



Designing and testing a method for scaling UWV's CSat\* measurement  
by combining multiple linear regression and factor analysis

*\* 'CSat = Customer Satisfaction, as measured bi-annually by Ipsos*

# Context

UWV coordinates the distribution of almost 20 billion euro spread over 600.000 clients annually.  
All forms of interaction are (becoming) digital-first.

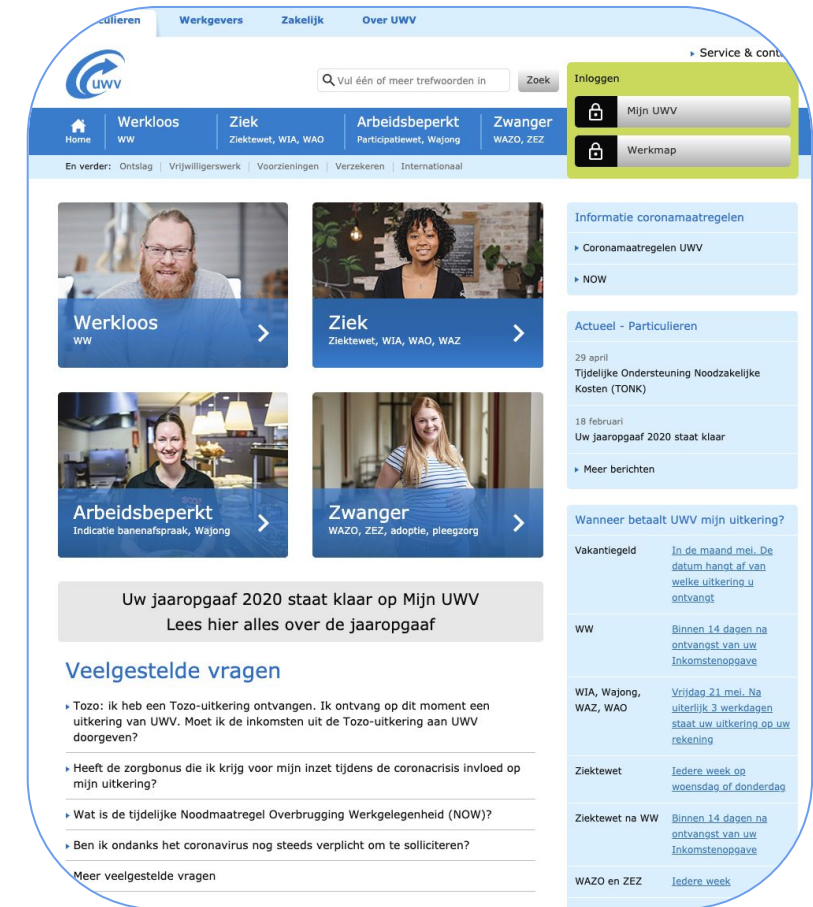
UWV is a governmental organization:

- Purchases, CR, CLTV or growth are no success metrics
- Primary KPI's are customer satisfaction and UX focused

## CSat score, UWV's #1 KPI

A costly, extensive, outsourced 62 item questionnaire performed twice a year by Ipsos.

- internal evaluation and monitoring of progress
- hard target set by dutch parliament





# Context

## Duration

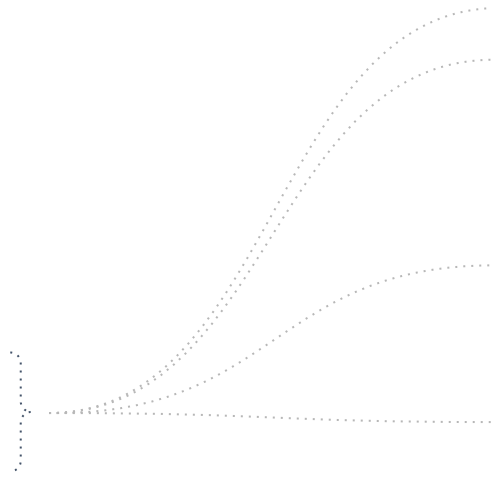
- 3 weeks (pilot)
- 4 weeks
- 11 weeks (3<sup>rd</sup> validation)
- *little dev time (bottleneck)*

## My role

- Self started project
- Excitement for xfn collab
- Planning & coordination
- Problem solving
- Implementation
- Validation
- Sharing & Visibility

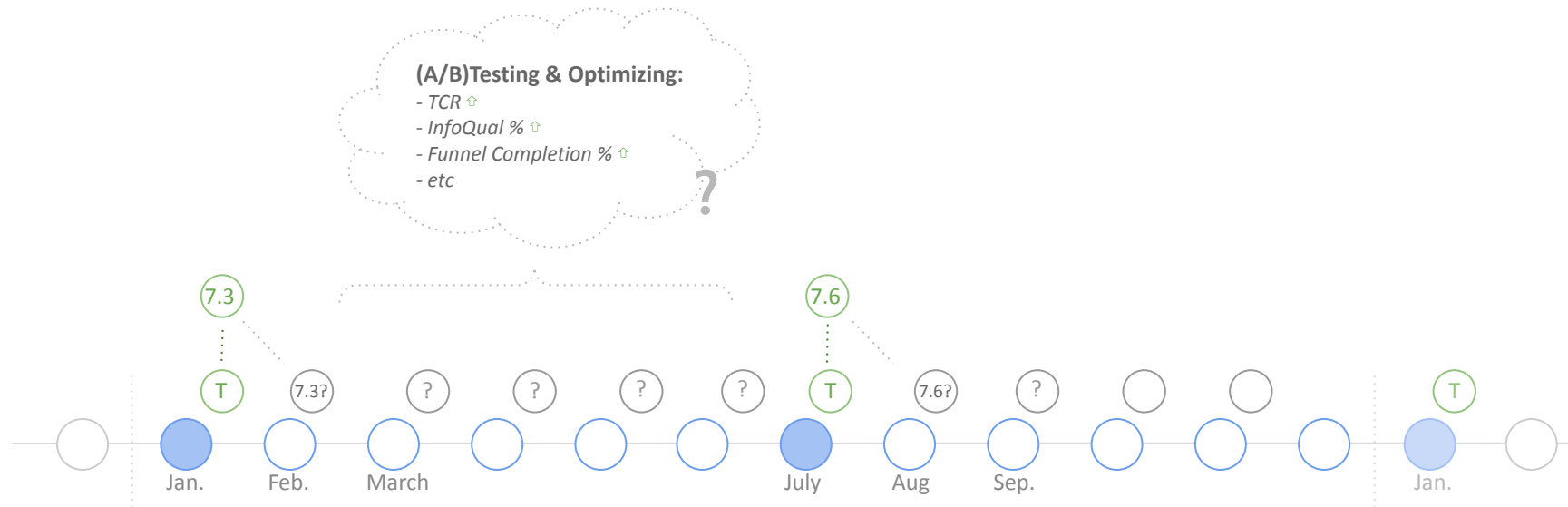
## Collaboration

- Front end developer
- Web Data Analysts
- Web Data Monitors
- UI designer
- Graphic Designer
- Product Owner
- Manager & Management Team
- Copywriters
- Content specialist
- Market researcher
- Portal Manager



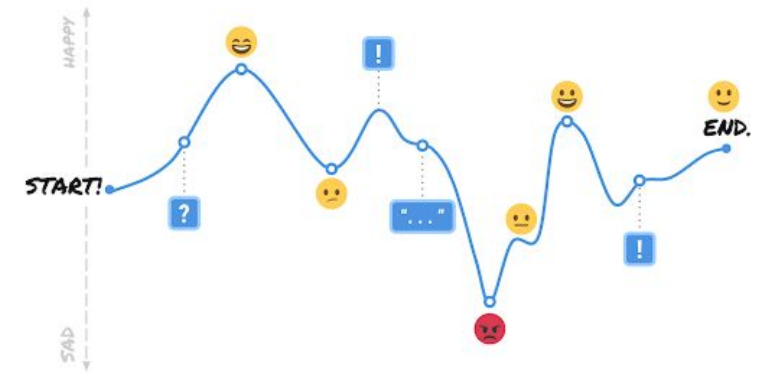
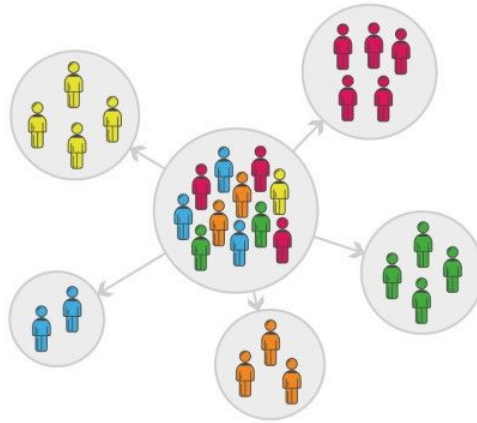
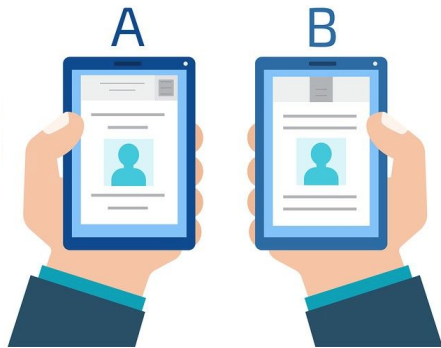
# Goals

- 1) Increase the cadence of CSat KPI without increasing costs



# Goals

- 2) Make the CSat KPI usable as OEC for website optimization (e.g. as guard rail metric for A/B testing)



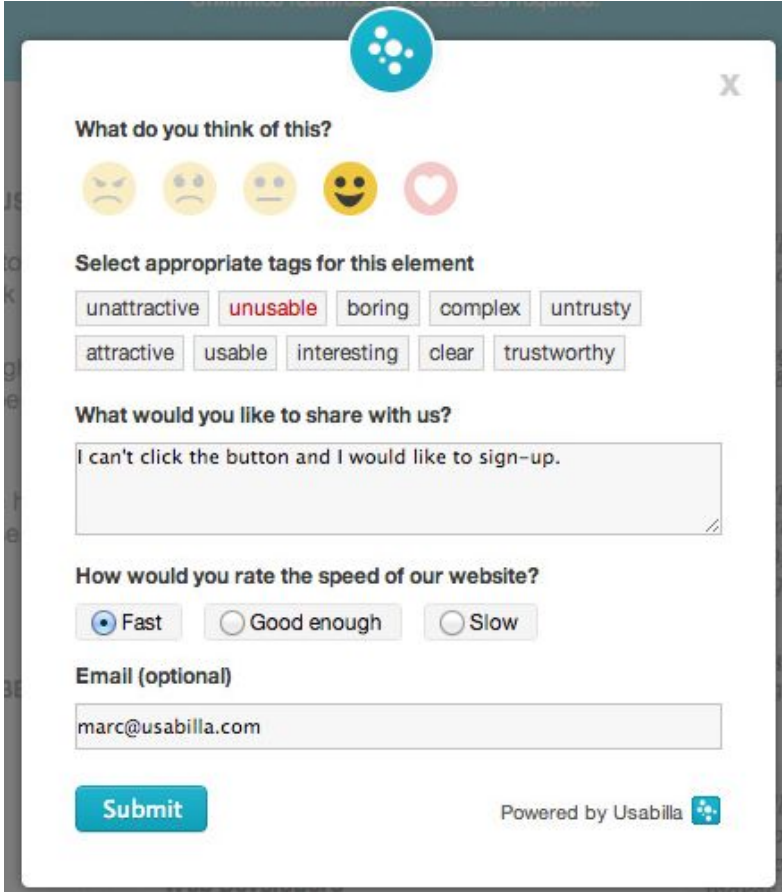
# Goals

- 3) Improve our understanding of customer satisfaction changes in relation to the dimension underlying customer satisfaction



# The main challenge

Limiting loss when reducing 62 items to a maximum of 8\*  
i.e. how to pick the 'most valuable' items?

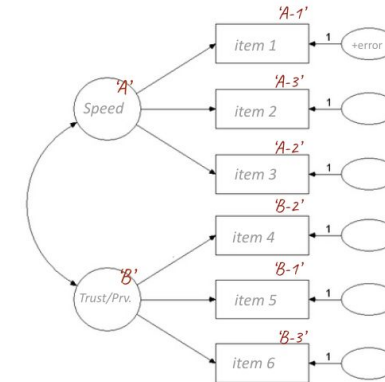
A screenshot of a Usabilla feedback widget. At the top is a blue circular icon with white dots. Below it is a close button 'X'. The first section is titled 'What do you think of this?' and features five emoji-like icons: two sad faces, one neutral face, one happy face, and one heart. The second section is titled 'Select appropriate tags for this element' and contains two rows of tag buttons: 'unattractive', 'unusable' (highlighted in red), 'boring', 'complex', 'untrustworthy' in the first row, and 'attractive', 'usable', 'interesting', 'clear', 'trustworthy' in the second row. The third section is titled 'What would you like to share with us?' and has a text input field containing the text 'I can't click the button and I would like to sign-up.' The fourth section is titled 'How would you rate the speed of our website?' and has three radio button options: 'Fast' (selected), 'Good enough', and 'Slow'. The fifth section is titled 'Email (optional)' and has a text input field containing 'marc@usabilla.com'. At the bottom left is a blue 'Submit' button, and at the bottom right is the text 'Powered by Usabilla' with the Usabilla logo.

Feedback

# Limiting loss in item reduction

picking the 'most valuable' items & remodelling

- 1) Exploratory factor analysis
  - a) first I identified dimensions underlying CSat
  - b) within each dimension I ranked items based on factor loading
  - c) finally I reduced dimensions based on explained variance



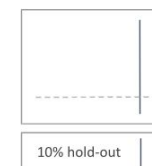
- 2) Multiple linear regression (with CSat score as dep. var)
  - I ranked items based on regression coefficient

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_i X_i$$

- 3) Reduce total items
  - I selected items based on combined rank scores from step 1 and 2

	DimFL_rank	RC_rank	Rank	Selected ?
item 1	A1	4	A5	Yes
item 2	A3	3	A6	No
item 3	A2	1	A3	Yes
item 4	B2	6	B8	No
item 5	B1	5	B6	Yes
item 6	B3	2	B5	Yes
i.n	..	..		

- 4) Validation & (pilot) Implementation
  - a) I used a 10% fold (hold-out from step 2) to validate the regression model
  - b) I compared 'predicted' CSat (from new data) against new CSat scores from Ipsos
  - c) I performed a final sanity check using a temporary surveyed single item CSat score





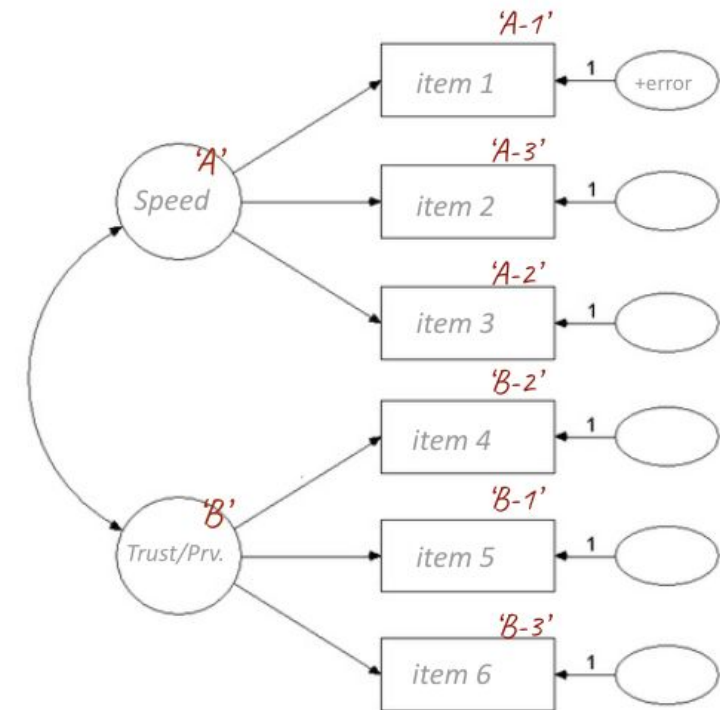
# Limiting loss in item reduction

picking the 'most valuable' items & remodelling

## 1) Exploratory factor analysis

- a) I identified dimensions underlying CSat using EFA
- & b) secondly I ranked items based on factor loadings

	<i>'A'</i>	<i>'B'</i>	<i>'C'</i>	
	Speed	Trust/Privacy	Info Qual	<i>'DimFL_Rank'</i>
<i>item 1</i>	.699	.093	..	<i>'A_1'</i>
<i>item 2</i>	.582	.111	..	<i>'A_3'</i>
<i>item 3</i>	.672	.043		<i>'A_2'</i>
<i>item 4</i>	.103	.586		<i>'B_2'</i>
<i>item 5</i>	.211	.621		<i>'B_1'</i>
<i>item 6</i>	.031	.574		<i>'B_3'</i>
<i>i..n</i>	..	..		



# Limiting loss in item reduction

picking the 'most valuable' items & remodelling

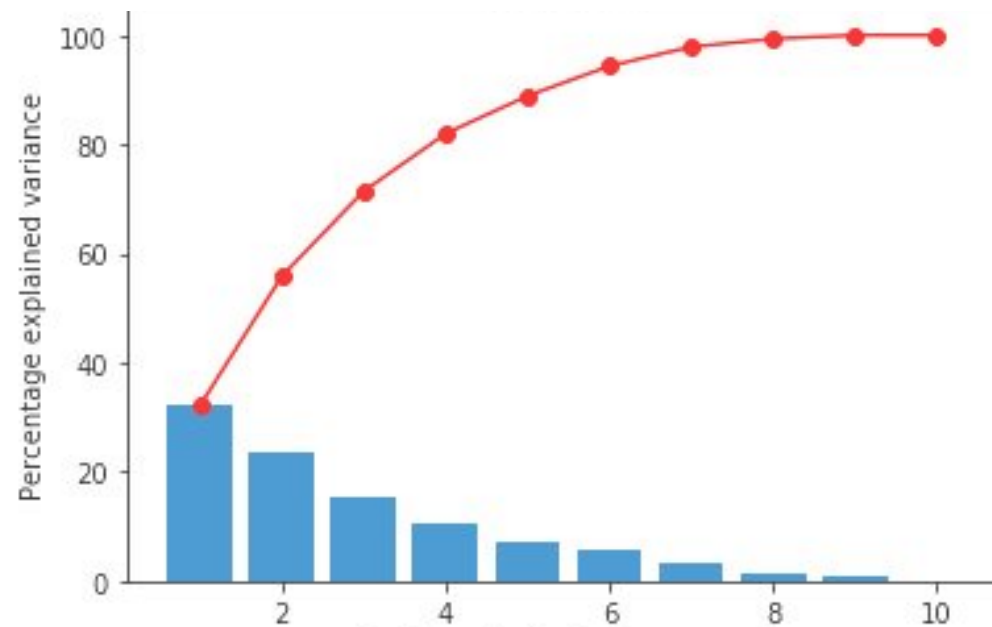
- 1) Exploratory factor analysis
- c) now I reduced dimensions based on explained variance

Total Variance Explained									
Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1 <i>'A'</i>	7.831	22.375	22.375	7.831	22.375	22.375	3.170	9.057	9.057
2 <i>'B'</i>	2.845	8.129	30.504	2.845	8.129	30.504	2.998	8.567	17.624
3 <i>'C'</i>	2.602	7.434	37.938	2.602	7.434	37.938	2.986	8.532	26.156
4 <i>'Etc'</i>	2.440	6.973	44.911	2.440	6.973	44.911	2.974	8.498	34.654
5	2.326	6.647	51.557	2.326	6.647	51.557	2.527	7.220	41.875
6	2.130	6.085	57.642	2.130	6.085	57.642	2.499	7.140	49.014
7	1.951	5.574	63.216	1.951	5.574	63.216	2.481	7.089	56.103
8	1.777	5.078	68.294	1.777	5.078	68.294	2.470	7.058	63.161
9	1.616	4.616	72.911	1.616	4.616	72.911	2.412	6.893	70.054
10	1.411	4.030	76.941	1.411	4.030	76.941	2.411	6.887	76.941
11	.719	2.055	78.996						

# Limiting loss in item reduction

picking the 'most valuable' items & remodelling

- 1) Exploratory factor analysis
- c) now I reduced dimensions based on explained variance

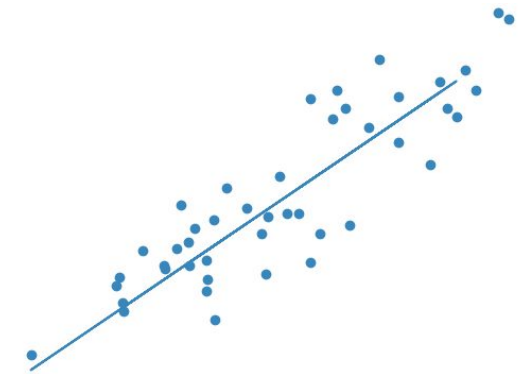


# Limiting loss in item reduction

picking the 'most valuable' items & remodelling

## 2) Multiple linear regression (with CSat score as dep. var)

As second step I ranked individual items based on regression coefficient



$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_i X_i$$

	Speed	Trust/Privacy	Info Qual	<u>CSat</u>	<i>RC_Rank</i>
<i>item 1</i>	.699	.093	..	.431	<i>'4'</i>
<i>item 2</i>	.582	.111	..	.532	<i>'3'</i>
<i>item 3</i>	.672	.043		.611	<i>'1'</i>
<i>item 4</i>	.103	.586		.123	<i>'6'</i>
<i>item 5</i>	.211	.621		.342	<i>'5'</i>
<i>item 6</i>	.031	.574		.604	<i>'2'</i>
<i>i..n</i>	..	..			

# Limiting loss in item reduction

picking the 'most valuable' items & remodelling

## 3) Reduce total items

Here I selected items based on combination of previous steps (rankings)

	DimFL_rank	RC_rank	<i>Combined_Rank</i>
<i>item 1</i>	A_1	4	<i>'A_5'</i>
<i>item 2</i>	A_3	3	<i>'A_6'</i>
<i>item 3</i>	A_2	1	<i>'A_3'</i>
<i>item 4</i>	B_2	6	<i>'B_8'</i>
<i>item 5</i>	B_1	5	<i>'B_6'</i>
<i>item 6</i>	B_3	2	<i>'B_5'</i>
<i>i..n</i>	..	..	



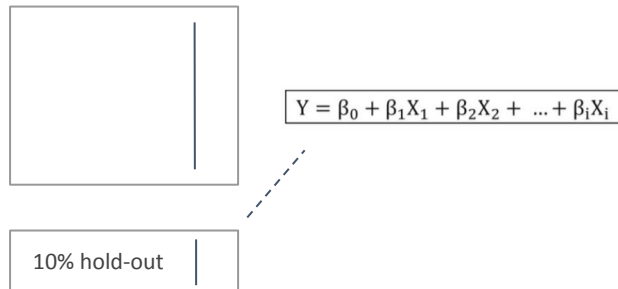
# Limiting loss in item reduction

picking the 'most valuable' items & remodelling

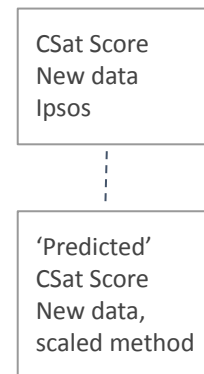
## 4) Validation

- I validated the foundations by using the 10% holdout, running this through the regression model and comparing the 'predicted' CSat score with the Ipsos CSat score
- After this I started (pilot) implementation using Usabilla VoC tooling, this let me compare the 'predicted' CSat scores of new data against new CSat scores from Ipsos
- A third and optional check\* can be performed by adding a 9th item, a 10 point scale item surveying overall satisfaction and comparing this with the models predicted CSat score.  
(will contain error as a fuzzy, multidimensional construct like overall customer satisfaction simply cannot reliably be measured through a single item construct)

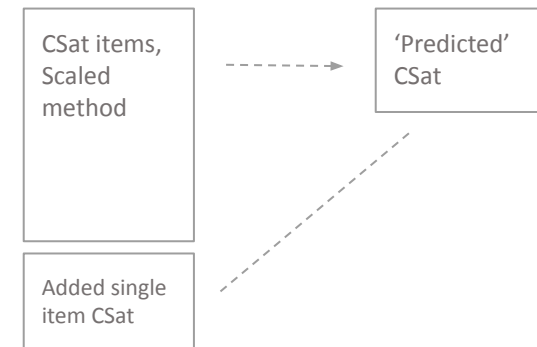
a)



b)



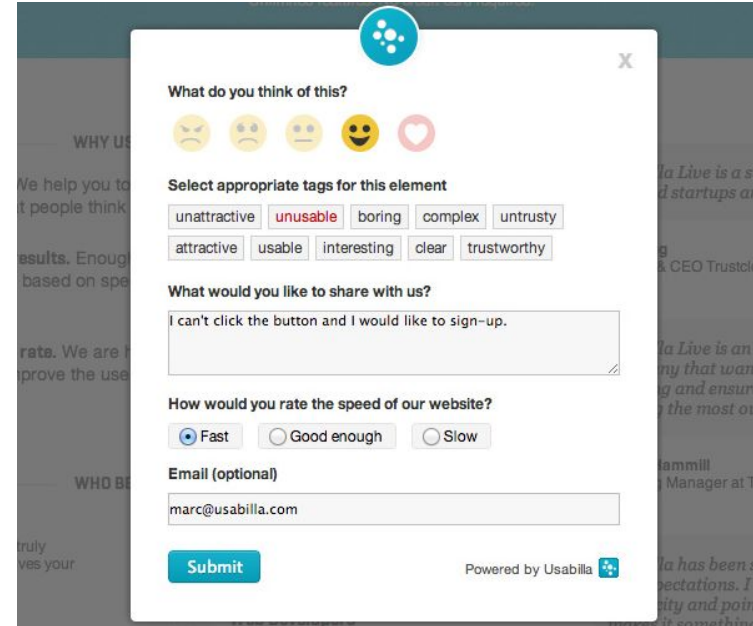
c)



# Limiting loss in item reduction

picking the 'most valuable' items & remodelling

- 5) Implementation
- Usabilla (front end)
  - A/B test
    - type
    - progress bar
    - button
- (response rate and attrition)
- 3% random sample (spread over 7 days)

A screenshot of a Usabilla feedback modal. The modal is white with a blue circular logo in the top left corner. It contains several sections: a rating section with five emoji icons (three sad, one happy, one heart); a tag selection section with buttons for 'unattractive', 'unusable', 'boring', 'complex', 'untrustworthy', 'attractive', 'usable', 'interesting', 'clear', and 'trustworthy'; a text input section for sharing feedback; a radio button section for rating website speed (Fast, Good enough, Slow); and an optional email input field. A blue 'Submit' button is at the bottom left, and 'Powered by Usabilla' is at the bottom right.

What do you think of this?

Select appropriate tags for this element

What would you like to share with us?

How would you rate the speed of our website?

Email (optional)

Submit

Powered by Usabilla

# Impact

Factor Analysis + multiple linear regression = continuous, KPI aligned, CSat insights

- predicted CSat score with 97% precision
  - 42% increased response rate
  - 65% lower attrition
- provided granularity and link to logged on-site behavior
  - underlying dimension of (changes in) customer satisfaction
  - differentiate customer journeys based on onsite behavior (cohorts & segments)
  - added value as optional guard rail metrics for larger A/B tests
- no more CSat surprises for management & xfn after 6 months of 'blind optimization'
  - Implemented as 'hero KPI' on UWV's Management Dashboard
  - Implemented as KPI on the optimization department's 24/7 live dashboard
- considered for adoption by other governmental departments  
(werk.nl & DUO.nl)

# Closing notes

## Concrete learnings

- A/B test your surveys!
- my mistake in communicating validity
  - > audience
- my mistake in planning
  - share progress frequently vs start & finish
  - timing: Ipsos hard date (validation)
- impact through others