

CUT AFTER THIS FROM PAPER
(INCLUDED IN COMPILE TO ENSURE SECTION REFERENCES
APPEAR PROPERLY IN MAIN TEXT)

Online Appendix for “Noise, Cognitive Function, and Worker Productivity” by
Joshua T. Dean

SUPPLEMENTARY FIGURES FOR ONLINE PUBLICATION

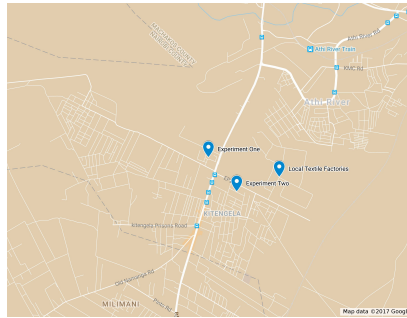


FIGURE A1. EXPERIMENT LOCATIONS AND SURROUNDING AREA

Note: This figure shows a map of the experiment locations and recruitment location in Kitengela, Kenya just outside of Nairobi. The experiment sites are less than a mile apart and close to the recruitment site at the gates of the local textile factories.



FIGURE A2. NOISE GENERATING ENGINE

Note: This figure shows a picture of the car engine used to generate noise in experiment one. The engine was borrowed from an automotive mechanic training facility at the TDC. A car engine was chosen as the noise source because it is representative of important sources of noise pollution – traffic noise and large industrial machines – and leaves minimal room for experimenter manipulation.

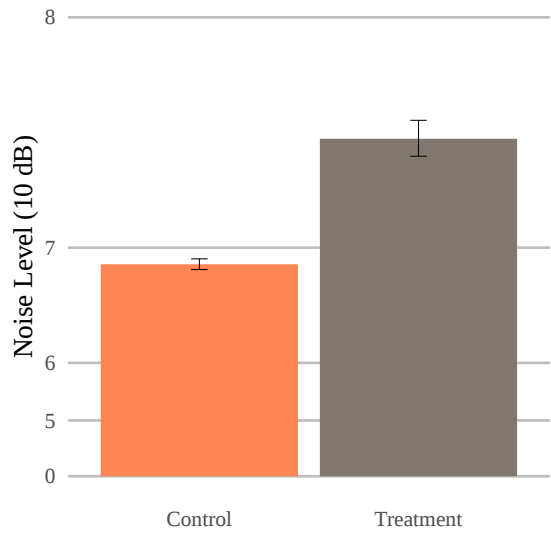


FIGURE A3. NOISE LEVEL BY TREATMENT STATUS

Note: This figure shows the average noise level in treatment and control. A noise increase of 10 dB is perceived as twice as loud by the human ear. Thus, for interpretability, all noise levels are reported as 10s of decibels and the y-axis is shown on a log scale. Treatment increased the noise level by the same amount as replacing a dishwasher running in the background with a vacuum cleaner.

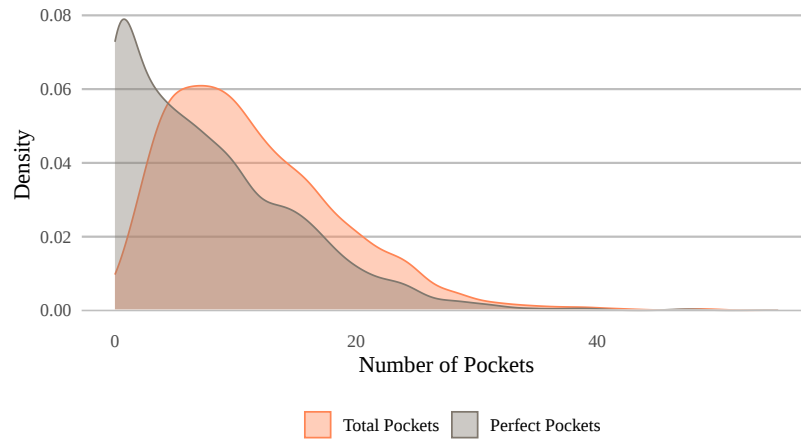


FIGURE A4. OUTPUT DENSITY

Note: This figure shows the density of the number of total and perfect pockets created in experiment one. The distribution is significantly skewed, but has zeros. Thus, to increase power I present the inverse hyperbolic sine-transformed versions of the outcome variables.

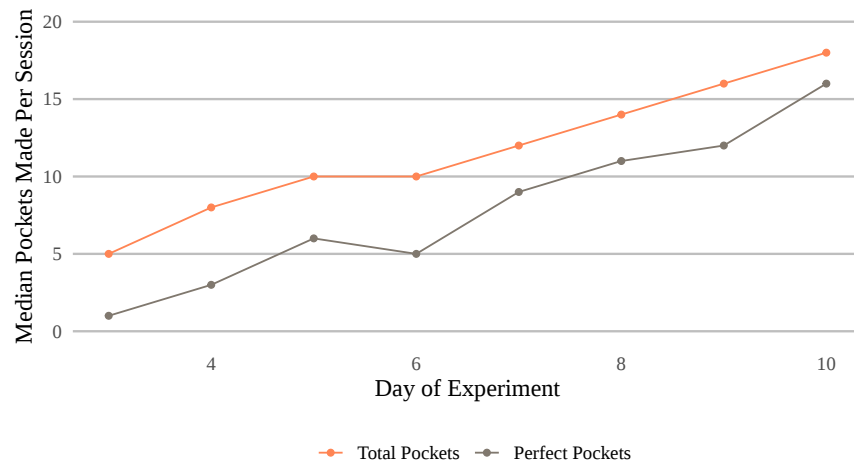


FIGURE A5. OUTPUT TRENDS

Note: This figure shows the median number of total and perfect pockets created in experiment one per day. The figure demonstrates the floor effect issues with the perfect pockets outcome during the first week.

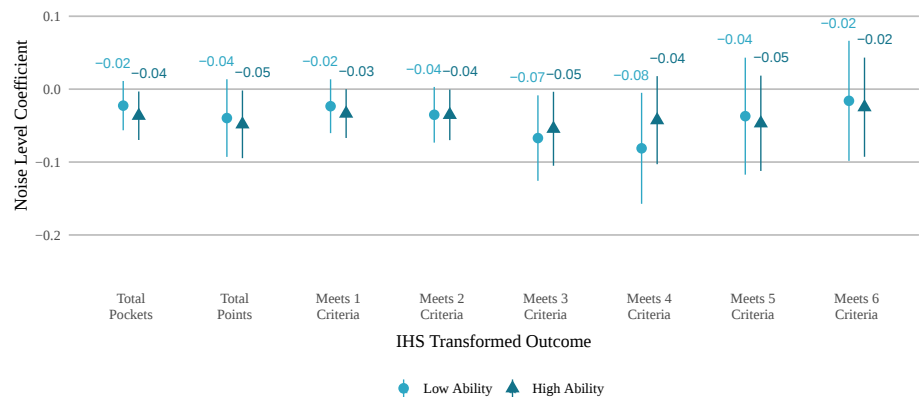


FIGURE A6. TREATMENT EFFECTS BY ABILITY

Note: This figure shows coefficient estimates and their 95% confidence intervals from a regression of productivity outcome variables on the treatment indicator, wage, session, person, and room fixed effects with standard errors clustered at the room by session level. Before estimation, within each session the sample was split by median performance in other control sessions. Treatment effects were estimated separately for the two groups in a stacked regression. The results show that the treatment effect is relatively constant across ability levels.

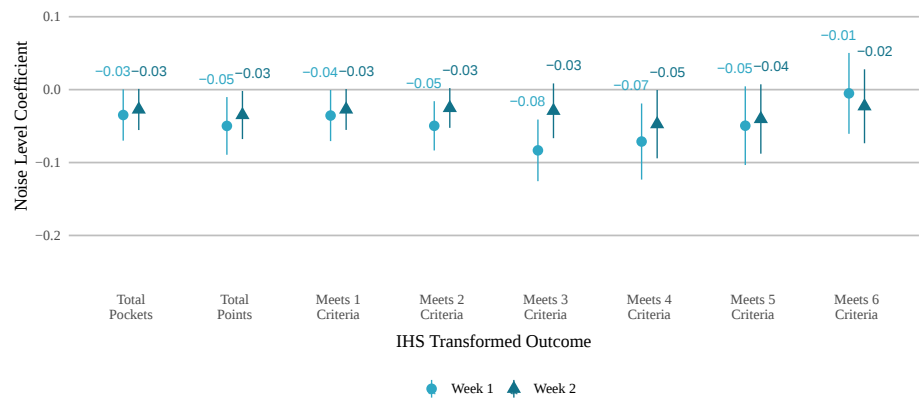


FIGURE A7. TREATMENT EFFECTS OVER TIME

Note: This figure shows coefficient estimates and their 95% confidence intervals from a regression of productivity outcome variables on the treatment indicator, wage, session, person, and room fixed effects with standard errors clustered at the room by session level. Before estimation, the sample was split into two groups by week. Treatment effects were estimated separately for the two groups in a stacked regression. The results show that the treatment effect is relatively constant across weeks.

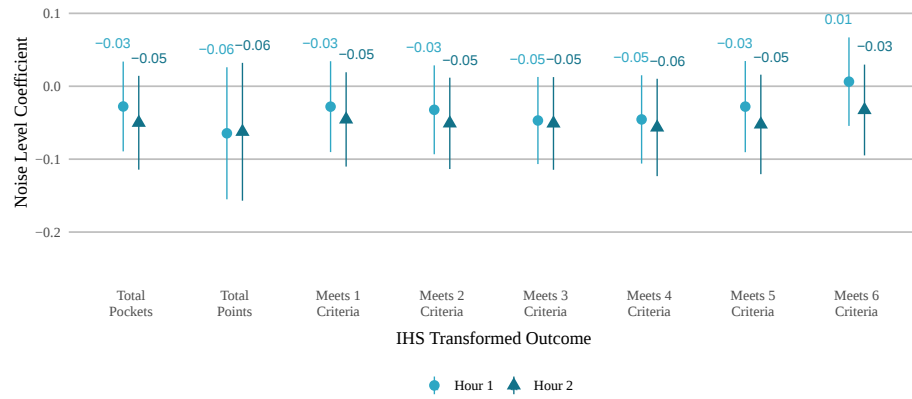


FIGURE A8. TREATMENT EFFECTS BY HOUR OF SESSION

Note: This figure shows coefficient estimates and their 95% confidence intervals from a regression of productivity outcome variables on the treatment indicator, wage, session, person, and room fixed effects with standard errors clustered at the room by session level. Before estimation the sample was split into two groups by hour within each session. Treatment effects were estimated separately for the two groups in a stacked regression. The results show that the treatment effect is relatively constant across hours.

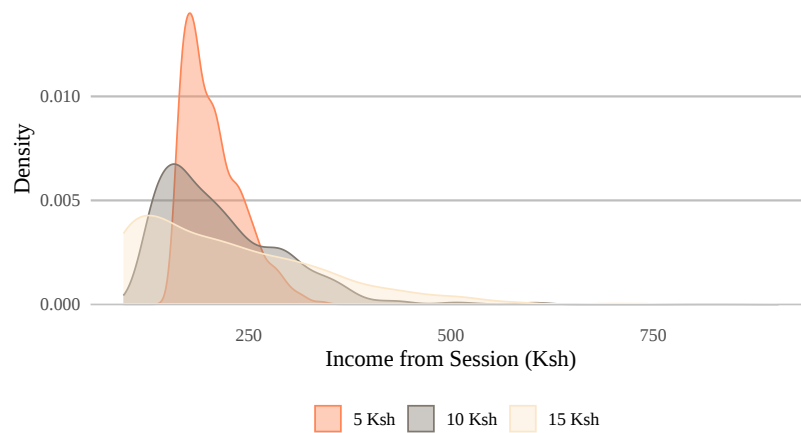


FIGURE A9. INCOME DENSITY BY PIECE RATE

Note: This figure shows the density of the earned income by piece rate in experiment one. Some subjects earned substantially more in the 15 Ksh condition which may explain the lack of impact of the piece rate on productivity.

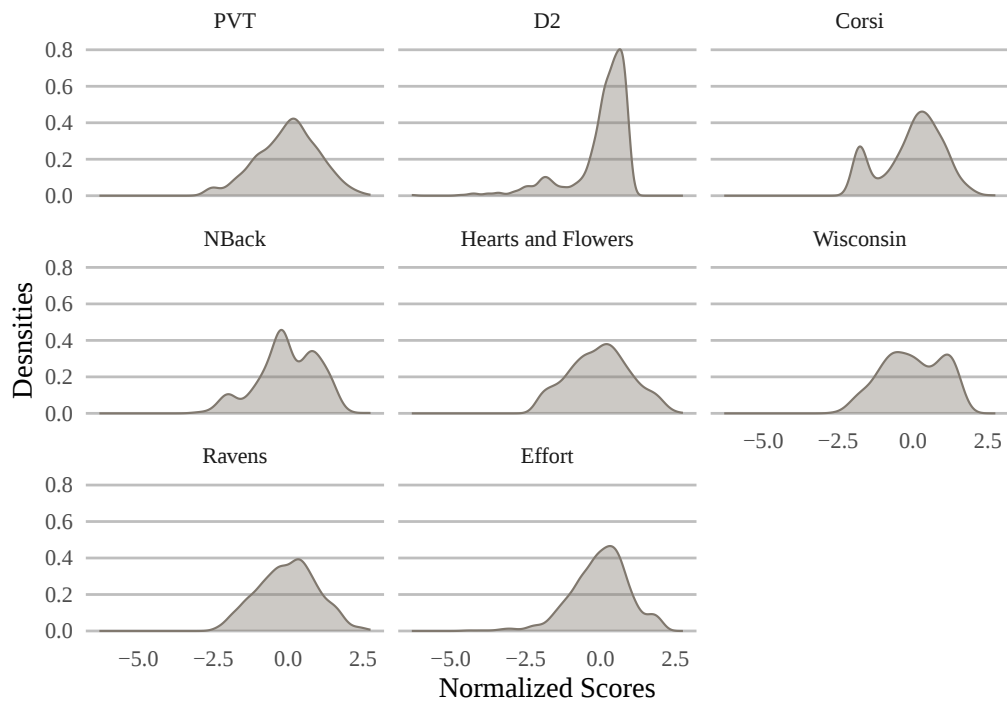


FIGURE A10. NORMALIZED TEST SCORE VARIATION

Note: This figure shows the density of the normalized scores for each measured outcome in experiment two. The results show that with the exception of d2, all metrics generate good variation and do not appear susceptible to ceiling or floor effects.

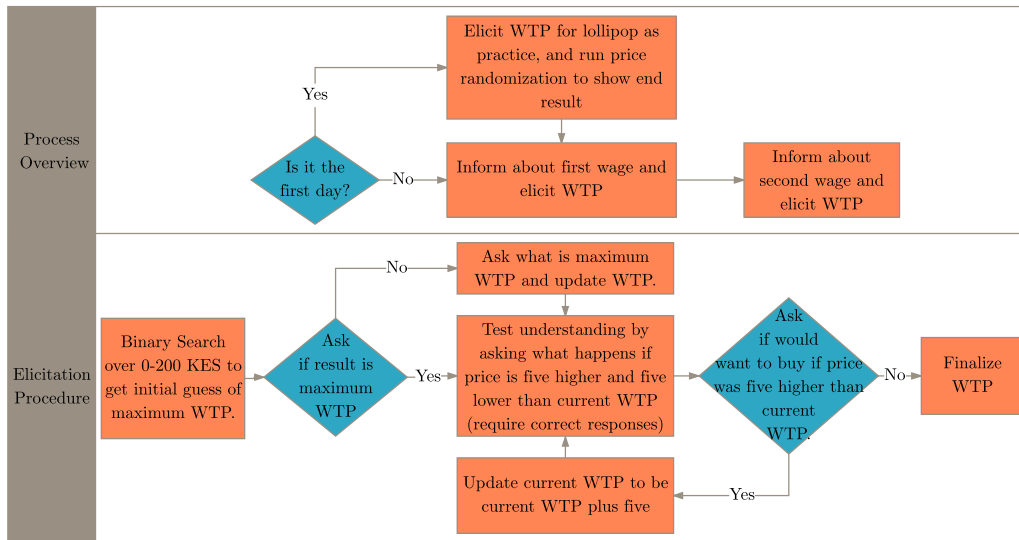


FIGURE A11. WILLINGNESS TO PAY OVERVIEW AND ELICITATION PROCEDURE

Note: The first panel of this figure shows the overview of the process of eliciting willingness to pay including when practices were conducted, and how information was timed with the elicitation. The second panel is a detailed view of how willingness to pay was elicited using a modified version of the procedure outlined in Berry et al. (2015).

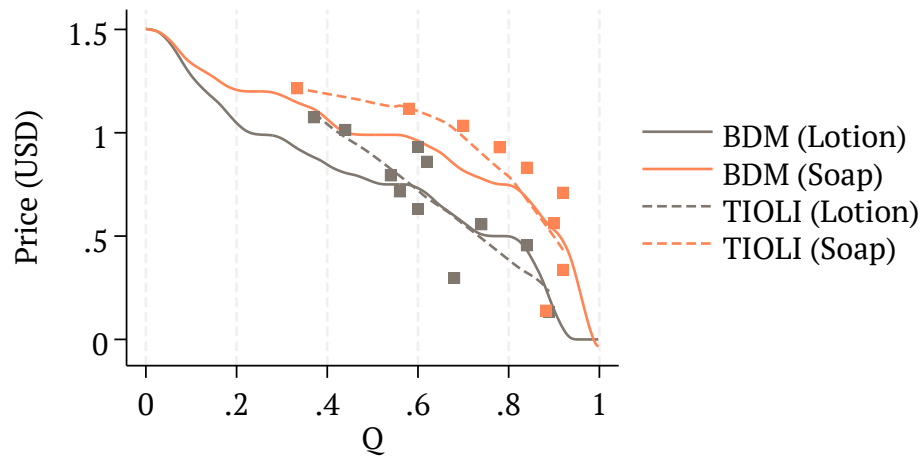
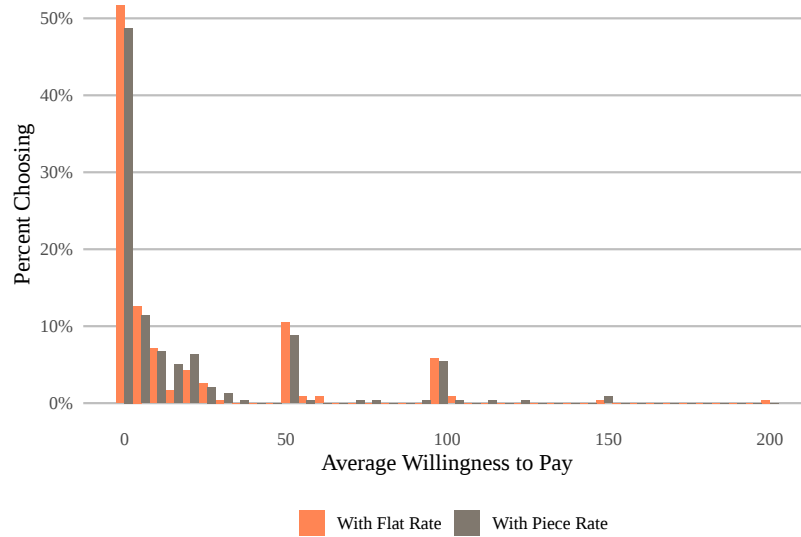
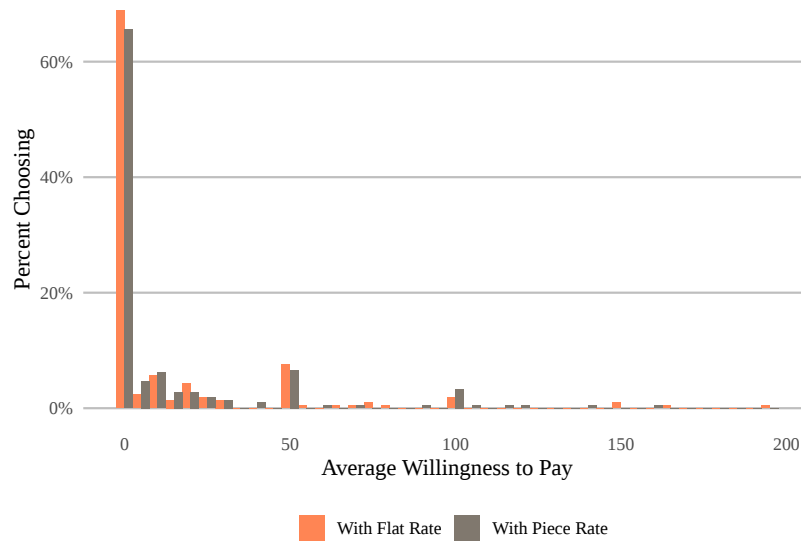


FIGURE A12. COMPARISON OF BDM AND TIOLI OFFERS FROM BERKOUWER AND DEAN (2022)

Note: The figure plots the BDM procedure validation results from Berkouwer and Dean (2022a). The authors use the same BDM elicitation procedure as this paper in a sample of 1000 households from Nairobi. However, instead of only having a single practice good, the authors randomly assign participants to practice the BDM for either lotion or soap. For the good not chosen as the practice good, the respondent is given a take it or leave it (TIOLI) offer at a randomly chosen price. This allows comparison of the demand curves elicited by the BDM script employed by the authors to those elicited by more traditional purchase decisions. The results show the BDM gives comparable demand curves.



A. EXPERIMENT ONE



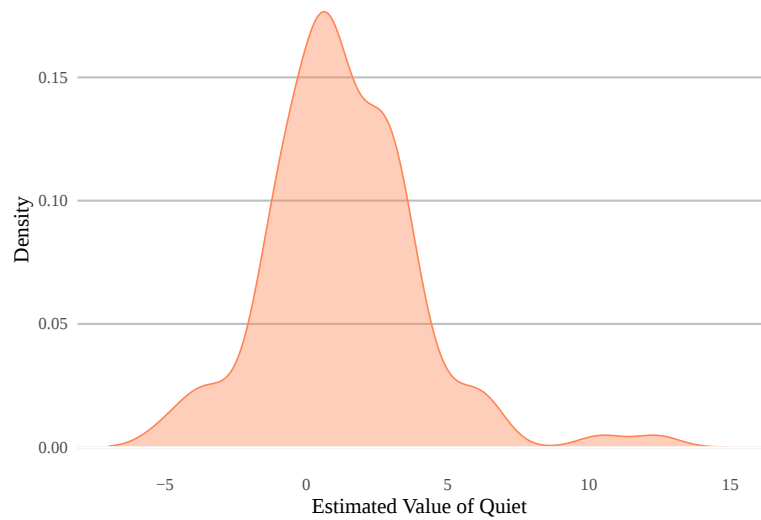
B. EXPERIMENT TWO

FIGURE A13. AVERAGE WILLINGNESS TO PAY BY COMPENSATION

Note: The figure shows the distribution of willingness to pay in order to work in the quiet room by whether the respondent was facing a piece rate or a flat rate. The first panel shows the results for the first experiment, while the second shows the results for the second. The results show that willingness to pay for quiet is generally quite low.



A. EXPERIMENT ONE



B. EXPERIMENT TWO

FIGURE A14. PRODUCTIVE VALUE OF QUIET

Note: The figure shows the densities of the hierarchical linear models' predicted income gains from working in the quiet room. The first panel shows the density from the first experiment and the second shows the density from the second. The model was used to obtain better predictions of the within-person treatment effect than would be obtained from simply taking the difference between treatment and control performance within person.

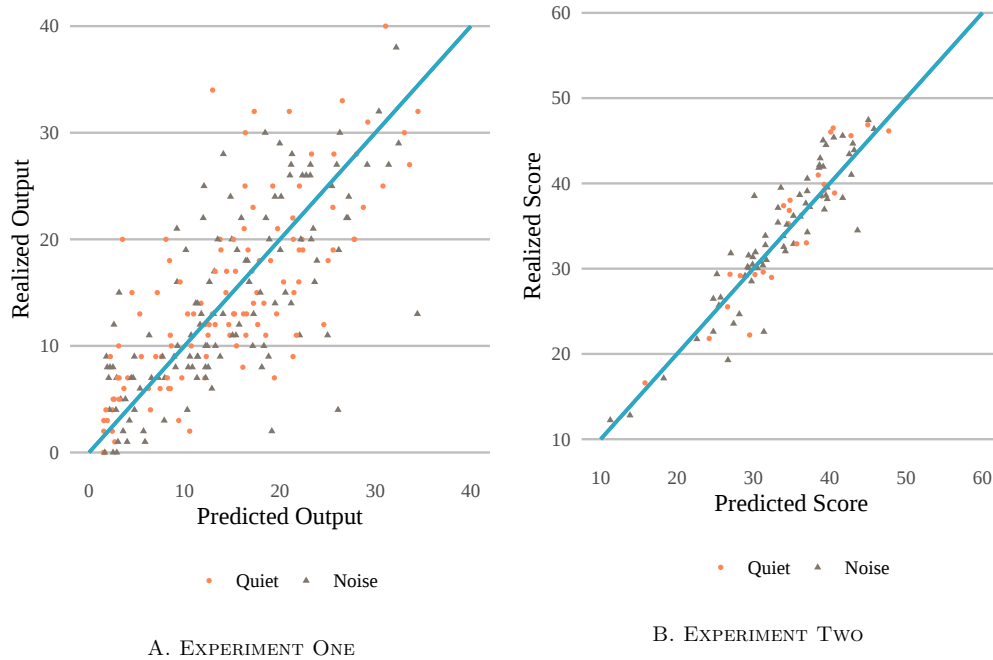
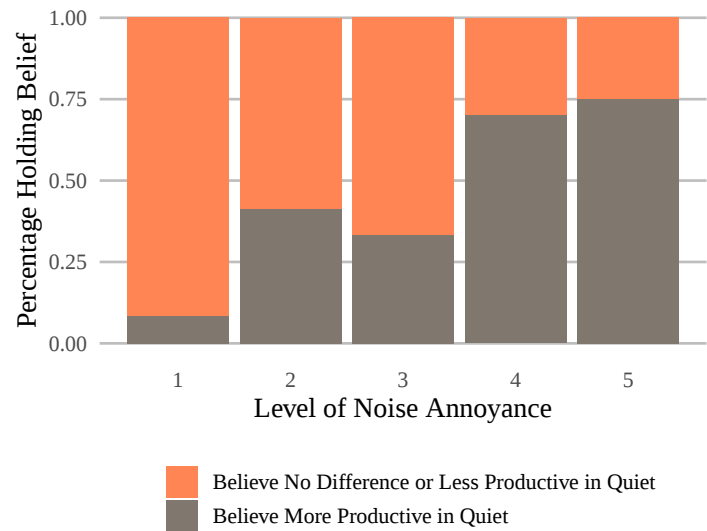
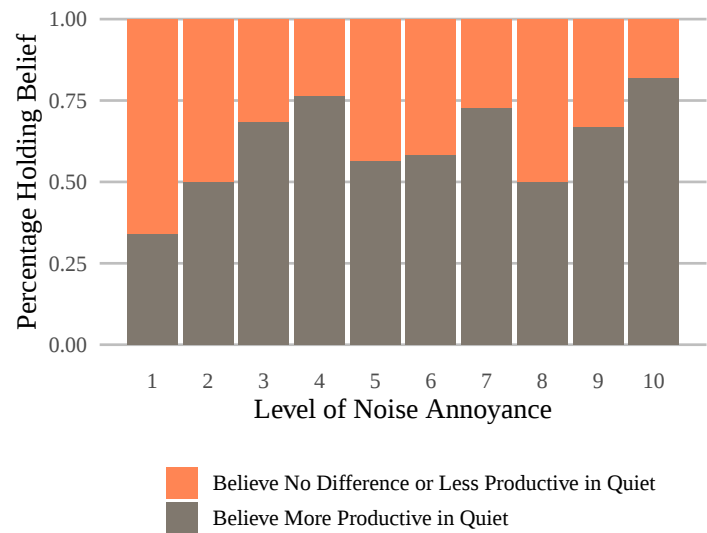


FIGURE A15. MODEL FIT

Note: The figure plots the hierarchical linear model's predictions of the respondents' output and scores in the willingness to pay sessions on the x-axis against the respondents' realized output and scores on the y-axis. Values are plotted separately depending on whether the individual ended up in quiet or in noise based on their willingness to pay. The solid line shows the 45 degree line or perfect prediction.



A. EXPERIMENT 1



B. EXPERIMENT 2

FIGURE A16. BELIEFS AND ANNOYANCE

Note: This figure shows the proportion of individuals who believe they are more productive in quiet for each level of stated annoyance with the noise level. The high level of correlation provides suggestive evidence that individuals do not actually understand the impact of noise on their productivity and are instead substituting in their annoyance level.

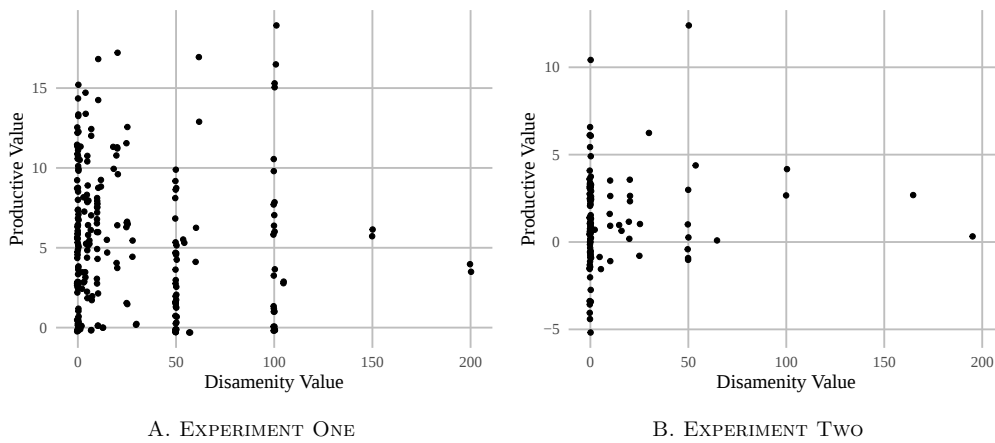


FIGURE A17. CORRELATION BETWEEN AMENITY AND PRODUCTIVE VALUE OF QUIET

Note: This figure plots the respondent's willingness to pay for quiet when facing a flat rate compensation scheme and what the model predicts is their productive value of quiet. Points are jittered to avoid overplotting. The results show that the two are essentially uncorrelated. This suggests that if respondents neglect the productive impact and sort simply on annoyance, this will change the composition of workers sorting into working in noise.

SUPPLEMENTARY TABLES FOR ONLINE PUBLICATION

TABLE B1—SAMPLE SUMMARY STATISTICS

	Experiment 1	Experiment 2	Total
Female	0.641 (0.482)	0.521 (0.501)	0.566 (0.496)
Age	28.84 (6.791)	26.07 (6.748)	27.11 (6.885)
High School or More	0.516 (0.502)	0.690 (0.464)	0.625 (0.485)
Typical Daily Wage	677.2 (725.8)	548.6 (633.4)	597.0 (671.5)
Days Worked Last Week	2.188 (2.528)	1.235 (2.130)	1.592 (2.330)
More Annoyed by Noise than Others	0.258 (0.439)	0.305 (0.462)	0.287 (0.453)
Participants	128	213	341

Note: This table presents summary statistics for each experiment sample. The main entries are the means of the variable in each row. Standard deviations are in parentheses below. The samples are relatively similar on demographic terms.

TABLE B2—EXPERIMENT ONE BALANCE

	(1) Female	(2) Age	(3) High School or More	(4) Typical Daily Wage	(5) Days Worked Last Week	(6) More Annoyed by Noise than Others
Treatment Mean	0.641	28.812	0.518	684.163	2.192	0.257
Control Mean	0.640	28.861	0.513	670.259	2.183	0.259
Two-Sided P-Value	0.962	0.889	0.846	0.673	0.915	0.936
Normalized Difference	0.006	−0.007	0.016	0.021	−0.006	−0.006
Observations	2560	2540	2560	2560	2560	2560
Room-Session Clusters	160	160	160	160	160	160
Participants	128	127	128	128	128	128

Note: This table assesses the balance of sample characteristics between treatment and control sessions. The first two rows display the average of the variable indicated in the column for individuals observed in treatment and control sessions, respectively. Row three shows the p-value from a regression of the variable on a treatment indicator with standard errors clustered at the room by session level. The normalized difference is the difference between the treatment and control means divided by the square root of the average of the treatment and control variances as defined by Imbens and Rubin (2015). The sample observed in treatment and control are almost identical. This is a result of the within-person randomization. The only reason balance does not hold exactly is due to small levels of attrition.

TABLE B3—EXPERIMENT ONE POISSON REGRESSIONS

	(1) Total Pockets	(2) Total Points Earned	(3) Pockets Meeting 1 Criterion	(4) Pockets Meeting 2 Criteria	(5) Pockets Meeting 3 Criteria	(6) Pockets Meeting 4 Criteria	(7) Pockets Meeting 5 Criteria	(8) Pockets Meeting 6 Criteria
<i>Marginal Effects</i>								
Treatment	−0.3094** (0.1210)	−1.7523** (0.7287)	−0.2961** (0.1212)	−0.3124*** (0.1206)	−0.3496*** (0.1290)	−0.3392** (0.1343)	−0.3000** (0.1381)	−0.1671 (0.1309)
<i>Poisson Coefficients</i>								
Treatment	−0.0269** (0.0105)	−0.0284** (0.0118)	−0.0258** (0.0106)	−0.0274*** (0.0106)	−0.0320*** (0.0118)	−0.0326** (0.0129)	−0.0306** (0.0141)	−0.0216 (0.0169)
Wage FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Session FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Person FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Room FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control Mean	11.527	61.849	11.493	11.421	10.965	10.446	9.832	7.692
Control Median	10	54	10	10	10	9	9	6
Observations	2447	2447	2447	2447	2447	2447	2447	2447
Room-Session Clusters	160	160	160	160	160	160	160	160
Participants	128	128	128	128	128	128	128	128

Note: This table shows the marginal effects and coefficients from poisson regressions of productivity outcome variables on a treatment indicator, wage, session, person, and room fixed effects with standard errors clustered at the room by session level. Respondents in treated rooms (those working with the background noise of a vacuum instead of a dishwasher) made approximately 3% fewer pockets.

TABLE B4—EXPERIMENT ONE FISHER P-VALUES

	Total Pockets	Total Points Earned	Pockets Meeting 1 Criterion	Pockets Meeting 2 Criteria	Pockets Meeting 3 Criteria	Pockets Meeting 4 Criteria	Pockets Meeting 5 Criteria	Pockets Meeting 6 Criteria
IHS Transformed	0.0145	0.017	0.0175	0.0065	0.0045	0.008	0.0555	0.588
Levels	0.0975	0.149	0.1185	0.0985	0.086	0.098	0.169	0.775

Note: This table shows the p-values from randomization inference. Treatment was randomly reassigned 2000 times using the original randomization code. Each outcome variable was then regressed on the reassigned treatment indicator, individual, session, room, and wage fixed effects. The true coefficient was then compared to the distribution of coefficients induced by reassignment in order to generate p-values. The inferences are similar to those reported in the main regressions.

TABLE B5—INTENSITY LEVEL BALANCE

	Experiment 1				Experiment 2			
	(1) Temperature	(2) CO ₂	(3) Humidity	(4) Second Half	(5) Temperature	(6) CO ₂	(7) Humidity	(8) Second Half
Intensity 1	27.17 (0.67)	642.67 (44.03)	39.24 (2.13)	0.50 (0.11)	25.10 (0.77)	894.75 (91.41)	45.87 (2.05)	0.40 (0.14)
Intensity 2	25.94 (0.65)	635.52 (42.63)	44.76 (2.06)	0.88 (0.11)	25.19 (0.88)	848.38 (103.64)	45.77 (2.32)	0.25 (0.16)
Intensity 3	27.78 (0.67)	636.00 (44.03)	39.05 (2.13)	0.38 (0.11)	22.19 (0.77)	872.54 (91.41)	52.26 (2.05)	0.80 (0.14)
Intensity 4	26.51 (0.70)	648.78 (45.58)	41.51 (2.20)	1.00 (0.12)	23.28 (0.88)	853.90 (103.64)	49.24 (2.32)	0.25 (0.16)
Intensity 5	26.16 (0.65)	653.34 (42.63)	44.17 (2.06)	0.62 (0.11)	23.54 (0.82)	909.17 (96.95)	48.13 (2.17)	1.00 (0.16)
Intensity 6	27.59 (0.67)	624.70 (44.03)	41.68 (2.13)	0.38 (0.11)	24.02 (0.73)	1003.74 (86.72)	50.90 (1.94)	0.60 (0.14)
Intensity 7	26.35 (0.70)	565.96 (45.58)	44.00 (2.20)	0.14 (0.12)	23.30 (0.82)	847.84 (96.95)	50.98 (2.17)	0.75 (0.16)
Intensity 8	26.96 (0.70)	617.94 (45.58)	42.52 (2.20)	0.75 (0.11)	26.37 (0.73)	897.77 (86.72)	38.83 (1.94)	0.40 (0.14)
Intensity 9	26.11 (0.65)	654.69 (42.63)	43.67 (2.06)	0.62 (0.11)	26.74 (0.82)	1033.98 (96.95)	40.79 (2.17)	0.00 (0.16)
Intensity 10	25.93 (0.70)	647.63 (45.58)	43.30 (2.20)	0.43 (0.12)	24.75 (0.82)	849.89 (96.95)	47.85 (2.17)	0.50 (0.16)
Room-Session Observations	149	149	149	154	84	84	84	88

Note: This table shows the observable differences in sessions by treatment intensity. The main entries in each row show the means of the variables listed at the top for sessions of a given decile of intensity. The standard errors of the means are in parentheses and are clustered at the session level. The results show no clear relationship between treatment intensity and any observable characteristic.

TABLE B6—FIRST STAGES

	Experiment 1		Experiment 2	
	(1) Noise Level	(2) Noise Level	(3) Noise Level	(4) Noise Level
Treatment	0.6745*** (0.0247)		0.9403*** (0.0368)	
Treated with Intensity 1		0.1529*** (0.0294)		0.3419*** (0.0363)
Treated with Intensity 2		0.3146*** (0.0057)		0.5649*** (0.0102)
Treated with Intensity 3		0.4171*** (0.0113)		0.7394*** (0.0150)
Treated with Intensity 4		0.4942*** (0.0041)		0.8635*** (0.0127)
Treated with Intensity 5		0.5725*** (0.0067)		0.9210*** (0.0075)
Treated with Intensity 6		0.6841*** (0.0071)		1.0675*** (0.0087)
Treated with Intensity 7		0.7620*** (0.0056)		1.1137*** (0.0052)
Treated with Intensity 8		0.8572*** (0.0117)		1.1803*** (0.0077)
Treated with Intensity 9		1.0970*** (0.0194)		1.2953*** (0.0239)
Treated with Intensity 10		1.4459*** (0.0460)		1.3966*** (0.0091)
Control Mean	6.892	6.892	7.202	7.202
F-Statistic	745	6699	651	23347
Observations	2512	2512	762	762
Room-Session Clusters	157	157	88	88
Participants	128	128	211	211

Note: This table reports coefficients of a regression of the noise level on the excluded instruments with standard errors clustered at the room by session level. Columns 1 and 3 use a single indicator for being in a treatment session. Columns 2 and 4 use separate indicators for each level of treatment intensity. F-statistics are for a joint test that the coefficients are zero. The results show that all instruments generate a strong first stage.

TABLE B7—IV EFFECT OF NOISE ON PRODUCTIVITY – TREATMENT INTENSITY INSTRUMENTS

	(1) Total Pockets	(2) Total Points Earned	(3) Pockets Meeting 1 Criterion	(4) Pockets Meeting 2 Criteria	(5) Pockets Meeting 3 Criteria	(6) Pockets Meeting 4 Criteria	(7) Pockets Meeting 5 Criteria	(8) Pockets Meeting 6 Criteria
<i>Inverse Hyperbolic Sine Transformation</i>								
Noise Level	−0.0646*** (0.0152)	−0.0751*** (0.0180)	−0.0651*** (0.0151)	−0.0661*** (0.0150)	−0.0905*** (0.0198)	−0.0975*** (0.0240)	−0.0861*** (0.0247)	−0.0549** (0.0252)
<i>Levels</i>								
Noise Level	−0.5253*** (0.1696)	−2.9572*** (0.9632)	−0.5192*** (0.1695)	−0.5187*** (0.1691)	−0.5466*** (0.1718)	−0.5671*** (0.1723)	−0.5205*** (0.1653)	−0.2851* (0.1492)
Wage FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Session FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Person FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Room FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control Mean-IHS	2.924	4.487	2.918	2.901	2.775	2.645	2.529	2.163
Control Median-Levels	10	54	10	10	10	9	9	6
Observations	2400	2400	2400	2400	2400	2400	2400	2400
Room-Session Clusters	157	157	157	157	157	157	157	157
Participants	128	128	128	128	128	128	128	128

Note: This table shows estimates from two-stage least squares regression of productivity outcome variables on the noise level, wage, session, person, and room fixed effects with standard errors clustered at the room by session level. The noise level is instrumented by a set of treatment indicators interacted with the session in order to capture variation in treatment intensity. The first panel shows the results for the inverse hyperbolic sine-transformed outcomes, while the second panel shows the results for the untransformed outcomes.

TABLE B8—QUALITY RESPONSE

	(1) Proportion Meeting 1 Criterion	(2) Proportion Meeting 2 Criteria	(3) Proportion Meeting 3 Criteria	(4) Proportion Meeting 4 Criteria	(5) Proportion Meeting 5 Criteria	(6) Proportion Meeting 6 Criteria
<i>Reduced Form Effect of Treatment</i>						
Treatment	0.0000 (0.0019)	-0.0025 (0.0035)	-0.0113* (0.0058)	-0.0133* (0.0073)	-0.0063 (0.0075)	0.0054 (0.0074)
<i>2SLS Effect of Noise - Treatment Indicator Instrument</i>						
Noise Level	-0.0009 (0.0024)	-0.0018 (0.0047)	-0.0117 (0.0079)	-0.0134 (0.0101)	-0.0030 (0.0105)	0.0131 (0.0107)
<i>2SLS Effect of Noise - Treatment Intensity Instruments</i>						
Noise Level	-0.0009 (0.0021)	0.0002 (0.0044)	-0.0137* (0.0078)	-0.0174* (0.0097)	-0.0100 (0.0104)	0.0022 (0.0099)
Wage FE	Yes	Yes	Yes	Yes	Yes	Yes
Session FE	Yes	Yes	Yes	Yes	Yes	Yes
Person FE	Yes	Yes	Yes	Yes	Yes	Yes
Room FE	Yes	Yes	Yes	Yes	Yes	Yes
Control Mean	0.995	0.980	0.904	0.833	0.765	0.574
Observations	2389	2389	2389	2389	2389	2389
Room-Session Clusters	157	157	157	157	157	157
Participants	128	128	128	128	128	128

Note: This table shows the impact of treatment on the proportion of pockets meeting or exceeding each quality threshold. The first panel shows the estimates from a regression of the proportions on a treatment indicator, wage, session, person, and room fixed effects with standard errors clustered at the room by session level. The second panel shows the estimates from a two-stage least squares regression of the proportions on the noise level, wage, session, person, and room fixed effects with standard errors clustered at the room by session level and the noise level instrumented with a treatment indicator. The third panel shows the estimates from a two-stage least squares regression of the proportions on the noise level, wage, session, person, and room fixed effects with standard errors clustered at the room by session level and the noise level instrumented by a set of treatment intensity indicators in order to capture variation in treatment intensity. The results show that there does not appear to have been a quality response to treatment.

TABLE B9—COMPARISON WITH OTHER ENVIRONMENTAL EFFECTS

Source	Setting	Stimulus	Stimulus Change	Productivity Effect
This Paper	Textile Production	Noise	Perceived Doubling	5%
Adhvaryu et al. (2016)	Textile Production	Temperature	0.81σ	1.3%
Zivin and Neidell (2012)	Agricultural Labor	Ozone Pollution	0.76σ	5.5%
Chang et al. (2016)	Call Center	Air Pollution	15%	0.35%
Chang et al. (2016)	Factory	Air Pollution	1σ	8%
He et al. (2019)	Manufacturing	Air Pollution	0.2σ	0%

Note: This table compares the effects of noise on productivity estimated in this experiment with the effects of other environmental factors on productivity estimated in the literature. While we should be cautious in interpreting results from different experiments with different kinds of environmental changes on different types of tasks, it appears the effects estimated in this paper are similarly sized to other environmental factors.

TABLE B10—HETEROGENEITY BY TYPICAL NOISE EXPOSURE

	(1) Total Pockets	(2) Total Points Earned	(3) Pockets Meeting 1 Criterion	(4) Pockets Meeting 2 Criteria	(5) Pockets Meeting 3 Criteria	(6) Pockets Meeting 4 Criteria	(7) Pockets Meeting 5 Criteria	(8) Pockets Meeting 6 Criteria
Treatment	−0.0312*** (0.0118)	−0.0426*** (0.0134)	−0.0315*** (0.0117)	−0.0381*** (0.0114)	−0.0569*** (0.0146)	−0.0604*** (0.0179)	−0.0450** (0.0182)	−0.0140 (0.0190)
Treatment × Typical Noise Exposure	0.0043 (0.0130)	0.0045 (0.0209)	0.0012 (0.0138)	0.0109 (0.0178)	−0.0063 (0.0253)	0.0302 (0.0253)	−0.0011 (0.0254)	0.0115 (0.0307)
Typical Noise Exposure	−1.4821*** (0.1626)	−1.5544*** (0.2477)	−1.4930*** (0.1612)	−1.4729*** (0.1527)	−1.4242*** (0.2070)	−1.4060*** (0.2353)	−1.4216*** (0.2895)	−1.9108*** (0.4348)
Wage FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Session FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Person FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Room FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control Mean	2.924	4.487	2.918	2.901	2.775	2.645	2.529	2.163
Observations	2447	2447	2447	2447	2447	2447	2447	2447
Room-Session Clusters	160	160	160	160	160	160	160	160
Participants	128	128	128	128	128	128	128	128

Note: This table shows estimates from an ordinary least squares regression of the inverse hyperbolic sine-transformed productivity outcome variables on a treatment indicator, a z-score of typical noise exposure, their interaction, wage, session, person, and room fixed effects with standard errors clustered at the room by session level. The typical noise exposure z-score is constructed from baseline data on the living conditions of respondents and their immediately previous paid employment. Given that respondents were all recruited while seeking employment at the same factories and all live in nearby locations, we should be cautious interpreting these results, but they do not suggest any heterogeneity in effect by prior noise exposure.

TABLE B11—LAGGED TREATMENT RESULTS

	(1) Total Pockets	(2) Total Points Earned	(3) Pockets Meeting 1 Criterion	(4) Pockets Meeting 2 Criteria	(5) Pockets Meeting 3 Criteria	(6) Pockets Meeting 4 Criteria	(7) Pockets Meeting 5 Criteria	(8) Pockets Meeting 6 Criteria
Treatment	−0.0163 (0.0172)	−0.0305 (0.0208)	−0.0168 (0.0171)	−0.0302* (0.0159)	−0.0649*** (0.0221)	−0.0554* (0.0282)	−0.0289 (0.0284)	0.0011 (0.0291)
Lagged Treatment	0.0000 (0.0213)	−0.0198 (0.0250)	−0.0043 (0.0214)	−0.0125 (0.0201)	−0.0406 (0.0291)	−0.0077 (0.0384)	0.0149 (0.0390)	0.0194 (0.0428)
Treatment × Lagged Treatment	−0.0504 (0.0354)	−0.0356 (0.0473)	−0.0515 (0.0366)	−0.0344 (0.0353)	0.0101 (0.0447)	−0.0033 (0.0571)	−0.0297 (0.0598)	−0.0023 (0.0637)
Wage FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Session FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Person FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Room FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control Mean	2.924	4.487	2.918	2.901	2.775	2.645	2.529	2.163
Observations	2209	2209	2209	2209	2209	2209	2209	2209
Room-Session Clusters	144	144	144	144	144	144	144	144
Participants	128	128	128	128	128	128	128	128

Note: This table shows estimates from an ordinary least squares regression of the inverse hyperbolic sine-transformed productivity outcome variables on a treatment indicator, a lagged treatment indicator, an interaction of the treatment indicator and the lagged treatment indicator, wage, session, person, and room fixed effects with standard errors clustered at the room by session level. The results are imprecise, but do not suggest that cumulative effects of noise exposure are important.

TABLE B12—DECISION TASK RESULTS

	(1) Net Savings	(2) Bought Bulk	(3) Worked Extra Hour
Treated 1st Session	0.3483 (39.9039)	0.0358 (0.0723)	−0.0241 (0.0942)
Treated 2nd Session	20.1870 (41.6946)	−0.0504 (0.0715)	0.0457 (0.0945)
Treated 3rd Session	10.7555 (33.9952)	−0.0782 (0.0717)	0.0013 (0.0930)
Day FE	Yes	No	No
Person FE	Yes	No	No
Outcome Mean	8.204	0.190	0.525
Outcome SD	694.611	0.394	0.501
Observations	733	126	120
Participant Clusters	128		

Note: This table shows regressions of the outcome variables from three decision tasks conducted at the end of selected days on indicators for whether the respondent was treated in the first, second, or third session on those days. Standard errors are clustered at the individual level. The first column shows the net amount saved (deposits less withdrawals) by respondents in an account with an interest rate of 1% per working day. The second column shows whether an individual chose to buy a 5 kg bag of maize flour when they had the opportunity to buy five 1 kg bags of flour at a lower cost. The final column shows whether the respondents elected to stay for an extra hour and work for a piece rate when offered the chance. All results show that noise exposure during the day does not seem to affect decisions taken later in quiet.

TABLE B13—EFFECT OF PIECE RATE ON PRODUCTIVITY

	(1) Total Pockets	(2) Total Points Earned	(3) Pockets Meeting 1 Criterion	(4) Pockets Meeting 2 Criteria	(5) Pockets Meeting 3 Criteria	(6) Pockets Meeting 4 Criteria	(7) Pockets Meeting 5 Criteria	(8) Pockets Meeting 6 Criteria
<i>Inverse Hyperbolic Sine Transformation</i>								
10 Ksh Piece Rate	0.0270* (0.0158)	0.0336 (0.0207)	0.0245 (0.0162)	0.0267 (0.0171)	0.0238 (0.0233)	0.0290 (0.0283)	0.0513* (0.0299)	0.0407 (0.0329)
15 Ksh Piece Rate	−0.0180 (0.0166)	−0.0239 (0.0233)	−0.0223 (0.0173)	−0.0104 (0.0179)	−0.0040 (0.0250)	−0.0028 (0.0296)	−0.0139 (0.0310)	−0.0120 (0.0341)
<i>Levels</i>								
10 Ksh Piece Rate	0.3352* (0.1719)	1.8285* (1.0068)	0.3199* (0.1722)	0.3328* (0.1711)	0.2738 (0.1745)	0.2986 (0.1824)	0.3547* (0.1824)	0.2487 (0.1847)
15 Ksh Piece Rate	0.1153 (0.1746)	0.9158 (1.0350)	0.1101 (0.1754)	0.1588 (0.1747)	0.1494 (0.1806)	0.2219 (0.1877)	0.1552 (0.1882)	0.1204 (0.1908)
Noise Condition	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Session FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Person FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Room FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
5 Ksh Mean-IHS	2.898	4.450	2.894	2.869	2.728	2.590	2.476	2.129
5 Ksh Median-Levels	10	51	10	10	9	9	8	6
Observations	2447	2447	2447	2447	2447	2447	2447	2447
Participants	128	128	128	128	128	128	128	128

Note: This table shows ordinary least squares regressions of inverse hyperbolic sine-transformed and untransformed productivity outcome variables on piece rate indicators, treatment indicators, session, person, and room fixed effects with robust standard errors. The results demonstrate that increasing the piece rate from 5 to 10 Ksh increased productivity by approximately 3%, but that there was no effect of the 15 Ksh piece rate condition, possibly due to income effects.

TABLE B14—EFFECT OF PIECE RATE ON PRODUCTIVITY IN QUIET

	(1) Total Pockets	(2) Total Points Earned	(3) Pockets Meeting 1 Criterion	(4) Pockets Meeting 2 Criteria	(5) Pockets Meeting 3 Criteria	(6) Pockets Meeting 4 Criteria	(7) Pockets Meeting 5 Criteria	(8) Pockets Meeting 6 Criteria
<i>Inverse Hyperbolic Sine Transformation</i>								
10 Ksh Piece Rate	0.0450* (0.0244)	0.0477 (0.0326)	0.0420* (0.0244)	0.0431* (0.0257)	0.0411 (0.0372)	0.0656 (0.0442)	0.0789* (0.0465)	0.0321 (0.0516)
15 Ksh Piece Rate	0.0018 (0.0236)	0.0217 (0.0328)	−0.0009 (0.0239)	0.0168 (0.0249)	0.0518 (0.0351)	0.0734* (0.0427)	0.0603 (0.0450)	0.0260 (0.0519)
<i>Levels</i>								
10 Ksh Piece Rate	0.6416** (0.2491)	3.4733** (1.4802)	0.6282** (0.2492)	0.6485*** (0.2487)	0.5432** (0.2570)	0.6733** (0.2683)	0.6315** (0.2676)	0.3486 (0.2824)
15 Ksh Piece Rate	0.2405 (0.2595)	1.7745 (1.5464)	0.2367 (0.2607)	0.2959 (0.2603)	0.3472 (0.2671)	0.4769* (0.2766)	0.3095 (0.2772)	0.1084 (0.2947)
Session FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Person FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Room FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
5 Ksh Mean-IHS	2.858	4.389	2.853	2.827	2.671	2.512	2.386	2.043
5 Ksh Median-Levels	9	48	9	9	9	8	7	5
Observations	1223	1223	1223	1223	1223	1223	1223	1223
Participants	128	128	128	128	128	128	128	128

Note: This table shows ordinary least squares regressions of inverse hyperbolic sine-transformed and untransformed productivity outcome variables on piece rate indicators for observations in quiet, session, person, and room fixed effects with robust standard errors.

TABLE B15—EFFECT OF PIECE RATE ON PRODUCTIVITY IN NOISE

	(1) Total Pockets	(2) Total Points Earned	(3) Pockets Meeting 1 Criterion	(4) Pockets Meeting 2 Criteria	(5) Pockets Meeting 3 Criteria	(6) Pockets Meeting 4 Criteria	(7) Pockets Meeting 5 Criteria	(8) Pockets Meeting 6 Criteria
<i>Inverse Hyperbolic Sine Transformation</i>								
10 Ksh Piece Rate	0.0086 (0.0228)	0.0183 (0.0298)	0.0078 (0.0234)	0.0175 (0.0255)	0.0122 (0.0335)	−0.0049 (0.0408)	0.0202 (0.0434)	0.0490 (0.0467)
15 Ksh Piece Rate	−0.0286 (0.0255)	−0.0809** (0.0378)	−0.0389 (0.0278)	−0.0274 (0.0285)	−0.0707* (0.0382)	−0.0973** (0.0447)	−0.1168** (0.0462)	−0.0659 (0.0493)
<i>Levels</i>								
10 Ksh Piece Rate	0.1556 (0.2560)	1.0339 (1.5083)	0.1456 (0.2568)	0.1742 (0.2546)	0.1438 (0.2580)	0.0747 (0.2718)	0.2250 (0.2738)	0.2706 (0.2788)
15 Ksh Piece Rate	0.2124 (0.2581)	1.1508 (1.5210)	0.1956 (0.2600)	0.2465 (0.2581)	0.1209 (0.2661)	0.1393 (0.2793)	0.1549 (0.2796)	0.2936 (0.2695)
Session FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Person FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Room FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
5 Ksh Mean-IHS	2.937	4.508	2.934	2.908	2.781	2.664	2.560	2.209
5 Ksh Median-Levels	10	54	10	10	10	10	9	6
Observations	1224	1224	1224	1224	1224	1224	1224	1224
Participants	128	128	128	128	128	128	128	128

Note: This table shows ordinary least squares regressions of inverse hyperbolic sine-transformed and untransformed productivity outcome variables on piece rate indicators for observations in noise, session, person, and room fixed effects with robust standard errors.

TABLE B16—MEASURES OF COGNITIVE FUNCTION

Domain	Task name	Ability task measures	Why ability is needed to sew
Attention	Psychomotor Vigilance	Ability to sustain focus	To avoid sewing off the edge or going past where the turn is supposed to be
	d2	Ability to ignore distractions	To focus on task while other things are going on around you
Working Memory	Reverse Corsi Block	Ability to store and manipulate information in your mind	To be able to keep in mind how elements will ultimately fit together
	N-Back	Ability to continuously update information	To be able to keep track of where you are in the task
Inhibitory Control	Hearts and Flowers	Ability to resist tempting impulses	To maintain control of sewing speed when surprised by something
Cognitive Flexibility	Wisconsin Card Sort	Ability to maintain multiple rules in memory and select which is most appropriate	To be able to switch from one element of the task to another (e.g. hemming to sewing the sides)
Higher-Level Reasoning	Raven's	Ability to recognize patterns and extrapolate	To identify potentially better methods for accomplishing the task

Note: This table contains descriptions of the cognitive tests used in the second experiment. The first column shows the domain of cognitive function that the test is designed to assess. The second column shows the name of each test. The third column shows the specific cognitive ability the test is designed to assess. The final column shows how this ability is potentially important in sewing.

TABLE B17—EXPERIMENT TWO BALANCE AND SUMMARY STATS

	(1) Female	(2) Age	(3) High School or More	(4) Typical Daily Wage	(5) Days Worked Last Week	(6) More Annoyed by Noise than Others
Treatment Mean	0.523	25.846	0.696	528.916	1.208	0.296
Control Mean	0.509	25.940	0.706	550.188	1.203	0.302
Two-Sided P-Value	0.716	0.868	0.768	0.670	0.973	0.852
Normalized Difference	0.028	−0.014	−0.021	−0.033	0.003	−0.014
Observations	759	759	762	755	759	762
Room-Session Clusters	88	88	88	88	88	88
Participants	210	210	211	209	210	211

Note: This table assesses balance of sample characteristics between treatment and control sessions. The first two rows display the average of the variable indicated in the column for individuals observed in treatment and control sessions, respectively. Row three shows the p-value from a regression of the variable on a treatment indicator with standard errors clustered at the room by session level. The normalized difference is the difference between the treatment and control means divided by the square root of the average of the treatment and control variances as defined by Imbens and Rubin (2015). The results show that those observed in treatment and control are well balanced on observable characteristics. This is due to the within-person design. The only lack of perfect balance comes from a small amount of attrition.

TABLE B18—EXPERIMENT TWO IV EFFECT OF NOISE – TREATMENT INDICATOR INSTRUMENT

	Cognitive Function Tests				Effort Task	
	(1) Normalized Sum of Scores	(2) Average of Normalized Scores	(3) PCA of Percent Correct and Reaction Time	(4) CFA of Percentage Correct and Reaction Time	(5) Key Presses	(6) Normalized Score
Noise Level	−0.0323*** (0.0113)	−0.0254*** (0.0083)	−0.0626*** (0.0150)	−0.0676*** (0.0175)	1.9391 (16.6155)	0.0041 (0.0355)
Wage FE	Yes	Yes	Yes	Yes	Yes	Yes
Session FE	Yes	Yes	Yes	Yes	Yes	Yes
Person FE	Yes	Yes	Yes	Yes	Yes	Yes
Room FE	Yes	Yes	Yes	Yes	Yes	Yes
Control Mean	0.000	0.000	−0.000	0.000	2192.013	−0.000
Observations	762	762	762	762	762	762
Room-Session Clusters	88	88	88	88	88	88
Participants	211	211	211	211	211	211

Note: This table shows estimates from a two-stage least squares regression of cognitive outcome variables on the noise level, wage, session, person, and room fixed effects with standard errors clustered at the room by session level. The noise level is instrumented by the treatment indicator. The first outcome is the normalized sum of points that participants earned on tests during a session. The second column normalizes first at the test-score level and averages across normalized scores within a session. The third outcome is the first component of a principal component analysis of percentage correct and reaction time estimated on each individual's first control session. The fourth column is my preferred outcome: the first factor of a common factor analysis of percentage correct and reaction time estimated on each individual's first control session. The last two columns show that there was no effect of the same noise change on the effort task.

TABLE B19—EXPERIMENT TWO IV EFFECT OF NOISE – TREATMENT INTENSITY INSTRUMENTS

	Cognitive Function Tests				Effort Task	
	(1) Normalized Sum of Scores	(2) Average of Normalized Scores	(3) PCA of Percent Correct and Reaction Time	(4) CFA of Percentage Correct and Reaction Time	(5) Key Presses	(6) Normalized Score
Noise Level	−0.0233** (0.0110)	−0.0168** (0.0082)	−0.0478*** (0.0154)	−0.0529*** (0.0182)	−2.8033 (15.8894)	−0.0060 (0.0339)
Wage FE	Yes	Yes	Yes	Yes	Yes	Yes
Session FE	Yes	Yes	Yes	Yes	Yes	Yes
Person FE	Yes	Yes	Yes	Yes	Yes	Yes
Room FE	Yes	Yes	Yes	Yes	Yes	Yes
Control Mean	0.000	0.000	−0.000	0.000	2192.013	−0.000
Observations	762	762	762	762	762	762
Room-Session Clusters	88	88	88	88	88	88
Participants	211	211	211	211	211	211

Note: This table shows estimates from a two-stage least squares regression of cognitive outcome variables on the noise level, wage, session, person, and room fixed effects with standard errors clustered at the room by session level. The noise level is instrumented by a set of treatment intensity indicators. The first outcome is the normalized sum of points that participants earned on tests during a session. The second column normalizes first at the test-score level and averages across normalized scores within a session. The third outcome is the first component of a principal component analysis of percentage correct and reaction time estimated on each individual's first control session. The fourth column is my preferred outcome: the first factor of a common factor analysis of percentage correct and reaction time estimated on each individual's first control session. The last two columns show that there was no effect of the same noise change on the effort task.

TABLE B20—EXPERIMENT TWO FISHER P-VALUES

Normalized Sum of Scores	Average of Normalized Scores	PCA of Percentage Correct and Reaction Time	CFA of Percentage Correct and Reaction Time	Key Presses	Normalized Score
0.1735	0.1115	0.026	0.0335	0.9135	0.9135

TABLE B21—IMPACTS OF NOISE ON NORMALIZED SCORES BY METRIC

	Attention		Working Memory		Inhibitory Control	Cognitive Flexibility	Higher Reasoning
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	PVT	D2	Corsi	NBack	Hearts and Flowers	Wisconsin	Ravens
Treatment	−0.0198 (0.0305)	−0.0010 (0.0221)	−0.0290 (0.0345)	−0.0238 (0.0222)	−0.0357 (0.0262)	−0.0478 (0.0308)	−0.0100 (0.0234)
Wage FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Session FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Person FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Room FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	762	762	762	762	762	762	762
Room-Session Clusters	88	88	88	88	88	88	88
Participants	211	211	211	211	211	211	211

Note: This table shows estimates from a regression of the normalized score on each test on the treatment indicator, wage, session, person, and room fixed effects with standard errors clustered at the room by session level. The results show that the effects of noise do not appear to be concentrated in any particular domain.

TABLE B22—RESPONSE TO PIECE RATE BY TYPICAL NOISE EXPOSURE

	Experiment 1			Experiment 2		
	(1)	(2)	(3)	(4)	(5)	(6)
	WTP	WTP Any	WTP COP	WTP	WTP Any	WTP COP
Piece Rate	0.3985 (1.4524)	0.0007 (0.0187)	0.7486 (2.3653)	0.3586 (1.6230)	0.0381* (0.0223)	−3.5316 (4.4504)
Typical Noise Exposure	−2.8163 (2.2146)	−0.0287 (0.0401)	−3.0122 (3.3072)	−1.7776 (2.6080)	−0.0066 (0.0311)	−5.4742 (7.0721)
Typical Noise Exposure × Piece Rate	0.4671 (1.2390)	−0.0149 (0.0156)	1.4874 (1.9054)	0.7829 (1.4614)	0.0050 (0.0170)	2.5921 (4.9477)
Day FE	Yes	Yes	Yes	No	No	No
Outcome Mean	17.697	0.538	32.906	13.392	0.316	42.373
Observations	476	476	256	420	420	138
Participant Clusters	120	120	85	210	210	80

Note: This table shows a regression of willingness to pay, an indicator for being willing to pay a positive amount, and willingness to pay for the subsample that are willing to pay a positive amount on an indicator for whether the respondent was facing a piece rate when the willingness to pay was elicited, a z-score of typical noise exposure and their interaction. The z-score is constructed using baseline variables on where the respondent lives and their immediately previous employment. Because in experiment one willingness to pay was elicited on two different days, the regressions include day fixed effects. Standard errors are clustered at the individual level. The results show that there is no significant impact of prior noise exposure on piece-rate responsiveness, and the point estimate suggests a 1 SD have a 0.5 or 0.8 Ksh greater response to the piece rate.

TABLE B23—EFFECTS OF BELIEFS IN EXPERIMENT ONE

	(1) WTP	(2) WTP Any	(3) WTP COP
Piece Rate	−0.2584 (1.5306)	0.0169 (0.0203)	−1.5658 (2.5481)
Believe More Productive in Quiet	6.2208 (5.7762)	0.2614*** (0.0760)	−3.0795 (7.9450)
Believe More Productive × Piece Rate	2.6918 (3.8755)	−0.0669 (0.0480)	7.1020 (5.4276)
Day FE	Yes	Yes	Yes
Outcome Mean	17.697	0.538	32.906
Observations	476	476	256
Participant Clusters	120	120	85

Note: This table shows the results of a regression of willingness to pay and an indicator for being willing to pay a positive amount on an indicator for whether an individual was facing a piece rate, whether they stated they were more productive in quiet and their interaction. Because willingness to pay was elicited over two days in experiment one, day fixed effects are also included. Standard errors are clustered at the individual level. The results are consistent with the more detailed belief data from experiment two.

COGNITIVE TASK DESCRIPTIONS

This appendix describes how the cognitive tasks were implemented and scored. All tasks were programmed on the python-based, open-source platform OpenSesame developed by Mathôt et al. (2012). During each trial session, respondents were seated at a desk and worked autonomously for approximately two hours. Tasks were presented in a random order on Windows touch-screen tablets with external keyboards attached at a resolution of 1280x768. During practice sessions, participants were instructed on the rules of each task, shown demonstrations, and given the opportunity to ask clarifying questions.

C1. Attention

PSYCHOMOTOR VIGILANCE. — The Psychomotor Vigilance Task is implemented following Basner and Dinges (2011). Respondents stare at a blank white screen while resting a finger on the spacebar. At random intervals between 2 and 10 seconds, a red counter appears (see Figure C1). When the counter appears, the respondent's job is to tap the spacebar as quickly as possible. In each session respondents completed 100 trials scored as follows:

- Pressing the spacebar while no counter is present results in an incorrect response, the screen flashes “FALSE START” and earns zero points.
- Responses faster than 100 ms are considered as anticipatory responses, counted as incorrect, and earn zero points.
- Responses slower than 500 ms are considered attentional lapses, counted as incorrect, and earn zero points.
- Following Basner and Dinges (2011), for each correct response participants earn points depending on their inverse response time according to the following scoring rule: $5000 \times \text{Inverse RT} - 10$.

The total score is then the average of the trial scores. For consistency with the other tests, in the common factor and principal component analyses response times are used rather than inverse response times.

D2. — The d2 task follows the general instructions outlined in Brickenkamp and Zillmer (1998) and Bates and Lemay Jr. (2004), but is modified for computer presentation. For each trial, eleven letters (either p or d) appear on the screen with between zero and two dashes above and zero and two dashes below for a total number of dashes between zero and four (see Figure C2). The respondent's job is to mark all of the d's with a total of two dashes by tapping the box below the letter. After 5106 ms, the trial ends. Until that time has elapsed, respondents can un-mark and re-mark letters as they please. Another set of eleven letters



FIGURE C1. PVT STIMULUS

Note: The figure shows a snapshot of a counter that appears in the PVT test displaying a time of 320 ms. The test is designed to assess attention. When the counter appears, respondents must press the space bar to stop it from counting up. The faster they press the space bar, the more points they earn.

appears after 500 ms. Respondents complete 100 trials. For every d with two dashes correctly marked, respondents earn one point. Respondents lose one point for marking anything else. Their score is total number of points earned divided by number of possible points.

C2. Working Memory

REVERSE CORSI BLOCK. — Implementation of the Reverse Corsi Block task follows Brunetti et al. (2014). For each trial, nine blue blocks appear in random locations on the screen. They take turns lighting up for 500 ms with 1000 ms between each flash. Respondents are then asked to tap the blocks in reverse order of how they lit up (see Figure C3). For each element in the sequence, if the respondent taps on the correct block, it turns green for 500 ms and the respondent can proceed to tap the next block in the sequence. If the respondent taps any other block, it flashes red and the respondent moves to the next trial. The first trial sequence contains two elements. For each sequence the respondent gets completely correct, the sequence length increases by one. For every sequence incorrect, the length decreases by one up to a minimum sequence length of two. Respondents complete 50 trials. The score is the average length of the sequences that respondents complete.

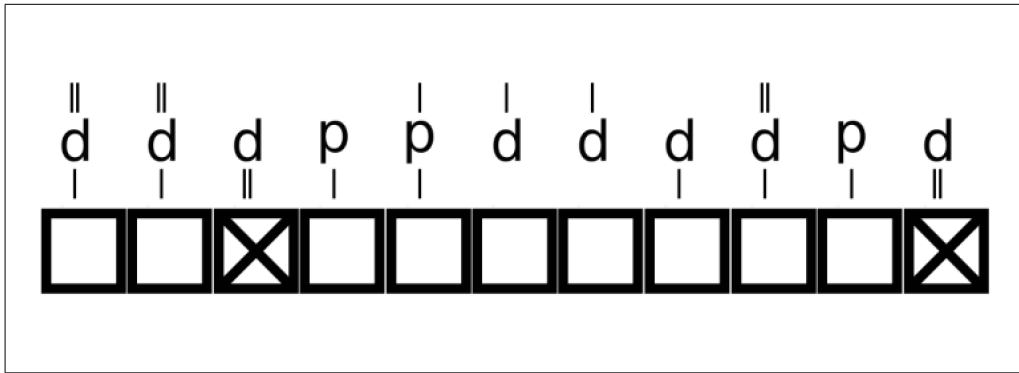


FIGURE C2. D2 STIMULI

Note: The figure shows an example of a trial from the d2 test. The test is designed to assess attention. Respondents see a series of d's and p's with up to two lines below and above. They must tap the boxes below all d's with a total of two dashes before the trial ends.

N-BACK. — Implementation of the N-Back task follows Wilhelm et al. (2013) with an “N” of two. For each trial, respondents see a sequence of twelve animal pictures. For each picture following the second, the respondents are required to tap either “MATCH” or “NO MATCH” depending on whether the image currently on screen matches the image shown two animals ago (see Figure C4). Each image is presented with a 2500 ms maximum response time and a 500 ms interstimulus interval. Each sequence is randomly determined by randomly drawing elements from a pool of ten images such that for each trial there is a 50% chance of the draw being a match. At the start of every session, respondents complete one practice trial sequence and then another 40 scored sequences. A respondent’s score is the percentage of responses correctly marked times 10.

C3. Inhibitory Control

HEARTS AND FLOWERS. — Implementation of the Hearts and Flowers task follows the “dots” task outlined by Davidson et al. (2006). Respondents see a fixation dot in the center of their screen with blue boxes on the left and right. Respondents then see a sequence of hearts and flowers appear on the boxes. For each trial, respondents must press either the “Q” or “P” key. When a heart appears, respondents must press the key on the same side as the heart. While when a flower appears, respondents must press the key on the opposite side (see Figure C5). During each session respondents complete the following:

- 1) 6 practice trials with only hearts.
- 2) 126 scored trials with only hearts.
- 3) 6 practice trials with only flowers.

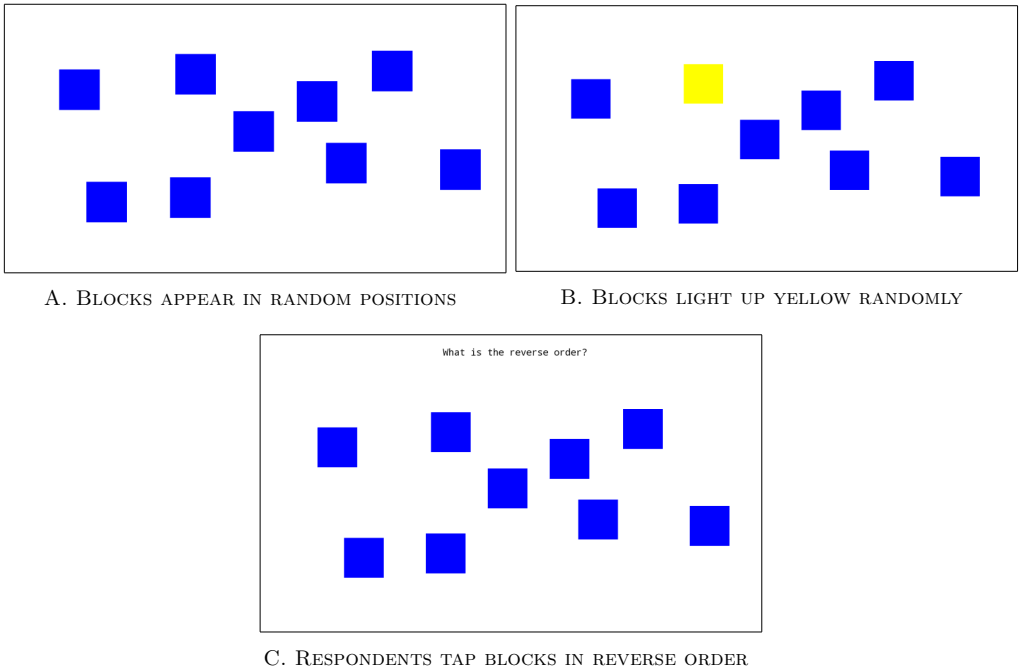


FIGURE C3. CORSI STIMULI

Note: This figure shows the three stages of the reverse corsi blocks test. The test is designed to measure working memory. First nine blocks appear in random positions. They then light up in a random sequence. Respondents must then tap the blocks in the reverse order of how they lit up. After each correct trial, the length of the sequence increases by one, and after every incorrect trial, the length of the sequence decreases by one down to a minimum of two elements.

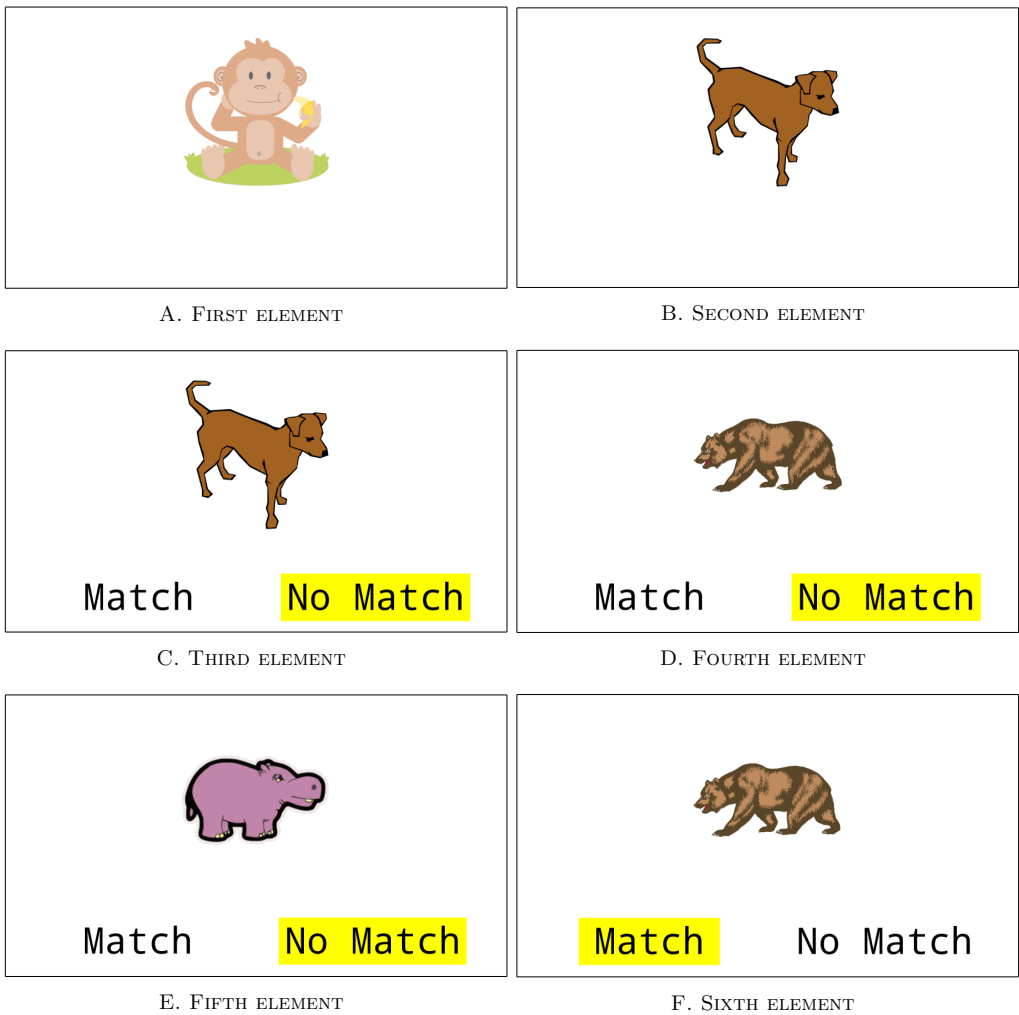


FIGURE C4. N-BACK STIMULI AND RESPONSES

Note: This figure shows an example of six elements from an N-back sequence. The test is designed to assess working memory. Respondents see a series of animals and must indicate whether the animal currently displayed matches the animal seen two elements previously.

- 4) 126 scored trials with only flowers.
- 5) 492 scored trials with both hearts and flowers.

Each stimulus times out after 750 ms and there is a 500 ms interstimulus interval. Trials are scored as follows:

- Responses faster than 200 ms are scored as incorrect, anticipatory responses and earn zero points.
- Trials where the incorrect key or no key is pressed are scored as incorrect and earn zero points.
- For each trial with a correct response, respondents earn points according to the following scoring rule that is linear in their response time: $10 \times \frac{RT - 750}{200 - 750}$

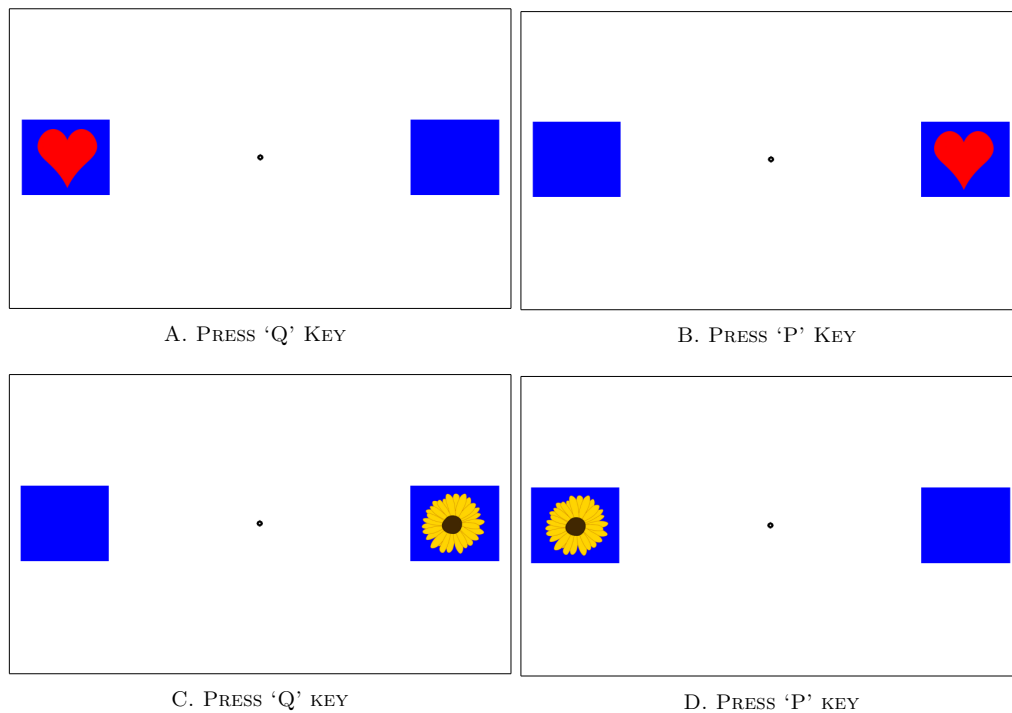


FIGURE C5. HEARTS AND FLOWERS POSSIBLE STIMULI AND RESPONSES

Note: The figure shows the four possible stimuli and responses for the hearts and flowers test. The test is designed to assess inhibitory control. Respondents see a series of hearts and flowers appear on the blocks. When a flower appears, the respondent must press the key on the opposite side of the keyboard. When a heart appears, the respondent must press the key on the same side of the keyboard.

C4. Cognitive Flexibility

WISCONSIN CARD SORT. — The Wisconsin Card Sort task follows the procedure originally outlined by Grant and Berg (1946) as modified for computer display by PsyToolkit (Stoet 2010, 2017). The respondent sees four response cards at the top of the screen and one question card (see Figure C6). Their job is to tap the response card that “matches” the question card. There are three possible matching rules:

- 1) Shape – In the example, the correct answer would be the fourth response card (four gold stars).
- 2) Color – In the example, the correct answer would be the third response card (three blue crosses).
- 3) Number – In the example, the correct answer would be the second response card (two green triangles).

Every ten trials a sorting rule is chosen at random. Respondents must figure out the sorting rule through trial and error. If the respondent taps the correct response card, the screen flashes “Correct!”. If the response card they tap is incorrect, the screen flashes “Wrong!”. Respondents complete 100 trials. Every incorrect trial earns zero points. Every correct trial is scored according to the following scoring rule linear in reaction time³⁸: $10 + 200 \times \frac{10}{30000 - 200} - RT \times \frac{10}{30000 - 200}$.

C5. Higher-Level Reasoning

RAVEN’S. — The Raven’s task follows the classic task described by Raven (2000) with supplemental matrices graciously provided by Heather Schofield based on Schofield (2014). Respondents see a matrix with a missing piece and a set of possible pieces (see Figure C7), and their job is to tap the piece that completes the pattern in the matrix. In each session, respondents completed ten original Raven’s progressive matrices alternating with ten supplemental matrices increasing in difficulty. For each incorrect response, respondents earned zero points. For each correct response, respondents earned points in a scoring rule linear in their reaction time³⁹: $10.0 + 200 \times \frac{10}{60000 - 200.0} - RT \times \frac{10}{60000 - 200}$.

38. Note that it is not traditional to score reaction time on this task because the task is typically presented with physical cards by an enumerator. I took advantage of the computer-based administration to collect reaction time and improve the measure’s sensitivity.

39. Reaction time is not traditionally collected on this test; however, I decided to take advantage of computer-based implementation and maximize the sensitivity of the test.

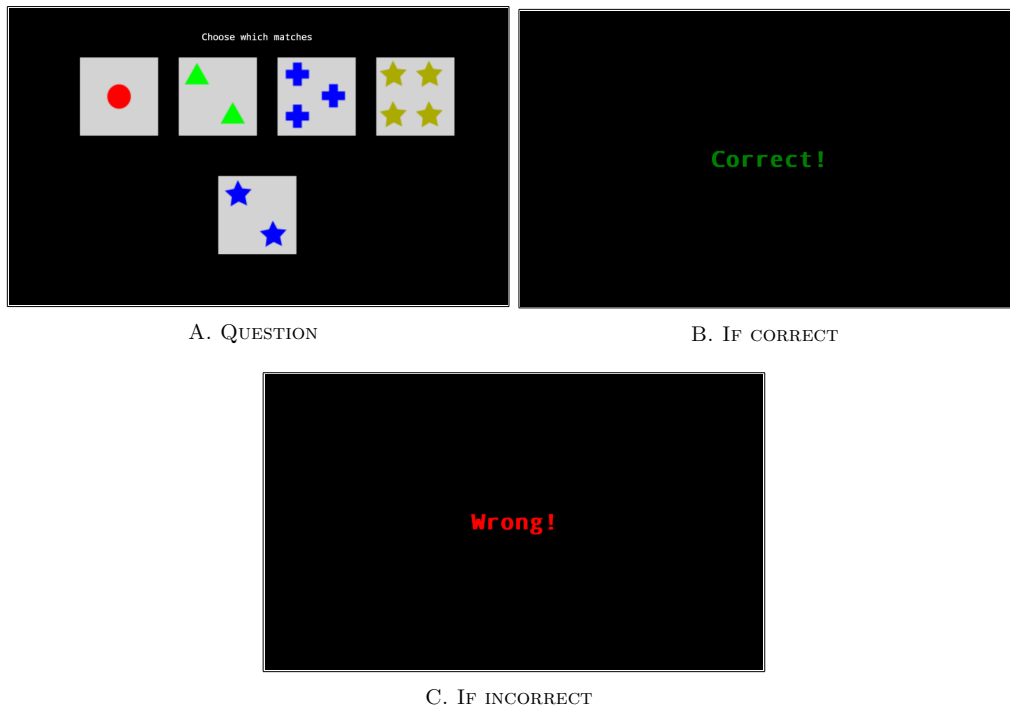


FIGURE C6. WISCONSIN STIMULI

Note: The figure shows the three key screens from the Wisconsin card sort test. The test is designed to assess cognitive flexibility. Respondents are shown a card at the bottom of the screen and are asked to choose which of four cards at the top of the screen it matches according to one of three possible sorting rules. Respondents are not told which of the rules is being used and must figure it out by trial and error. Every ten trials the sorting rule changes.

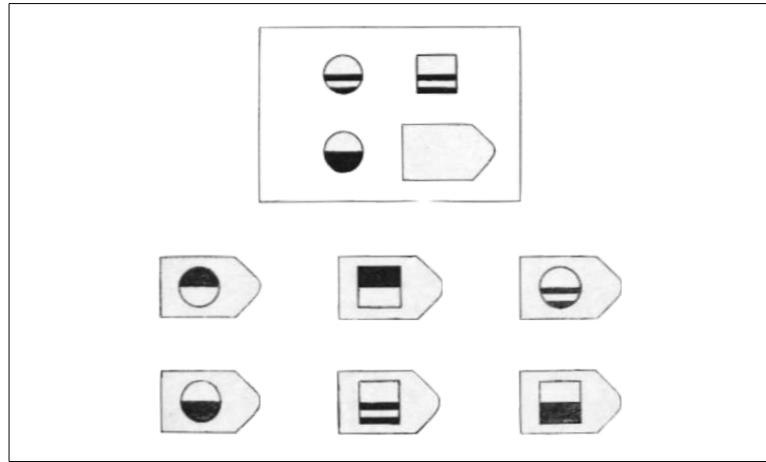


FIGURE C7. RAVEN'S STIMULUS

Note: This figure shows an example of a trial from the Raven's matrices test. The test is designed to assess higher-reasoning skills. The respondent sees a pattern of shapes with one missing. They must choose which of the possible answers completes the pattern.

C6. *Effort*

EFFORT TASK. — The effort task is implemented following DellaVigna and Pope (2018). Respondents have 10 minutes to alternate pressing the 'a' and 'b' keys. For each complete alternation, a progress bar on the screen increases by one hash mark. At increments of 50, the bar resets and respondents are reminded of their total score (see Figure C8). Respondents earn one point for every 300 alternations.

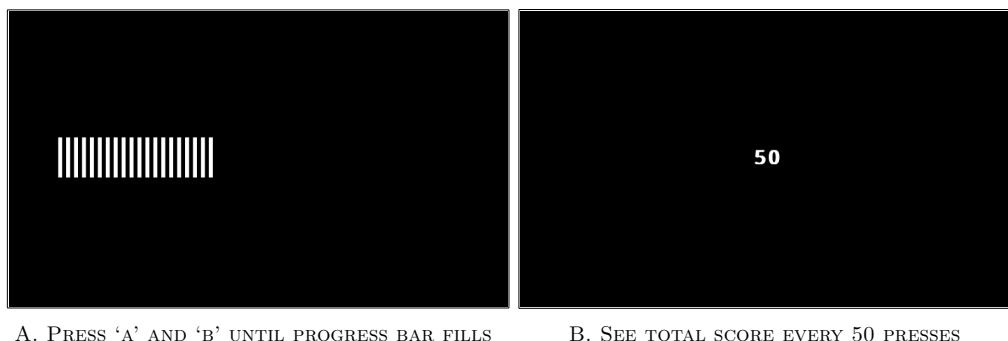


FIGURE C8. EFFORT STIMULI

Note: The figure shows the two key screens from the effort task. The respondent must alternate pressing the 'a' and 'b' keys. As they do, a progress bar (shown on the left) begins to fill up. After every 50 completed alternations the respondent sees his/her score (shown on the right), and the progress bar resets to zero.

WILLINGNESS TO PAY SCRIPT FOR ONLINE PUBLICATION

"I'm sure you've noticed these two weeks that sometimes a noisy engine is outside of the rooms. For each of the first two practice sessions tomorrow, we are going to give you the chance to pay in order to work in a room without the engine outside. However, the price for working in the quieter room has not yet been decided. It will be determined for each session by a game of chance. You will not have to pay anything more than you want to, and you might even get it for less! Here's how this will work:

For each session, you and I will figure out the highest price that you are willing to pay to work in the quieter room. Then tonight our computer will randomly decide the price. If the price is higher than you said you are willing to pay, when you come tomorrow you will be in the room with the engine outside. If the price is lower than what you said you were willing to pay, the randomly chosen price will be deducted from your pay for that session, and you will work in the quieter room.

Since this is complicated, we will first make a plan for which prices you would like to pay to work in the quieter room. I will ask you whether you would be willing to pay several prices in order to be in the quieter room and you will tell me yes or no. After we are done, you will not be able to change your plan. Do you understand?"

PRE-REGISTRATION, DEVIATIONS AND PILOTING

While this experiment was pre-registered at the AEA trial registry under ID AEARCTR-0001500 (Dean 2018), I did not file a pre-analysis plan. In the pre-registration, filed before the first experiment, I specified the primary outcomes

would be “Productivity in producing the practice good, performance on cognitive tests, decisions made in three real stakes decision tasks, and willingness to pay for quiet.”, but did not specify the functional forms for these outcomes. I present all functional forms considered except for $\log(1 + n)$ which was replaced with the inverse hyperbolic sign transformation because it is more standard and similarly solves the issue of skewness. No heterogeneity analyses were pre-specified and should be treated as exploratory.

I originally planned to simultaneously collect cognitive function data during the first experiment, but this did not work. Specifically, I did not want to take too much time away from sewing and only included one short test within each session. This resulted in substantial treatment imbalance across both test domains and baseline abilities. To rectify this situation, I conducted the second experiment presented in this paper. An effort task was not originally going to be among the cognitive function measures, but was included in the second experiment as a method of disentangling the relative contributions of cognitive function and other mechanisms. Specifically, the hope was that a null effect on the effort task would provide additional validity to the exclusion restriction in an IV using noise exposure to estimate the effect of cognitive function on productivity. Following critical feedback, this is no longer contained in the paper. Additionally, belief elicitation were added to the second experiment after seeing the lack of response in willingness to pay to the wage structure in the first experiment.

I ran one pilot with 32 subjects for the first experiment and no pilots for the second. I used the effect size from this pilot to determine the sample size for the first experiment. For the second, I used simulations based on the variability in the first experiment.