Codebook of dataset.csv

The first row contains the column titles. Then each row represents one subject.

Column description:

The first columns are subject-specific information

- SubjectId: subject id, a random number < 200000 for TP treatment, >200000 for control
- Group: treatment (tp or control)
- Age
- Gender (1="Male" and 2="Female")
- Nationality
- Knowledge: self-reported knowledge of the AEX (from 1="not at all" to 5="very well")

The other columns are matching probability-specific information. We elicited matching probabilities for 6 events, in three parts (0, 1, and 2; part 0 = training) and replicated the elicitation of 2 matching probabilities for each part. The data have the following structure:

For i=0 to 2

For j in {1,2,3,12,23,13}

- part**i**m**j** (e.g. part0m1) is the matching probability of event j in part i. They are given in %, i.e., between 0 and 100.
- part**i**m**j**time is the time spent in seconds by the subject for matching probability of event j in part i
- part**i**m**j**missing is 0 if the matching probability was submitted on time, 1 otherwise (can only be 1 for subjects in tp treatment and with i=1)
- part**i**m**j**rank indicates the order of appearance of the event in the experiment. Ranks for i=0 are between 1 and 8, those for i=1 are between 9 and 16, and those for i=2 are between 17 and 24.

End for

End for

To assess answer consistency, 6 matching probabilities were elicited twice. They are referred by the name of the initial matching probability, followed by “b”. These are columns:

part0m1b, part0m23b, part1m2b, part1m12b, part2m3b, part2m13b

These columns are followed by the columns indicating the time, whether the matching probability is missing, and the rank.