Climpact: A User Study of Perceived Carbon Footprint

Victor Kristof[1], Robin Zbinden[1], Valentin Quelquejay-Leclère[1], Blanche Dalimier[2], Alexis Barrou[2], Edouard Cattin[2], Lucas Maystre[1], Jérôme Payet[2], Matthias Grossglauser[1], Patrick Thiran[1]

victor.kristof@epfl.ch

[1] **I**nformation and **N**etwork **Dy**namics Lab (indy.epfl .ch)

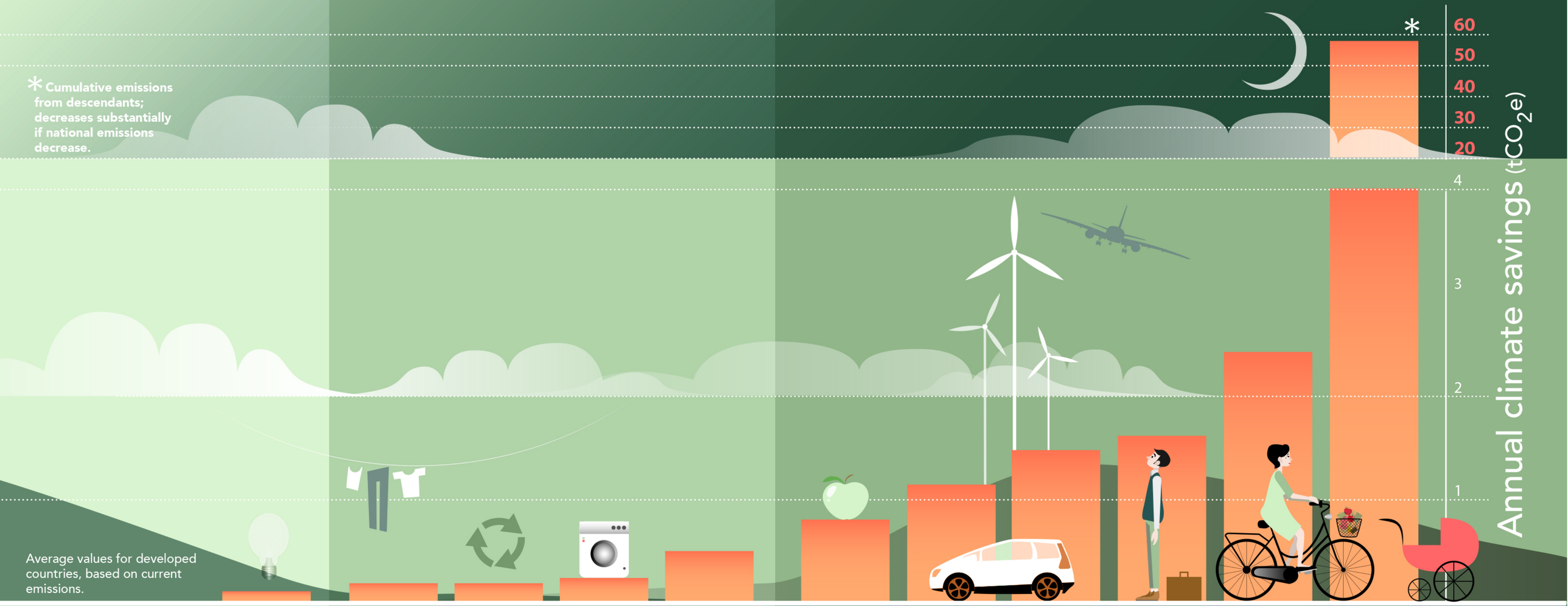
[2] CYCLECO Life Cycle Assessment (cycleco.eu)

Data Science in Climate and Climate Impact Research

August 21, 2020

EPFL

Personal choices to reduce your contribution to climate change



Low Impact

< 0.2 tCO_2e

Moderate Impact

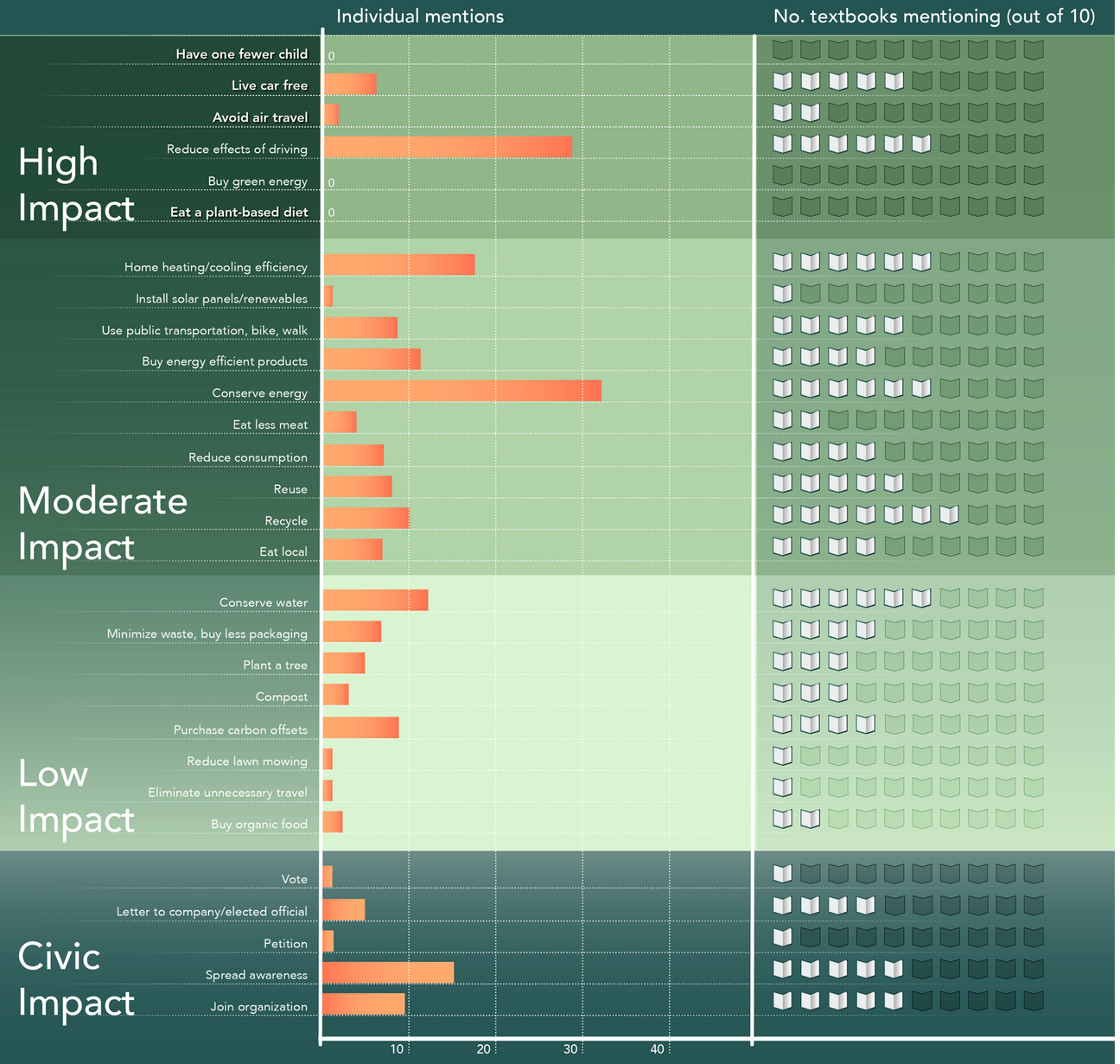
0.8-0.2 tCO_2e

High Impact

> 0.8 tCO_2e

Seth Wynes & Kimberly Nicholas, 2017, Environmental Research Letters

Recommended climate actions in 10 Canadian high school textbooks



Seth Wynes & Kimberly Nicholas, 2017, Environmental Research Letters

High impact, poorly documented

(Except for car-related actions)

Moderate impact, well documented

Low impact, well documented

This Work

Are people actually **poorly educated** about the impact of their actions?

Understand how people **perceive** the carbon footprint of their **actions**

How much does **flying** / **eating meat** / **lighting a house** emit?

Except for **experts**, it is very difficult to estimate our **absolute** carbon footprint

To **make decisions** about daily actions, we need the **relative** carbon footprint

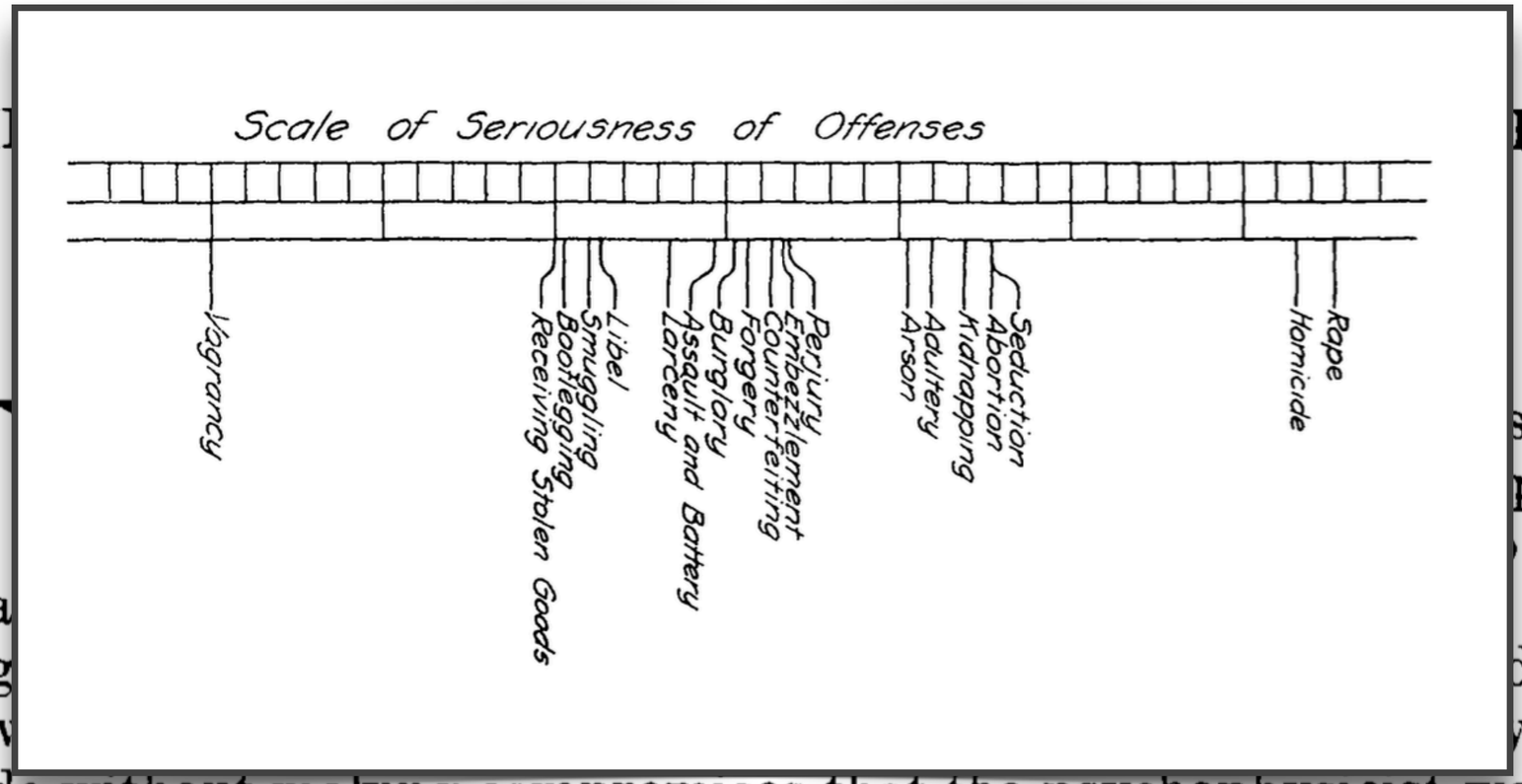
Psychometrics [Thurstone 1927]

THE METHOD OF PAIRED COMPARISONS FOR SOCIAL VALUES *

L. L. THURSTONE
UNIVERSITY OF CHICAGO AND INSTITUTE FOR JUVENILE RESEARCH

THIS is an attempt to apply the ideas of psychophysical measurement in the field of social values. Some of the psychophysical methods have been applied in a crude way to the measurement of educational products such as handwriting and English composition, and it seems feasible to apply the same ideas as well to social values although the attempt cannot readily be made without making compromises that the psychophysicist would not tolerate. The application of the principles of psychophysical measurement to educational products has been made with more or less similar logical handicaps but these do not seem to have dis-

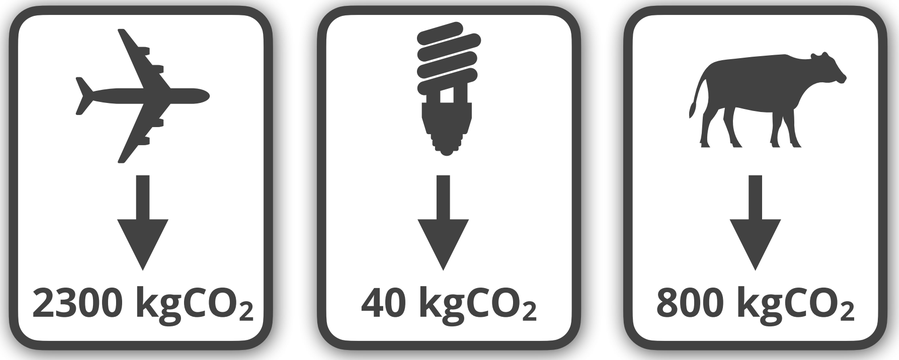
Psychometrics [Thurstone 1927]



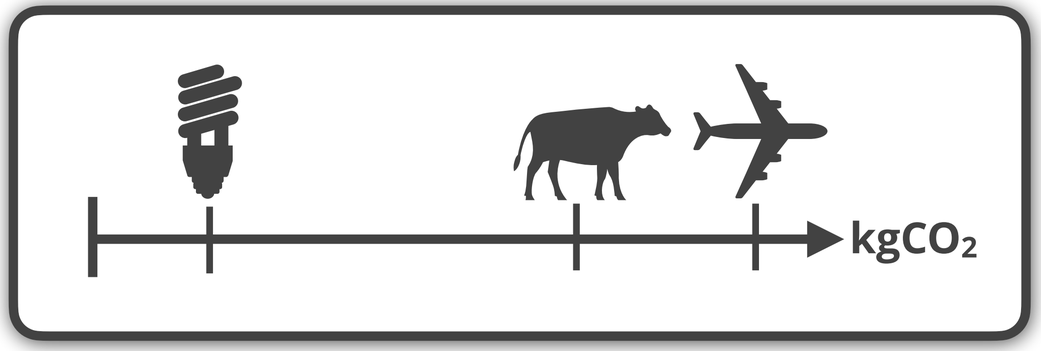
made without making compromises that the psychophysicist would not tolerate. The application of the principles of psychophysical measurement to educational products has been made with more or less similar logical handicaps but these do not seem to have dis-

Ranking from Pairwise Comparisons

Difficult task

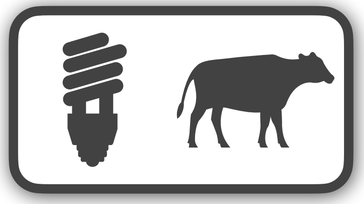


Easy computation

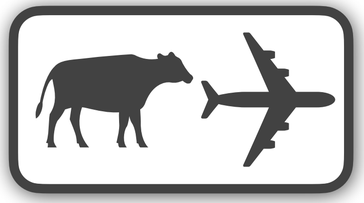


Easy task

→ 100 times more



→ 10 times more



→ 2 times more

Difficult computation...

...made possible *via* a statistical model of **pairwise comparisons!**

Statistical Model of Comparisons

Let $\mathscr{A}$ be a set of M **actions** and let (i, j, y) be a **triplet** encoding that action i has an **impact ratio** of $y \in \mathbf{R}_{>0}$ over j.

$$\left(\text{✈}, \text{💡}, y = 100\right)$$

Information: relative order of magnitude

Given some **parameters** $w_i, w_j \in \mathbf{R}$ representing the « *log* » *carbon footprint* of actions i and j, we posit

$$100 = \frac{\text{✈}}{\text{💡}} \quad \Longleftrightarrow \quad y = \frac{\tilde{w}_i}{\tilde{w}_j} = \frac{\exp w_i}{\exp w_j} \quad \text{where } w_i = \log \tilde{w}_i$$

$$y = \frac{\exp w_i}{\exp w_j} \xrightarrow{\log(\,\cdot\,)} \log y = w_i - w_j + \epsilon = \boldsymbol{x}^\intercal \boldsymbol{w} + \epsilon$$

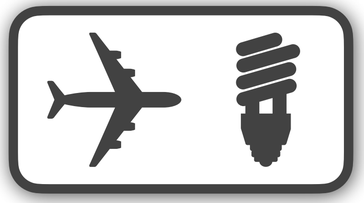
with $\epsilon \sim \mathcal{N}(0, \sigma_n^2)$, assuming comparisons are **noisy**

$$\log y \sim \mathcal{N}(\boldsymbol{x}^\intercal \boldsymbol{w}, \sigma_n^2)$$

comparison vector in $\mathbf{R}^M$ *selecting* the pairs of actions

$$\boldsymbol{x} = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ -1 \\ \vdots \\ 0 \end{bmatrix} \begin{matrix} \\ \\ i \\ \\ j \\ \\ \end{matrix}$$

Reminder:



→ **100 times more**

We estimate the **global perception** from **relative comparisons**

Estimating the Global Perception

For a dataset of N independent triplets, the **likelihood** of the model is

$$p(\boldsymbol{y} \mid \boldsymbol{X}, \boldsymbol{w}) = \prod_{i=1}^{N} p(y_i \mid \boldsymbol{x}_i^\intercal \boldsymbol{w}, \sigma_n^2) = \mathcal{N}(\boldsymbol{X}\boldsymbol{w}, \sigma_n^2 \mathbf{I})$$

Hyperparameters

- prior mean $\boldsymbol{\mu} \in \mathbf{R}^M$
- prior covariance $\boldsymbol{\Sigma}_p \in \mathbf{R}^{M \times M}$

Assuming a **Gaussian prior** for the parameters $\boldsymbol{w} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}_p)$, we compute the **posterior distribution** as

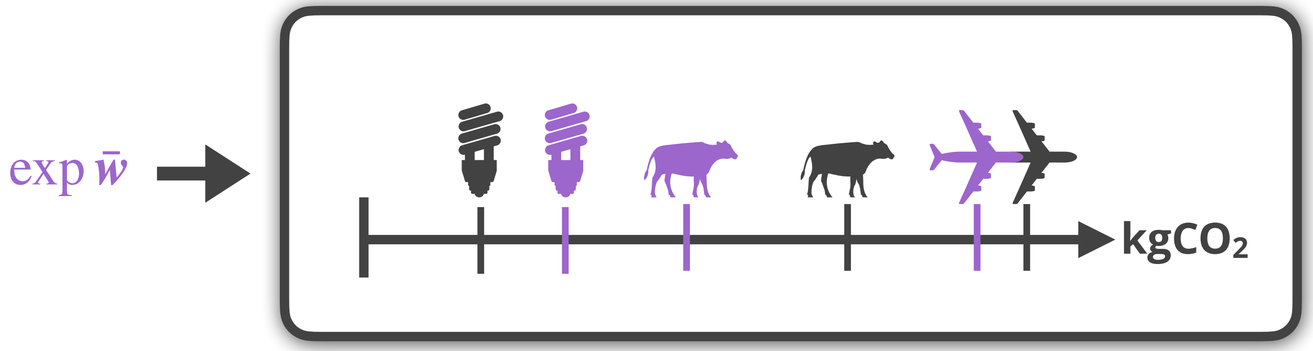
$$p(\boldsymbol{w} \mid \boldsymbol{X}, \boldsymbol{y}) = \frac{p(\boldsymbol{y} \mid \boldsymbol{X}, \boldsymbol{w}) p(\boldsymbol{w})}{p(\boldsymbol{y} \mid \boldsymbol{X})} = \mathcal{N}(\bar{\boldsymbol{w}}, \boldsymbol{\Sigma})$$

$$\boldsymbol{\Sigma} = \left(\sigma_n^{-2} \boldsymbol{X}^\intercal \boldsymbol{X} + \boldsymbol{\Sigma}_p^{-1} \right)^{-1}$$

$$\bar{\boldsymbol{w}} = \boldsymbol{\Sigma} \left(\sigma_n^{-2} \boldsymbol{X}^\intercal \boldsymbol{y} + \boldsymbol{\Sigma}_p^{-1} \boldsymbol{\mu} \right)$$

$\boldsymbol{\Sigma}$ used for **active learning**

$\exp \bar{\boldsymbol{w}}$ gives the **perception**



Enriching the Model: Perception Bias

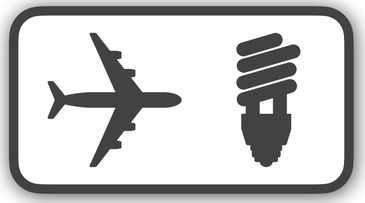
We want to capture the **perception bias** of users and actions into the model

Users:
- Age
- Gender
- Education

Actions:
- Category
- Source of energy
- Duration

$$y = \frac{\exp w_i}{\exp w_j} \longrightarrow y = \frac{\exp\left(w_i + \sum_k b_{ik}^{(u)}\right)}{\exp\left(w_j + \sum_k b_{jk}^{(u)}\right)},$$ where the bias $b_{ik}^{(u)} \in \mathbf{R}$ depends on user u and on action i

Example:



$$y = \frac{\exp\left(\text{✈} + \text{✈}_{\text{♀}} + \text{« Transport »} \right)}{\exp\left(\text{💡} + \text{♀}_{\text{💡}} + \text{« Housing »} \right)}$$

if the user u is a female participant

These assumptions enable **flexibility** and **interpretability** of the model!

Active Learning

We can use the **covariance matrix** of the **posterior distribution** of the model to (smartly) **select pairs of actions**.

Recall: $p(\boldsymbol{w} \mid \boldsymbol{X}, \boldsymbol{y}) = \mathcal{N}(\bar{\boldsymbol{w}}, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma} = \left(\sigma_n^{-2}\boldsymbol{X}^\intercal\boldsymbol{X} + \boldsymbol{\Sigma}_p^{-1}\right)^{-1}$

$\boldsymbol{\Sigma}$ used for **active learning**

As proposed by [MacKay* 1992], we want to select the pair of actions that is **maximally informative** about the values that the model parameters $\boldsymbol{w}$ should take. This is obtained by maximizing the **total information gain**:

$$\Delta S = S_N - S_{N+1} = \frac{1}{2}\log\left(1 + \sigma_n^2 \boldsymbol{x}^\intercal \boldsymbol{\Sigma}_N \boldsymbol{x}\right), \quad \text{where } \boldsymbol{\Sigma}_N = [\sigma_{ij}^2]_{i,j=1}^{M}$$

Entropy of multivariate Gaussian

i.e., all possible comparisons $\boldsymbol{x} = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ -1 \\ \vdots \\ 0 \end{bmatrix} \begin{matrix} \\ \\ i \\ \\ j \\ \\ \\ \end{matrix}$

To maximize ΔS, we maximize $\boldsymbol{x}^\intercal\boldsymbol{\Sigma}_N\boldsymbol{x}$ for all possible $\boldsymbol{x}$ in our dataset. We seek, therefore, to find

$$(i^\star, j^\star) = \underset{i,j}{\operatorname{argmax}}\{\sigma_{ii}^2 + \sigma_{jj}^2 - 2\sigma_{ij}^2\}$$

Very **fast** to compute for our model!

We can **actively** select the next pair of actions

* Yes, the same MacKay who wrote the book *Sustainable Energy – Without The Hot Air* !

Dataset of Actions

Take the train on a 1000-km round-trip
The train is a high-speed train with 360 seats. The seat-occupancy rate is 55% (average rate for these types of trains). We count the CO2 emissions per passenger.
Carbon footprint: 17 kgCO2-equivalent

Eat eggs and dairy products for one year
The production of eggs and dairy products (milk, cheese, etc.) emits CO2 because of water and land consumption, animal methane, and fossil fuel consumption for transportation and heating. We consider an average citizen consuming 50 kg of eggs and dairy products per year.
Carbon footprint: 100 kgCO2-equivalent

Light your house with incandescent bulbs
Incandescent bulbs emit CO2 because they consume electricity to generate light. The electricity is consumed from a grid with average CO2 rate.
Carbon footprint: 239 kgCO2-equivalent

Fly in first class for a 12000-km round-trip
The plane is a standard aircraft for long-distance flights with 390 seats. The seat-occupancy rate is close to 100%. We count the CO2 emissions per passenger. Passengers flying in first class use more space than passengers in economy.
Carbon footprint: 9000 kgCO2-equivalent

A total of **18 actions** covering **3 categories** (housing, transportation, and food)

New dataset of **50+ actions** covering **5 categories** (goods and services)

New Actions

These features can be integrated into the model to capture **action biases**!

Round-trip in train from Lausanne to Zurich

The train is an SBB long-distance IC train. The seat occupancy rate is 28 % (392 passengers). SBB trains run on electricity. They have a service life of 40 years. The travel distance is 348 km. Emissions include rail construction/dismantling, train maintenance, SBB's HV power generation, train station and train construction/dismantling. Emissions are in kg of CO2 eq for one passenger.

Carbon footprint: **2.35 kgCO2-equivalent**

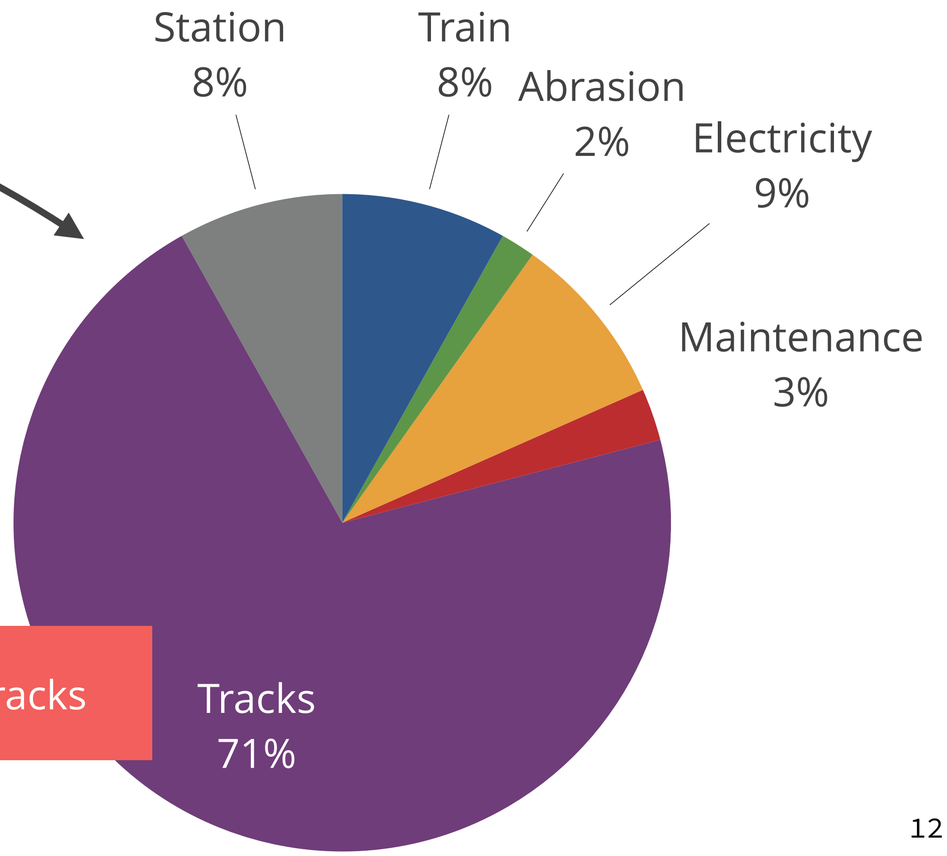
Perimeter:

- Production & dismantlement of train
- Production & dismantlement of tracks
- Electricity source
- Maintenance
- Train station

Functional Unit: Ensure the transportation of people in train

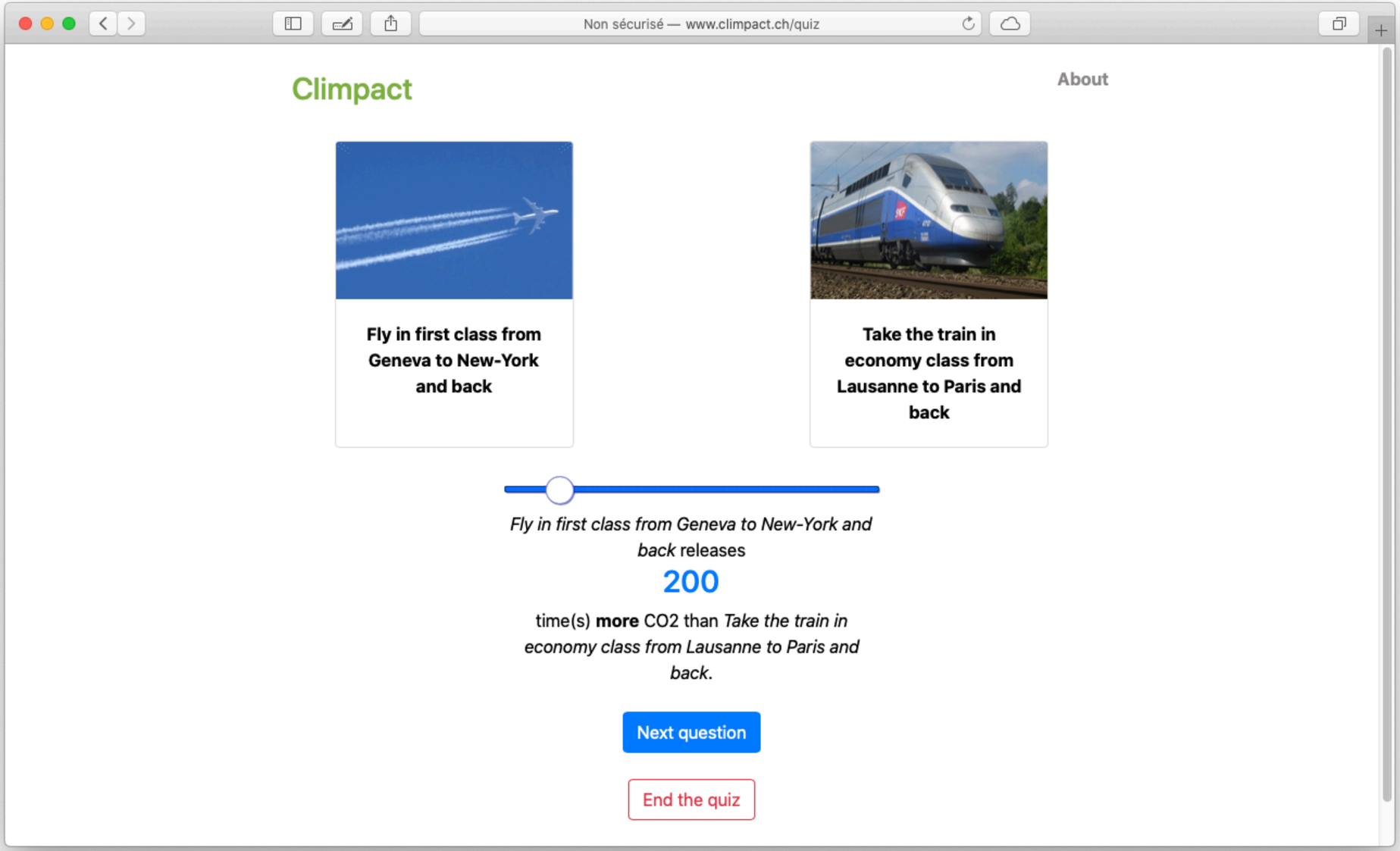
Methodology: Bottom-Up LCA

Data: *Ecoinvent* Database

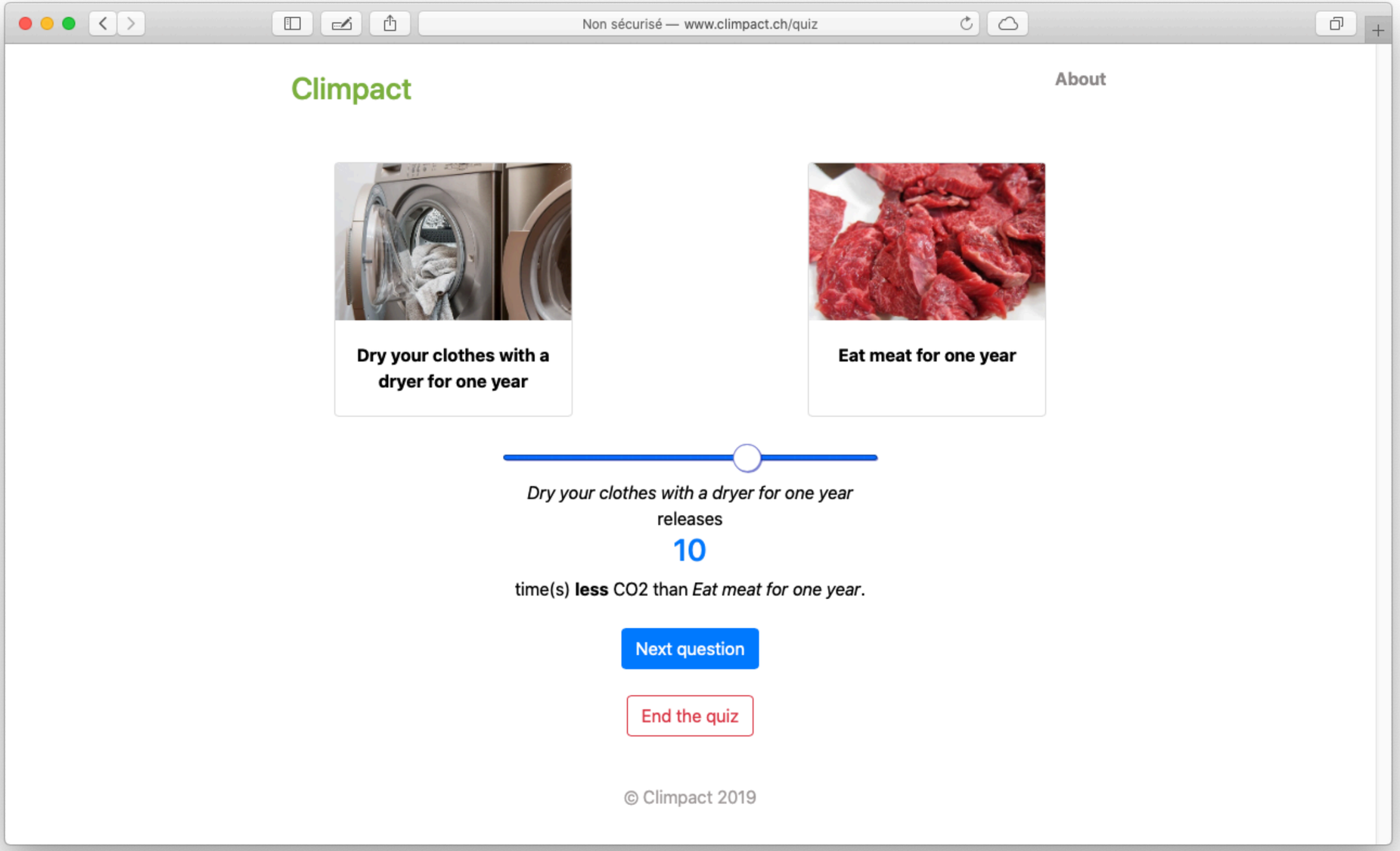


Because of **steel** and **concrete** required to lay tracks

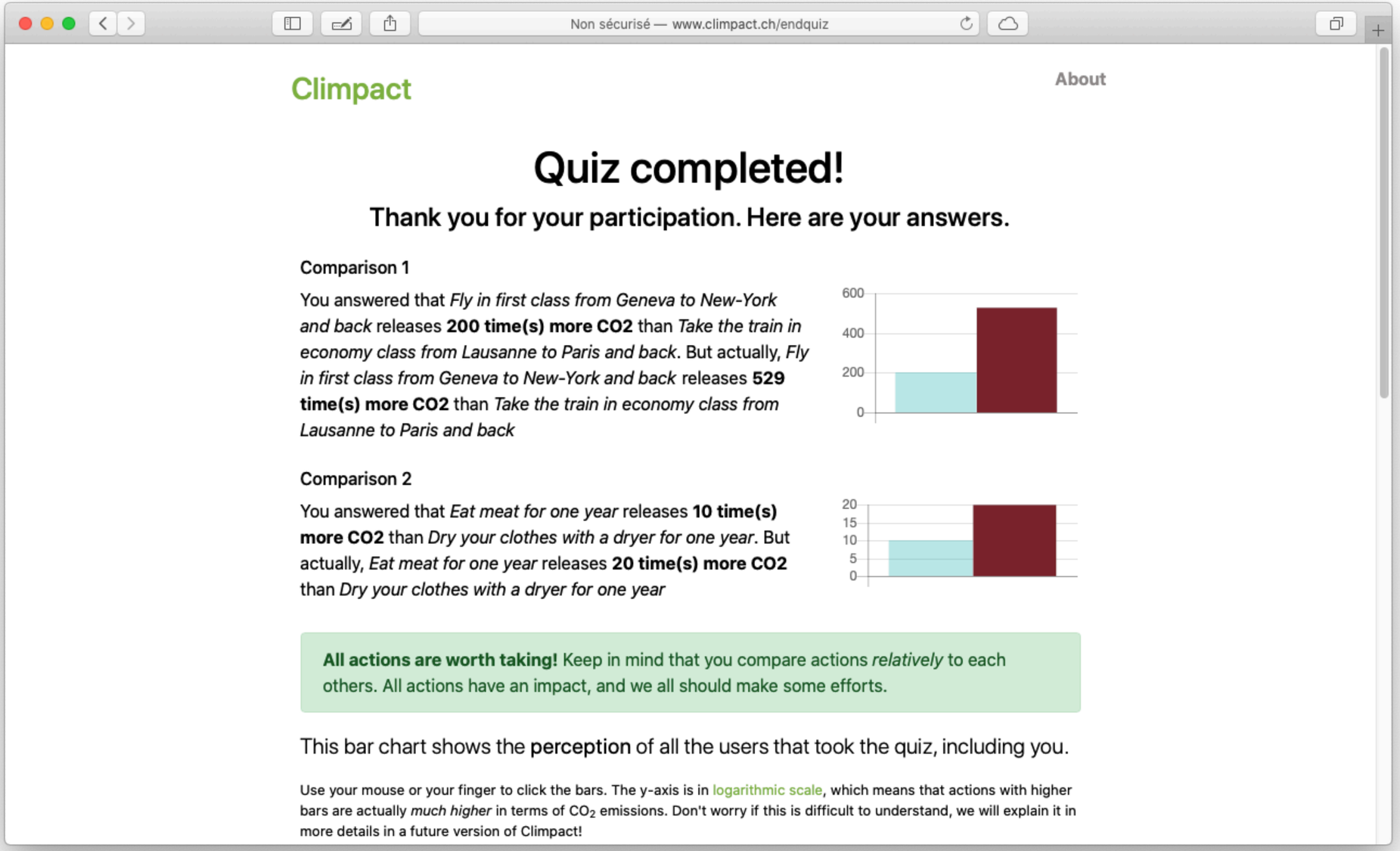
Climpact.ch: Collecting the Data



Climpact.ch: Collecting the Data

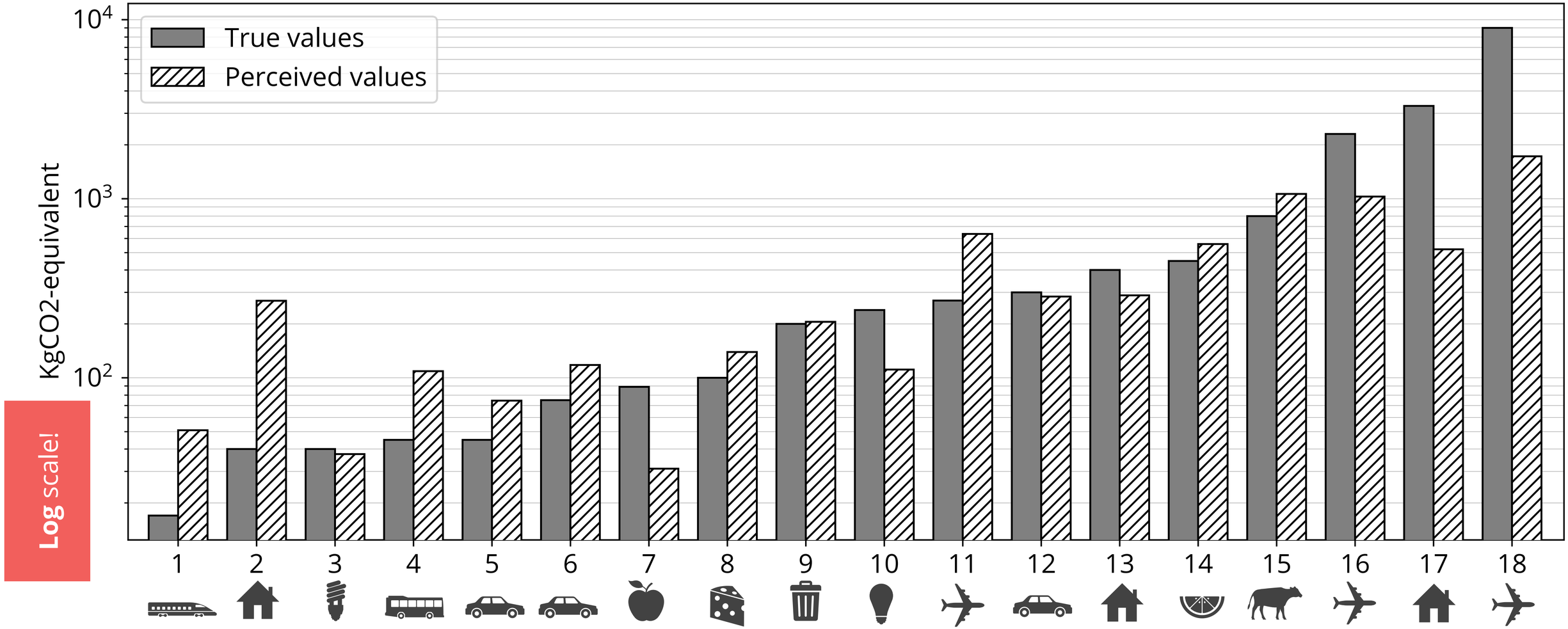


Climpact.ch: Collecting the Data



Results

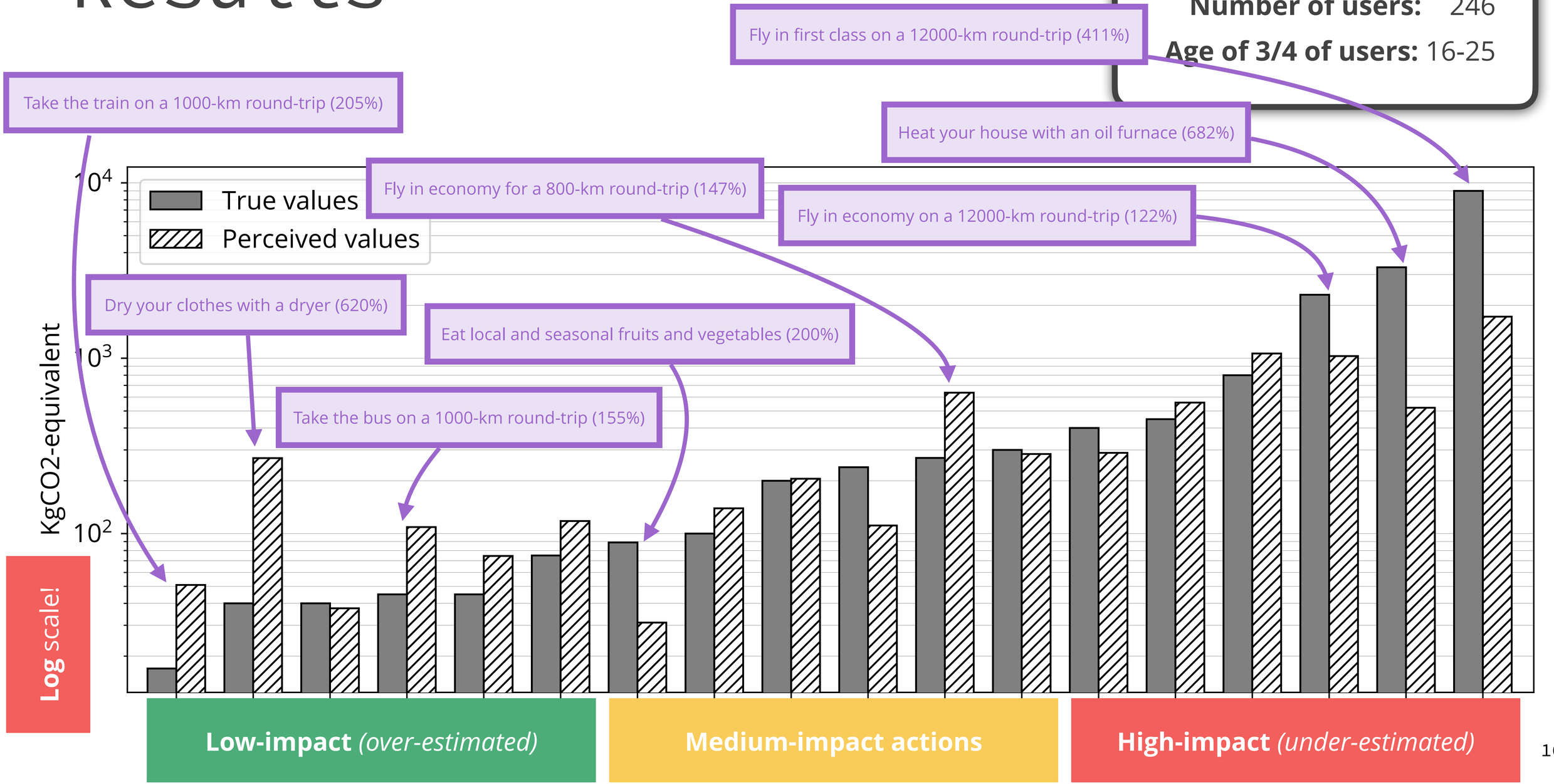
Number of answers: 3102
Number of users: 246
Age of 3/4 of users: 16-25



Log scale!

Results

Number of answers: 3102
Number of users: 246
Age of 3/4 of users: 16-25



Gender Bias

$$y = \frac{\exp\left(\text{✈} + \text{🚺✈} \right)}{\exp\left(\text{💡} + \text{🚺💡} \right)}$$

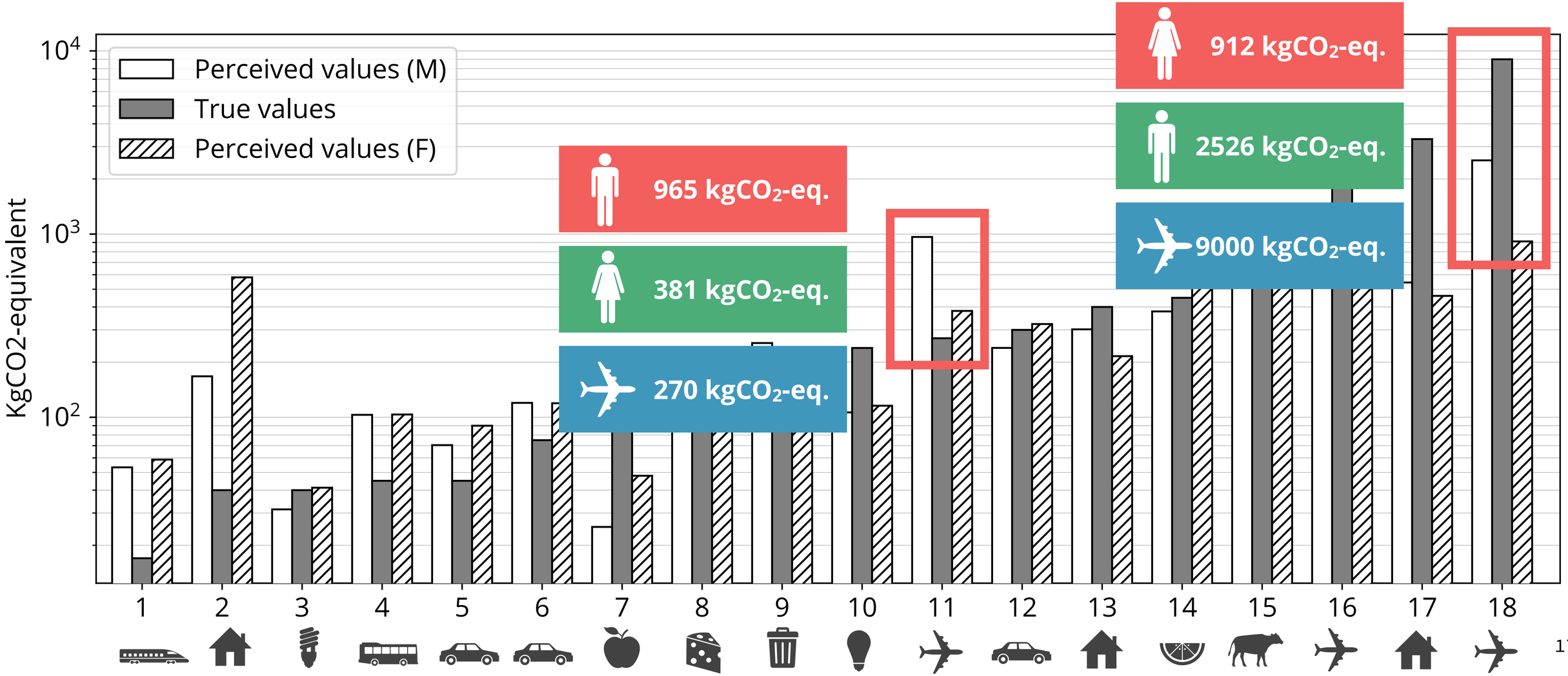
143 🚹

96 🚺

Number of answers: 2905

Number of users: 239

Age of 3/4 of users: 16-25



Limitations and Ongoing Work

The model is currently rather **simple**	→	Include **more features** (derived from new actions)
Active learning is equivalent to **uniform** selection	→	Make the data collection even more **efficient**
The online platform is very **basic**	→	Integrate the **new actions** and **visualization tools**
Data collected over a **small, biased** population	→	Open the platform to the **general public**

Can we move from active learning to **active teaching...**?

Thank you!

https://climpact.ch

(But please don’t share it further!)

Read our **paper:**

https://infoscience.epfl.ch/record/275472

Or scan this code

Connect with me on Twitter!

@VictorKristof

EPFL

indy.epfl.ch

Or reach out by email!

victor.kristof@epfl.ch