Description of Supplemental Materials

Contents

Supplemental Materials for Section 3 (Taxonomy) p. 2

Supplemental Materials for Section 4 (Interface) pp. 3 - 5

Supplemental Materials for Section 5 (User Study) pp. 6 - 10

Readers are referred to the online version of this document as it is easier to navigate through the supplemental materials online. The online version of this document can be found here: https://github.com/duospreadsheet/supplemental

Supplemental Materials for Section 3 (Taxonomy)

Supplemental materials for Section 3 can be found in the "Taxonomy" folder.

1. Codebook

The definition of codes are in "codebook.xlsx" in the "Taxonomy" folder.

2. Scenarios and Corresponding Sources

The scenarios for soliciting pairwise comparison questions from workers on MTurk and their sources are provided in "scenarios for soliciting questions from crowdworkers.pdf" in the "Taxonomy" folder.

3. Code Assigned by Coders

The "Codes Assigned" folder under the "Taxonomy" folder contains the codes assigned for the four dimensions by the two coders.

4. Code Statistics

The summary statistics and distributions of the codes are documented in "code statistics.pdf" in the "Taxonomy" folder.

Supplemental Materials for Section 4 (Interface)

Supplemental materials for Section 4 can be found in the "Interface" folder.

1. *Final Prototype (Sloppy Rules)*

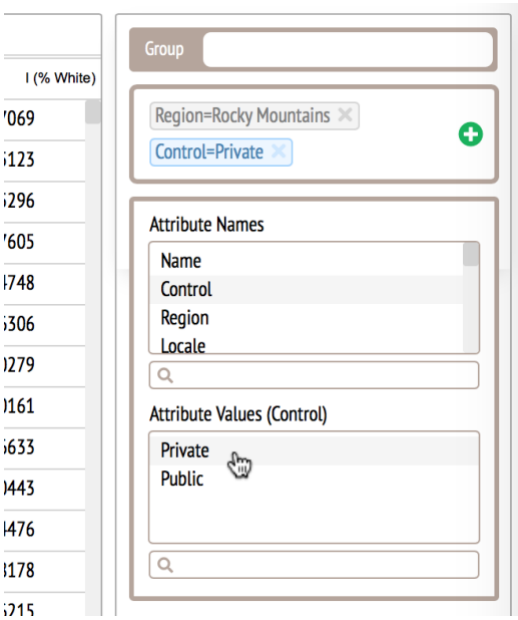
"duo demo.mp4" under the "Interface" folder is a video introducing various aspects of final prototype that allows users to use sloppy rules for specifying pairwise comparison questions. The final prototype (only tested in Chrome) can be found here: Link

The system provides the sloppy rule menu that allows users to specify pairwise comparison rules using the sloppy syntax.

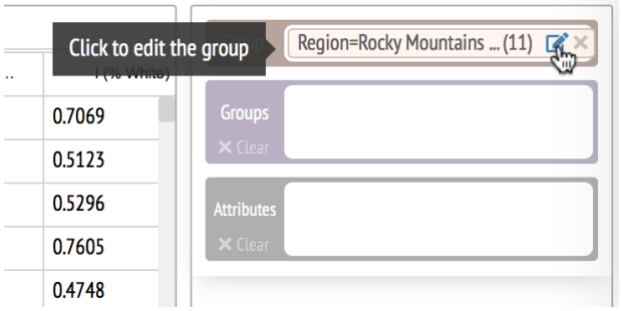
2. *First Prototype (GUI)*

The first prototype (only tested in Chrome) can be found here: Link

This version allows users to specify pairwise comparison rules using point and click. The menu only allows users to enter base rules.



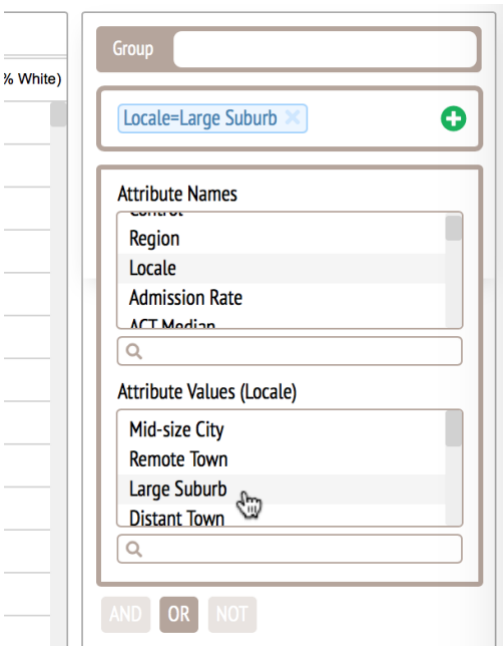
To enter inclusion and exclusion rules, users are required to use the edit button to open an editor. The complexity of the interface caused confusion to our pilot study participants.



3. Second Prototype (GUI)

The second prototype (only tested in Chrome) can be found here: Link

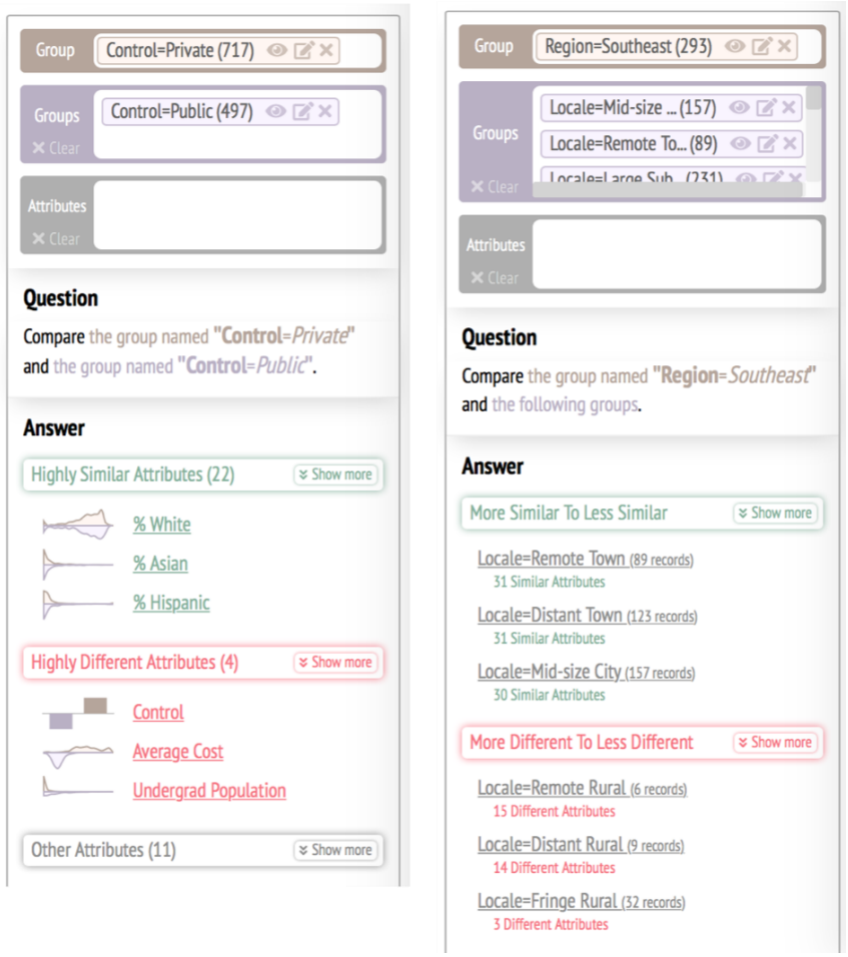
This version again allows users to specify pairwise comparison rules using point and click but this time, users can directly enter base rules, inclusion rules and exclusion rules using the menu.



We compared this interface with the natural language interface in an early pilot and the participant commented that the GUI helped "articulating" and "think about" pairwise comparison questions but the interface was still too complicated -- the interface complexity counteracted with the ease of articulation. This finding impelled us to improve the design and eventually resulted in the sloppy rules.

4. Logistic Regression Models for Classifying Attributes

For “single” comparisons, the system classifies attributes into highly similar attributes, highly different attributes and other attributes. For “repeated” comparisons, the system sorts the groups based on their number of similar and different attributes. The attributes are classified using logistic regression models.



Supplemental materials for Section 3 can be found in the “Classifying Attributes” folder under the “Interface” folder. “README - logistic Regression Models for Classifying Attributes.pdf” inside the folder provides a description of the logistic regression models. The “data’ folder under the “Classifying Attributes” folder contains the training data.

Supplemental Materials for Section 5 (User Study)

Supplemental materials for Section 4 can be found in the "User Study" folder.

1. Study Materials

The "Study Materials" folder under the "User Study" folder contains all the study materials.

To reduce learning effect, the presentation order of the two interfaces were counterbalanced. The following were the steps the participants followed for a participant who uses the natural language interface first.

- Fill out the consent form
  - **Not in the study material folder**
- Read the sample tasks
  - **"sample tasks.pdf" in the "Study Materials" folder**
- Watch a video about the visual language
  - **"Introduction to the Visual Language.mp4" in the "Study Materials" folder**
- Watch a video about the natural language interface
  - **"Introduction to the Natural Language Interface.mp4" in the "Study Materials" folder**
- Read the attributes in the car dataset used in the practice tasks
  - **"attribute list cars.pdf" in the "Study Materials" folder**
- Perform three practice tasks using the sloppy rule interface
  - **"practice tasks.pdf" in the "Study Materials" folder**
- Read the attributes in the college dataset used in the test tasks
  - **"attribute list college.pdf" in the "Study Materials" folder**
- Perform four test tasks (set 1) using the sloppy rule interface
  - **"test tasks set 1.pdf" in the "Study Materials" folder**
- Watch a video about the sloppy rule interface
  - **"Introduction to the Sloppy Rule Interface.mp4" in the "Study Materials" folder**
- Read the attributes in the car dataset used in the practice tasks
  - **"attribute list cars.pdf" in the "Study Materials" folder**
- Perform three practice tasks using the sloppy rule interface
  - **"practice tasks.pdf" in the "Study Materials" folder**
- Read the attributes in the college dataset used in the test tasks
  - **"attribute list college.pdf" in the "Study Materials" folder**
- Perform four test tasks (set 2) using the sloppy rule interface
  - **"test tasks set 2.pdf" in the "Study Materials" folder**
- Fill out the questionnaire
  - **"end-of-study questionnaire.pdf" in the "Study Materials" folder**
- Being interviewed on the relative pros and cons of the interfaces

2. Tutorial Videos

"Introduction to the Visual Language.mp4" in the "Study Materials" folder is a video that introduces the visual language.

"Introduction to the Natural Language Interface.mp4" in the "Study Materials" folder is a video that introduces the natural language interface.

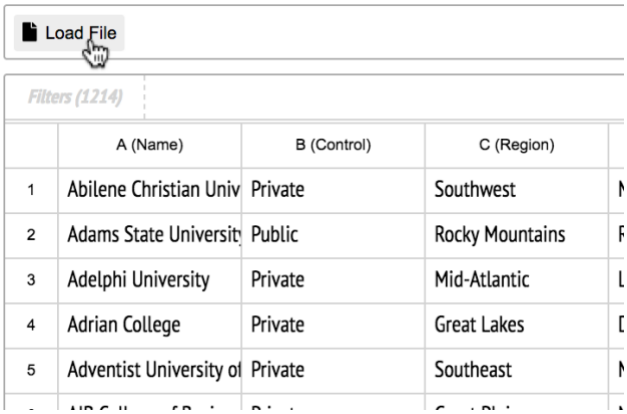
"Introduction to the Sloppy Rule Interface.mp4" in the "Study Materials" folder is a video that introduces the sloppy rule interface.

3. Natural Language Interface Used in the Study

The natural language interface for the study (only tested in Chrome, may take around a minute to load) can be found here: Link

The code of the natural language system can be found here (need Flask and Python 3 to run the server): Link

Here, we describe how the natural language interface was used during the user study. Let's suppose a participant is working with the practice tasks. At the beginning, the experimenter loads "oneToOnePractice.csv" using the load button in the interface. The data have to be loaded for the system to classify queries correctly.



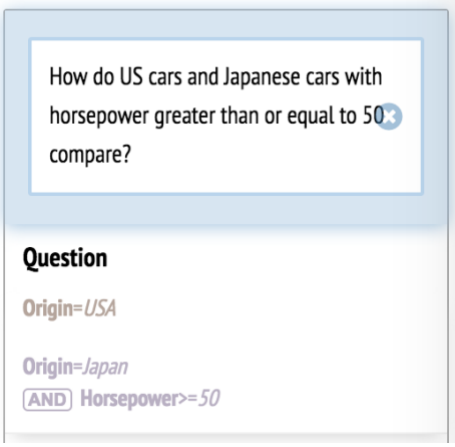
The participant can starting working on the three practice tasks using the natural language interface. The following is the first practice task.

Question 1

You should see this in the interface if you get it right:

If the query is correct, the system displays "**YOU GOT IT RIGHT!**"

If something are missing from the query, the system shows its interpretation to the participant so the participant can correct themselves.



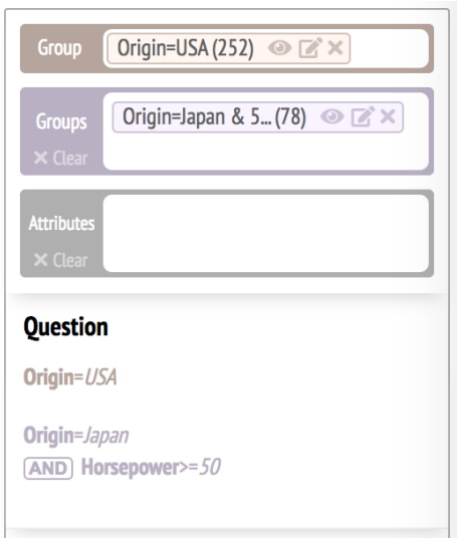
Similarly, if a participant works with test tasks (set 1) using the natural language interface, the experimenter first loads "oneToOneSetOne.csv". When a participant works with test tasks (set 2) using the natural language interface, the experimenter first loads "oneToOneSetTwo.csv".

Although the natural language interface is highly limited, it serves the purpose of the user study.

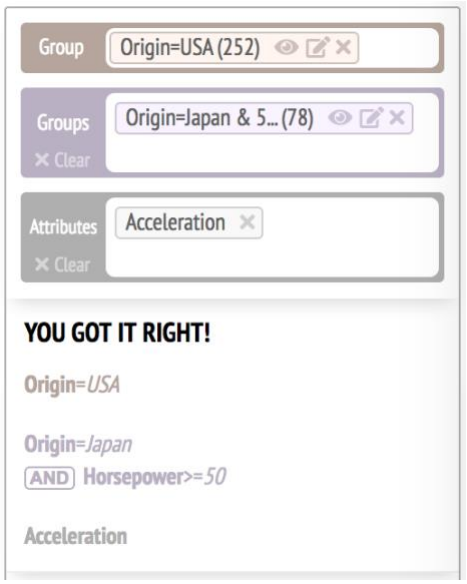
4. *Sloppy Rule Interface Used in the Study*

The sloppy rule interface for the study (only tested in Chrome) can be found here: Link
The code of the sloppy rule interface can be found here: Link

The sloppy rule interface was modified to show its interpretation of a query as participants enter pairwise comparison rules.



It also displays "**YOU GOT IT RIGHT!**" if a query is correct.



5. *Natural Language Queries Entered by the Participants*

While the participants were using the natural language interface, they were presented with either test tasks (set 1) or test tasks (set 2). They were asked to translate the visual language into an English sentence.

"participants' queries for the natural language interface.xlsx" in the "Raw Data" folder under the "User Study" folder contains the natural language queries entered by the participants.

6. Data Analysis Scripts

The "Analysis" folder under the "User Study" contains the data analysis scripts.

There are three dependent measures in the user study:

Completion Time

- Two-way repeated measures ANOVA was conducted to test for interaction between of interface and level of difficulty in predicting completion time.
    - **"Two-Way Repeated Measures ANOVA" folder in the "Analysis" folder**
- Post-hoc Wilcoxon signed-rank tests were conducted to test whether the time differences between the sloppy rule condition and natural language condition were significant across different level of difficulties.
    - **"Wilcoxon Signed-Rank Tests" folder in the "Analysis" folder**

Ease of Specification

This dependent variable corresponds to Q1 in "Task-Related Survey" of the questionnaire: "Which interface made it easier to articulate the above visual language (in other words, which interface made it easier to specify the above pairwise comparison questions)?"

- We computed the Spearman correlation between the rating and level of difficulty.
    - **"Spearman's Correlation" folder in the "Analysis" folder**
- Sign tests were used to test whether the rating is significantly different from 4 (neutral) for each level of difficulty.
    - **"Sign Tests" folder in the "Analysis" folder**

Preference

This dependent variable corresponds to Q2 in "Task-Related Survey" of the questionnaire: "Which interface would you prefer to perform the above tasks?"

- Binomial tests were conducted to test whether the number of participants who preferred sloppy rules is significantly greater than the number of participants who preferred natural language for each level of difficulty.
    - **"Binomial Tests" folder in the "Analysis" folder**