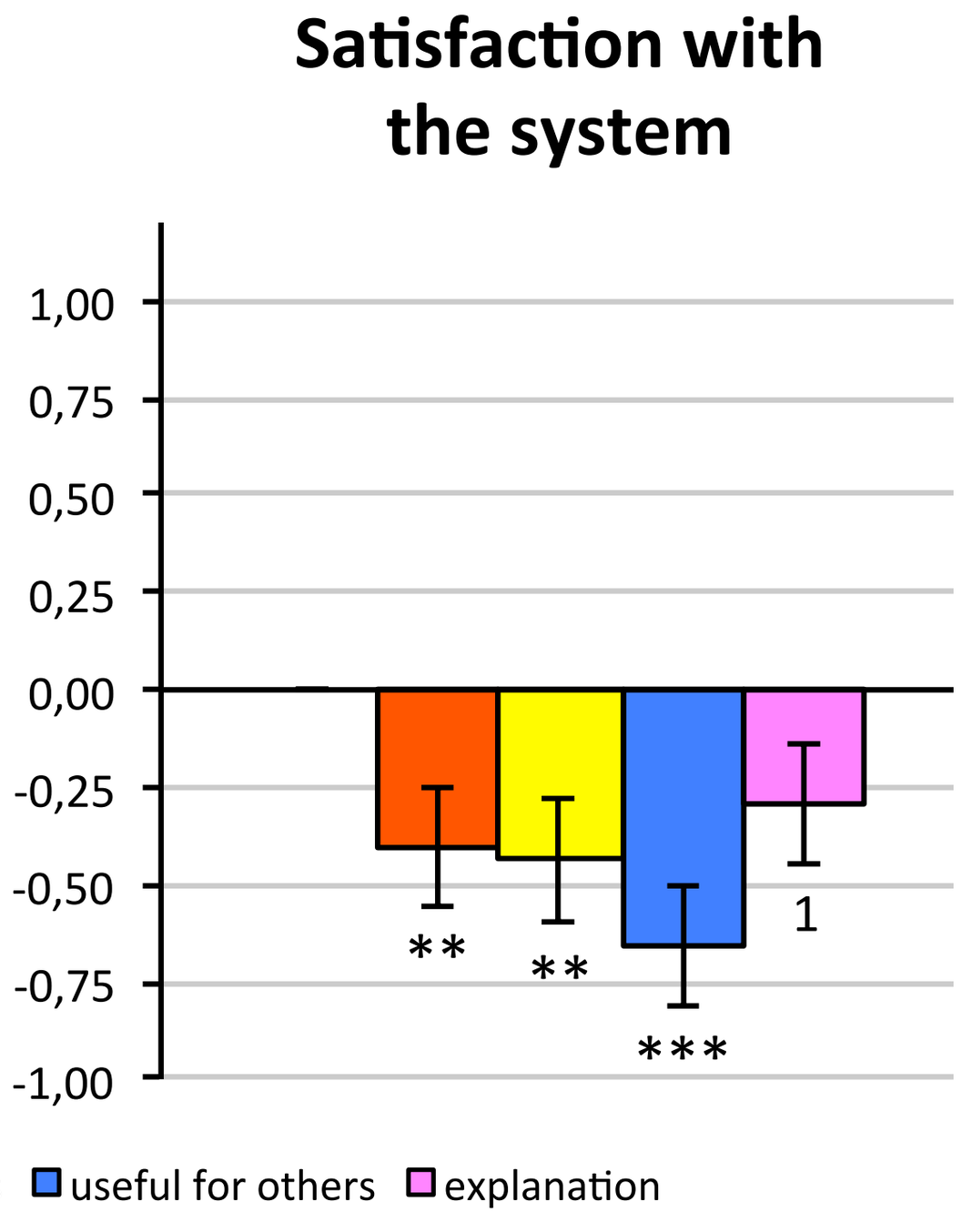


Results

Satisfaction with the system:

6 items, e.g. "Overall, I'm satisfied with the system"

Lower for any justification!



Why this doesn't work

What is the "**right**" **direction** of a nudge?

- **More disclosure:** better personalization, but some may feel tricked.
- **More private:** less threat, but harder to enjoy the benefits of disclosure.
- Going for the **average** (e.g. "smart default"): impossible, because people vary too much.

Some solutions

Before we move to IoT

Fixing transparency

Is this detailed enough?

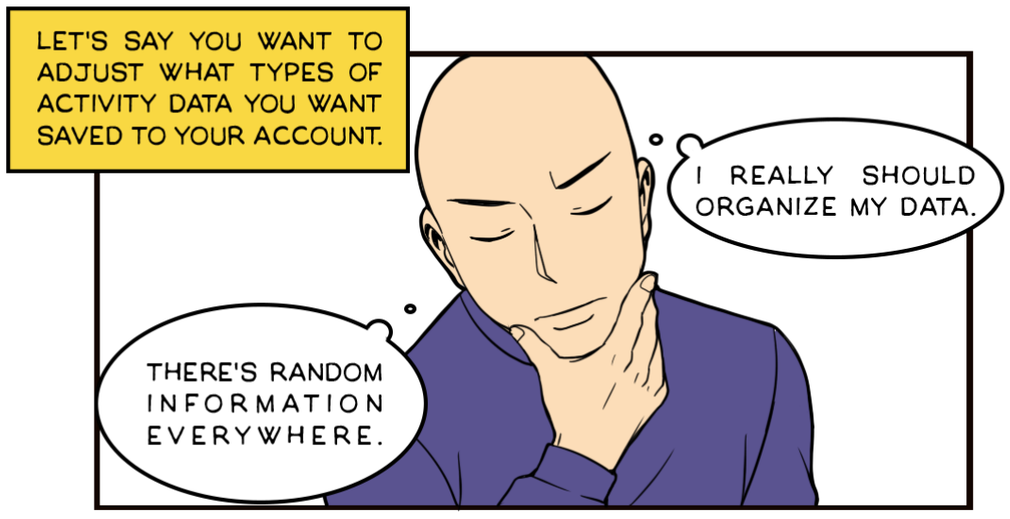
"Review and update your Google activity controls to decide what types of data, such as videos you've watched on YouTube or past searches, you would like saved with your account when you use Google services. You can also visit these controls to manage whether certain activity is stored in a cookie or similar technology on your device when you use our services while signed-out of your account."

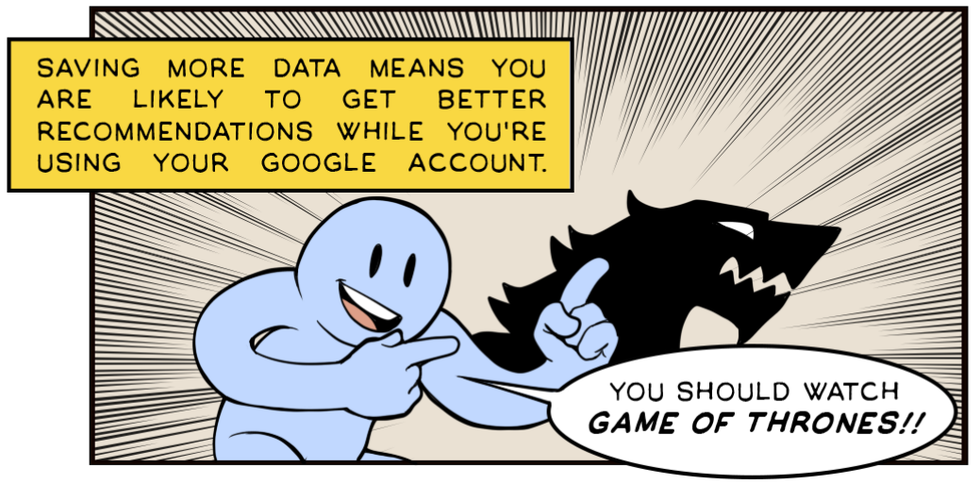
Fixing transparency

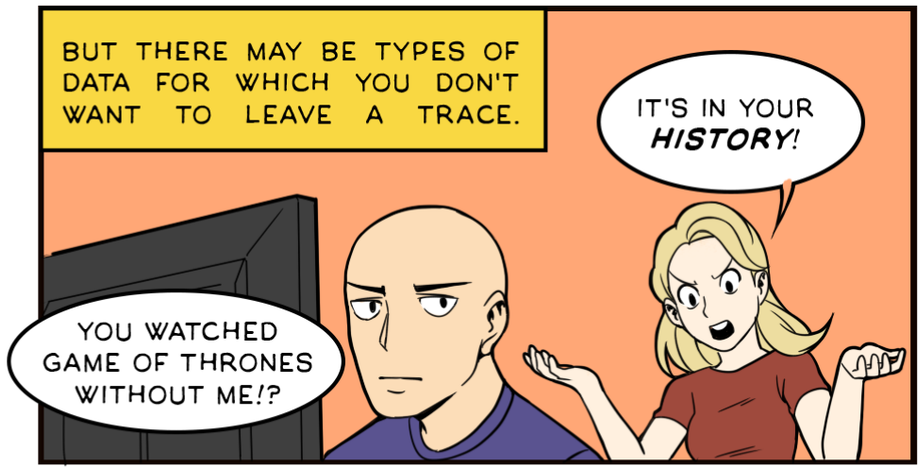
Would you read this?

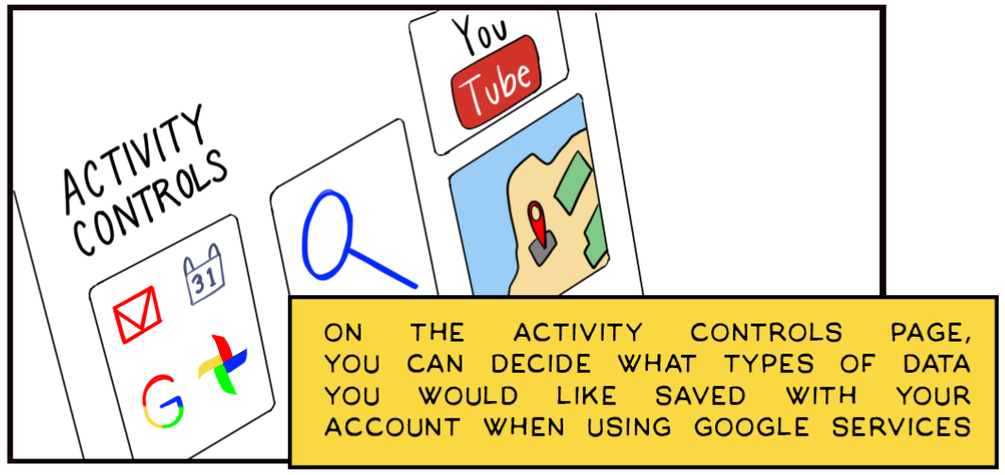
"Let's say you want to adjust what types of activity data you would like saved with your account. Saving more data means that you are likely to get better recommendations while you're using your Google account. But there may be types of data for which you don't want to leave a trace. On the Activity Controls page you can decide what types of data you would like saved with your account when you use Google services. These data include videos you've watched on Youtube, past searches in Google, places you go, and information from your Android devices. You can pause the tracking of this information (note that this does not delete any existing data!) or turn it back on (this may improve Google's services). You can also have a look at the actual activity, and delete items one by one. Finally, you can also manage whether certain activity is stored in a cookie or similar technology on your device when you use our services while signed-out of your login."

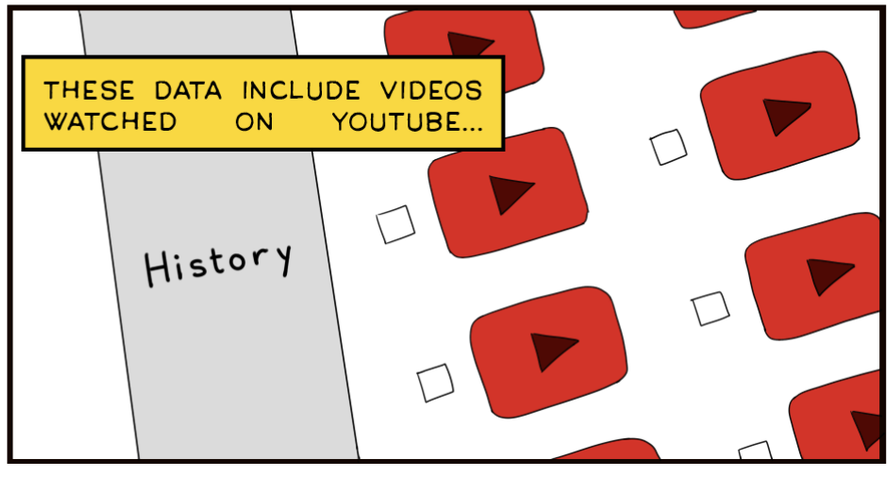
How about now?

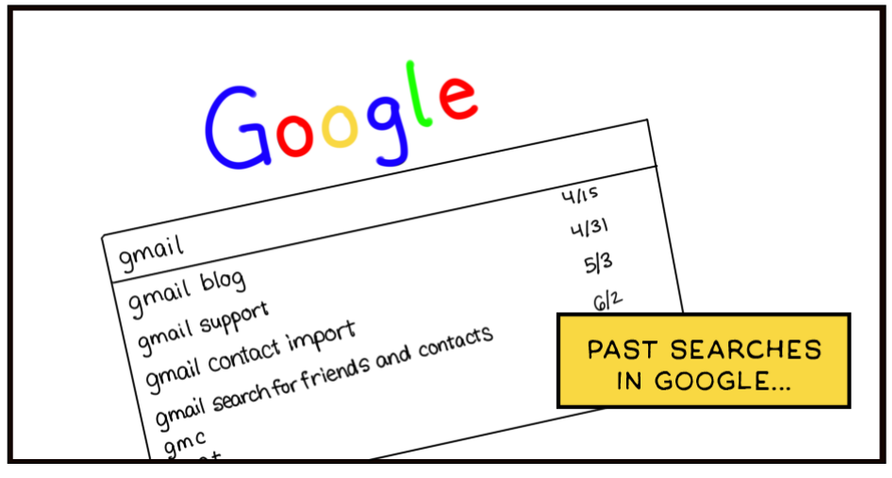












History

gmail
gmail blog
gmail support
gmail contact import
gmail search for friends and contacts
gmc

4/15
4/31
5/3
6/2

PAST SEARCHES IN GOOGLE...

PLACES YOU GO...

AND INFORMATION FROM YOUR ANDROID DEVICES.

YOUR INFO

YOU USE YOUR PHONE TOO MUCH!

YOU CAN PAUSE THE TRACKING OF THIS INFORMATION...

Google

Computer Science Program

(Will not be saved in history)

gmail blog
gmail support
gmail contact import
gmail search for friends and contacts

4/15
4/31
5/3
6/2
6/7

(NOTE THAT THIS DOES NOT DELETE ANY EXISTING DATA!)

...OR TURN IT BACK ON. THIS MAY IMPROVE GOOGLE'S SERVICES!

Google

Computer Science Program
Computer Science Program at Clemson
Computer Science Degree
Computers for Sale

8/15
8/8
10/3

Suggested:
Computer Science Program Requirements...
Computer Science Program Advice...

YOU CAN ALSO HAVE A LOOK AT THE ACTUAL ACTIVITY, AND DELETE ITEMS ONE BY ONE.

THEN SHE WON'T KNOW I SAW IT WITHOUT HER!

I'LL JUST DELETE GAME OF THRONES FROM MY HISTORY BEFORE MY GIRL-FRIEND COMES BACK!

FINALLY, EVEN WHILE SIGNED OUT OF YOUR ACCOUNT, WHEN USING OUR SERVICE...

THIS COOKIE STORES:
BROWSING HISTORY
LOGINS
PREVIOUSLY VISITED WEBSITES

...YOU CAN CHOOSE WHETHER ACTIVITY IS STORED IN A COOKIE OR SIMILAR TECHNOLOGY ON YOUR DEVICE.

Comics for privacy

Currently running a study testing comics vs. text at various levels of detail

Hypothesis: Comics can make privacy notices **inviting**, **engaging**, **comprehensible**, and **memorable**

Extra advantage for people with dyslexia or **low literacy** (target user group for a future study)

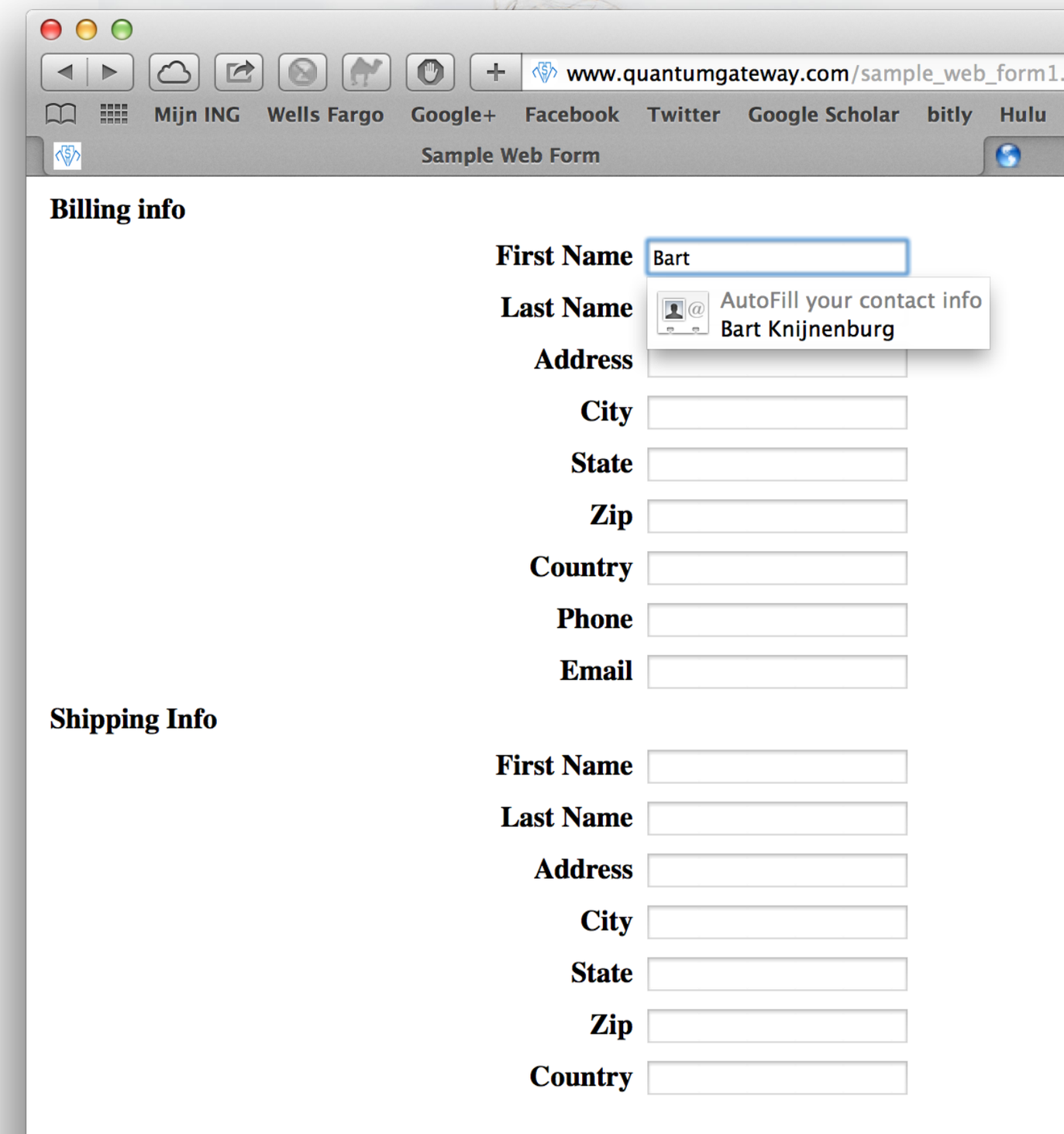
Knijnenburg and Cherry, SOUPS 2016

Fixing control

Modern browsers offer an **auto-completion feature** that reduces the effort of filling out web forms

These tools may cause users to **complete more fields** than they intended

They make it so easy to submit a fully completed form that users may **skip weighing benefits and risk**



FormFiller

Create a Profile

Please create your profile by entering your information below.
Note that FormFiller will store the information locally on your device, and only for the duration of this study. We will never submit any forms automatically or disclose this information to others without your active involvement.

About you:

First name: Last name:

Gender:

Age:

Address:

City: State: Zip:

E-mail:

About you:

First name: Last name:

Gender:

Age:

Address:

City: State: Zip:

E-mail:

Phone:

Tastes and Preferences:

Favorite movie:

Favorite band/artist:

Favorite food:

Favorite weekend pastime:

Last holiday location:

Political views:

Work and education:

Current/previous job:

Sector:

Employment status:

Work experience (yrs):

Income level:

Highest completed degree:

Computer skills

Health and lifestyle:

Overall health:

Dietary restrictions:

Number of doctor visits last month:

Weight (lbs):

Birth control usage (you or your partner):

Medical conditions:

- Diabetes
- Hypertension
- Respiratory (COPD etc.)
- High cholesterol

Study Procedures

FormFiller

Create a Profile

Please create your profile by entering your information below.
Note that FormFiller will store the information locally on your device, and only for the duration of this study. We will never submit any forms automatically or disclose this information to others without your active involvement.

About you:

First name: Last name:

Gender:

Age:

Address:

City: State: Zip:

E-mail:

Phone:

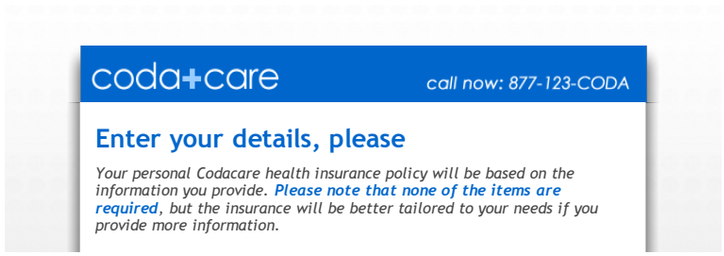
Tastes and Preferences:

Favorite movie:

Favorite band/artist:

Favorite food:

Favorite weekend pastime:



Fixing control

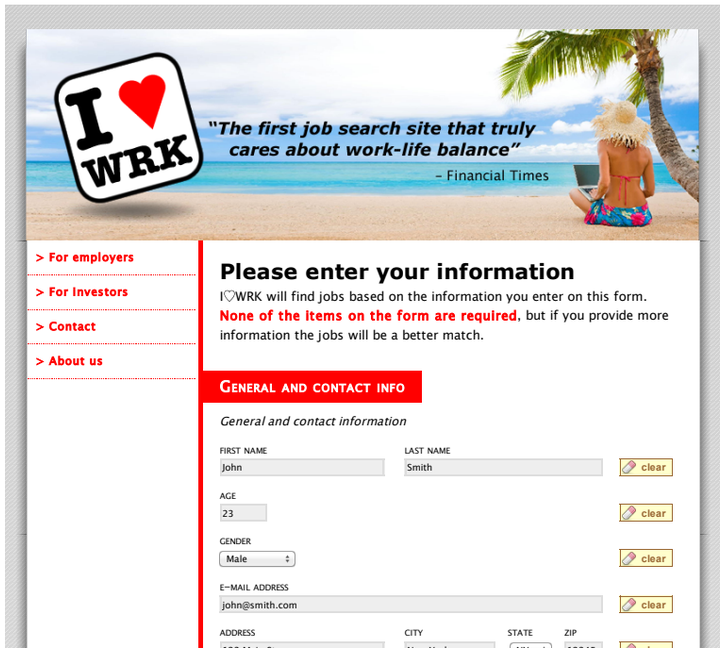
Each site corresponds to a particular type of info:

blogging community = personal interest items

job search website = job skills items

health insurer = health record items

They requested all the info, not just the relevant stuff!

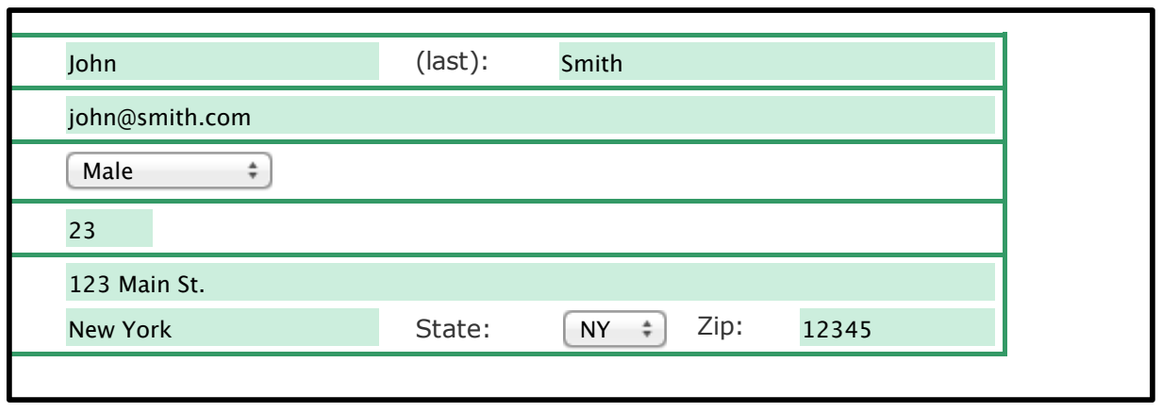


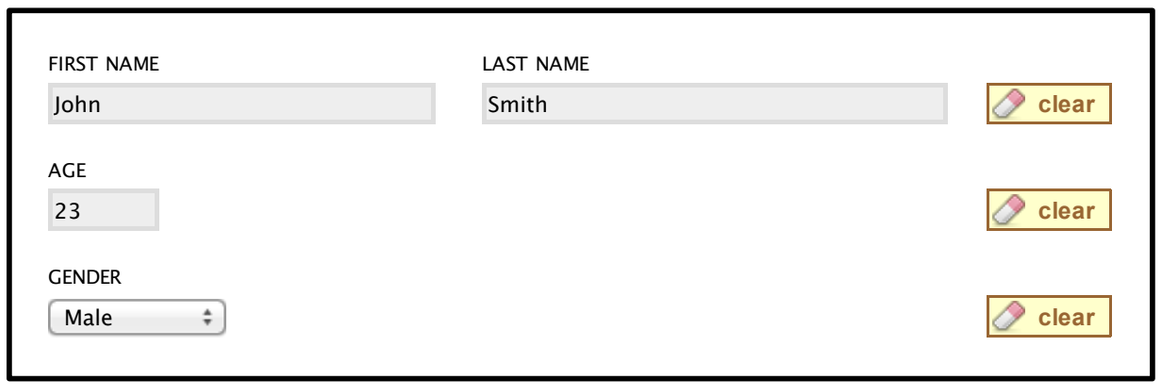
Fixing control

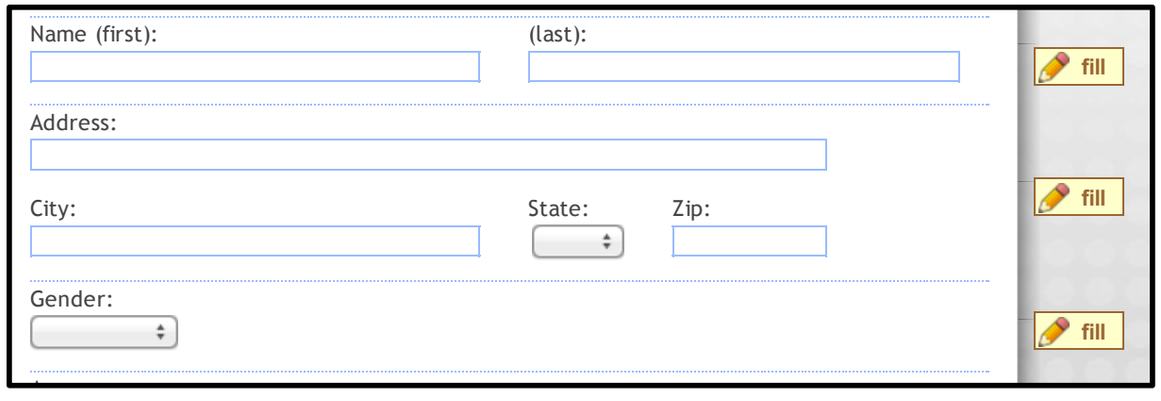
We introduce **two new efficacy-increasing designs**

We compare three tools:

- **Auto FormFiller:** auto-fills fields, users can remove manually
- **Remove FormFiller:** click to remove each field
- **Add FormFiller:** click to fill each field



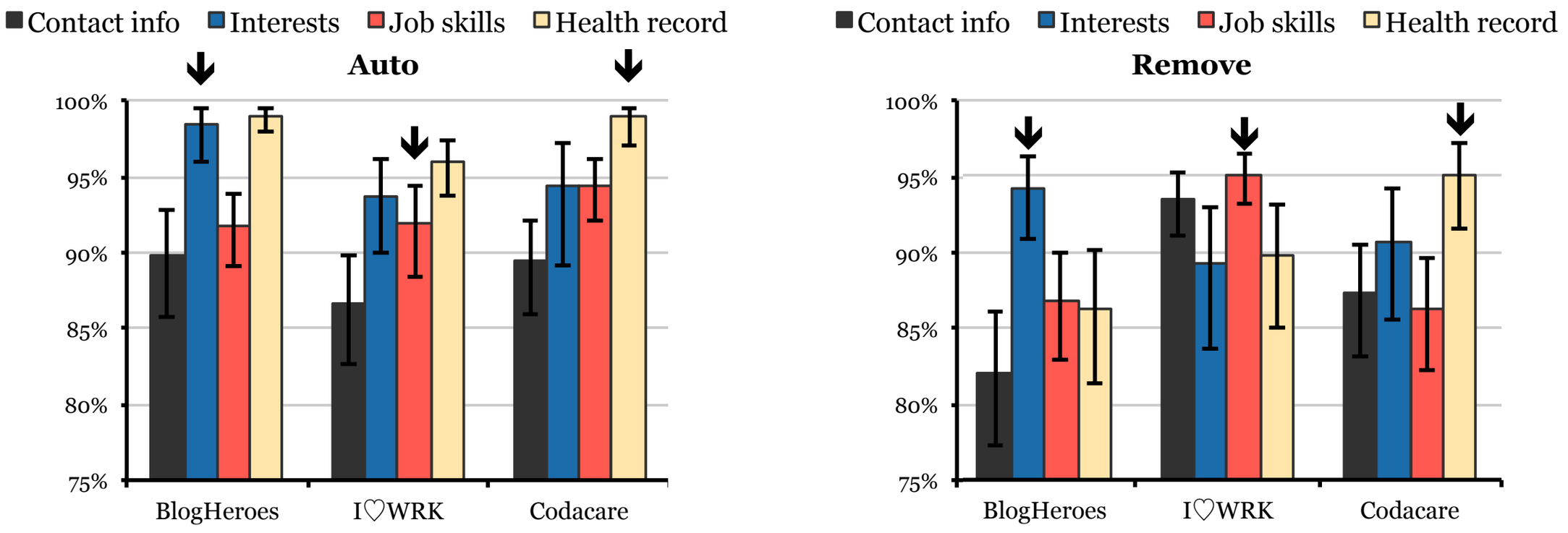




Fixing control

Disclosure was **not purpose-specific** for users of the **Auto FormFiller**

Disclosure was **purpose-specific** for users of the **Remove and Add FormFillers**.



Knijnenburg et al., ICIS 2013

Fixing nudges

Idea: **user-tailored privacy:** figure out what people want, then nudge them in that direction

Step 1: find determinants of privacy calculus

Characteristics of the **user**

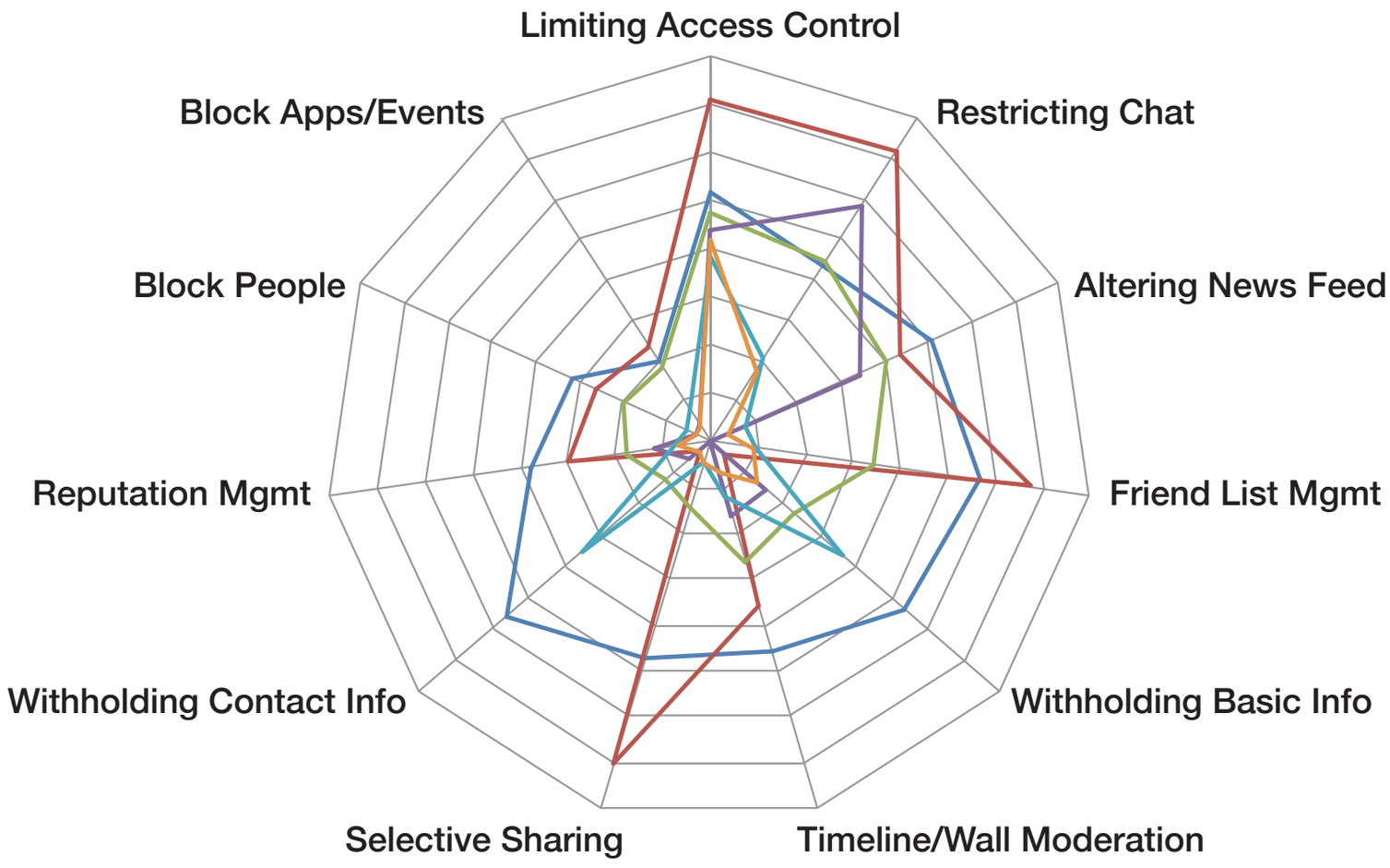
What **information** is being requested

The **recipient** of the information

Step 2: Adapt the nudge to the context

Privacy profiles

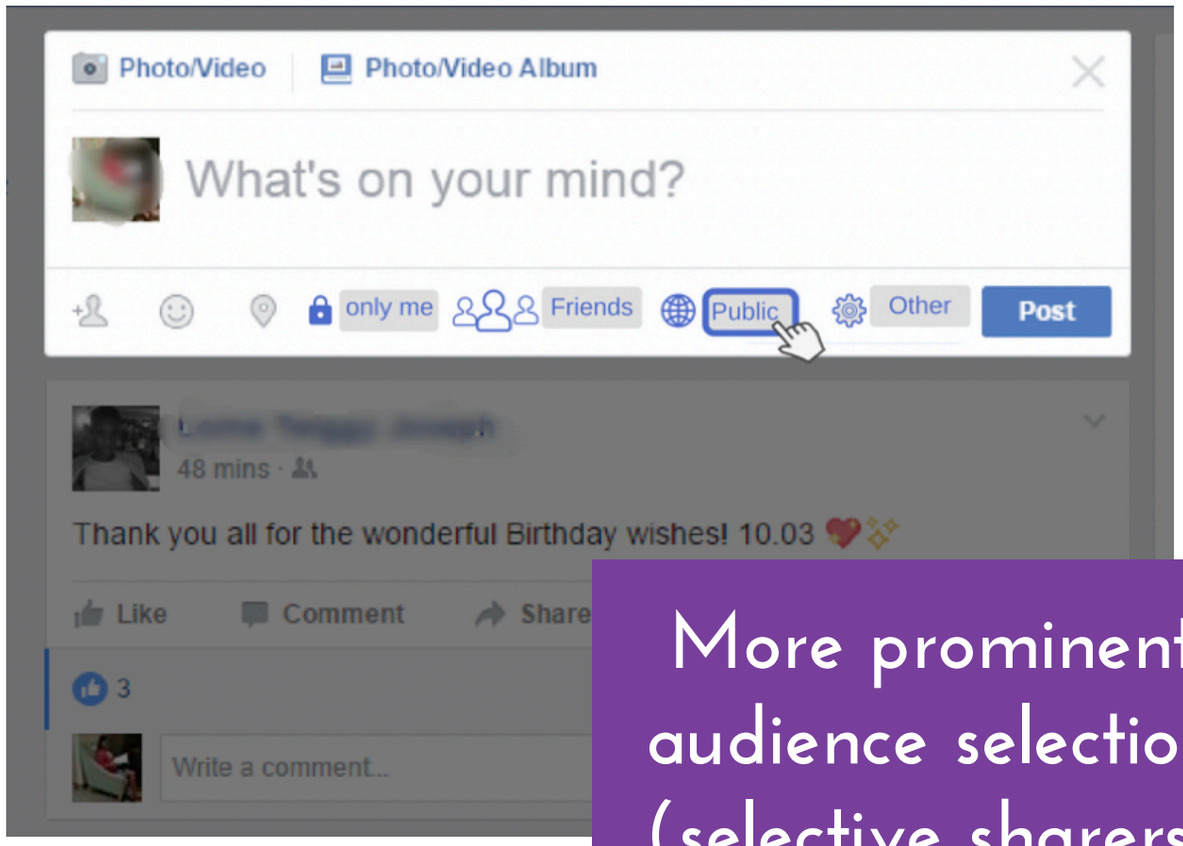
Privacy Maximizers | Selective Sharers | Privacy Balancers | Time Savers/Consumers | Self-Censors | Privacy Minimalists



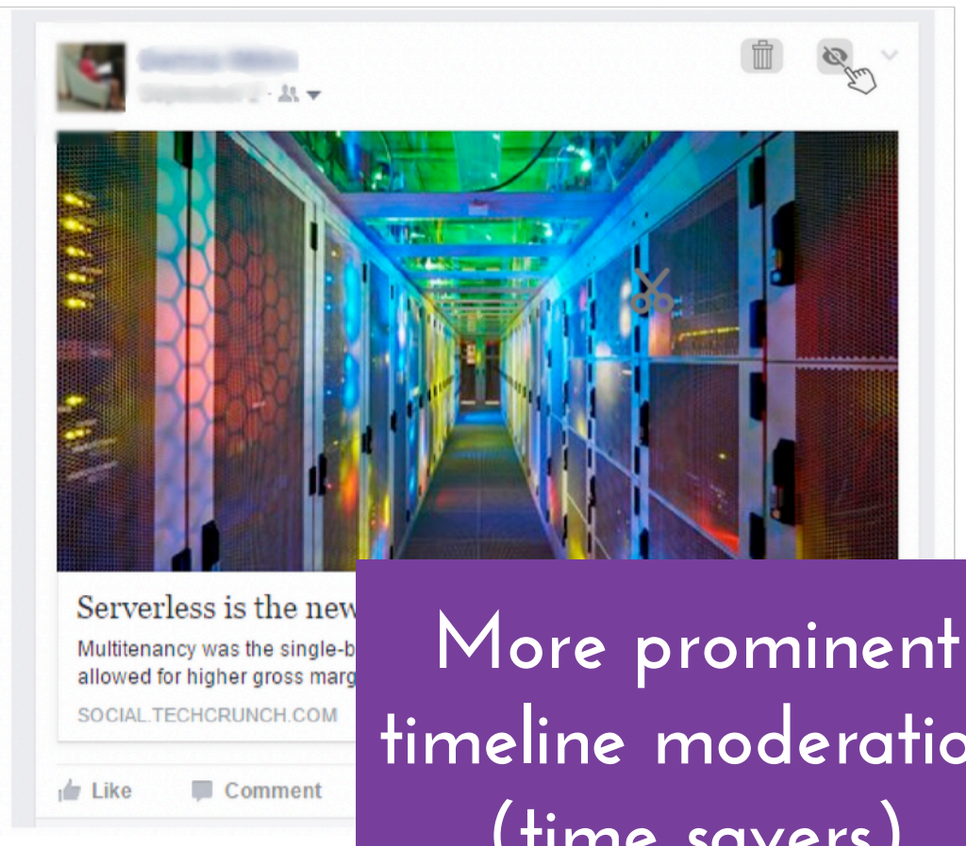
Winsniewski, Knijnenburg, and Lipford, IJHCS 2017

see www.usabart.nl/chart

Adapt to profile



More prominent audience selection (selective sharers)



More prominent timeline moderation (time savers)

Wilkinson et al., USEC 2017

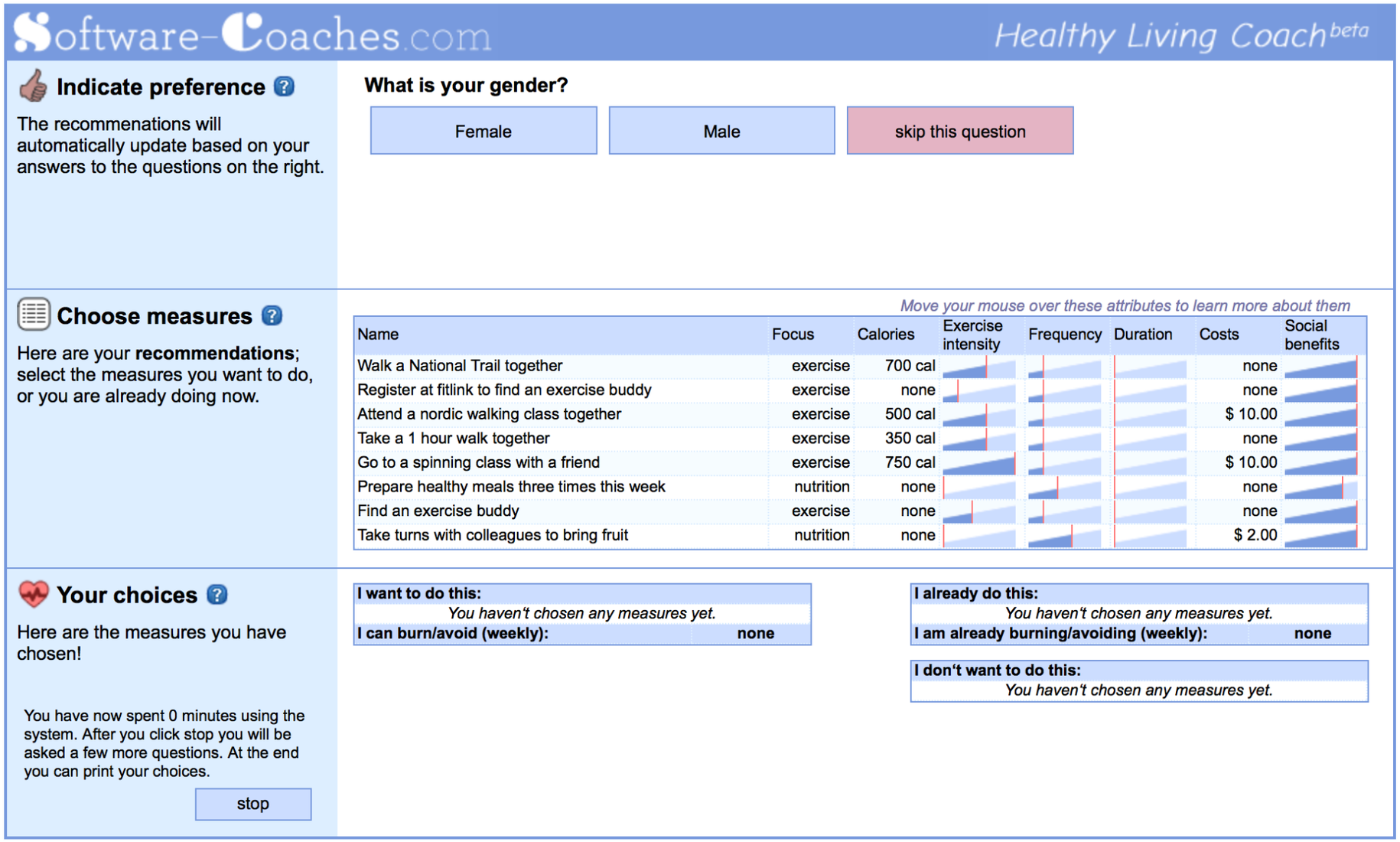
Adapt the nudge

Adaptive justifications:

What if we gave different types of users different types of justifications?

Knijnenburg and Kobsa, IUI 2013

Adaptive request order



Knijnenburg, dissertation 2015

Adaptive request order

$$u_o = \sum_{r_{oa}} \frac{v_r}{d_{an}} \quad \text{where} \quad d_{an} = abs\left(w_{an} - \bar{w}_n\right) + .0001$$

$$p_{ni} = \frac{e^{\beta_n - \delta_i}}{1 + e^{\beta_n - \delta_i}}$$

Useful-ness

$$r_i = \begin{cases} u_i & if \quad \delta_i < \alpha, \\ -\delta_i & if \quad \delta_i > \alpha. \end{cases}$$

Sensi-tivity

Trade-off

Tendency

$$\beta_n = \text{mean}_n(\delta) + \sqrt{1 + \text{var}_n(\delta)/2.9} * \ln\left(\frac{|D_n|}{|L_n| - |D_n|}\right) \quad \text{and} \quad \alpha_n^H = \beta_n - 1.5$$

The Internet of Things

...and our solution to its privacy problems

Iot privacy decisions

Study: 2,800 public IoT-related scenarios + decisions from 200 participants

Manipulate scenarios along 5 dimensions

Example scenario: "A device of a friend (who) records your video to detect your presence (what). This happens continuously (when), while you are at someone else's place (where), for your safety (why)."

Questions

If this situation happens, would you like to be notified about it? If Yes, just once or every time it happens?

If you had a choice to allow/reject, what would you choose? Just this time or always?

How comfortable/uncomfortable do you feel about this scenario?

How risky or safe is this situation?

How appropriate do you consider this situation?

Predict allow/reject

Let's say we create a settings interface...

How often are we correct if our default setting is always "yes"?

Only 28% of the time!

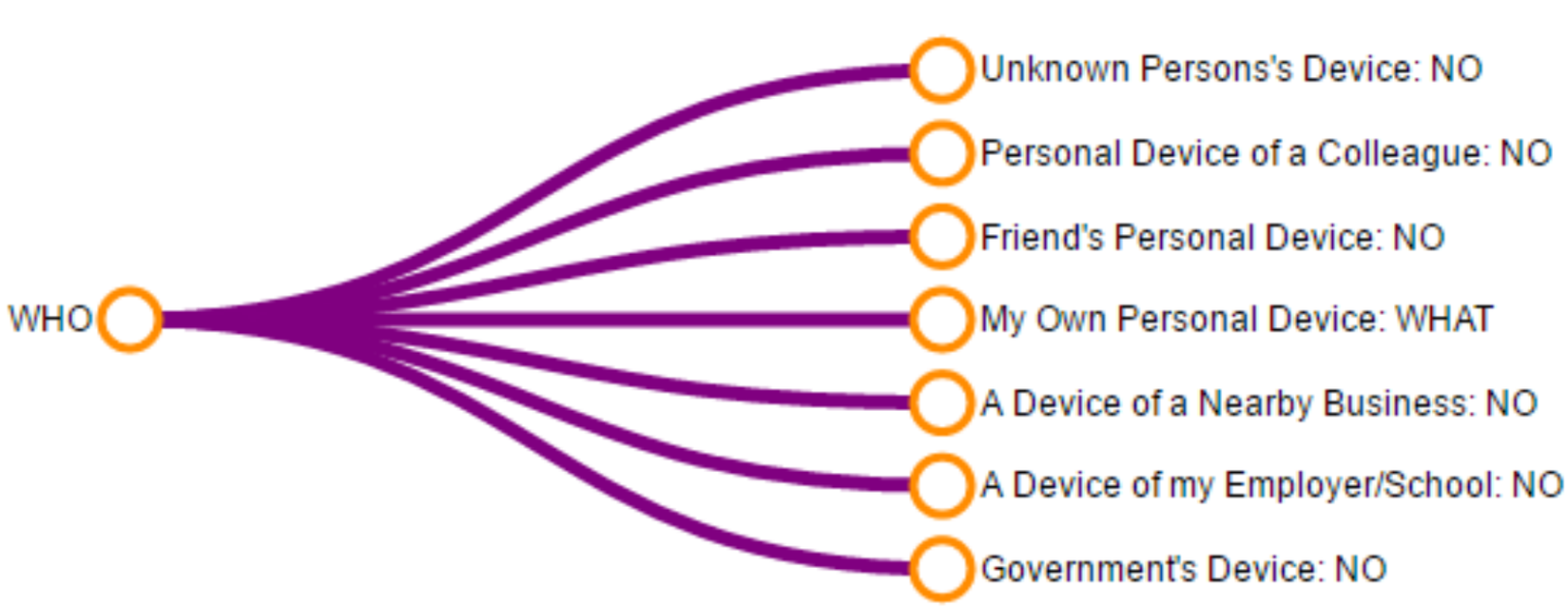
What if our default setting is always "no"?

72% of the time, but this is kind of useless...

Can we do better than this?

Machine learning

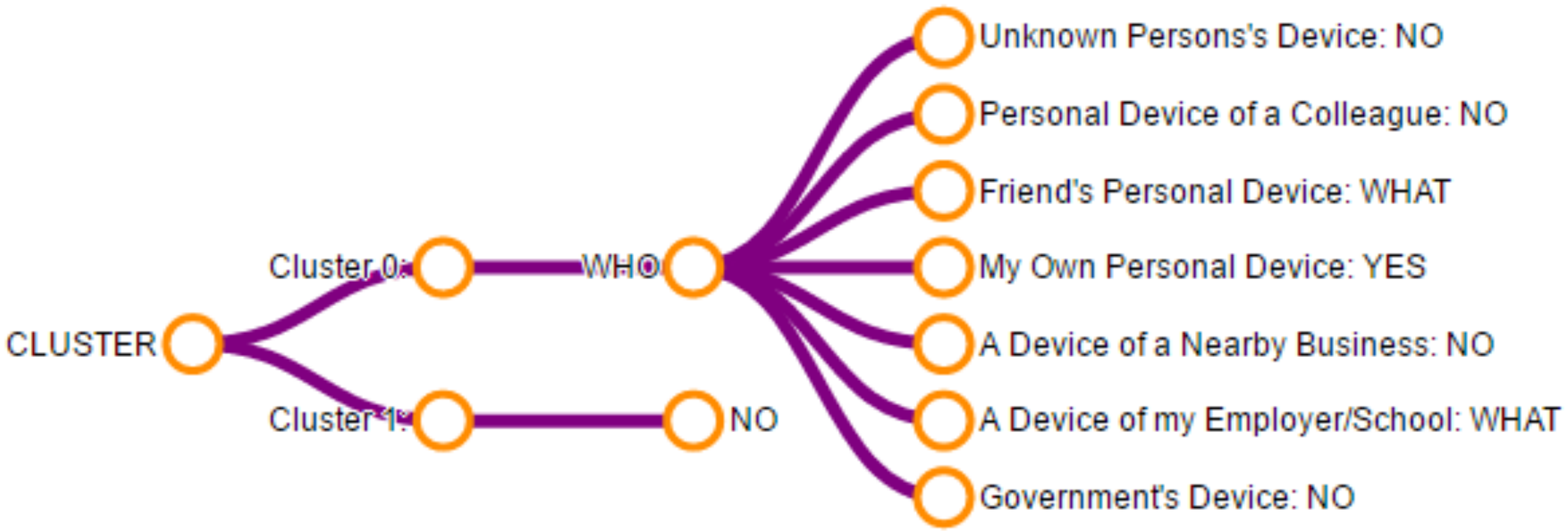
What if we make our best guess, given the 5 parameters (who, what, where, when, why)?



Correct 75% of the time!

Advanced ML

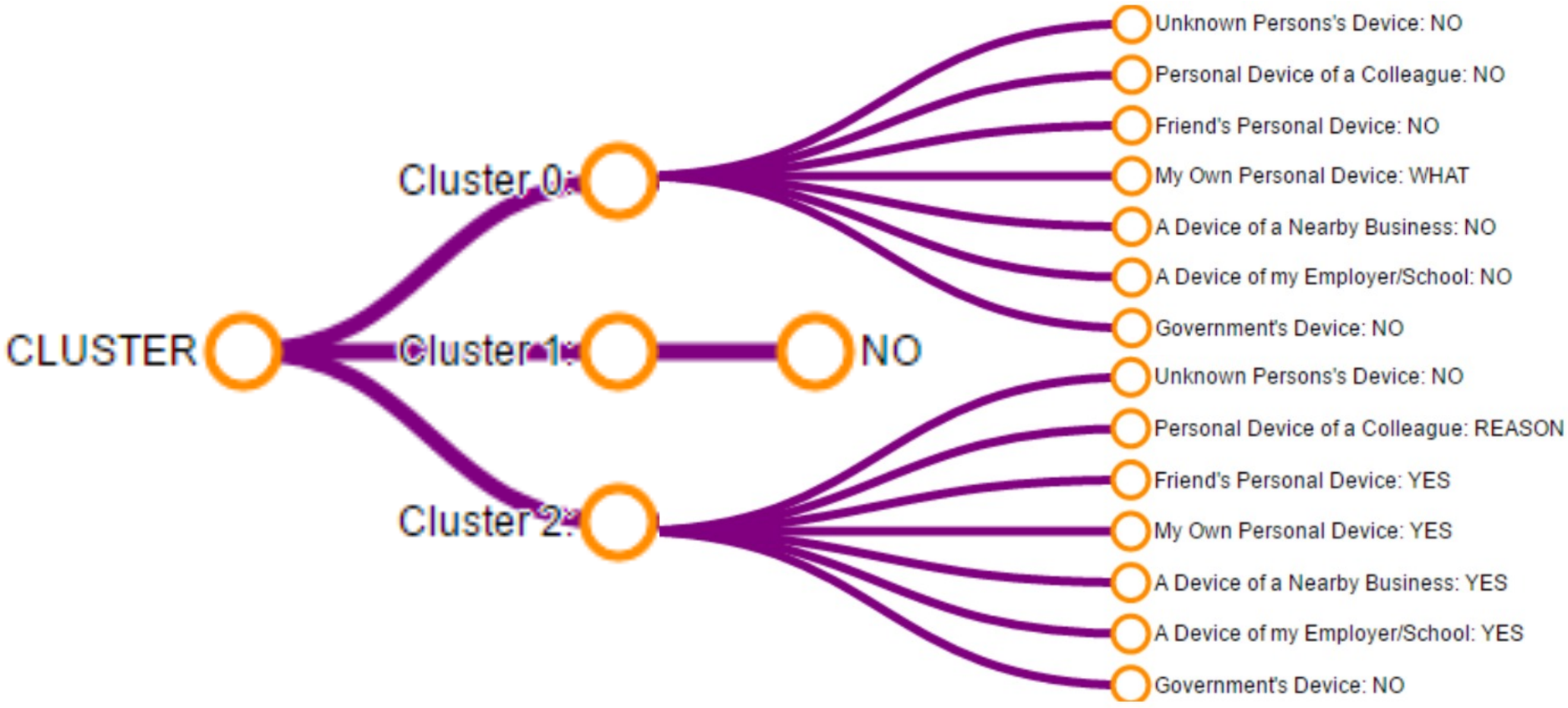
What if we first divide our participants based on their overall attitude towards the scenarios?



Correct 77% of the time!

Advanced ML

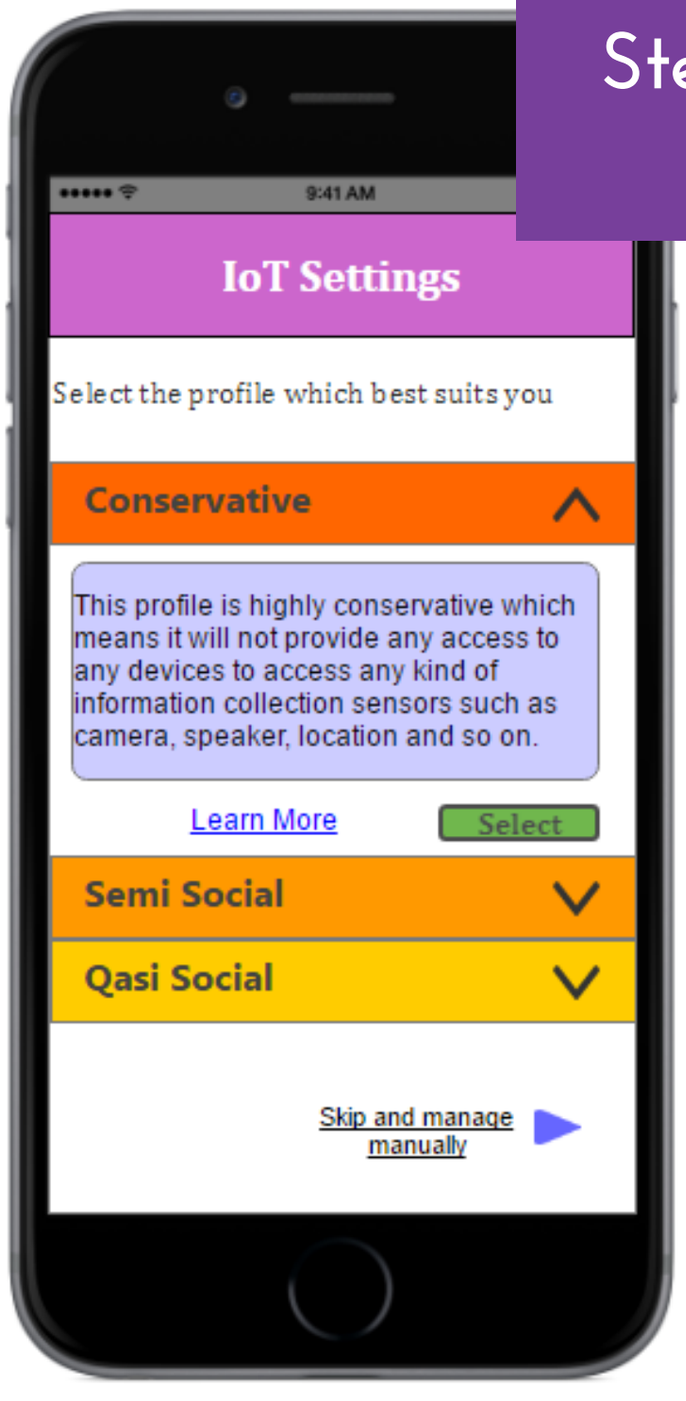
What if we divide our participants on the fly?



Correct 82% of the time!

User Interface

Step 1: choose a profile



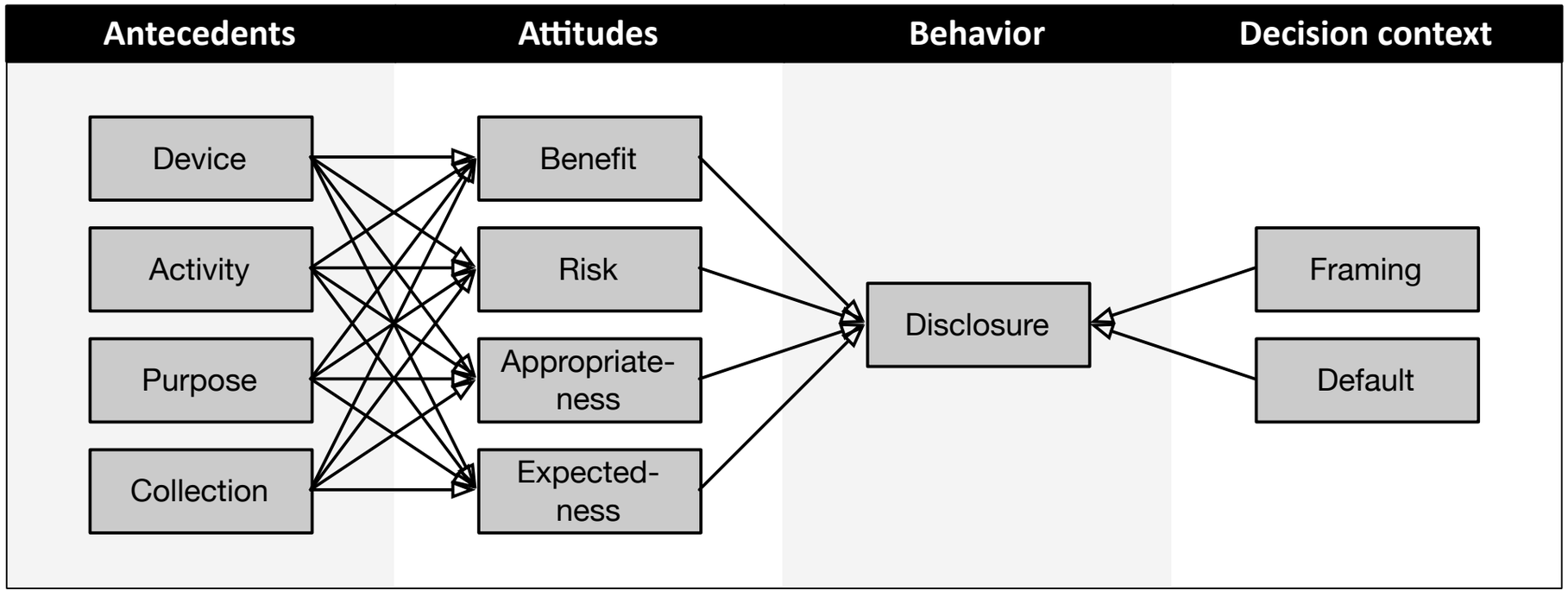
Step 2: adjust the default settings

Future work: household IoT

Scenarios (8x12x4x12 mixed fractional factorial design)			
Device	**Activity**	**Purpose**	**Collection**
Your smart home security system	uses information from your smart home security system[1]	to detect your presence in the house.	The data is not stored.
Your smart refrigerator	uses information from your smart refrigerator	to detect where you are in the house.	The data is stored locally and used to optimize the service.
Your smart HVAC system	uses information from your smart HVAC system	to automate its operations.	The data is stored locally and used to give you insight into your behavior.
Your smart washing machine	uses information from your smart washing machine	to give you timely alerts.	The data is stored locally and used to recommend you other [brand] services.
Your smart lighting system	uses information from your smart lighting system		The data is stored on [brand] servers and used to optimize the service.
Your smart microwave	uses information from your smart microwave		The data is stored on [brand] servers and used to give you insight into your behavior.
Your smart TV	uses information from your smart TV		The data is stored on [brand] servers and used to recommend you other [brand] services.
Your smart alarm clock	uses information from your smart alarm clock		The data is stored on [brand] servers and sold to advertisers.
	uses a location sensor		The data is stored in the cloud and used to optimize the service.
	uses a camera		The data is stored in the cloud and used to give you insight into your behavior.
	uses a microphone		The data is stored in the cloud and used to recommend you other [brand] services.
	connects to your phone/watch		The data is stored in the cloud and sold to advertisers.

Table-1 – Scenarios are generated by selecting one row from each column.

Future work: household IoT



Study mechanisms

Multi-level structural modeling

More machine learning

Eye tracking

Live test with prototype interface

General conclusion

What have we learned?

Societal impact

User-tailored privacy:

Relieves some of the burden of controlling privacy, while at the same time respecting each individual's preferences

Provides **realistic empowerment**: the right amount of transparency and the right amount of control

Refrains from making moral judgments about what the "right" level of privacy should be

The best way forward to support people's privacy decisions!

Thanks!

Mentioned papers

Knijnenburg et al, CHI 2013: Preference-based Location Sharing: Are More Privacy Options Really Better?

Knijnenburg and Kobsa, TiiS 2013: Making Decisions about Privacy: Information Disclosure in Context-Aware Recommender Systems.

Knijnenburg and Cherry, SOUPS 2016: Comics as a Medium for Privacy Notices.

Knijnenburg et al., ICIS 2013: Counteracting the Negative Effect of Form Auto-completion on the Privacy Calculus.

Wisniewski et al., IJHCS 2017: Making privacy personal: Characterizing social network users by their privacy proficiency and management strategies.

Wilkinson et al., USEC 2017: User-Tailored Privacy by Design.

Knijnenburg and Kobsa, IUI 2013: Helping Users with Information Disclosure Decisions: Potential for Adaptation.

Knijnenburg, dissertation 2015: A user-tailored approach to privacy decision support.