**Appendix A – Descriptive Statistics**

|  |
| --- |
|  |
|  | FULL SAMPLE | CONTROL | T1:MajSec | T2:MajPub | T3:DelSec | T4:DelPub |
|   | Mean | sd | Mean | sd | Mean | sd | Mean | sd | Mean | sd | Mean | sd |
| Female | .63 |   | .63 |   | .63 |   | .67 |   | .57 |   | .67 |   |
| Age | 33.83 | 11.38 | 33.29 | 11.29 | 33.08 | 11.72 | 34.34 | 11.97 | 35.06 | 11.75 | 34.13 | 10.32 |
| Education | 10.25 | 2.16 | 10.29 | 2.08 | 10.06 | 2.18 | 10.29 | 2.18 | 10.30 | 2.16 | 10.26 | 2.32 |
| Ever Married | .65 |   | .63 |   | .67 |   | .66 |   | .63 |   | .66 |   |
| No. of Children | 2.24 | 1.91 | 2.25 | 1.94 | 2.19 | 1.79 | 2.28 | 1.87 | 2.30 | 2.06 | 2.16 | 1.88 |
| Coll. Action | .74 |   | .70 |   | .80+ |   | .76 |   | .71 |   | .78 |   |
| *Ability (Performance on Individual Task)* |  |  |  |  |  |  |  |  |  |  |
|  Letter Counting | 5.01 | 2.85 | 5.37 | 2.94 | 4.84 | 2.76 | 4.68+ | 2.88 | 4.67+ | 2.61 | 5.04 | 2.91 |
|  Sliders | 6.31 | 4.13 | 6.51 | 3.95 | 6.81 | 4.59 | 5.80 | 3.83 | 6.23 | 3.98 | 5.90 | 4.47 |
|  Stroop | 6.85 | 5.06 | 6.82 | 5.35 | 6.67 | 4.88 | 6.59 | 5.06 | 7.09 | 4.66 | 7.16 | 5.00 |
| *Preferred Task (Individual)* |   |  |   |  |   |  |   |  |  |   |   |
|  Letter Counting | .46 |   | .46 |   | .53 |   | .44 |   | .41 |   | .46 |   |
|  Sliders | .20 |   | .21 |   | .17 |   | .27 |   | .17 |   | .19 |   |
|  Stroop | .34 |   | .33 |   | .30 |   | .29 |   | .42 |   | .36 |   |
| *Selected Task (Team)* |   |  |   |  |   |  |   |  |  |   |   |
|  Letter Counting | .43 |   | .36 |   | .61\*\*\* |   | .67\*\*\* |   | .22\* |   | .39 |   |
|  Sliders | .15 |   | .26 |   | .11\*\* |   | .06\*\*\* |   | .11\*\* |   | .06\*\*\* |   |
|  Stroop | .42 |   | .38 |   | .28+ |   | .28+ |   | .67\*\*\* |   | .56\*\* |   |
| *Contributions (Performance on Team Task)* |  |  |  |  |  |  |  |  | .01 | .92 |
|  Raw | 46.72 | 20.72 | 44.87 | 21.88 | 47.70 | 20.39 | 43.78 | 19.80 | 48.74+ | 20.38 | 50.96\* | 18.89 |
| Standardized | .00 | 1.00 | -.06 | 1.05 | .10 | 1.06 | -.08 | 1.00 | .01 | .92 | .12 | .88 |
| Observations | 570 |  | 210 |  | 90 |  | 90 |  | 90 |  | 90 |  |

Significance stars indicate results of ttest comparing each treatment group to the control

+ *p* < 0.1, \* *p* < 0.05, \*\* *p* < 0.01, \*\*\* *p* < 0.001

**Appendix B – Statistical Power Calculations**

This study included 570 experimental subjects, and randomization occurred at the level of the individual. The application of the treatment, however, required sorting subjects into 5-person teams after treatment assignment. Even though the individual-level random assignment ensures that there is no pre-existing intra-group correlation, the experience of the treatment itself will likely co-vary for the members of the 5-person team group each individual was randomly assigned to, leading to intra-group correlation in the outcomes of the study. Thus, the statistical power calculations presented below are based on the formula for cluster-randomized designs.

Throughout my calculations, I assume standard values for statistical power equal to 0.80 and alpha equal to 0.05. The value I use for ρ, the intracluster correlation, is 0.03, and is based on the observed within group dependence on the primary outcome variable in the realized data. R2 takes the value of 0.24 which is the explanatory power of covariates in the realized data. The number of individuals per cluster, n, is also fixed at 5. The remaining inputs to the power calculation are: J, the number of clusters, and P, the proportion of the sample that belongs to the treatment group. Because of the nested design of this study, these vary depending on the particular comparison being made in each analysis.

These calculations yield an MDES as high as 0.39σ in the case of comparing two single treatments (not a comparison emphasized in the results of this paper) and as low as 0.26σ in the case of comparing pooled treatment groups to the control (the main analysis presented in this paper). The MDES is only slightly higher (0.27σ) when comparing two pooled treatments to each other (as is done whenever I directly compare the deliberation treatment to the majority rule treatment). The table below shows MDES conditional on comparison type.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **single treatment vs. control** | **2 pooled treatments vs. control** | **single treatment vs. single treatment** | **Pooled treatments vs. each other** |
| **J** | 60 | 78 | 36 | 72 |
| **P** | 0.3 | 0.46 | 0.5 | 0.5 |
| **MDES** | **0.33σ** | **0.26σ** | **0.39σ** | **0.27σ** |

Following Cohen (1988), effect sizes below 0.3 are considered small, so this sample size allows us to identify most moderate effects, if they exist, with a probability of 80%. I may not be able to rule out the existence of small effects, but given the additional cost of deliberative decision-making processes in practice, we may conclude that only moderately sized effects would be sufficient to single-handedly justify their use – with the caveat that there may still be other, normative reasons why deliberation is worthwhile.

**Appendix C – Main Analyses: Individual Performance on Team Task (linear regression)**

|  |  |  |  |
| --- | --- | --- | --- |
|  | All Treatment Groups |  | Collapsed Treatments |
|  | (1) | (2) |  | (3) | (4) |
|  |  |  |  |  |  |
| Majority Rule  | 2.120 | 2.714 | Majority Rule | 0.835 | 1.512 |
| + Secret Ballot | (2.623) | (2.440) |  | (2.068) | (1.852) |
|  |  |  |  |  |  |
| Majority Rule | -0.453 | 0.377 |  |  |  |
| + Show of Hands | (2.363) | (1.959) |  |  |  |
|  |  |  |  |  |  |
| Deliberation | 3.165 | 4.207+ | Deliberation | 4.946\* | 5.881\*\* |
| + Secret Ballot | (2.305) | (2.290) |  | (2.242) | (2.017) |
|  |  |  |  |  |  |
| Deliberation | 6.725\* | 7.557\*\* |  |  |  |
| + Show of Hands | (3.143) | (2.633) |  |  |  |
|  |  |  |  |  |  |
| SizeSession | -1.214\* | -1.103\* |  | -1.188\*\* | -1.080\*\* |
|  | (0.464) | (0.435) |  | (0.416) | (0.395) |
| Control Variables |  |  |  |  |  |
| Female |  | -6.047\*\*\* |  |  | -6.054\*\*\* |
|  |  | (1.758) |  |  | (1.754) |
| Age |  | -0.645\*\*\* |  |  | -0.648\*\*\* |
|  |  | (0.0880) |  |  | (0.0877) |
| Education |  | 1.621\*\*\* |  |  | 1.603\*\*\* |
|  |  | (0.444) |  |  | (0.446) |
| Married |  | -3.627+ |  |  | -3.543+ |
|  |  | (1.970) |  |  | (1.959) |
| No. of Children |  | 0.450 |  |  | 0.427 |
|  |  | (0.611) |  |  | (0.608) |
| Collective Action |  | -1.049 |  |  | -0.894 |
|  |  | (1.984) |  |  | (1.965) |
| KnewOthers |  | 0.277 |  |  | 0.190 |
|  |  | (3.971) |  |  | (4.184) |
| Prop.Female |  | 4.061 |  |  | 4.445 |
|  |  | (3.787) |  |  | (3.826) |
| numCoethnic |  | -0.109 |  |  | -0.0478 |
|  |  | (0.662) |  |  | (0.654) |
| Constant | 62.45\*\*\* | 69.10\*\*\* |  | 62.06\*\*\* | 68.61\*\*\* |
|  | (7.167) | (8.346) |  | (6.498) | (7.935) |
| *R*2 | 0.025 | 0.261 |  | 0.022 | 0.258 |
| Observations | 570 | 570 |  | 570 | 570 |

Standard errors in parentheses

+ *p* < 0.1, \* *p* < 0.05, \*\* *p* < 0.01, \*\*\* *p* < 0.001

**Appendix D – Main Analyses: Individual Performance on Team Task (poisson models)**

|  |  |  |  |
| --- | --- | --- | --- |
|  | All Treatment Groups |  | Collapsed Treatments |
|  | (1) | (2) |  | (3) | (4) |
|  |  |  |  |  |  |
| Majority Rule  | 0.0457 | 0.0564 | Majority Rule | 0.0180 | 0.0300 |
| + Secret Ballot | (0.0563) | (0.0526) |  | (0.0456) | (0.0410) |
|  |  |  |  |  |  |
| Majority Rule | -0.0114 | 0.00326 |  |  |  |
| + Show of Hands | (0.0533) | (0.0443) |  |  |  |
|  |  |  |  |  |  |
| Deliberation | 0.0673 | 0.0917+ | Deliberation | 0.104\* | 0.126\*\* |
| + Secret Ballot | (0.0492) | (0.0491) |  | (0.0473) | (0.0428) |
|  |  |  |  |  |  |
| Deliberation | 0.140\* | 0.159\*\* |  |  |  |
| + Show of Hands | (0.0635) | (0.0530) |  |  |  |
|  |  |  |  |  |  |
| SizeSession | -0.0240\*\* | -0.0207\* |  | -0.0234\*\* | -0.0202\*\* |
|  | (0.00859) | (0.00814) |  | (0.00758) | (0.00729) |
| Controls |  |  |  |  |  |
| Female |  | -0.120\*\* |  |  | -0.121\*\* |
|  |  | (0.0372) |  |  | (0.0372) |
| Age |  | -0.0155\*\*\* |  |  | -0.0156\*\*\* |
|  |  | (0.00237) |  |  | (0.00236) |
| Education |  | 0.0372\*\*\* |  |  | 0.0367\*\*\* |
|  |  | (0.0105) |  |  | (0.0106) |
| Married |  | -0.0567 |  |  | -0.0549 |
|  |  | (0.0419) |  |  | (0.0416) |
| No. of Children |  | 0.0121 |  |  | 0.0116 |
|  |  | (0.0157) |  |  | (0.0155) |
| Collective Action |  | -0.0173 |  |  | -0.0143 |
|  |  | (0.0417) |  |  | (0.0414) |
| KnewOthers |  | 0.0108 |  |  | 0.00836 |
|  |  | (0.0769) |  |  | (0.0818) |
| Prop.Female |  | 0.0911 |  |  | 0.0969 |
|  |  | (0.0795) |  |  | (0.0810) |
| numCoethnic |  | -0.00411 |  |  | -0.00265 |
|  |  | (0.0138) |  |  | (0.0137) |
| Constant | 4.150\*\*\* | 4.259\*\*\* |  | 4.141\*\*\* | 4.251\*\*\* |
|  | (0.134) | (0.176) |  | (0.120) | (0.169) |
| Observations | 570 | 570 |  | 570 | 570 |

Standard errors in parentheses

+ *p* < 0.1, \* *p* < 0.05, \*\* *p* < 0.01, \*\*\* *p* < 0.001

**Appendix E – Decision Quality vs. Behavior Change (Effort)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) |
|  | Contributions  | Ability | Effort | Decision Quality |
|  |  |  |  |  |
| Majority Rule | 1.512 | -0.113 | 1.896 | 0.207 |
|  | (1.852) | (0.461) | (1.258) | (0.296) |
|  |  |  |  |  |
| Deliberation | 5.881\*\* | 1.340\*\* | 1.338 | 0.955\*\* |
|  | (2.017) | (0.473) | (1.411) | (0.300) |
|  |  |  |  |  |
| Ability |  |  | 3.391\*\*\* |  |
|  |  |  | (0.161) |  |
|  |  |  |  |  |
| Controls? | Yes | Yes | Yes | Yes |
|  |  |  |  |  |
| Constant | 68.61\*\*\* | 6.809\*\*\* | 45.52\*\*\* | -1.715 |
|  | (7.935) | (1.773) | (5.566) | (1.057) |
| Model Type | ols(appendix C, model 4) | ols | ols | probit (by team) |
| Observations | 570 | 570 | 570 | 114 |

Standard errors in parentheses

+ *p* < 0.1, \* *p* < 0.05, \*\* *p* < 0.01, \*\*\* *p* < 0.001

**Appendix F – Effects of Deliberation relative to Majority Rule group**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) |
|  | Contributions | Ability | Contributions | Decision Quality |
|  |  |  |  |  |
| Deliberation | 4.393\* | 1.456\*\* | -0.331 | 0.750\* |
|  | (1.888) | (0.480) | (1.370) | (0.311) |
|  |  |  |  |  |
| Ability |  |  | 3.245\*\*\* |  |
|  |  |  | (0.198) |  |
|  |  |  |  |  |
| Controls? | Yes | Yes | Yes | Yes |
|  |  |  |  |  |
| Constant | 74.00\*\*\* | 8.864\*\*\* | 45.24\*\*\* | -1.367 |
|  | (8.961) | (1.995) | (6.306) | (1.182) |
| Model Type | ols(appendix C, model 4) | ols | ols | probit (by team) |
| Observations | 360 | 360 | 360 | 72 |

Standard errors in parentheses

+ *p* < 0.1, \* *p* < 0.05, \*\* *p* < 0.01, \*\*\* *p* < 0.001

**Note:** These regressions are identical to those presented in Appendix E, except that they exclude all observations from the control group. Thus, the reference category in these models is the majority rule treatment group only.The coefficients ondeliberation can therefore be interpreted asthe increase in each outcome that is achieved by using deliberation instead of majority rule voting**.**

**Appendix G – Effect of Treatments on Secondary Outcomes**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) |
|  | Preference Alignment | Preference Change | Perceived Fairness | Acceptance of Outcomes |
|  |  |  |  |  |
| Majority Rule | 0.379\*\* | 0.00816 | 0.0749 | -0.219\* |
|  | (0.130) | (0.167) | (0.120) | (0.0966) |
|  |  |  |  |  |
| Deliberation | 0.333\*\* | 0.416\*\* | 0.313\*\* | 0.297\*\* |
|  | (0.127) | (0.139) | (0.0965) | (0.0895) |
|  |  |  |  |  |
| Final Preference |  |  |  | 0.399\*\*\* |
|  Matches Team |  |  |  | (0.0924) |
|  |  |  |  |  |
| Controls? | Yes | Yes | Yes | Yes |
|  |  |  |  |  |
| Constant | -1.858\*\*\* | -0.288 | 4.944\*\*\* | 3.019\*\*\* |
|  | (0.446) | (0.615) | (0.308) | (0.455) |
| Model Type | probit | probit | ols | ols |
| Observations | 570 | 570 | 570 | 570 |

Standard errors in parentheses

+ *p* < 0.1, \* *p* < 0.05, \*\* *p* < 0.01, \*\*\* *p* < 0.001

**Appendix H – Effect of Intermediate Outcomes on Contributions**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) |
|  | Contributions | Contributions | Contributions | Contributions |
| Preference Alignment | 6.520\*\*\* |  |  |  |
|  | (1.651) |  |  |  |
|  |  |  |  |  |
| Preference Change |  | 0.0232 |  |  |
|  |  | (2.026) |  |  |
|  |  |  |  |  |
| Perceived Fairness |  |  | 1.068 |  |
|  |  |  | (0.799) |  |
|  |  |  |  |  |
| Acceptance of |  |  |  | 1.645\* |
|  Outcomes |  |  |  | (0.702) |
|  |  |  |  |  |
| Final Preference |  |  |  | 5.093\*\* |
|  Matches Team |  |  |  | (1.884) |
|  |  |  |  |  |
| Controls? | Yes | Yes | Yes | Yes |
|  |  |  |  |  |
| Constant | 71.42\*\*\* | 70.57\*\*\* | 65.18\*\*\* | 63.58\*\*\* |
|  | (8.018) | (8.003) | (8.700) | (7.945) |
| Observations | 570 | 570 | 570 | 570 |

Standard errors in parentheses

+ *p* < 0.1, \* *p* < 0.05, \*\* *p* < 0.01, \*\*\* *p* < 0.001

Note: All the models in this table take the same form as Appendix C, Model 4, other than the addition of each intermediate outcome.

**Appendix J – Effect of Intermediate Outcomes on Effort (Contributions controlling for Ability)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) |
|  | Contributions | Contributions | Contributions | Contributions |
|  |  |  |  |  |
| Ability | 3.464\*\*\* | 3.436\*\*\* | 3.393\*\*\* | 3.351\*\*\* |
|  | (0.160) | (0.157) | (0.157) | (0.156) |
|  |  |  |  |  |
| Preference Alignment | -1.637 |  |  |  |
|  | (1.135) |  |  |  |
|  |  |  |  |  |
| Preference Change |  | 3.249\* |  |  |
|  |  | (1.436) |  |  |
|  |  |  |  |  |
| Perceived Fairness |  |  | 0.433 |  |
|  |  |  | (0.708) |  |
|  |  |  |  |  |
| Acceptance of |  |  |  | 0.874 |
|  Outcomes |  |  |  | (0.529) |
|  |  |  |  |  |
| Final Preference |  |  |  | 1.084 |
|  Matches Team |  |  |  | (1.353) |
|  |  |  |  |  |
| Controls? | Yes | Yes | Yes | Yes |
|  |  |  |  |  |
| Constant | 45.57\*\*\* | 44.64\*\*\* | 44.10\*\*\* | 43.43\*\*\* |
|  | (5.354) | (5.291) | (6.491) | (5.401) |
| Observations | 570 | 570 | 570 | 570 |

Standard errors in parentheses

+ *p* < 0.1, \* *p* < 0.05, \*\* *p* < 0.01, \*\*\* *p* < 0.001

Note: All the models in this table take the same form as Appendix E, Model 3, other than the addition of each intermediate outcome.

**Appendix K – Explanatory Note: Deviations from PAP & Multiple Hypothesis Testing**

The experimental design and primary analysis described in this paper were pre-registered as a Pre-Analysis Plan (PAP) in the EGAP (Evidence in Governance and Politics) repository [*Link to PAP redacted for blind review*].The main analysis described in the PAP uses tasks completed as the main outcome variable, with treatment assignment as the explanatory variables, and ability as a key control to distinguish between the behavior change and decision quality mechanisms. That main analysis provides the results presented in Figure 2 and Appendices C & D, and described under the subheading “Primary Analysis”.This was the primary motivation for the design of the study, and can be considered confirmatory. All other results should be considered exploratory.

The PAP also lists some secondary analyses, but explicitly classifies them as exploratory in nature. These were included in the hopes that they might help further explain the results (unknown at the time) of the primary analysis. As a result, after seeing the results of the primary analysis, some of the exploratory analyses put forth in the PAP were not ultimately included, because their theoretical relevance was based on the behavior change (effort) mechanism exclusively. These pre-registered exploratory analyses are described under the subheading “Exploratory Analysis” and presented in Figure 4 and Appendix G. In addition, one additional exploratory analysis (conducted with the team as the unit of analyses) was conducted in order to help further clarify the decision quality mechanism. That additional group-level analysis was not included in the PAP (which involved only individual-level analyses), but was added after I realized how helpful it would be in helping to clarify the main result. The table below summarizes differences between exploratory analyses proposed in the PAP and those included in the analysis:

|  |  |  |  |
| --- | --- | --- | --- |
| **Exploratory Outcome Variable** | **Theoretical Justification** | **Appears in PAP?** | **Appears in Final Analysis?** |
| Decision Quality(group-level best choice) | Helps explain Decision Quality mechanism | No | Yes |
| Preference Alignment | Helps explain Decision Quality mechanism | Yes | Yes |
| Preference Change | Helps explain both Effort and Decision Quality | Yes | Yes |
| Acceptance of Outcomes (Agreement with Results) | Helps explain both Effort and Decision Quality  | Yes | Yes |
| Perceived Fairness | Helps explain both Effort and Decision Quality | Yes | Yes |
| Belief in Teammate Effort | Helps explain Effort mechanism | Yes | No |
| Complementary Public Good Investment | Related to Effort mechanism | Yes | No |
| Willingness to Participate Again | Related to Effort mechanism | Yes | No |

Finally, the results presented under the subheading “Suggestive Findings” were conducted in a subsequent round of analysis, upon the advice of a discussant, after the main pre-registered analyses had already been presented. Thus, the additional (not pre-registered) analyses presented in Appendix H & J are presented as merely suggestive, but are included in the paper for their value in pointing to fruitful areas of future research.

**Appendix L – Adjustments for Multiple Hypothesis Testing**

Given that multiple exploratory analyses were conducted, subsequent to the pre-registered main analysis, multiple hypothesis testing may be a concern for these exploratory results. Below I apply a sequential Bonferroni correction procedure (Rice 1989, Holm 1979) to the exploratory results presented in this paper. Erring on the side of caution, I use 8 as the number of tests, which includes all exploratory outcomes mentioned in either the PAP or the final paper, even though some were not warranted given the results of the main analysis (as described in Appendix K). The table below shows the results of the Bonferroni-corrected significance tests, which do not change any of the results presented in the main text of the paper.

|  |  |  |  |
| --- | --- | --- | --- |
| **Exploratory Outcome Variable** | **p-value (deliberation coefficient)** | **Sequential Bonferroni-corrected** ⍺ **value** | **Statistically significant after adjustment?** |
| Decision Quality(group-level best choice) | p=0.001 | 0.05/8 = 0.00625 | Yes |
| Acceptance of Outcomes (Agreement with Results) | p=0.001 | 0.05/7 = 0.00714 | Yes |
| Perceived Fairness | p=0.002 | 0.05/6 = 0.00833 | Yes |
| Preference Change | p=0.003 | 0.05/5 = 0.01 | Yes |
| Preference Alignment | p=0.009 | 0.05/4 = 0.0125 | Yes |

Just as the exploratory analyses were subsequent to and conditional on the findings of the primary analysis, some further suggestive results are presented in the paper which are subsequent to and conditional on the findings of the exploratory analyses. These “suggestive findings,” presented under the subheading of that name and in Appendix H & J, are subjected to corrections for multiple hypothesis testing in the table below:

|  |  |  |
| --- | --- | --- |
|  | **Effect on Contributions****(Decision Quality)** | **Effect on Effort** **(Contributions w/control for ability)** |
| **Explanatory Variable** | **p-value**  | **Bonferroni-corrected** ⍺ **values** | **Significant after adjustment?** | **p-value** | **Bonferroni-corrected** ⍺ **values** | **Significant after adjustment?** |
| Preference Alignment | p=0.000 | 0.05/4 = 0.0125 | Yes | p=0.152 | 0.05/2 = 0.025 | No |
| Acceptance of Outcomes  | p=0.021 | 0.05/3 = 0.0167 | No | p=0.101 | 0.05/3 = 0.0167 | No |
| Perceived Fairness | p=0.184 | 0.05/2 = 0.025 | No | p=0.542 | 0.05/1 = 0.05 | No |
| Preference Change | p=0.991 | 0.05/1 = 0.05 | No | p=0.026 | 0.05/4 = 0.0125 | No |

There is one key difference between these results and what is presented in the main text of the paper based on the simple significance test in Appendix J: The effect of preference change on effort does not hold up to corrections for multiple hypothesis testing. As emphasized in the main text, this result should thus be viewed as having only weak support, and future confirmatory research would be needed in order to make any conclusive statements about them.