**Appendix 3. Statistical Appendix**

**1. Ease of identifiability of behavior change method**

Table A1. Tests of between-subject effects on Ease of identifiability ratings

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Source | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial η2 |
| Intercept | 34857659.98 | 1 | 34857659.98 | 24982.29 | < 0.001 | 0.936 |
| Transparency | 228409.53 | 1 | 228409.53 | 163.70 | < 0.001 | 0.088 |
| Argument | 32829.54 | 2 | 16414.77 | 11.76 | < 0.001 | 0.014 |
| Designer | 2728.24 | 2 | 1364.12 | 0.98 | 0.376 | 0.001 |
| Argument \* Transparency | 1831.76 | 2 | 915.88 | 0.66 | 0.519 | 0.001 |
| Designer \* Transparency | 2575.79 | 2 | 1287.89 | 0.92 | 0.398 | 0.001 |
| Argument \* Designer | 9382.86 | 4 | 2345.72 | 1.68 | 0.152 | 0.004 |
| Argument \* Designer \* Transparency | 6816.342 | 4 | 1704.09 | 1.22 | 0.300 | 0.003 |
| Error | 2367815.55 | 1697 | 1395.30 |  |  |  |

Multivariate tests showed a large main effect of Context on Ease of identifiability, *F* (4, 1964) = 119.96, *p* < .001, Wilks’ Λ =.784, partial η2 = .216, suggesting that there were differences between at least one pair of contexts. Post-hoc pairwise comparisons revealed that the means of the Ease of identifiability ratings in each context differed (all *p* < 0.001, apart from the difference between Smoking and Finance, which was *p* = 0.001, and all lower than the Bonferroni adjusted significance level of *p* = 0.005). The behavior change methods were most easily identifiable in the context of Diet (*M* = 72.01, *SE* = .54), followed by Alcohol (*M* = 67.88, *SE* = .65), Exercise (*M* = 62.86, *SE* = .64), Smoking (*M* = 60.00, *SE* = 0.66), and Finance (*M* = 57.41, *SE* = 0.70).

Table A2. Multivariate within-subject tests for Ease of identifiability ratings

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Source | Wilks’ *Λ* | F | df | Error df | Sig. | Partial η2 |
| Ease of identifiability | .78 | 116.96 | 4 | 1694 | < 0.001 | .216 |
| Ease of identifiability \*Transparency | .90 | 48.32 | 4 | 1694 | < 0.001 | .102 |
| Ease of identifiability \*Argument | .99 | 2.60 | 8 | 3388 | .008 | .006 |
| Ease of identifiability \*Designer | 1.00 | .88 | 8 | 3388 | .53 | .002 |
| Ease of identifiability  \*Argument \* Transparency | .99 | 1.56 | 8 | 3388 | .132 | .004 |
| Ease of identifiability  \*Designer \* Transparency | .99 | 1.63 | 8 | 3388 | .230 | .003 |
| Ease of identifiability \*Argument \* Designer | .99 | 1.32 | 16 | 5175.89 | .540 | .004 |
| Ease of identifiability \*Argument \* Designer \* Transparency | .98 | 1.14 | 16 | 5175.89 | .307 | .003 |

*Test of within-subject effects on Ease of identifiability ratings*

Since Mauchly’s test of sphericity was violated and the Greenhouse-Geisser estimate was > .75, a Huynh-Feldt estimate was used.

There was a small main effect of Context on Ease of Identifiability, *F* (3.86, 6547.72) = 125.42, *p* < .001, partial η2 =.069. There was also a small interaction effect between Ease of Identifiability and Transparency, *F* (3.86, 6547.72) = 49.17, *p* < .001, partial η2 = .028, and a very small interaction effect between Ease of Identifiability and Experiment, *F* (7.37, 6547.72) = 2.57, *p* < .009, partial η2 = .003.

**2. Effectiveness**

Table A3. Tests of between-subject effects on Effectiveness ratings

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Source | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial η2 |
| Intercept | 182330.51 | 1 | 182330.51 | 15232.36 | < 0.001 | 0.900 |
| Transparency | 1032.46 | 1 | 1032.46 | 86.25 | < 0.001 | 0.048 |
| Argument | 123.32 | 2 | 61.66 | 5.15 | 0.006 | 0.006 |
| Designer | 87.69 | 2 | 43.85 | 3.66 | 0.026 | 0.004 |
| Argument \* Transparency | 37.87 | 2 | 18.94 | 1.58 | 0.206 | 0.002 |
| Designer \* Transparency | 3.63 | 2 | 1.81 | 0.15 | 0.859 | 0.000 |
| Argument \* Designer | 13.92 | 4 | 3.48 | 0.29 | 0.884 | 0.001 |
| Argument \* Designer \* Transparency | 77.51 | 4 | 19.38 | 1.62 | 0.167 | 0.004 |
| Error | 20313.00 | 1697 | 11.97 |  |  |  |

Multivariate tests showed a large main effect of Context on Effectiveness, *F* (4, 1964) = 224.61, *p* < .001, Wilks’ Λ =.653, partial η2 = .347, suggesting that there were differences between at least one pair of contexts. Post-hoc pairwise comparisons revealed that ratings of the effectiveness of the BIs in each context differed (all *p* < 0.001, apart from the difference between Exercise and Smoking which was *p* = 0.001, and between Alcohol and Finance which was *p* = 0.002, and all lower than the Bonferroni adjusted significance level of *p* = 0.005). BIs were considered most likely to have a positive effect on behavior in the context of Diet (*M* = 5.47, *SE* = .54), followed by Exercise (*M* = 5.02, *SE* = .56), Smoking (*M* = 4.80, *SE* = .58), Alcohol (*M* = 4.02, *SE* = 0.55), and Finance (*M* = 3.84, *SE* = 0.53).

Table A4. Multivariate tests for Effectiveness

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Source | Wilks’ *Λ* | F | df | Error df | Sig. | Partial η2 |
| Effectiveness | .65 | 224.61 | 4 | 1694 | < 0.001 | .347 |
| Effectiveness \*Transparency | .92 | 38.98 | 4 | 1694 | < 0.001 | .084 |
| Effectiveness\*Argument | .98 | 3.90 | 8 | 3388 | .000 | .009 |
| Effectiveness\*Designer | 1.00 | .81 | 8 | 3388 | .594 | .002 |
| Effectiveness\*Argument \* Transparency | .99 | 2.97 | 8 | 3388 | .003 | .007 |
| Effectiveness\*Designer \* Transparency | .99 | 2.10 | 8 | 3388 | .033 | .005 |
| Effectiveness\*Argument \* Designer | .99 | 1.17 | 16 | 5175.89 | .280 | .003 |
| Effectiveness\*Argument \* Designer \* Transparency | .98 | 1.98 | 16 | 5175.89 | .011 | .005 |

*Test of within-subject effects on Effectiveness rating*

Since Mauchly’s test of sphericity was violated and the Greenhouse-Geisser estimate was > .75, a Huynh-Feldt estimate was used.

There was a medium main effect of Context on Effectiveness, *F* (3.99, 6773.46) = 230.56, *p* < .001, partial η2 =.120. There was also a small interaction effect between Effectiveness and Transparency, *F* (3.99, 6773.46) = 36.55, *p* < .001, partial η2 = .021, and a very small interaction effect between Effectiveness and Experiment, *F* (7.98, 6773.46) = 3.90, *p* < .001, partial η2 = .005, a very small three-way interaction between Effectiveness, Experiment, and Transparency, *F* (7.99, 6249.97) = 3.01, *p* =.002, partial η2 = .004, and a very small four-way interaction between Effectiveness, Experiment, Transparency, and Agent *F* (15.97, 6249.97) = 2.00, *p* =.010, partial η2 = .005.

**3. Acceptability**

Table A5. Tests of between-subject effects for Acceptability

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Source | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial η2 |
| Intercept | 349202.56 | 1 | 349202.56 | 34026.52 | < 0.001 | 0.952 |
| Transparency | 2545.96 | 1 | 2545.96 | 248.08 | < 0.001 | 0.128 |
| Argument | 215.97 | 2 | 107.99 | 10.52 | < 0.001 | 0.012 |
| Designer | 73.82 | 2 | 36.91 | 3.60 | 0.028 | 0.004 |
| Argument \* Transparency | 54.73 | 2 | 27.36 | 2.67 | 0.070 | 0.003 |
| Designer \* Transparency | 18.37 | 2 | 9.18 | 0.90 | 0.409 | 0.001 |
| Argument \* Designer | 18.97 | 4 | 4.74 | 0.46 | 0.764 | 0.001 |
| Argument \* Designer \* Transparency | 89.26 | 4 | 22.31 | 2.17 | 0.070 | 0.005 |
| Error | 17415.73 | 1697 | 10.26 |  |  |  |

Table A6. Multivariate tests for Acceptability

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Source | Wilks’ *Λ* | F | df | Error df | Sig. | Partial η2 |
| Acceptability | .50 | 432.40 | 4 | 1694 | < 0.001 | .505 |
| Acceptability \*Transparency | .78 | 122.94 | 4 | 1694 | < 0.001 | .225 |
| Acceptability\*Argument | .98 | 4.44 | 8 | 3388 | < 0.001 | .010 |
| Acceptability\*Designer | 1.00 | 1.04 | 8 | 3388 | .401 | .002 |
| Acceptability\*Argument \* Transparency | .99 | 2.91 | 8 | 3388 | .003 | .007 |
| Acceptability\*Designer \* Transparency | 1.00 | .55 | 8 | 3388 | .823 | .00 |
| Acceptability \*Argument \* Designer | .99 | 1.47 | 16 | 5175.89 | .100 | .003 |
| Acceptability \*Argument \* Designer \* Transparency | .98 | 1.33 | 16 | 5175.89 | .169 | .003 |

Post-hoc simple effects tests revealed that there was an effect of Argument on Acceptability ratings in the contexts of Alcohol, *F* (2, 1697) = 6.18, *p* = .002, partial η2 = .007, and Finance, *F* (2, 1697) = 112.66, *p* < .001, partial η2 = .026, but not in Exercise, *F* (2, 1697) = 0.99, *p* = .37, partial η2 = .001, Diet, *F* (2, 1697) = 3.81, *p* = .022, partial η2 = .004, or Smoking, *F* (2, 1697) = 2.28, *p* = .103, partial η2 = .003 (with Bonferroni adjusted significance level of *p =* 0.01). Pairwise comparisons then showed that in the contexts of Alcohol and Finance, Positive Arguments had higher Acceptability ratings than Positive + Negative and Negative (Finance both *p* < 0.001; Alcohol Experiments 1 & 2 differed at *p =* 0.001 and Experiments 2 & 3 at *p =* 0.004), but there was no significant difference between Positive + Negative and Negative (Finance, *p =* 0.32, Alcohol, *p =* 0.65). The means on which post-hoc tests were conducted are presented in Table A7.

Table A7: Means for Acceptability by Argument, for Alcohol and Finance Contexts

| Condition | Mean | Standard Error |
| --- | --- | --- |
| Finance Positive Argument | *M =* 5.24 | *SE =* .099 |
| Finance Positive + Negative Argument | *M =* 4.52 | *SE =* .090 |
| Finance Negative Argument | *M =* 4.39 | *SE = .*090 |
| Alcohol Positive Argument | *M =* 6.79 | *SE =* .095 |
| Alcohol Positive + Negative Argument | *M =* 6.37 | *SE =* .086 |
| Alcohol Negative Argument | *M =* 6.42 | *SE =* .087 |

*Tests of within-subject effects on Acceptability*

Since Mauchly’s test of sphericity was violated and the Greenhouse-Geisser estimate was > .75, a Huynh-Feldt estimate was used.

There was a large main effect of Context on Acceptability, *F* (3.68, 6249.97) = 610.42, *p* < .001, partial η2 = .26. However, this main effect needs to be interpreted in the light of the significant interaction effects: the small interaction effect between acceptability and transparency, *F* (7.37, 6249.97) = 4.44, *p* < .001, partial η2 = .093; the significant but negligible interaction effect between acceptability and experiment, *F* (7.37, 6249.97) = 4.44, *p* < .001, partial η2 = .006, and a very small three-way interaction between acceptability, experiment, and transparency, *F* (7.37, 6249.97) = 3.40, *p* <.001, partial η2 = .004.