Supplemental Results Appendix and PAP Report ("Supplement"): Destructive Behavior, Judgment, and Economic Decision-making under Thermal Stress

Ingvild Almås, Maximilian Auffhammer, Tessa Bold,
Ian Bolliger, Aluma Dembo, Solomon Hsiang,
Shuhei Kitamura, Edward Miguel, and Robert Pickmans

Not for publication

Contents

F   Pre-specified analyses

F.1   Description

Adhering strictly to the pre-analysis plan, Supplement Section F.2 features the same specifications noted earlier, with covariates found to be unbalanced across treatment arms included in the regression specifications. To determine these covariates, we performed Student's t-test to check for balance across treatment arms on: 1) self-reported weight; 2) self-reported height; 3) age; 4) gender; 5) area of origin (Xlab) or ethnicity (Busara Center); 6) father's occupation status; 7) mother's occupation status; 8) combined parental income; 9) father's highest level of education; and 10) mother's highest level of education.[1]. As shown in Appendix Table C.1, self-reported weight was found to be unbalanced at the .01 level in Nairobi. Thus, the heterogeneity specification and Nairobi specification include participant self-reported weight with imputed values from site sample means for missing values, as well as an indicator variable for whether weight data was not entered by the participant.

As pre-specified, for each specification within Supplement Section F.2, multiple hypothesis testing adjustments are performed on two sets of p-values: 1) the set associated with Heat, across primary outcomes; 2) within the heterogeneity specification, separately for each primary outcome, the set associated with the interaction between Heat and Male and the interaction between Heat and Nairobi.[2] Although the California and Nairobi specifications were not pre-specified, they have adjustments in an analogous fashion to the main specification. Exact p-values for treatment effects associated with primary outcomes and generated through 10,000 permutations, as pre-specified, are included, as are multiple testing adjusted FDR q-value significance levels based on those p-values. At the end of Supplement Section F.2 is a table (Table F.2.13) containing results from tests of overall significance on the 10 primary outcomes for which this is possible (thus, excluding the outcomes for Patience and Time Inconsistency), as well as p-values derived from randomization inference.[3] Although the tests were only pre-specified for the set of coefficients associated with Heat in the main specification, the set of coefficients associated with the interaction between Male and Heat in the heterogeneity specification, and the set of coefficients associated with the interaction between Nairobi and Heat in the heterogeneity specification, this table also includes results for the set of coefficients associated with Heat in the heterogeneity specification as well as for the sets of coefficients associated with Heat from the site-specific specifications, which was not pre-specified.

---

[1]To facilitate interpretation and to be able to conduct the balance check as intended, several of these variables were recast. Education for both the mother and father became indicator variables for having university education, occupation status for both the mother and father became indicator variables for taking on either self-employment or wage work, and combined parental income became an indicator for parents earning $24,000 a year (PPP adjusted). Residency status and ethnicity were site-specific questions, with the former only asked in California and the latter in Nairobi.

[2]For the Charitable donation outcome, the latter set also includes the interaction between Heat and being matched with an ingroup charity. Additionally, although the pre-analysis plan specified that the risk measure outcome would be a categorical variable indicating coin choice from menu A, both interpretation and accounting for multiple hypothesis testing motivated the primary outcome for risk to be defined as the variance of the coin toss in menu A.

[3]Table F.2.14 contains results from test of overall significance, excluding the Joy of Destruction.

Supplement Section F.3 follows a similar format to Supplement Section F.2, but instead examines additional outcomes. As pre-specified, these outcomes do not feature any adjustments for multiple hypothesis testing. Moreover, no joint-tests of significance are calculated. In the pre-analysis plan, an alternative specification featuring additional covariates was also laid out for estimating the Fairness outcome, and was to be included if the alternative specification led to far more precise estimates. Supplement Section F.4 uses the same four preferred specifications as above but also includes an indicator for being part of the "high" (=1) or "low" (=0) group as well as the number of correct sliders.

F.2 Results with primary outcomes

Table F.2.1: Precision task

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	0.516	0.700	0.590	0.367
	(0.049)*	(0.106)	(0.083)	(0.346)
	[0.384]	[0.875]	[0.457]	[1.000]
	{0.051}	{0.109}	{0.080}	{0.355}
	⟨0.379⟩	⟨0.836⟩	⟨0.418⟩	⟨1.000⟩
Male	1.974	1.991	1.892	1.791
	(0.000)**	(0.000)**	(0.000)**	(0.000)**
Male × Heat		-0.256		
		(0.593)		
		[1.000]		
		{0.596}		
		⟨1.000⟩		
Nairobi	-11.19	-11.11		
	(0.000)**	(0.000)**		
Nairobi × Heat		-0.104		
		(0.837)		
		[1.000]		
		{0.838}		
		⟨1.000⟩		
Weight		0.0131		0.0517
		(0.262)		(0.017)*
Weight originally missing		-0.298		-0.182
		(0.485)		(0.853)
Outcome control mean	17.95	17.95	23.63	12.99
R-squared	0.549	0.550	0.0630	0.0347
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Exact p-values are included in curly brackets, and Multiple testing adjusted FDR q-value significance levels based on those p-values are included in triangle brackets. "Weight" refers to self-reported weight (in kg) and "Weight originally missing" is an indicator variable for whether weight data was not entered by the participant. The precision task is also known as the slider task. The outcome in this table is the number of correct sliders made in three minutes. Final earnings from the precision task are based off either being weakly above (high) or below (low) the median within treatment cohort. The median pair is randomly assigned to high or low. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table F.2.2: Fairness

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-0.0167	-0.0200	-0.0245	-0.00966
	(0.170)	(0.272)	(0.114)	(0.597)
	[0.618]	[0.875]	[0.457]	[1.000]
	{0.176}	{0.278}	{0.114}	{0.599}
	⟨0.654⟩	⟨0.836⟩	⟨0.459⟩	⟨1.000⟩
Male	-0.0407	-0.0300	-0.0783	-0.00492
	(0.001)**	(0.091)	(0.000)**	(0.783)
Male × Heat		-0.0117		
		(0.611)		
		[1.000]		
		{0.600}		
		⟨1.000⟩		
Nairobi	0.0337	0.0224		
	(0.012)*	(0.217)		
Nairobi × Heat		0.0185		
		(0.439)		
		[1.000]		
		{0.444}		
		⟨1.000⟩		
Weight		-0.000596		-0.000216
		(0.282)		(0.825)
Weight originally missing		0.00288		0.0119
		(0.896)		(0.795)
Outcome control mean	0.318	0.318	0.309	0.326
R-squared	0.0102	0.0111	0.0323	0.000637
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Exact p-values are included in curly brackets, and Multiple testing adjusted FDR q-value significance levels based on those p-values are included in triangle brackets. "Weight" refers to self-reported weight (in kg) and "Weight originally missing" is an indicator variable for whether weight data was not entered by the participant. Fairness here refers to the real effort dictator game, where the level of endowment is determined by the number of correct sliders made in the precision task. The outcome in this table is the share of joint earnings (2400 tokens in the high group, 1200 tokens in the low group) that each participant desires to give to the anonymous partner. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table F.2.3: Risk-taking

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-1.835	2.747	-5.553	-0.768
	(0.929)	(0.926)	(0.858)	(0.978)
	[1.000]	[0.876]	[0.850]	[1.000]
	{0.881}	{0.928}	{0.862}	{0.979}
	⟨1.000⟩	⟨0.910⟩	⟨0.889⟩	⟨1.000⟩
Male	163.6	156.6	316.8	13.32
	(0.000)**	(0.000)**	(0.000)**	(0.657)
Male × Heat		-19.68		
		(0.659)		
		[1.000]		
		{0.656}		
		⟨1.000⟩		
Nairobi	-116.3	-112.9		
	(0.000)**	(0.001)**		
Nairobi × Heat		8.760		
		(0.848)		
		[1.000]		
		{0.848}		
		⟨1.000⟩		
Weight		2.068		1.824
		(0.052)		(0.327)
Weight originally missing		-2.641		58.83
		(0.948)		(0.483)
Outcome control mean	366.3	366.3	409.0	329.1
R-squared	0.0396	0.0417	0.112	0.00258
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Exact p-values are included in curly brackets, and Multiple testing adjusted FDR q-value significance levels based on those p-values are included in triangle brackets. "Weight" refers to self-reported weight (in kg) and "Weight originally missing" is an indicator variable for whether weight data was not entered by the participant. The outcome in this table is the variance of the coin toss from menu A, in tokens. Note that the expected value is not constant across each coin, so that the outcome does not capture the trade-off between expected value and variance. Note also that under this approach, Coin 7 will be as good as Coin 5, even though Coin 5 strictly dominates Coin 7 with expected utility. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table F.2.4: Rational choice violation I

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-0.00431	-0.00277	-0.00455	-0.00409
	(0.561)	(0.818)	(0.687)	(0.678)
	[1.000]	[0.876]	[0.846]	[1.000]
	{0.568}	{0.826}	{0.679}	{0.693}
	⟨1.000⟩	⟨0.910⟩	⟨0.740⟩	⟨1.000⟩
Male	0.00268	0.00299	0.00540	0.000123
	(0.725)	(0.809)	(0.627)	(0.992)
Male × Heat		-0.00403		
		(0.804)		
		[1.000]		
		{0.813}		
		⟨1.000⟩		
Nairobi	0.0000607	0.00158		
	(0.993)	(0.890)		
Nairobi × Heat		0.000988		
		(0.950)		
		[1.000]		
		{0.951}		
		⟨1.000⟩		
Weight		0.000248		0.0000118
		(0.545)		(0.987)
Weight originally missing		0.0181		0.00351
		(0.314)		(0.904)
Outcome control mean	0.0277	0.0277	0.0275	0.0279
R-squared	0.000259	0.00130	0.000499	0.000182
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Exact p-values are included in curly brackets, and Multiple testing adjusted FDR q-value significance levels based on those p-values are included in triangle brackets. "Weight" refers to self-reported weight (in kg) and "Weight originally missing" is an indicator variable for whether weight data was not entered by the participant. The outcome in this table is an indicator of transitivity violation using both menus A & B. A transitivity violation comes from choosing two coins in the interior region of the intersection of both menus, where it is not the case that it can be said that one preferring coin A to coin B and then preferring coin B to coin C implies that one prefers coin A to coin C. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-0.00005823	-0.00005823	-0.00060946	0.00096101
	(0.854)	(0.854)	(0.093)	(0.279)
	[1.000]	[0.876]	[0.457]	[1.000]
	{0.859}	{0.859}	{0.080}	{0.291}
	⟨1.000⟩	⟨0.910⟩	⟨0.418⟩	⟨1.000⟩
Heat (Male)		-0.0002442		
		(0.702)		
		[0.541]		
		{0.705}		
		⟨0.537⟩		
Heat (Nairobi)		0.00157046		
		(0.102)		
		[0.256]		
		{0.104}		
		⟨0.264⟩		
Outcome control mean	0.9938358	0.9938358	0.9937076	0.9939217
Observations	6612	6612	3200	3412

Note: Standard errors clustered at the individual level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. For (1) and (2), "Heat" refers to the difference between the aggregate delta statistics for the treatment and control groups, divided by the square root of the sum of squared standard errors. For (2), "Heat (Male)" refers to the difference between the aggregate delta statistics for male treatment and control groups, subtracting the difference between the aggregate delta statistics for female treatment and control groups, and then dividing by the square root of the sum of squared standard errors. "Heat (Nairobi)" refers to the difference between the aggregate delta statistics for Nairobi treatment and control groups, subtracting the difference between the aggregate delta statistics for California treatment and control groups, and then dividing by the square root of the sum of squared standard errors. (3) and (4) carry out a similar estimation to "Heat" for (1) and (2) but for the California sample and Nairobi sample, respectively. Exact p-values are included in curly brackets, and Multiple testing adjusted FDR q-value significance levels based on those p-values are included in triangle brackets. Given that the non-linear least squares specification used to measure patience does not incorporate interaction terms, we instead use p-values from analogous differences in this table for multiple testing adjustments. Also note that the non-linear least squares specification does not allow for controls for self-reported weight, and thus the results (aside from the q-values) are similar to those in Appendix Table D.1.5. The outcome in this table is the aggregate δ statistic from the non-linear least squares specification featured in Andreoni et al. (2015), carried out at the individual-choice level. δ is the daily discount factor between two future days. For comparability with regression results, individuals who did not respond "Female" or "Male" to the gender survey question were dropped (2% of the sample). Before estimation of aggregate parameters, individuals who never altered their decision from a specific corner solution in all convex time budgets were dropped (as they provided insufficient variation for the calculation of utility parameters) as were individuals who exhibited generalized axiom of revealed preference (GARP) violations. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table F.2.6: Time inconsistency

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	0.00048029	0.00048029	-0.00878209	0.0179143
	(0.961)	(0.961)	(0.489)	(0.487)
	[1.000]	[0.876]	[0.850]	[1.000]
	{0.960}	{0.960}	{0.494}	{0.479}
	⟨1.000⟩	⟨0.910⟩	⟨0.740⟩	⟨1.000⟩
Heat (Male)		-0.00824147		
		(0.673)		
		[1.000]		
		{0.675}		
		⟨1.000⟩		
Heat (Nairobi)		0.02669638		
		(0.353)		
		[1.000]		
		{0.348}		
		⟨1.000⟩		
Outcome control mean	1.001042	1.001042	0.9886241	1.05313
Observations	6612	6612	3200	3412

Note: Standard errors clustered at the individual level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. For (1) and (2), "Heat" refers to the difference between the aggregate beta statistics for the treatment and control groups, divided by the square root of the sum of squared standard errors. For (2), "Heat (Male)" refers to the difference between the aggregate beta statistics for male treatment and control groups, then subtracting the difference between the aggregate beta statistics for female treatment and control groups, and then dividing by the square root of the sum of squared standard errors. "Heat (Nairobi)" refers to the difference between the aggregate beta statistics for Nairobi treatment and control groups, subtracting the difference between the aggregate beta statistics for California treatment and control groups, and then dividing by the square root of the sum of squared standard errors. (3) and (4) carry out a similar estimation to "Heat" for (1) and (2) but for the California sample and Nairobi sample, respectively. Note that the effects presented above are multiplied by -1, so that a positive difference reflects more time inconsistency. Exact p-values are included in curly brackets, and Multiple testing adjusted FDR q-value significance levels based on those p-values are included in triangle brackets. Given that the non-linear least squares specification used to measure patience does not incorporate interaction terms, we instead use p-values from analogous differences in this table for multiple testing adjustments. Also note that the non-linear least squares specification does not allow for controls for self-reported weight, and thus the results (aside from the q-values) are similar to those in Appendix Table D.1.6. The outcome in this table is the aggregate β statistic from the non-linear least squares specification featured in Andreoni et al. (2015), carried out at the individual-choice level. β measures present bias, and values less than 1 denote time inconsistency. For comparability with regression results, individuals who did not respond "Female" or "Male" to the gender survey question were dropped (2% of the sample). Before estimation of aggregate parameters, individuals who never altered their decision from a specific corner solution in all convex time budgets were dropped (as they provided insufficient variation for the calculation of utility parameters) as were individuals who exhibited generalized axiom of revealed preference (GARP) violations. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table F.2.7: Trust

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	0.00780	-0.0362	-0.0215	0.0332
	(0.617)	(0.147)	(0.377)	(0.113)
	[1.000]	[0.875]	[0.759]	[0.603]
	{0.623}	{0.155}	{0.377}	{0.112}
	⟨1.000⟩	⟨0.836⟩	⟨0.740⟩	⟨0.594⟩
Male	0.0699	0.0537	0.0582	0.0812
	(0.000)**	(0.020)*	(0.020)*	(0.000)**
Male × Heat		0.0358		
		(0.224)		
		[0.289]		
		{0.231}		
		⟨0.282⟩		
Nairobi	-0.137	-0.165		
	(0.000)**	(0.000)**		
Nairobi × Heat		0.0464		
		(0.162)		
		[0.289]		
		{0.166}		
		⟨0.282⟩		
Weight		-0.000348		-0.000229
		(0.667)		(0.851)
Weight originally missing		-0.0623		-0.106
		(0.018)*		(0.045)*
Outcome control mean	0.417	0.417	0.495	0.348
R-squared	0.0465	0.0514	0.00804	0.0261
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Exact p-values are included in curly brackets, and Multiple testing adjusted FDR q-value significance levels based on those p-values are included in triangle brackets. "Weight" refers to self-reported weight (in kg) and "Weight originally missing" is an indicator variable for whether weight data was not entered by the participant. The outcome in this table is the share of endowed tokens (out of 600) entrusted to the other person in the first round of the trust game. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table F.2.8: Public contribution

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-8.923	-37.37	-36.88	13.39
	(0.647)	(0.253)	(0.213)	(0.610)
	[1.000]	[0.875]	[0.520]	[1.000]
	{0.592}	{0.251}	{0.213}	{0.614}
	⟨1.000⟩	⟨0.836⟩	⟨0.474⟩	⟨1.000⟩
Male	38.11	23.20	9.059	60.31
	(0.067)	(0.423)	(0.762)	(0.045)*
Male × Heat		2.465		
		(0.947)		
		[0.900]		
		{0.950}		
		⟨0.902⟩		
Nairobi	-170.5	-198.0		
	(0.000)**	(0.000)**		
Nairobi × Heat		48.25		
		(0.223)		
		[0.807]		
		{0.221}		
		⟨0.825⟩		
Weight		1.372		0.707
		(0.218)		(0.639)
Weight originally missing		-146.1		-190.9
		(0.000)**		(0.000)**
Outcome control mean	529.7	529.7	629.6	442.5
R-squared	0.0383	0.0471	0.00184	0.0149
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Exact p-values are included in curly brackets, and Multiple testing adjusted FDR q-value significance levels based on those p-values are included in triangle brackets. "Weight" refers to self-reported weight (in kg) and "Weight originally missing" is an indicator variable for whether weight data was not entered by the participant. Public contribution here refers to the public goods game. The outcome in this table is the amount of tokens (out of 1200) put into the shared fund. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table F.2.9: Fluid intelligence

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	0.0150	-0.0102	0.000858	0.0259
	(0.061)	(0.368)	(0.919)	(0.045)*
	[0.384]	[0.875]	[0.850]	[0.332]
	{0.056}	{0.372}	{0.969}	{0.047}*
	⟨0.379⟩	⟨0.836⟩	⟨0.939⟩	⟨0.346⟩
Male	0.0176	0.00575	0.0207	0.01000
	(0.045)*	(0.662)	(0.012)*	(0.508)
Male × Heat		0.0272		
		(0.095)		
		[0.236]		
		{0.091}		
		⟨0.241⟩		
Nairobi	-0.127	-0.138		
	(0.000)**	(0.000)**		
Nairobi × Heat		0.0201		
		(0.207)		
		[0.236]		
		{0.216}		
		⟨0.241⟩		
Weight		-0.000205		0.000950
		(0.622)		(0.266)
Weight originally missing		0.00892		0.0275
		(0.508)		(0.398)
Outcome control mean	0.862	0.862	0.935	0.799
R-squared	0.121	0.124	0.00787	0.00770
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Exact p-values are included in curly brackets, and Multiple testing adjusted FDR q-value significance levels based on those p-values are included in triangle brackets. "Weight" refers to self-reported weight (in kg) and "Weight originally missing" is an indicator variable for whether weight data was not entered by the participant. Fluid intelligence is measured through Raven's Progressive Matrices. The outcome in this table is the share of six matrices answered correctly. Each puzzle answered correctly yields an Airtime Voucher worth 50 KSh (or an Amazon Gift Card worth 1 dollar in the California sample), which provides the earnings to be used for the next module. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table F.2.10: Joy of Destruction

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	0.0234	-0.0203	-0.0409	0.0801
	(0.069)	(0.161)	(0.004)**	(0.000)**
	[0.384]	[0.875]	[0.053]	[0.001]$^{++}$
	{0.069}	{0.165}	{0.003}**	{0.000}**
	⟨0.379⟩	⟨0.836⟩	⟨0.038⟩$^+$	⟨0.003⟩$^{++}$
Male	0.0369	0.0662	0.0252	0.0506
	(0.001)**	(0.000)**	(0.106)	(0.003)**
Male × Heat		-0.0517		
		(0.016)*		
		[0.009]$^{++}$		
		{0.017}*		
		⟨0.009⟩$^{++}$		
Nairobi	0.113	0.0436		
	(0.000)**	(0.014)*		
Nairobi × Heat		0.134		
		(0.000)**		
		[0.001]$^{++}$		
		{0.000}**		
		⟨0.001⟩$^{++}$		
Weight		-0.000469		-0.000701
		(0.384)		(0.459)
Weight originally missing		-0.00764		-0.0429
		(0.713)		(0.255)
Outcome control mean	0.110	0.110	0.0771	0.139
R-squared	0.0643	0.0809	0.0138	0.0279
Observations	1859	1859	864	995

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Exact p-values are included in curly brackets, and Multiple testing adjusted FDR q-value significance levels based on those p-values are included in triangle brackets. "Weight" refers to self-reported weight (in kg) and "Weight originally missing" is an indicator variable for whether weight data was not entered by the participant. The outcome in this table is the share of the anonymous partner's earned Airtime Vouchers (Amazon Gift Cards in the California sample) destroyed by the participant. The earned Airtime Vouchers and Amazon Gift Cards resulted from the number of Raven's Progressive Matrices answered correctly in the previous module, where one puzzle answered correctly yielded one Airtime Voucher or one Amazon Gift Card. Airtime Vouchers were worth 50 KSh each, while Amazon Gift Cards were worth 1 dollar each. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table F.2.11: Cognitive reflection

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-0.00228	-0.0271	-0.0228	0.0130
	(0.837)	(0.163)	(0.217)	(0.320)
	[1.000]	[0.875]	[0.520]	[1.000]
	{0.867}	{0.169}	{0.179}	{0.310}
	⟨1.000⟩	⟨0.836⟩	⟨0.474⟩	⟨1.000⟩
Male	0.0845	0.0852	0.177	-0.00963
	(0.000)**	(0.000)**	(0.000)**	(0.509)
Male × Heat		0.00885		
		(0.720)		
		[0.563]		
		{0.721}		
		⟨0.563⟩		
Nairobi	-0.245	-0.266		
	(0.000)**	(0.000)**		
Nairobi × Heat		0.0382		
		(0.117)		
		[0.306]		
		{0.123}		
		⟨0.299⟩		
Weight		-0.000632		0.00176
		(0.377)		(0.034)*
Weight originally missing		0.00461		-0.0141
		(0.864)		(0.513)
Outcome control mean	0.325	0.325	0.455	0.212
R-squared	0.178	0.180	0.0773	0.00698
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Exact p-values are included in curly brackets, and Multiple testing adjusted FDR q-value significance levels based on those p-values are included in triangle brackets. "Weight" refers to self-reported weight (in kg) and "Weight originally missing" is an indicator variable for whether weight data was not entered by the participant. The outcome in this table is the share of questions (out of 5) from the Cognitive Reflection Test answered correctly. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table F.2.12: Charitable donation

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-4.443	-57.22	-12.67	7.391
	(0.887)	(0.194)	(0.679)	(0.892)
	[1.000]	[0.875]	[0.846]	[1.000]
	{0.899}	{0.190}	{0.641}	{0.892}
	⟨1.000⟩	⟨0.836⟩	⟨0.740⟩	⟨1.000⟩
Male	7.254	23.74	-58.73	98.65
	(0.820)	(0.601)	(0.104)	(0.058)
Male × Heat		2.706		
		(0.964)		
		[1.000]		
		{0.965}		
		⟨1.000⟩		
Nairobi	323.8	281.9		
	(0.000)**	(0.000)**		
Nairobi × Heat		47.98		
		(0.437)		
		[0.776]		
		{0.435}		
		⟨0.782⟩		
Matched with ingroup charity	-9.405	-71.45	31.79	-86.01
	(0.771)	(0.123)	(0.409)	(0.136)
Matched with ingroup charity × Heat		116.9		
		(0.101)		
		[0.434]		
		{0.095}		
		⟨0.460⟩		
Earnings in tokens	0.00356	0.00399	-0.0182	0.0303
	(0.734)	(0.703)	(0.117)	(0.090)
Weight		-2.349		-5.168
		(0.090)		(0.098)
Weight originally missing		-142.1		-172.2
		(0.000)**		(0.062)
Outcome control mean	410.4	410.4	245.0	561.8
R-squared	0.0657	0.0712	0.00695	0.0122
Observations	1806	1806	861	945

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Exact p-values are included in curly brackets, and Multiple testing adjusted FDR q-value significance levels based on those p-values are included in triangle brackets. "Weight" refers to self-reported weight (in kg) and "Weight originally missing" is an indicator variable for whether weight data was not entered by the participant. The outcome in this table is the amount of tokens earned in the experiment that is donated to the randomly selected charity. In Nairobi, "Matched with ingroup charity" is an indicator taking on a value of one if a

participant is matched to a charity associated with her ethnicity, and 0 otherwise. In California, "Matched with ingroup charity" is an indicator taking on a value of one if a participant has resided in the San Francisco Bay Area for five years or more and is matched with a charity in the San Francisco Bay Area. "Earnings in tokens" captures the amount of tokens earned in the experiment. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table F.2.13: Tests of overall significance on 10 primary outcomes

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	14.14 (0.167) {0.224}	10.39 (0.407) {0.408}	17.89 (0.057) {0.139}	33.61 (0.000)** {0.004}
Male × Heat		13.02 (0.223) {0.296}		
Nairobi × Heat		44.95 (0.000)** {0.000}**		

Note: Standard errors clustered at the session level. Wald statistics are presented, as are p-values from the the Wald test for joint hypotheses in parentheses and exact p-values generated through 10,000 permutations, as pre-specified, in curly brackets. As prespecified, this table contains results from tests of overall significance on the 10 primary outcomes for which this is possible (thus, excluding the outcomes for Patience and Time Inconsistency). Although the tests were only prespecified for the set of coefficients associated with Heat in (1), the set of coefficients associated with the interaction between Male and Heat in (2), and the set of coefficients associated with the interaction between Nairobi and Heat in (2), this table also includes results for the set of coefficients associated with Heat in (2) as well as for the sets of coefficients associated with Heat from the site-specific specifications (3) and (4), which were not prespecified. * $p < .05$, ** $p < .01$

Table F.2.14: Tests of overall significance on 9 primary outcomes (also excluding Joy of Destruction)

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	10.01	7.78	8.13	10.20
	(0.350)	(0.556)	(0.521)	(0.335)
	{0.406}	{0.540}	{0.621}	{0.426}
Male × Heat		6.32		
		(0.708)		
		{0.762}		
Nairobi × Heat		7.10		
		(0.627)		
		{0.673}		

Note: Standard errors clustered at the session level. Wald statistics are presented, as are p-values from the the Wald test for joint hypotheses in parentheses and exact p-values generated through 10,000 permutations, as pre-specified, in curly brackets. This table contains results from tests of overall significance on the 9 primary outcomes, excluding the outcomes for the Joy of Destruction, Patience, and Time Inconsistency. Although the tests were only prespecified for the set of coefficients associated with Heat in (1), the set of coefficients associated with the interaction between Male and Heat in (2), and the set of coefficients associated with the interaction between Nairobi and Heat in (2), this table also includes results for the set of coefficients associated with Heat in (2) as well as for the sets of coefficients associated with Heat from the site-specific specifications (3) and (4), which were not prespecified. * $p < .05$, ** $p < .01$

F.3 Results with additional outcomes

Table F.3.1: Rational choice violation II

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-0.0119	-0.0152	-0.00701	-0.0153
	(0.313)	(0.461)	(0.660)	(0.375)
Male	0.000669	-0.0121	-0.0311	0.0337
	(0.955)	(0.527)	(0.055)	(0.055)
Male × Heat		0.0210		
		(0.401)		
Nairobi	0.0471	0.0554		
	(0.000)**	(0.003)**		
Nairobi × Heat		-0.0153		
		(0.503)		
Weight		0.000268		-0.000780
		(0.662)		(0.478)
Weight originally missing		-0.00558		-0.0477
		(0.838)		(0.413)
Outcome control mean	0.0896	0.0896	0.0618	0.1138
R-squared	0.00771	0.00822	0.00446	0.00409
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. "Weight" refers to self-reported weight (in kg) and "Weight originally missing" is an indicator variable for whether weight data was not entered by the participant. The outcome in this table is an indicator for having made a first order stochastic dominance (FOSD) violation by having selected coin 7 in menu A in the risk preference module. * $p < .05$, ** $p < .01$

Table F.3.2: Rational choice violation III

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	0.00628	-0.0126	-0.00000483	0.00921
	(0.458)	(0.106)	(0.999)	(0.538)
Male	0.00603	-0.0135	-0.00106	0.00455
	(0.487)	(0.260)	(0.879)	(0.766)
Male × Heat		0.0322		
		(0.035)*		
Nairobi	0.0491	0.0513		
	(0.000)**	(0.000)**		
Nairobi × Heat		0.00269		
		(0.866)		
Weight		0.000479		0.00164
		(0.150)		(0.055)
Weight originally missing		0.0290		0.0299
		(0.137)		(0.539)
Outcome control mean	0.0330	0.0330	0.0092	0.0539
R-squared	0.0188	0.0228	0.0000296	0.00490
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. "Weight" refers to self-reported weight (in kg) and "Weight originally missing" is an indicator variable for whether weight data was not entered by the participant. The outcome in this table is an indicator for having made a violation of the Generalized Axiom of Revealed Preference (GARP) between menu A and B or between menu C and D in the time preference module. * $p < .05$, ** $p < .01$

Table F.3.3: Patience, non-parametric, menu A

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	0.0318	0.0477	0.0341	0.0439
	(0.673)	(0.689)	(0.744)	(0.683)
Male	-0.113	-0.0571	-0.0553	-0.120
	(0.119)	(0.595)	(0.595)	(0.248)
Male × Heat		-0.0416		
		(0.794)		
Nairobi	-1.127	-1.164		
	(0.000)**	(0.000)**		
Nairobi × Heat		0.0162		
		(0.916)		
Weight		-0.00467		-0.00939
		(0.216)		(0.133)
Weight originally missing		-0.188		-0.285
		(0.277)		(0.368)
Outcome control mean	4.471	4.471	5.087	3.934
R-squared	0.117	0.118	0.000398	0.00594
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. "Weight" refers to self-reported weight (in kg) and "Weight originally missing" is an indicator variable for whether weight data was not entered by the participant. The outcome in this table is the choice from Menu A, where choice 1 (with value 1) equals 840 tokens today and 0 tokens in 3 weeks, and choice 6 (with value 6) equals 0 tokens today and 1200 tokens in 3 weeks (the trade-off is fixed between 1 token today or 1.43 tokens in 3 weeks). * $p < .05$, ** $p < .01$

Table F.3.4: Patience, non-parametric, menu B

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-0.116	-0.165	-0.128	-0.0993
	(0.151)	(0.255)	(0.294)	(0.355)
Male	-0.411	-0.437	-0.456	-0.353
	(0.000)**	(0.000)**	(0.000)**	(0.002)**
Male × Heat		0.0891		
		(0.586)		
Nairobi	1.393	1.377		
	(0.000)**	(0.000)**		
Nairobi × Heat		0.00540		
		(0.973)		
Weight		-0.00245		-0.00304
		(0.519)		(0.660)
Weight originally missing		-0.0617		0.234
		(0.717)		(0.503)
Outcome control mean	2.803	2.803	2.119	3.399
R-squared	0.131	0.131	0.0164	0.0125
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. "Weight" refers to self-reported weight (in kg) and "Weight originally missing" is an indicator variable for whether weight data was not entered by the participant. The outcome in this table is the choice from Menu B, where choice 1 (with value 1) equals 1020 tokens today and 0 tokens in 3 weeks, and choice 6 (with value 6) equals 0 tokens today and 1020 tokens in 3 weeks (the trade-off is fixed between 1 token today or 1 token in 3 weeks). * $p < .05$, ** $p < .01$

Table F.3.5: Patience, non-parametric, menu C

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-0.126	-0.0984	-0.176	-0.0784
	(0.120)	(0.483)	(0.172)	(0.442)
Male	-0.201	-0.0540	-0.0231	-0.352
	(0.014)*	(0.665)	(0.835)	(0.005)**
Male × Heat		-0.205		
		(0.260)		
Nairobi	-1.305	-1.426		
	(0.000)**	(0.000)**		
Nairobi × Heat		0.159		
		(0.357)		
Weight		-0.00611		-0.00284
		(0.150)		(0.639)
Weight originally missing		-0.311		-0.175
		(0.089)		(0.623)
Outcome control mean	4.290	4.290	5.039	3.637
R-squared	0.143	0.147	0.00256	0.0128
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. "Weight" refers to self-reported weight (in kg) and "Weight originally missing" is an indicator variable for whether weight data was not entered by the participant. The outcome in this table is the choice from Menu C, where choice 1 (with value 1) equals 840 tokens in 3 weeks and 0 tokens in 7 weeks, and choice 6 (with value 6) equals 0 tokens in 3 weeks and 1200 tokens in 7 weeks (the trade-off is fixed between 1 token in 3 weeks or 1.43 tokens in 7 weeks). * $p < .05$, ** $p < .01$

Table F.3.6: Patience, non-parametric, menu D

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-0.0168	-0.127	-0.118	0.0860
	(0.818)	(0.268)	(0.230)	(0.420)
Male	-0.196	-0.160	-0.179	-0.164
	(0.007)**	(0.141)	(0.063)	(0.143)
Male × Heat		0.0162		
		(0.911)		
Nairobi	1.492	1.374		
	(0.000)**	(0.000)**		
Nairobi × Heat		0.196		
		(0.174)		
Weight		-0.00541		-0.00949
		(0.142)		(0.128)
Weight originally missing		0.00635		-0.0406
		(0.973)		(0.903)
Outcome control mean	2.519	2.519	1.799	3.148
R-squared	0.169	0.171	0.00468	0.00621
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. "Weight" refers to self-reported weight (in kg) and "Weight originally missing" is an indicator variable for whether weight data was not entered by the participant. The outcome in this table is the choice from Menu D, where choice 1 (with value 1) equals 1020 tokens in 3 weeks and 0 tokens in 7 weeks, and choice 6 (with value 6) equals 0 tokens in 3 weeks and 1020 tokens in 7 weeks (the trade-off is fixed between 1 token in 3 weeks or 1 token in 7 weeks). * $p < .05$, ** $p < .01$

Table F.3.7: Trustworthiness

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	0.00855	0.0290	0.00878	0.00619
	(0.407)	(0.058)	(0.507)	(0.697)
Male	-0.0258	-0.000542	-0.0336	-0.0218
	(0.022)*	(0.973)	(0.039)*	(0.156)
Male × Heat		-0.0468		
		(0.034)*		
Nairobi	0.100	0.0949		
	(0.000)**	(0.000)**		
Nairobi × Heat		0.00900		
		(0.671)		
Weight		-0.000258		-0.000194
		(0.651)		(0.830)
Weight originally missing		-0.00365		0.00558
		(0.867)		(0.881)
sharesentto	0.115	0.115	0.201	0.00875
	(0.000)**	(0.000)**	(0.000)**	(0.782)
sharesent	0.271	0.272	0.188	0.383
	(0.000)**	(0.000)**	(0.000)**	(0.000)**
Outcome outcome mean	0.245	0.245	0.227	0.262
R-squared	0.150	0.152	0.155	0.176
Observations	1580	1580	731	849

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. "Weight" refers to self-reported weight (in kg) and "Weight originally missing" is an indicator variable for whether weight data was not entered by the participant. The outcome in this table is the share of tokens (received and multiplied by 3) given back to the anonymous partner in the trust game. sharesentto captures the share of 600 tokens that the recipient received from the anonymous partner in the first round. sharesent captures how much the share of 600 tokens sent that the individual sent to her anonymous partner in the first round. * $p < .05$, ** $p < .01$

Table F.3.8: Trust measure

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-0.0288 (0.819)	0.0900 (0.629)	0.105 (0.534)	-0.192 (0.298)
Male	-0.142 (0.280)	-0.253 (0.197)	-0.504 (0.005)**	0.0609 (0.772)
Male × Heat		0.0500 (0.830)		
Nairobi	0.388 (0.004)**	0.526 (0.006)**		
Nairobi × Heat		-0.287 (0.271)		
Weight		0.00922 (0.190)		0.0254 (0.028)*
Weight originally missing		-0.683 (0.007)**		-1.087 (0.040)*
Outcome control mean	6.220	6.220	5.963	6.443
R-squared	0.00475	0.00988	0.00971	0.0114
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. "Weight" refers to self-reported weight (in kg) and "Weight originally missing" is an indicator variable for whether weight data was not entered by the participant. The outcome in this table is the number chosen from a 0-to-10 scale that asked how well does the following statement described one as a person: as long as I am not convinced otherwise, I assume that people have only the best intentions (where 0 means does not describe one at all, and a 10 means describes one perfectly). * $p < .05$, ** $p < .01$

Table F.3.9: Correct beliefs about others' contributions

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-0.000439	-0.0300	0.00687	-0.00329
	(0.982)	(0.376)	(0.822)	(0.902)
Male	0.00693	-0.0414	0.00145	0.0234
	(0.687)	(0.117)	(0.952)	(0.366)
Male × Heat		0.0922		
		(0.011)*		
Nairobi	0.00303	0.0238		
	(0.884)	(0.422)		
Nairobi × Heat		-0.0374		
		(0.367)		
Weight		0.000309		-0.00237
		(0.737)		(0.118)
Weight originally missing		0.0166		-0.0687
		(0.648)		(0.271)
Outcome control mean	0.212	0.212	0.206	0.218
R-squared	0.000101	0.00330	0.0000742	0.00337
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. "Weight" refers to self-reported weight (in kg) and "Weight originally missing" is an indicator variable for whether weight data was not entered by the participant. The outcome in this table is an indicator for having guessed correctly about one other person's contribution in the group towards the fund in the public goods module. * $p < .05$, ** $p < .01$

Table F.3.10: Time spent on Cognitive Reflection Test

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	1.188 (0.338)	1.816 (0.421)	2.043 (0.354)	0.414 (0.751)
Male	-2.885 (0.017)*	-2.816 (0.099)	-7.421 (0.000)**	1.252 (0.373)
Male × Heat		0.540 (0.816)		
Nairobi	8.389 (0.000)**	8.843 (0.000)**		
Nairobi × Heat		-1.678 (0.538)		
Weight		-0.0481 (0.471)		-0.00741 (0.919)
Weight originally missing		-3.442 (0.114)		-6.697 (0.081)
Outcome control mean	167.8	167.8	163.3	171.8
R-squared	0.0287	0.0305	0.0172	0.00531
Observations	1866	1866	862	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. "Weight" refers to self-reported weight (in kg) and "Weight originally missing" is an indicator variable for whether weight data was not entered by the participant. The outcome in this table is the total amount of time spent on the Cognitive Reflection Test, where participants were allowed 3 minutes in total to complete 5 questions. * $p < .05$, ** $p < .01$

Table F.3.11: Happiness

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-0.0791 (0.208)	0.0415 (0.668)	0.0490 (0.579)	-0.180 (0.042)*
Male	0.174 (0.005)**	0.149 (0.129)	0.208 (0.025)*	0.179 (0.040)*
Male × Heat		0.0188 (0.886)		
Nairobi	1.127 (0.000)**	1.242 (0.000)**		
Nairobi × Heat		-0.248 (0.071)		
Weight		0.00150 (0.698)		-0.00764 (0.071)
Weight originally missing		-0.246 (0.067)		-0.451 (0.050)*
Outcome control mean	5.313	5.313	4.625	5.914
R-squared	0.180	0.184	0.00618	0.0168
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. "Weight" refers to self-reported weight (in kg) and "Weight originally missing" is an indicator variable for whether weight data was not entered by the participant. The outcome in this table is the number chosen from a 1-7 scale that asked how one felt towards the end of the experiment, with 1 being sad and 7 being happy. * p < .05, ** p < .01

Table F.3.12: Alertness

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-0.420 (0.000)**	-0.579 (0.000)**	-0.530 (0.000)**	-0.308 (0.002)**
Male	0.389 (0.000)**	0.332 (0.003)**	0.729 (0.000)**	0.120 (0.248)
Male × Heat		0.115 (0.436)		
Nairobi	1.817 (0.000)**	1.710 (0.000)**		
Nairobi × Heat		0.181 (0.230)		
Weight		-0.000564 (0.892)		-0.00947 (0.099)
Weight originally missing		-0.235 (0.122)		-0.296 (0.324)
Outcome control mean	5.247	5.247	4.281	6.090
R-squared	0.303	0.306	0.0739	0.0175
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. "Weight" refers to self-reported weight (in kg) and "Weight originally missing" is an indicator variable for whether weight data was not entered by the participant. The outcome in this table is the number chosen from a 1-7 scale that asked how one felt towards the end of the experiment, with 1 being tired and 7 being alert. * p < .05, ** p < .01

F.4 Fairness, alternate specification

Table F.4.1: Fairness (alternate specification)

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-0.0168	-0.0204	-0.0245	-0.0102
	(0.169)	(0.262)	(0.111)	(0.579)
Male	-0.0371	-0.0308	-0.0710	-0.00722
	(0.006)**	(0.084)	(0.000)**	(0.690)
Male × Heat		-0.0121		
		(0.599)		
Nairobi	0.0458	0.0364		
	(0.034)*	(0.144)		
Nairobi × Heat		0.0189		
		(0.427)		
Weight	-0.000577	-0.000601	-0.000348	-0.000280
	(0.300)	(0.277)	(0.605)	(0.773)
Weight originally missing	0.00318	0.00339	-0.00388	0.0118
	(0.886)	(0.878)	(0.873)	(0.798)
High	-0.0123	-0.0126	-0.0259	0.00628
	(0.341)	(0.330)	(0.126)	(0.742)
Number of correct sliders	0.00131	0.00133	0.000309	0.000986
	(0.359)	(0.352)	(0.900)	(0.583)
Outcome control mean	0.318	0.318	0.309	0.326
R-squared	0.0114	0.0118	0.0356	0.00161
Observations	1878	1878	874	1004

Standard errors clustered at the session level. Per-comparison p-values in parentheses. Fairness here refers to the real effort dictator game, where the level of endowment is affected by the number of correct sliders made in the precision task. The outcome in this table is the share of joint earnings (2400 tokens in the high group, 1200 tokens in the low group) that each participant desires to give to the anonymous partner. These specifications follow the robustness check mentioned in the pre-analysis plan, including in the regressions whether an indicator for being in the high precision group ("High", as opposed to being "Low"), which yielded different endowments for the real effort dictator game. The number of sliders correctly chosen in the precision task ("Number of correct sliders") was also included, as any effect from being in the high or low group may be due to precision, and not stakes per se. * $p < .05$, ** $p < .01$

G   Robustness checks

G.1   Description

Although the pre-analysis plan specified that the risk measure outcome would be a categorical variable indicating coin choice from menu A, both interpretation and accounting for multiple hypothesis testing motivated the primary outcome for risk to be defined as the variance of the coin toss in menu A. In Supplement Section G.2 we investigate the effect of the temperature treatment on risk where 1) we assume a mean-variance utility function, and then use the backed out lambda parameter as the outcome variable with OLS, and 2) use a categorical indicator of choice with an ordered logit and ordered probit estimation. However, *Heat* is not statistically significant at conventional levels in any of the tables presented.

The Joy of Destruction outcome presented throughout the analysis is the share of vouchers or gift cards destroyed. However, observing whether there is an effect of *any* voucher or gift card destruction is also interesting and may be more comparable to other studies. Supplement Table G.3.1 repeats the analysis with the dependent variable as whether the participant engaged in any destruction. The results mirror that of the original Joy of Destruction outcome definition, with participants 17 percentage points more likely to engage in any destruction as a result of treatment.

Given the strong results in Nairobi along ethnic lines, one might consider whether the share of ethnicities within a room affects one's performance in the Joy of Destruction. Supplement Table G.4.1 examines how the ethnic composition of others in the room relative to oneself may affect the Joy of Destruction result, by ethnic group (ethnic Luo, ethnic Other, ethnic Kikuyu). Each specification is carried out over a different reference ethnic group. The coefficient on the treatment effect is large (14 percentage points) for the ethnic Luo subsample, but it is not statistically significant, likely due to low power. The average treatment effect of heat for the ethnic Other group (11 percentage points) is statistically significant.

Given the history of ethnic-based politics in Kenya and the fact that the experiment took place during a heavily-contested election period (see Appendix Section A.1 for details), Supplement Section G.5 includes indicators for ethnicity (ethnic Kikuyu or ethnic Other) and their interactions for all selected outcomes, with ethnic Luos as the reference group. Here it is clear that ethnicity does not differentially change the treatment effect for any primary outcomes examined except for the Joy of Destruction. Ethnic Kikuyus also report being less alert by the end of the experiment as a result of the treatment.

In order to gauge whether there may have been a general experimenter demand effect in California, we include within our specifications an indicator for whether the participant noted temperature as an object of study within the post-experiment survey, as well as its interaction with the heat treatment. We carry out this analysis for the pooled sample as well as the California and Nairobi samples. As mentioned in the main text, there is no systematic interaction between awareness of the underlying temperature treatment and the treatment effect for the primary outcomes of interest (see Supplement Section G.6), suggesting that the effect of heat is unlikely to be due to a general experimenter demand effect.

Additionally, we study whether gender composition in the lab affect responses to heat exposure. The majority of participants in Nairobi were male, while the majority of participants in California were female (details shown Appendix Table C.1). and a relevant question

is weather this may explain some of the differences observed between the Nairobi and the California site. We test this hypothesis by including in the main and heterogeneity specification the share of other participants in the room who are male, as well as its interaction with treatment (shown in Supplement Section G.7). In the main specification, an increase in the share of others being male increases destruction in the treatment rooms, but with the introduction of interactions in the heterogeneity model, the effect is no longer statistically significant. Within the California sample, an increase in the share of others being male reduces the share of tokens given to the partner in the treatment room by 19 percentage points. Finally, an increase in the share of others being male lowers self-reported happiness in the treatment room.

Supplement Section G.8 includes session-fixed effects in order to account for the importance of unobserved heterogeneity. In the bottom part of the tables we include the F-statistic from the joint F-test on the fixed effects in a certain specification. We find that the results of the analysis do not substantively change compared to the results in the Appendix Section D.

Supplement Section G.9 includes variables for outdoor weather. "Outside temperature (pooled)" captures the demeaned (relative to the pooled sample) average outside temperature in Celsius for the day of the session. "Outside relative humidity (pooled)" captures the demeaned (relative to the pooled sample) average outside relative humidity in percentage for the day of the session. "Outside temperature (site)" captures the demeaned (relative to the site sample) average site-specific outside temperature in Celsius for the day of the session. Finally, "Outside relative humidity (site)" captures the demeaned (relative to the site sample outside relative humidity) average site-specific outside relative humidity in percentage for the day of the session. Our specifiations also include the above variables interacted with treatment status. The treatment effects observed are similar to those in Appendix Section D, with outside weather seeming to explain little of the variation. This may be due to adaptation mechanisms available to students (i.e., they may be in air-conditioned buildings before heading over to the laboratories). The only instances in which there is a significant interaction effect is in Charitable donation, where the temperature treatment lowers the amount of tokens donated at higher daily mean outdoor temperatures in the heterogeneity specification and lowers the amount of tokens donated at higher daily mean outdoor relative humidities in the Nairobi specification. The temperature treatment also lowers alertness in Nairobi at higher daily mean outdoor temperatures.

G.2 Risk outcome: alternative specifications

Table G.2.1: Risk, lambda

	(1) Pooled
Heat	-0.000100
	(0.653)
Male	-0.000633
	(0.006)**
Nairobi	0.000778
	(0.000)**
Outcome control mean	0.00387
R-squared	0.0113
Observations	1721

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. The risk outcome in this table is backed out lambda parameter coming from a mean-variance utility function using risk menu A, where coins 1, 2 , 3, 4, 5, and 6 are associated with lambda parameters of 0.0002314814815, 0.0003086419753, 0.000462962963, 0.0009259259259, 0.001851851852, and 0.01, respectively. Note that this approach results in linearity of the outcome variable across coin 1 through coin 6. Choosing coin 7 is first-order stochastically dominated and so participants choosing coin 7 in risk menu A are dropped from this calculation. * $p < .05$, ** $p < .01$

Table G.2.2: Risk, ordered logit model

		Marginal effects						
	(1) Risk	(2) Coin 1	(3) Coin 2	(4) Coin 3	(5) Coin 4	(6) Coin 5	(7) Coin 6	(8) Coin 7
Heat	-0.0748	0.00624	0.00296	0.00327	0.00547	-0.000683	-0.0117	-0.00555
	(0.367)	(0.366)	(0.368)	(0.369)	(0.369)	(0.425)	(0.366)	(0.369)
Male	-0.462	0.038	0.0181	0.0200	0.0337	-0.00330	-0.0717	-0.0349
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.162)	(0.000)**	(0.000)
Nairobi	0.549	-0.047	-0.0219	-0.0238	-0.0389	0.00604	0.0850	0.0404
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.041)*	(0.000)**	(0.000)**
Outcome control mean	4.613							
Pseudo R-squared	0.0085							
Observations	1878							

Note: Standard errors generated via bootstrap (with 1000 replications). Per-comparison p-values in parentheses. This table features results from both the ordered logit model on the coin toss chosen (from a set of seven) in risk menu A in (1), as well as the marginal effects for each coin toss in risk menu A in (2) through (8). * $p < .05$, ** $p < .01$

Table G.2.3: Risk, ordered probit model

	(1) Risk	Marginal effects						
		(2) Coin 1	(3) Coin 2	(4) Coin 3	(5) Coin 4	(6) Coin 5	(7) Coin 6	(8) Coin 7
Heat	-0.0339	0.00552	0.00212	0.00215	0.00328	-0.00037	-0.00763	-0.00507
	(0.475)	(0.474)	(0.476)	(0.478)	(0.477)	(0.505)	(0.475)	(0.476)
Male	-0.277	0.0448	0.0172	0.0175	0.0270	-0.00237	-0.0619	-0.0422
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.189)	(0.000)**	(0.000)**
Nairobi	0.328	-0.0545	-0.0204	-0.0205	-0.0309	0.00438	0.0735	0.0485
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.047)*	(0.000)**	(0.000)**
Outcome control mean	4.613							
Pseudo R-squared	0.0090							
Observations	1878							

Note: Standard errors generated via bootstrap (with 1000 replications). Per-comparison p-values in parentheses. This table features results from both the ordered probit model on the coin toss chosen (from a set of seven) in risk menu A in (1), as well as the marginal effects for each coin toss in risk menu A in (2) through (8). * $p < .05$, ** $p < .01$

G.3 Joy of Destruction: engagement in any destruction

Table G.3.1: Joy of Destruction (any destruction)

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	0.0721	-0.00100	-0.0366	0.166
	(0.000)**	(0.967)	(0.075)	(0.000)**
Male	0.0422	0.0858	0.00364	0.0770
	(0.033)*	(0.002)**	(0.862)	(0.016)*
Male × Heat		-0.0881		
		(0.020)*		
Nairobi	0.287	0.174		
	(0.000)**	(0.000)**		
Nairobi × Heat		0.225		
		(0.000)**		
Outcome control mean	0.215	0.215	0.111	0.306
R-squared	0.126	0.142	0.00403	0.0348
Observations	1859	1859	864	995

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. The outcome in this table is an indicator variable for any destruction of the anonymous partner's earned Airtime Vouchers (Amazon Gift Cards in the California sample) destroyed by the participant. The earned Airtime Vouchers and Amazon Gift Cards resulted from the number of Raven's Progressive Matrices answered correctly in the previous module, where one puzzle answered correctly yielded one Airtime Voucher or one Amazon Gift Card. Airtime Vouchers were worth 50 KSh each, while Amazon Gift Cards were worth 1 dollar each. * $p < .05$, ** $p < .01$

G.4 Joy of Destruction: share of ethnicities in room

Table G.4.1: Joy of Destruction, by ethnicity

	(1) Luo	(2) Other	(3) Kikuyu
Heat	0.138	0.112	0.00501
	(0.394)	(0.016)*	(0.975)
Male	0.0443	0.0481	0.0818
	(0.293)	(0.058)	(0.034)*
Share Kikuyu	0.238	-0.0104	
	(0.155)	(0.899)	
Share Kikuyu × Heat	-0.121	0.00376	
	(0.666)	(0.979)	
Share Other	-0.0182		0.125
	(0.830)		(0.200)
Share Other × Heat	0.0726		-0.107
	(0.727)		(0.627)
Share Luo		0.0355	-0.0522
		(0.734)	(0.676)
Share Luo × Heat		-0.149	0.131
		(0.275)	(0.533)
Outcome control mean	0.119	0.155	0.135
R-squared	0.0813	0.0287	0.0436
Observations	195	549	193

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. The outcome in this table is the share of the anonymous partner's earned Airtime Vouchers (Amazon Gift Cards in the California sample) destroyed by the participant. The earned Airtime Vouchers and Amazon Gift Cards resulted from the number of Raven's Progressive Matrices answered correctly in the previous module, where one puzzle answered correctly yielded one Airtime Voucher or one Amazon Gift Card. Airtime Vouchers were worth 50 KSh each, while Amazon Gift Cards were worth 1 dollar each. "Share Kikuyu" captures the share of participants in the room, excluding the participant herself or himself, who identify as Kikuyu, out of those who identify with an ethnicity. "Share Other" captures the share of participants in the room, excluding the participant herself or himself, who identify as an ethnicity that is neither Kikuyu nor Luo, out of those who identify with an ethnicity. "Share Luo" captures the share of participants in the room, excluding the participant herself or himself, who identify as Luo, out of those who identify with an ethnicity. Note that (1) conditions on the Luo sample, (2) conditions on the non-Luo and non-Kikuyu sample, and (3) conditions on the Kikuyu sample. * $p < .05$, ** $p < .01$

G.5 Selected outcomes: by ethnicity (Nairobi only)

Table G.5.1: Outcomes, by ethnicity (Nairobi sample only)

	(1) Precision	(2) Fairness	(3) Risk-taking	(4) RCV I	(5) Trust	(6) PC	(7) FI	(8) JoD	(9) CR	(10) Charity	(11) Happiness	(12) Alertness
Heat	-0.569 (0.438) [1.000]	-0.00899 (0.839) [1.000]	-36.34 (0.450) [1.000]	-0.0245 (0.256) [0.887]	0.0564 (0.186) [0.887]	33.48 (0.489) [1.000]	0.0133 (0.615) [1.000]	0.147 (0.002)** [0.031]+	0.0327 (0.234) [0.887]	-173.1 (0.135) [0.887]	-0.0672 (0.725)	0.0316 (0.877)
Male	2.111 (0.000)**	-0.00845 (0.633)	37.36 (0.170)	-0.00152 (0.893)	0.0852 (0.000)**	64.73 (0.037)*	0.0176 (0.248)	0.0531 (0.002)**	0.00631 (0.676)	72.55 (0.164)	0.124 (0.151)	0.102 (0.302)
Matched										-117.0 (0.089)		
Earnings										0.0268 (0.138)		
Other eth	-0.914 (0.167)	-0.0357 (0.276)	-45.70 (0.324)	-0.00808 (0.715)	0.0111 (0.769)	-7.040 (0.868)	-0.0366 (0.130)	0.0360 (0.257)	0.0387 (0.130)	-160.6 (0.060)	0.0966 (0.562)	0.230 (0.218)
Other eth × Heat	1.013 (0.261)	0.0250 (0.625)	83.89 (0.171)	0.0307 (0.242)	0.0000848 (0.999)	4.846 (0.933)	0.0218 (0.482)	-0.0637 (0.190)	-0.0225 (0.498)	209.6 (0.073)	-0.0979 (0.630)	-0.368 (0.101)
Kikuyu	0.828 (0.323)	0.00871 (0.829)	-15.48 (0.776)	0.00227 (0.930)	0.0369 (0.424)	-12.30 (0.827)	0.0357 (0.234)	0.0294 (0.417)	0.0722 (0.016)*	-119.1 (0.253)	0.166 (0.372)	0.329 (0.089)
Kikuyu × Heat	1.268 (0.253)	-0.0153 (0.799)	6.279 (0.935)	-0.0116 (0.671)	-0.0508 (0.335)	-21.06 (0.768)	0.0142 (0.716)	-0.172 (0.003)**	0.00336 (0.933)	246.9 (0.115)	-0.258 (0.301)	-0.696 (0.021)*
Outcome control mean	13.11	0.319	321.5	0.0297	0.342	433.5	0.799	0.144	0.212	561.8	5.907	6.091
R-squared	0.0427	0.00372	0.00427	0.00490	0.0263	0.00870	0.0224	0.0425	0.0166	0.0120	0.00906	0.0202
Observations	945	945	945	945	945	945	945	937	945	945	945	945

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat for each primary outcome variable in Supplement Tables G.5.1 and G.5.2. "Precision" refers to the precision task, Fairness refers to Fairness, "Risk-taking" refers to Risk-taking, "RCV I" refers to Rational choice violation I, "Trust" refers to Trust, "PC" refers to Public contribution, "FI" refers to Fluid intelligence, "JoD" refers to Joy of Destruction, "CR" refers to Cognitive reflection, "Charity" refers to Charitable donation, "Happiness" refers to Happiness, and "Alertness" refers to Alertness. In Nairobi, "Matched" is an indicator taking on a value of one if a participant is matched to a charity associated with her ethnicity, and 0 otherwise. In California, "Matched" is an indicator taking on a value of one if a participant has resided in the San Francisco Bay Area for five years or more and is matched with a charity in the San Francisco Bay Area. "Earnings" is earnings in tokens from the experiment. "Other eth" is an indicator variable for the individual self-identifying as an ethnicity that is not Kikuyu or Luo. "Kikuyu" is an indicator variable for the individual self-identifying as Kikuyu. Thus, the treatment effect on Heat is for the Luo population. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.5.2: Time preference parameters (Nairobi sample only)

	(1) Patience	(2) Time inconsistency
Heat	0.0011791	0.00237012
	(0.203)	(0.929)
	[0.887]	[1.000]
Outcome control mean	0.9941125	1.045024
Observations	3220	3220

Note: Standard errors in parentheses, clustered at the individual level. Per comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. "Heat" refers to the difference between the aggregate beta or delta statistics for treatment and control individuals, divided by the square root of the sum of squared standard errors. Multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat for each primary outcome variable in this table and in Supplement Table G.5.1. Note that the effect for time inconsistency is multiplied by -1, so that a positive difference reflects more time inconsistency. Also note that the non-linear least squares specification does not allow one to control for ethnicity, and thus the results (aside from the q-values) are similar in construction to those in Appendix Tables D.1.5 and D.1.6. The outcomes in this table are the aggregate δ (Patience) and β (Time inconsistency) statistics from the non-linear least squares specification featured in Andreoni et al. (2015), carried out at the individual-choice level. β measures present bias, and values less than 1 denote time inconsistency. δ is the daily discount factor between two future days. For comparability with regression results, individuals who did not respond "Female" or "Male" to the gender survey question were dropped (2% of the sample) as well as those who did not self-identify with a listed ethnicity in the demographic survey. Before estimation of aggregate parameters, individuals who never altered their decision from a specific corner solution in all convex time budgets were dropped (as they provided insufficient variation for the calculation of utility parameters) as were individuals who exhibited generalized axiom of revealed preference (GARP) violations. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

G.6 Primary outcomes: suspecting treatment

Table G.6.1: Outcomes, controlling for suspecting temperature and its interaction with treatment

	(1) Precision	(2) Altruism	(3) Risk-taking	(4) RCV I	(5) Trust	(6) PC	(7) FI	(8) JoD	(9) CR	(10) Charity
Heat	0.640	-0.0224	12.88	-0.00963	0.00611	1.044	0.0220	0.0540	0.000104	48.17
	(0.094)	(0.335)	(0.668)	(0.458)	(0.790)	(0.974)	(0.119)	(0.016)*	(0.995)	(0.391)
Male	2.257	-0.0410	156.2	0.0157	0.0627	27.42	0.0423	0.0182	0.0532	-36.78
	(0.000)**	(0.042)*	(0.000)**	(0.184)	(0.003)**	(0.354)	(0.004)**	(0.277)	(0.013)*	(0.446)
Nairobi	-11.42	0.0266	-139.0	-0.00412	-0.175	-167.3	-0.124	0.141	-0.245	330.5
	(0.000)**	(0.234)	(0.000)**	(0.746)	(0.000)**	(0.001)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
Matched										-38.79
										(0.447)
Earnings										0.0181
										(0.246)
Suspect	0.401	-0.0809	16.69	0.0150	-0.128	25.85	0.0158	0.0781	-0.0135	-68.33
	(0.552)	(0.130)	(0.810)	(0.678)	(0.003)**	(0.741)	(0.552)	(0.122)	(0.694)	(0.502)
Suspect × Heat	-0.794	0.0740	-92.69	-0.0315	0.105	-50.35	-0.00220	-0.0880	0.0353	108.2
	(0.377)	(0.224)	(0.254)	(0.417)	(0.042)*	(0.586)	(0.941)	(0.140)	(0.425)	(0.414)
R-squared	0.534	0.0113	0.0403	0.00528	0.0732	0.0350	0.115	0.0770	0.187	0.0598
Observations	896	896	896	896	896	896	896	892	896	864

Note: Standard errors clustered at the session level. Per-comparison p values in parentheses. "Precision" refers to the precision task, "Fairness" refers to Fairness, "Risk-taking" refers to Risk-taking, "RCV I" refers to Rational choice violation I, "Trust" refers to Trust, "PC" refers to Public contribution, "FI" refers to Fluid intelligence, "JoD" refers to Joy of Destruction, "CR" refers to Cognitive reflection, and "Charity" refers to Charitable donation. In Nairobi, "Matched" is an indicator taking on a value of one if a participant is matched to a charity associated with her ethnicity, and 0 otherwise. In California, "Matched" is an indicator taking on a value of one if a participant has resided in the San Francisco Bay Area for five years or more and is matched with a charity in the San Francisco Bay Area. "Earnings" is earnings in tokens from the experiment. "Suspect" is an indicator for whether the participant indicated temperature in the debriefing survey as something the experiment was studying. * $p < .05$, ** $p < .01$

Table G.6.2: Outcomes, controlling for suspecting temperature and its interaction with treatment, California

	(1) Precision	(2) Altruism	(3) Risk-taking	(4) RCV I	(5) Trust	(6) PC	(7) FI	(8) JoD	(9) CR	(10) Charity
Heat	-0.164	-0.0391	67.82	-0.0210	-0.0683	-31.10	-0.0150	-0.0308	-0.0362	177.2
	(0.791)	(0.171)	(0.138)	(0.411)	(0.101)	(0.600)	(0.288)	(0.240)	(0.377)	(0.018)*
Male	2.256	-0.109	358.3	0.0342	0.0750	7.401	0.0214	0.0185	0.176	-136.9
	(0.000)**	(0.000)**	(0.000)**	(0.076)	(0.048)*	(0.871)	(0.161)	(0.471)	(0.000)**	(0.023)*
Matched										4.026
										(0.943)
Earnings										-0.0334
										(0.056)
Suspect	-0.521	-0.0229	136.3	-0.0423	-0.153	48.20	0.0196	0.0593	-0.00276	281.2
	(0.517)	(0.708)	(0.307)	(0.040)*	(0.015)*	(0.690)	(0.472)	(0.549)	(0.969)	(0.097)
Suspect × Heat	1.098	0.0210	-208.9	0.0353	0.193	-99.86	0.00434	-0.0748	0.0490	-417.7
	(0.273)	(0.760)	(0.144)	(0.203)	(0.007)**	(0.459)	(0.884)	(0.450)	(0.481)	(0.028)*
R-squared	0.0808	0.0640	0.147	0.0183	0.0255	0.00715	0.0158	0.0186	0.0825	0.0537
Observations	329	329	329	329	329	329	329	327	329	326

Note: Standard errors clustered at the session level. Per-comparison p values in parentheses. "Precision" refers to the precision task, "Fairness" refers to Fairness, "Risk-taking" refers to Risk-taking, "RCV I" refers to Rational choice violation I, "Trust" refers to Trust, "PC" refers to Public contribution, "FI" refers to Fluid intelligence, "JoD" refers to Joy of Destruction, "CR" refers to Cognitive reflection, and "Charity" refers to Charitable donation. In California, "Matched" is an indicator taking on a value of one if a participant has resided in the San Francisco Bay Area for five years or more and is matched with a charity in the San Francisco Bay Area. "Earnings" is earnings in tokens from the experiment. "Suspect" is an indicator for whether the participant indicated temperature in the debriefing survey as something the experiment was studying. * $p < .05$, ** $p < .01$

Table G.6.3: Outcomes, controlling for suspecting temperature and its interaction with treatment, Nairobi

	(1) Precision	(2) Altruism	(3) Risk-taking	(4) RCV I	(5) Trust	(6) PC	(7) FI	(8) JoD	(9) CR	(10) Charity
Heat	0.975	-0.0141	-18.01	-0.00467	0.0376	17.40	0.0391	0.0938	0.0131	-6.038
	(0.044)*	(0.649)	(0.630)	(0.762)	(0.171)	(0.655)	(0.044)*	(0.002)**	(0.502)	(0.935)
Male	2.303	0.00376	22.00	0.00408	0.0566	38.40	0.0560	0.0157	-0.0283	27.32
	(0.000)**	(0.888)	(0.542)	(0.791)	(0.027)*	(0.336)	(0.011)*	(0.472)	(0.125)	(0.709)
Matched										-85.01
										(0.337)
Earnings										0.0557
										(0.017)*
Suspect	0.947	-0.114	-60.60	0.0485	-0.115	12.08	0.0136	0.0873	-0.0247	-261.5
	(0.326)	(0.121)	(0.444)	(0.389)	(0.052)	(0.905)	(0.724)	(0.126)	(0.446)	(0.011)*
Suspect × Heat	-2.603	0.109	-34.44	-0.0757	0.0404	54.75	0.0157	0.00433	0.0575	546.8
	(0.099)	(0.213)	(0.736)	(0.206)	(0.550)	(0.666)	(0.738)	(0.961)	(0.315)	(0.003)**
R-squared	0.0434	0.00653	0.00450	0.00575	0.0222	0.00450	0.0288	0.0347	0.00836	0.0283
Observations	567	567	567	567	567	567	567	565	567	538

Note: Standard errors clustered at the session level. Per-comparison p values in parentheses. "Precision" refers to the precision task, "Fairness" refers to Fairness, "Risk-taking" refers to Risk-taking, "RCV I" refers to Rational choice violation I, "Trust" refers to Trust, "PC" refers to Public contribution, "FI" refers to Fluid intelligence, "JoD" refers to Joy of Destruction, "CR" refers to Cognitive reflection, and "Charity" refers to Charitable donation. In Nairobi, "Matched" is an indicator taking on a value of one if a participant is matched to a charity associated with her ethnicity, and 0 otherwise. "Earnings" is earnings in tokens from the experiment. "Suspect" is an indicator for whether the participant indicated temperature in the debriefing survey as something the experiment was studying. * $p < .05$, ** $p < .01$

G.7 Selected outcomes: accounting for share of male participants in room

Table G.7.1: Precision task

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-0.110	0.0788	0.0484	-0.699
	(0.851)	(0.900)	(0.946)	(0.531)
	[1.000]	[1.000]	[1.000]	[1.000]
Male	1.947	2.175	1.842	2.063
	(0.000)**	(0.000)**	(0.000)**	(0.000)**
Male × Heat		-0.458		
		(0.316)		
		[1.000]		
Nairobi	-11.24	-11.01		
	(0.000)**	(0.000)**		
Nairobi × Heat		-0.458		
		(0.420)		
		[0.496]		
Share of male participants	-0.378	-0.657	0.00950	-1.097
	(0.587)	(0.354)	(0.992)	(0.288)
Share of male participants × Heat	1.168	1.730	1.352	1.762
	(0.220)	(0.089)	(0.311)	(0.257)
Outcome control mean	17.95	17.95	23.63	12.99
R-squared	0.550	0.550	0.0670	0.0312
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Within Supplement Section G.7, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes. Additionally, within Supplement Section G.7, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes. The precision task is also known as the slider task. The outcome in this table is the number of correct sliders made in three minutes. Final earnings from the production task are based off either being weakly above (high) or below (low) the median within treatment cohort. The median pair is randomly assigned to high or low. "Share of male participants" captures the share of participants in the room, excluding the participant herself or himself, who identify as male, out of those who identify as either male or female. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.7.2: Fairness

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	0.00877	0.00513	-0.0188	0.0718
	(0.738)	(0.842)	(0.494)	(0.179)
	[1.000]	[1.000]	[1.000]	[1.000]
Male	-0.0383	-0.0365	-0.0745	-0.00442
	(0.002)**	(0.042)*	(0.000)**	(0.786)
Male × Heat		-0.00355		
		(0.883)		
		[1.000]		
Nairobi	0.0382	0.0218		
	(0.011)*	(0.247)		
Nairobi × Heat		0.0327		
		(0.223)		
		[0.496]		
Share of male participants	0.00373	0.0149	-0.0388	0.0594
	(0.917)	(0.690)	(0.385)	(0.277)
Share of male participants × Heat	-0.0476	-0.0699	-0.0147	-0.126
	(0.291)	(0.187)	(0.796)	(0.106)
Outcome control mean	0.318	0.318	0.309	0.326
R-squared	0.0115	0.0124	0.0351	0.00490
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Within Supplement Section G.7, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes. Additionally, within Supplement Section G.7, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes. Fairness here refers to the real effort dictator game, where the level of endowment is determined by the number of correct sliders made in the precision task. The outcome in this table is the share of joint earnings (2400 tokens in the high group, 1200 tokens in the low group) that each participant desires to give to the anonymous partner. "Share of male participants" captures the share of participants in the room, excluding the participant herself or himself, who identify as male, out of those who identify as either male or female. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.7.3: Risk-taking

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-9.822	-5.218	50.93	-105.1
	(0.826)	(0.905)	(0.348)	(0.172)
	[1.000]	[1.000]	[1.000]	[1.000]
Male	166.5	177.3	311.5	30.83
	(0.000)**	(0.000)**	(0.000)**	(0.282)
Male × Heat		-21.67		
		(0.635)		
		[1.000]		
Nairobi	-110.4	-114.9		
	(0.000)**	(0.002)**		
Nairobi × Heat		9.068		
		(0.853)		
		[0.620]		
Share of male participants	-34.11	-35.98	125.4	-155.1
	(0.559)	(0.568)	(0.120)	(0.042)*
Share of male participants × Heat	14.79	18.83	-139.7	164.0
	(0.827)	(0.812)	(0.189)	(0.115)
Outcome control mean	366.3	366.3	409.0	329.1
R-squared	0.0399	0.0400	0.115	0.00534
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Within Supplement Section G.7, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes. Additionally, within Supplement Section G.7, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes. The outcome in this table is the variance of the coin toss from menu A, in tokens. Note that the expected value is not constant across each coin, so that the outcome does not capture the trade-off between expected value and variance. Note also that under this approach, Coin 7 will be as good as Coin 5, even though Coin 5 strictly dominates Coin 7 with expected utility. "Share of male participants" captures the share of participants in the room, excluding the participant herself or himself, who identify as male, out of those who identify as either male or female.
* $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.7.4: Rational choice violation I

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-0.000763 (0.960) [1.000]	-0.000362 (0.981) [1.000]	-0.00402 (0.847) [1.000]	0.00575 (0.799) [1.000]
Male	0.00184 (0.818)	0.00336 (0.784)	0.00577 (0.609)	-0.00231 (0.841)
Male × Heat		-0.00305 (0.856) [1.000]		
Nairobi	-0.00162 (0.816)	-0.00314 (0.784)		
Nairobi × Heat		0.00304 (0.854) [0.620]		
Share of male participants	0.0109 (0.560)	0.0113 (0.557)	-0.00391 (0.886)	0.0261 (0.329)
Share of male participants × Heat	-0.00658 (0.789)	-0.00730 (0.792)	-0.00137 (0.975)	-0.0151 (0.646)
Outcome control mean	0.0277	0.0277	0.0275	0.0279
R-squared	0.000448	0.000482	0.000556	0.00122
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Within Supplement Section G.7, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes. Additionally, within Supplement Section G.7, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes. The outcome in this table is an indicator of transitivity violation using both menus A & B. A transitivity violation comes from choosing two coins in the interior region of the intersection of both menus, where it is not the case that it can be said that one preferring coin A to coin B and then preferring coin B to coin C implies that one prefers coin A to coin C. "Share of male participants" captures the share of participants in the room, excluding the participant herself or himself, who identify as male, out of those who identify as either male or female. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.7.5: Patience

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-0.00005823 (0.854) [1.000]	-0.00005823 (0.854) [1.000]	-0.00060946 (0.093) [1.000]	0.00096101 (0.279) [1.000]
Heat (Male)		-0.0002442 (0.702) [1.000]		
Heat (Nairobi)		0.00157046 (0.102) [0.437]		
Outcome control mean	0.9938358	0.9938358	0.9937076	0.9939217
Observations	6612	6612	3200	3412

Note: Standard errors clustered at the individual level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. For (1) and (2), "Heat" refers to the difference between the aggregate delta statistics for the treatment and control groups, divided by the square root of the sum of squared standard errors. For (2), "Heat (Male)" refers to the difference between the aggregate delta statistics for male treatment and control groups, subtracting the difference between the aggregate delta statistics for female treatment and control groups, and then dividing by the square root of the sum of squared standard errors. "Heat (Nairobi)" refers to the difference between the aggregate delta statistics for Nairobi treatment and control groups, subtracting the difference between the aggregate delta statistics for California treatment and control groups, and then dividing by the square root of the sum of squared standard errors. (3) and (4) carry out a similar estimation to "Heat" for (1) and (2) but for the California sample and Nairobi sample, respectively. Within Supplement Section G.7, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes. Additionally, within Supplement Section G.7, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes. Given that the non-linear least squares specification used to measure patience does not incorporate interaction terms, we instead use p-values from analogous differences in this table for multiple testing adjustments. Also note that the non-linear least squares specification does not allow one to control for share of male participants, and thus the results (aside from the q-values) are similar to that in Appendix Table D.1.5. The outcome in this table is the aggregate δ statistic from the non-linear least squares specification featured in Andreoni et al. (2015), carried out at the individual-choice level. δ is the daily discount factor between two future days. For comparability with regression results, individuals who did not respond "Female" or "Male" to the gender survey question were dropped (2% of the sample). Before estimation of aggregate parameters, individuals who never altered their decision from a specific corner solution in all convex time budgets were dropped (as they provided insufficient variation for the calculation of utility parameters) as were individuals who exhibited generalized axiom of revealed preference (GARP) violations. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.7.6: Time inconsistency

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	0.00048029 (0.961) [1.000]	0.00048029 (0.961) [1.000]	-0.00878209 (0.489) [1.000]	0.0179143 (0.487) [1.000]
Heat (Male)		-0.00824147 (0.673) [1.000]		
Heat (Nairobi)		0.02669638 (0.353) [0.496]		
Outcome control mean	1.001042	1.001042	0.9886241	1.05313
Observations	6612	6612	3200	3412

Note: Standard errors clustered at the individual level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. For (1) and (2), "Heat" refers to the difference between the aggregate beta statistics for the treatment and control groups, divided by the square root of the sum of squared standard errors. For (2), "Heat (Male)" refers to the difference between the aggregate beta statistics for male treatment and control groups, then subtracting the difference between the aggregate beta statistics for female treatment and control groups, and then dividing by the square root of the sum of squared standard errors. "Heat (Nairobi)" refers to the difference between the aggregate beta statistics for Nairobi treatment and control groups, subtracting the difference between the aggregate beta statistics for California treatment and control groups, and then dividing by the square root of the sum of squared standard errors. (3) and (4) carry out a similar estimation to "Heat" for (1) and (2) but for the California sample and Nairobi sample, respectively. Note that the effects presented above are multiplied by -1, so that a positive difference reflects more time inconsistency. Within Supplement Section G.7, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes. Additionally, within Supplement Section G.7, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes. Given that the non-linear least squares specification used to measure patience does not incorporate interaction terms, we instead use p-values from analogous differences in this table for multiple testing adjustments. Also note that the non-linear least squares specification does not allow one to control for share of male participants, and thus the results (aside from the q-values) are similar to that in Appendix Table D.1.6. The outcome in this table is the aggregate β statistic from the non-linear least squares specification featured in Andreoni et al. (2015), carried out at the individual-choice level. β measures present bias, and values less than 1 denote time inconsistency. For comparability with regression results, individuals who did not respond "Female" or "Male" to the gender survey question were dropped (2% of the sample). Before estimation of aggregate parameters, individuals who never altered their decision from a specific corner solution in all convex time budgets were dropped (as they provided insufficient variation for the calculation of utility parameters) as were individuals who exhibited generalized axiom of revealed preference (GARP) violations. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.7.7: Trust

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	0.0232	0.000859	0.0541	0.0383
	(0.465)	(0.979)	(0.162)	(0.517)
	[1.000]	[1.000]	[1.000]	[1.000]
Male	0.0705	0.0473	0.0619	0.0758
	(0.000)**	(0.030)*	(0.013)*	(0.000)**
Male × Heat		0.0466		
		(0.117)		
		[0.698]		
Nairobi	-0.136	-0.170		
	(0.000)**	(0.000)**		
Nairobi × Heat		0.0686		
		(0.068)		
		[0.437]		
Share of male participants	0.00999	0.0465	0.0383	0.0379
	(0.796)	(0.260)	(0.571)	(0.454)
Share of male participants × Heat	-0.0287	-0.102	-0.188	-0.00761
	(0.565)	(0.084)	(0.030)*	(0.923)
Outcome control mean	0.417	0.417	0.495	0.348
R-squared	0.0467	0.0508	0.0147	0.0225
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Within Supplement Section G.7, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes. Additionally, within Supplement Section G.7, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes. The outcome in this table is the share of endowed tokens (out of 600) entrusted to the other person in the first round of the trust game. "Share of male participants" captures the share of participants in the room, excluding the participant herself or himself, who identify as male, out of those who identify as either male or female. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.7.8: Public contribution

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-8.765	-19.27	-11.14	43.69
	(0.842)	(0.671)	(0.844)	(0.555)
	[1.000]	[1.000]	[1.000]	[1.000]
Male	37.79	34.85	13.18	58.94
	(0.056)	(0.173)	(0.662)	(0.020)*
Male × Heat		5.992		
		(0.873)		
		[1.000]		
Nairobi	-171.1	-202.5		
	(0.000)**	(0.000)**		
Nairobi × Heat		62.87		
		(0.137)		
		[0.437]		
Share of male participants	2.983	27.61	-21.25	66.51
	(0.962)	(0.659)	(0.795)	(0.471)
Share of male participants × Heat	-0.283	-49.37	-64.48	-43.45
	(0.997)	(0.534)	(0.563)	(0.692)
Outcome control mean	529.7	529.7	629.6	442.5
R-squared	0.0383	0.0395	0.00311	0.00782
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Within Supplement Section G.7, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes. Additionally, within Supplement Section G.7, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes. Public contribution here refers to the public goods game. The outcome in this table is the amount of tokens (out of 1200) put into the shared fund. "Share of male participants" captures the share of participants in the room, excluding the participant herself or himself, who identify as male, out of those who identify as either male or female. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.7.9: Fluid intelligence

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	0.00932	-0.00222	0.00249	0.0384
	(0.564)	(0.893)	(0.885)	(0.277)
	[1.000]	[1.000]	[1.000]	[1.000]
Male	0.0171	0.00225	0.0214	0.0125
	(0.045)*	(0.856)	(0.009)**	(0.396)
Male × Heat		0.0298		
		(0.068)		
		[0.605]		
Nairobi	-0.128	-0.140		
	(0.000)**	(0.000)**		
Nairobi × Heat		0.0245		
		(0.166)		
		[0.437]		
Share of male participants	-0.00121	0.0150	-0.00667	0.0243
	(0.958)	(0.539)	(0.809)	(0.538)
Share of male participants × Heat	0.0106	-0.0221	-0.00414	-0.0170
	(0.707)	(0.485)	(0.907)	(0.730)
Outcome control mean	0.862	0.862	0.935	0.799
R-squared	0.121	0.124	0.00826	0.00618
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Within Supplement Section G.7, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes. Additionally, within Supplement Section G.7, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes. Fluid intelligence is measured through Raven's Progressive Matrices. The outcome in this table is the share of six matrices answered correctly. Each puzzle answered correctly yields an Airtime Voucher worth 50 KSh (or an Amazon Gift Card worth 1 dollar in the California sample), which provides the earnings to be used for the next module. "Share of male participants" captures the share of participants in the room, excluding the participant herself or himself, who identify as male, out of those who identify as either male or female. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.7.10: Joy of Destruction

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-0.0482 (0.032)* [0.632]	-0.0479 (0.026)* [0.462]	-0.0699 (0.008)** [0.105]	0.0550 (0.164) [1.000]
Male	0.0332 (0.004)**	0.0635 (0.000)**	0.0252 (0.118)	0.0398 (0.016)*
Male × Heat		-0.0610 (0.006)** [0.080]		
Nairobi	0.106 (0.000)**	0.0481 (0.008)**		
Nairobi × Heat		0.116 (0.000)** [0.001]++		
Share of male participants	-0.0363 (0.241)	-0.00818 (0.795)	-0.0314 (0.507)	0.0328 (0.405)
Share of male participants × Heat	0.133 (0.001)**	0.0783 (0.053)	0.0717 (0.138)	0.0372 (0.520)
Outcome control mean	0.110	0.110	0.0771	0.139
R-squared	0.0710	0.0831	0.0159	0.0293
Observations	1859	1859	864	995

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Within Supplement Section G.7, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes. Additionally, within Supplement Section G.7, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes. The outcome in this table is the share of the anonymous partner's earned Airtime Vouchers (Amazon Gift Cards in the California sample) destroyed by the participant. The earned Airtime Vouchers and Amazon Gift Cards resulted from the number of Raven's Progressive Matrices answered correctly in the previous module, where one puzzle answered correctly yielded one Airtime Voucher or one Amazon Gift Card. Airtime Vouchers were worth 50 KSh each, while Amazon Gift Cards were worth 1 dollar each. "Share of male participants" captures the share of participants in the room, excluding the participant herself or himself, who identify as male, out of those who identify as either male or female. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.7.11: Cognitive reflection

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-0.0151	-0.0234	0.0140	-0.0346
	(0.533)	(0.343)	(0.706)	(0.300)
	[1.000]	[1.000]	[1.000]	[1.000]
Male	0.0857	0.0807	0.177	-0.00158
	(0.000)**	(0.000)**	(0.000)**	(0.907)
Male × Heat		0.0102		
		(0.689)		
		[1.000]		
Nairobi	-0.242	-0.262		
	(0.000)**	(0.000)**		
Nairobi × Heat		0.0390		
		(0.142)		
		[0.437]		
Share of male participants	-0.0234	-0.00655	0.0369	-0.0384
	(0.462)	(0.854)	(0.492)	(0.364)
Share of male participants × Heat	0.0239	-0.00979	-0.0913	0.0774
	(0.541)	(0.831)	(0.266)	(0.089)
Outcome control mean	0.325	0.325	0.455	0.212
R-squared	0.178	0.180	0.0787	0.00438
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Within Supplement Section G.7, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes. Additionally, within Supplement Section G.7, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes. The outcome in this table is the share of questions (out of 5) from the Cognitive Reflection Test answered correctly. "Share of male participants" captures the share of participants in the room, excluding the participant herself or himself, who identify as male, out of those who identify as either male or female. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.7.12: Charitable donation

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	8.274	-38.82	18.73	36.40
	(0.891)	(0.555)	(0.763)	(0.769)
	[1.000]	[1.000]	[1.000]	[1.000]
Male	2.601	2.250	-46.37	46.44
	(0.931)	(0.957)	(0.185)	(0.330)
Male × Heat		3.134		
		(0.957)		
		[1.000]		
Nairobi	314.6	289.6		
	(0.000)**	(0.000)**		
Nairobi × Heat		53.23		
		(0.397)		
		[0.496]		
Matched with ingroup charity	-8.524	-64.52	29.57	-74.01
	(0.791)	(0.158)	(0.436)	(0.193)
Matched with ingroup charity × Heat		118.6		
		(0.096)		
Earnings in tokens	0.00330	0.00349	-0.0184	0.0272
	(0.752)	(0.738)	(0.114)	(0.129)
Share of male participants	53.90	59.26	-112.6	221.1
	(0.548)	(0.501)	(0.254)	(0.116)
Share of male participants × Heat	-23.65	-41.30	-79.84	-54.65
	(0.852)	(0.745)	(0.555)	(0.794)
Outcome control mean	410.4	410.4	245.0	561.8
R-squared	0.0660	0.0675	0.0130	0.0121
Observations	1806	1806	861	945

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Within Supplement Section G.7, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes. Additionally, within Supplement Section G.7, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes. The outcome in this table is the amount of tokens earned in the experiment that is donated to the randomly selected charity. In Nairobi, "Matched with ingroup charity" is an indicator taking on a value of one if a participant is matched to a charity associated with her ethnicity, and 0 otherwise. In California, "Matched with ingroup charity" is an indicator taking on a value of one if a participant has resided in the San Francisco Bay Area for five years

or more and is matched with a charity in the San Francisco Bay Area. "Earnings in tokens" captures the amount of tokens earned in the experiment. "Share of male participants" captures the share of participants in the room, excluding the participant herself or himself, who identify as male, out of those who identify as either male or female. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.7.13: Happiness

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	0.215	0.215	0.300	0.0304
	(0.092)	(0.095)	(0.067)	(0.903)
Male	0.174	0.139	0.221	0.119
	(0.005)**	(0.142)	(0.017)*	(0.161)
Male × Heat		0.0698		
		(0.590)		
Nairobi	1.123	1.191		
	(0.000)**	(0.000)**		
Nairobi × Heat		-0.136		
		(0.383)		
Share of male participants	0.291	0.257	0.112	0.369
	(0.052)	(0.116)	(0.687)	(0.045)*
Share of male participants × Heat	-0.548	-0.483	-0.625	-0.338
	(0.009)**	(0.056)	(0.109)	(0.309)
Outcome control mean	5.313	5.313	4.625	5.914
R-squared	0.183	0.184	0.0112	0.0126
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. The outcome in this table is the number chosen from a 1-7 scale that asked how one felt towards the end of the experiment, with 1 being sad and 7 being happy. "Share of male participants" captures the share of participants in the room, excluding the participant herself or himself, who identify as male, out of those who identify as either male or female. * $p < .05$, ** $p < .01$

Table G.7.14: Alertness

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-0.396 (0.018)*	-0.471 (0.007)**	-0.380 (0.079)	-0.266 (0.347)
Male	0.390 (0.000)**	0.318 (0.003)**	0.730 (0.000)**	0.0659 (0.497)
Male × Heat		0.145 (0.331)		
Nairobi	1.820 (0.000)**	1.696 (0.000)**		
Nairobi × Heat		0.247 (0.133)		
Share of male participants	0.00842 (0.964)	0.134 (0.465)	0.147 (0.630)	0.109 (0.608)
Share of male participants × Heat	-0.0432 (0.860)	-0.296 (0.283)	-0.373 (0.344)	-0.0859 (0.813)
Outcome control mean	5.247	5.247	4.281	6.090
R-squared	0.304	0.305	0.0748	0.0134
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. The outcome in this table is the number chosen from a 1-7 scale that asked how one felt towards the end of the experiment, with 1 being tired and 7 being alert. "Share of male participants" captures the share of participants in the room, excluding the participant herself or himself, who identify as male, out of those who identify as either male or female. * $p < .05$, ** $p < .01$

G.8 Selected outcomes: including session fixed effects

Table G.8.1: Precision task

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	0.610	0.881	0.684	0.544
	(0.028)*	(0.047)*	(0.057)	(0.191)
	[0.512]	[1.000]	[0.454]	[0.872]
Male	1.979	2.219	1.684	2.267
	(0.000)**	(0.000)**	(0.000)**	(0.000)**
Male × Heat		-0.479		
		(0.335)		
		[1.000]		
Nairobi	-7.833	-7.833		
	(0.000)**	(0.000)**		
Nairobi × Heat		-0.0238		
		(0.965)		
		[1.000]		
Outcome control mean	17.95	17.95	23.63	12.99
R-squared	0.638	0.638	0.348	0.181
F-statistic on F.E.	71.17	61.14	124.64	77.12
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Within Supplement Section G.8, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes. Additionally, within Supplement Section G.8, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes. The precision task is also known as the slider task. The outcome in this table is the number of correct sliders made in three minutes. Final earnings from the precision task are based off either being weakly above (high) or below (low) the median within treatment cohort. The median pair is randomly assigned to high or low. In comparison to the specifications in Section A, these specifications feature session fixed effects, with the F-statistic from the joint F-test on the fixed effects included at the bottom of the table. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.8.2: Fairness

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-0.0167 (0.200) [0.861]	-0.0203 (0.302) [1.000]	-0.0240 (0.151) [0.590]	-0.0107 (0.583) [1.000]
Male	-0.0300 (0.024)*	-0.0261 (0.181)	-0.0600 (0.002)**	-0.00103 (0.954)
Male × Heat		-0.00786 (0.762) [1.000]		
Nairobi	-0.0139 (0.000)**	-0.0139 (0.000)**		
Nairobi × Heat		0.0147 (0.571) [1.000]		
Outcome control mean	0.318	0.318	0.309	0.326
R-squared	0.129	0.129	0.106	0.147
F-statistic on F.E.	1072.42	88.72	370.23	19.13
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Within Supplement Section G.8, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes. Additionally, within Supplement Section G.8, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes. Fairness here refers to the real effort dictator game, where the level of endowment is determined by the number of correct sliders made in the precision task. The outcome in this table is the share of joint earnings (2400 tokens in the high group, 1200 tokens in the low group) that each participant desires to give to the anonymous partner. In comparison to the specifications in Section A, these specifications feature session fixed effects, with the F-statistic from the joint F-test on the fixed effects included at the bottom of the table. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.8.3: Risk-taking

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	0.128	0.153	-5.093	5.918
	(0.995)	(0.996)	(0.879)	(0.840)
	[1.000]	[1.000]	[1.000]	[1.000]
Male	163.4	172.6	290.6	40.01
	(0.000)**	(0.000)**	(0.000)**	(0.226)
Male × Heat		-18.53		
		(0.704)		
		[1.000]		
Nairobi	150.0	150.0		
	(0.000)**	(0.000)**		
Nairobi × Heat		18.47		
		(0.704)		
		[1.000]		
Outcome control mean	366.3	366.3	409.0	329.1
R-squared	0.137	0.137	0.194	0.0931
F-statistic on F.E.	1201.15	105.40	78.88	69.15
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Within Supplement Section G.8, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes. Additionally, within Supplement Section G.8, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes. The outcome in this table is the variance of the coin toss from menu A, in tokens. Note that the expected value is not constant across each coin, so that the outcome does not capture the trade-off between expected value and variance. Note also that under this approach, Coin 7 will be as good as Coin 5, even though Coin 5 strictly dominates Coin 7 with expected utility. In comparison to the specifications in Section A, these specifications feature session fixed effects, with the F-statistic from the joint F-test on the fixed effects included at the bottom of the table. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.8.4: Rational choice violation I

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-0.00408	-0.000709	-0.00433	-0.00386
	(0.602)	(0.955)	(0.718)	(0.711)
	[1.000]	[1.000]	[1.000]	[1.000]
Male	-0.000612	0.00392	0.000622	-0.00181
	(0.946)	(0.770)	(0.955)	(0.899)
Male × Heat		-0.00910		
		(0.623)		
		[1.000]		
Nairobi	-4.10e-14	-3.84e-14		
	(0.189)	(0.203)		
Nairobi × Heat		0.00284		
		(0.869)		
		[1.000]		
Outcome control mean	0.0277	0.0277	0.0275	0.0279
R-squared	0.0684	0.0686	0.0705	0.0666
F-statistic on F.E.	0.14	89.00	0.07	0.07
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Within Supplement Section G.8, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes. Additionally, within Supplement Section G.8, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes. The outcome in this table is an indicator of transitivity violation using both menus A & B. A transitivity violation comes from choosing two coins in the interior region of the intersection of both menus, where it is not the case that it can be said that one preferring coin A to coin B and then preferring coin B to coin C implies that one prefers coin A to coin C. In comparison to the specifications in Section A, these specifications feature session fixed effects, with the F-statistic from the joint F-test on the fixed effects included at the bottom of the table. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.8.5: Patience

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-0.00005823 (0.854) [1.000]	-0.00005823 (0.854) [1.000]	-0.00060946 (0.093) [0.516]	0.00096101 (0.279) [0.872]
Heat (Male)		-0.0002442 (0.702) [1.000]		
Heat (Nairobi)		0.00157046 (0.102) [1.000]		
Outcome control mean	0.9938358	0.9938358	0.9937076	0.9939217
Observations	6612	6612	3200	3412

Note: Standard errors in parentheses, clustered at the individual level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. For (1) and (2), "Heat" refers to the difference between the aggregate delta statistics for the treatment and control groups, divided by the square root of the sum of squared standard errors. For (2), "Heat (Male)" refers to the difference between the aggregate delta statistics for male treatment and control groups, subtracting the difference between the aggregate delta statistics for female treatment and control groups, and then dividing by the square root of the sum of squared standard errors. "Heat (Nairobi)" refers to the difference between the aggregate delta statistics for Nairobi treatment and control groups, subtracting the difference between the aggregate delta statistics for California treatment and control groups, and then dividing by the square root of the sum of squared standard errors. (3) and (4) carry out a similar estimation to "Heat" for (1) and (2) but for the California sample and Nairobi sample, respectively. Within Supplement Section G.8, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes. Additionally, within Supplement Section G.8, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes. Given that the non-linear least squares specification used to measure patience does not incorporate interaction terms, we instead use p-values from analogous differences in this table for multiple testing adjustments. Also note that the non-linear least squares specification does not allow for fixed effects, and thus the results (aside from the q-values) are similar to that in Appendix Table D.1.5. The outcome in this table is the aggregate δ statistic from the non-linear least squares specification featured in Andreoni et al. (2015), carried out at the individual-choice level. δ is the daily discount factor between two future days. For comparability with regression results, individuals who did not respond "Female" or "Male" to the gender survey question were dropped (2% of the sample). Before estimation of aggregate parameters, individuals who never altered their decision from a specific corner solution in all convex time budgets were dropped (as they provided insufficient variation for the calculation of utility parameters) as were individuals who exhibited generalized axiom of revealed preference (GARP) violations. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.8.6: Time inconsistency

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	0.00048029	0.00048029	-0.00878209	0.0179143
	(0.961)	(0.961)	(0.489)	(0.487)
	[1.000]	[1.000]	[1.000]	[1.000]
Heat (Male)		-0.00824147		
		(0.673)		
		[1.000]		
Heat (Nairobi)		0.02669638		
		(0.353)		
		[1.000]		
Outcome control mean	1.001042	1.001042	0.9886241	1.05313
Observations	6612	6612	3200	3412

Note: Standard errors in parentheses, clustered at the individual level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. For (1) and (2), "Heat" refers to the difference between the aggregate beta statistics for the treatment and control groups, divided by the square root of the sum of squared standard errors. For (2), "Heat (Male)" refers to the difference between the aggregate beta statistics for male treatment and control groups, then subtracting the difference between the aggregate beta statistics for female treatment and control groups, and then dividing by the square root of the sum of squared standard errors. "Heat (Nairobi)" refers to the difference between the aggregate beta statistics for Nairobi treatment and control groups, subtracting the difference between the aggregate beta statistics for California treatment and control groups, and then dividing by the square root of the sum of squared standard errors. (3) and (4) carry out a similar estimation to "Heat" for (1) and (2) but for the California sample and Nairobi sample, respectively. Note that the effects presented above are multiplied by -1, so that a positive difference reflects more time inconsistency. Within Supplement Section G.8, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes. Additionally, within Supplement Section G.8, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes. Given that the non-linear least squares specification used to measure patience does not incorporate interaction terms, we instead use p-values from analogous differences in this table for multiple testing adjustments. Also note that the non-linear least squares specification does not allow for fixed effects, and thus the results (aside from the q-values) are similar to that in Appendix Table D.1.6. The outcome in this table is the aggregate β statistic from the non-linear least squares specification featured in Andreoni et al. (2015), carried out at the individual-choice level. β measures present bias, and values less than 1 denote time inconsistency. For comparability with regression results, individuals who did not respond "Female" or "Male" to the gender survey question were dropped (2% of the sample). Before estimation of aggregate parameters, individuals who never altered their decision from a specific corner solution in all convex time budgets were dropped (as they provided insufficient variation for the calculation of utility parameters) as were individuals who exhibited generalized axiom of revealed preference (GARP) violations. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.8.7: Trust

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	0.00755	-0.0335	-0.0177	0.0292
	(0.648)	(0.222)	(0.501)	(0.164)
	[1.000]	[1.000]	[1.000]	[0.872]
Male	0.0764	0.0566	0.0806	0.0719
	(0.000)**	(0.019)*	(0.006)**	(0.001)**
Male × Heat		0.0392		
		(0.217)		
		[1.000]		
Nairobi	-0.341	-0.341		
	(0.000)**	(0.000)**		
Nairobi × Heat		0.0368		
		(0.294)		
		[1.000]		
Outcome control mean	0.417	0.417	0.495	0.348
R-squared	0.160	0.163	0.106	0.163
F-statistic on F.E.	386.53	82.91	12.40	66.77
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Within Supplement Section G.8, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes. Additionally, within Supplement Section G.8, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes. The outcome in this table is the share of endowed tokens (out of 600) entrusted to the other person in the first round of the trust game. In comparison to the specifications in Section A, these specifications feature session fixed effects, with the F-statistic from the joint F-test on the fixed effects included at the bottom of the table. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.8.8: Public contribution

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-11.07	-41.33	-35.60	9.816
	(0.591)	(0.239)	(0.263)	(0.716)
	[1.000]	[1.000]	[0.669]	[1.000]
Male	45.37	37.88	37.98	52.15
	(0.037)*	(0.168)	(0.290)	(0.041)*
Male × Heat		14.64		
		(0.709)		
		[1.000]		
Nairobi	-308.3	-308.3		
	(0.000)**	(0.000)**		
Nairobi × Heat		41.49		
		(0.328)		
		[1.000]		
Outcome control mean	529.7	529.7	629.6	442.5
R-squared	0.188	0.189	0.126	0.195
F-statistic on F.E.	760.98	104.20	213.72	112.52
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Within Supplement Section G.8, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes. Additionally, within Supplement Section G.8, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes. Public contribution here refers to the public goods game. The outcome in this table is the amount of tokens (out of 1200) put into the shared fund. In comparison to the specifications in Section A, these specifications feature session fixed effects, with the F-statistic from the joint F-test on the fixed effects included at the bottom of the table. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.8.9: Fluid intelligence

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	0.0159	-0.00707	0.00252	0.0274
	(0.061)	(0.567)	(0.782)	(0.045)*
	[0.512]	[1.000]	[1.000]	[0.334]
Male	0.0143	0.00237	0.0190	0.00959
	(0.120)	(0.856)	(0.019)*	(0.559)
Male × Heat		0.0238		
		(0.177)		
		[1.000]		
Nairobi	-0.111	-0.111		
	(0.000)**	(0.000)**		
Nairobi × Heat		0.0189		
		(0.265)		
		[1.000]		
Outcome control mean	0.862	0.862	0.935	0.799
R-squared	0.203	0.205	0.111	0.0957
F-statistic on F.E.	100.27	24.18	1616.39	96.67
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Within Supplement Section G.8, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes. Additionally, within Supplement Section G.8, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes. Fluid intelligence is measured through Raven's Progressive Matrices. The outcome in this table is the share of six matrices answered correctly. Each puzzle answered correctly yields an Airtime Voucher worth 50 KSh (or an Amazon Gift Card worth 1 dollar in the California sample), which provides the earnings to be used for the next module. In comparison to the specifications in Section A, these specifications feature session fixed effects, with the F-statistic from the joint F-test on the fixed effects included at the bottom of the table. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.8.10: Joy of Destruction

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	0.0218	-0.0241	-0.0428	0.0767
	(0.116)	(0.128)	(0.006)**	(0.000)**
	[0.628]	[1.000]	[0.075]	[0.004]++
Male	0.0295	0.0520	0.0264	0.0312
	(0.023)*	(0.005)**	(0.159)	(0.078)
Male × Heat		-0.0466		
		(0.048)*		
		[1.000]		
Nairobi	0.0667	0.0667		
	(0.000)**	(0.000)**		
Nairobi × Heat		0.131		
		(0.000)**		
		[0.001]++		
Outcome control mean	0.110	0.110	0.0771	0.139
R-squared	0.153	0.168	0.0885	0.128
F-statistic on F.E.	167.45	26.82	25.55	8.45
Observations	1859	1859	864	995

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Within Supplement Section G.8, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes. Additionally, within Supplement Section G.8, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes. The outcome in this table is the share of the anonymous partner's earned Airtime Vouchers (Amazon Gift Cards in the California sample) destroyed by the participant. The earned Airtime Vouchers and Amazon Gift Cards resulted from the number of Raven's Progressive Matrices answered correctly in the previous module, where one puzzle answered correctly yielded one Airtime Voucher or one Amazon Gift Card. Airtime Vouchers were worth 50 KSh each, while Amazon Gift Cards were worth 1 dollar each. In comparison to the specifications in Section A, these specifications feature session fixed effects, with the F-statistic from the joint F-test on the fixed effects included at the bottom of the table. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.8.11: Cognitive reflection

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-0.00253	-0.0306	-0.0222	0.0153
	(0.832)	(0.149)	(0.267)	(0.265)
	[1.000]	[1.000]	[0.669]	[0.872]
Male	0.0849	0.0762	0.175	-0.00286
	(0.000)**	(0.000)**	(0.000)**	(0.857)
Male × Heat		0.0171		
		(0.518)		
		[1.000]		
Nairobi	-0.167	-0.167		
	(0.000)**	(0.000)**		
Nairobi × Heat		0.0350		
		(0.179)		
		[1.000]		
Outcome control mean	0.325	0.325	0.455	0.212
R-squared	0.274	0.276	0.164	0.160
F-statistic on F.E.	13.45	343.11	94.96	15.96
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Within Supplement Section G.8, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes. Additionally, within Supplement Section G.8, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes. The outcome in this table is the share of questions (out of 5) from the Cognitive Reflection Test answered correctly. In comparison to the specifications in Section A, these specifications feature session fixed effects, with the F-statistic from the joint F-test on the fixed effects included at the bottom of the table. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.8.12: Charitable donation

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-6.906	-48.66	-14.82	-0.802
	(0.838)	(0.305)	(0.653)	(0.989)
	[1.000]	[1.000]	[1.000]	[1.000]
Male	3.056	8.704	-18.94	27.39
	(0.922)	(0.842)	(0.605)	(0.590)
Male × Heat		-8.804		
		(0.891)		
		[1.000]		
Nairobi	248.4	126.8		
	(0.000)**	(0.067)		
Nairobi × Heat		41.75		
		(0.527)		
		[1.000]		
Matched with ingroup charity	-13.64	-64.06	10.67	-53.30
	(0.705)	(0.204)	(0.806)	(0.408)
Matched with ingroup charity × Heat		106.5		
		(0.175)		
Earnings in tokens	-0.0127	-0.0122	-0.0229	-0.000982
	(0.234)	(0.251)	(0.057)	(0.958)
Outcome control mean	410.4	410.4	245.0	561.8
R-squared	0.195	0.196	0.0964	0.161
F-statistic on F.E.	97.44	118.89	54.90	188.86
Observations	1806	1806	861	945

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Within Supplement Section G.8, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes. Additionally, within Supplement Section G.8, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes. The outcome in this table is the amount of tokens earned in the experiment that is donated to the randomly selected charity. In Nairobi, matched with ingroup charity is an indicator taking on a value of one if a participant is matched to a charity associated with her ethnicity, and 0 otherwise. In California, "Matched with ingroup charity" is an indicator taking on a value of one if a participant has resided in the San Francisco Bay Area for five years or more and is matched with a charity in the San Francisco Bay Area. "Earnings in tokens" captures the amount of tokens earned in the experiment. In comparison to the specifications in Section A, these specifications feature session fixed effects, with the F-statistic from the joint F-test on the fixed effects included at the bottom of the table. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.8.13: Happiness

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-0.0707	0.0107	0.0557	-0.178
	(0.285)	(0.918)	(0.551)	(0.055)
Male	0.172	0.119	0.277	0.0722
	(0.015)*	(0.268)	(0.008)**	(0.448)
Male × Heat		0.108		
		(0.438)		
Nairobi	1.333	1.333		
	(0.000)**	(0.000)**		
Nairobi × Heat		-0.259		
		(0.070)		
Outcome control mean	5.313	5.313	4.625	5.914
R-squared	0.263	0.265	0.112	0.106
F-statistic on F.E.	2128.30	132.59	13.68	301.03
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. The outcome in this table is the number chosen from a 1-7 scale that asked how one felt towards the end of the experiment, with 1 being sad and 7 being happy. In comparison to the specifications in Section A, these specifications feature session fixed effects, with the F-statistic from the joint F-test on the fixed effects included at the bottom of the table. * $p < .05$, ** $p < .01$

Table G.8.14: Alertness

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-0.412 (0.000)**	-0.597 (0.000)**	-0.524 (0.000)**	-0.313 (0.003)**
Male	0.385 (0.000)**	0.302 (0.013)*	0.768 (0.000)**	0.0127 (0.913)
Male × Heat		0.166 (0.310)		
Nairobi	2.667 (0.000)**	2.667 (0.000)**		
Nairobi × Heat		0.177 (0.278)		
Outcome control mean	5.247	5.247	4.281	6.090
R-squared	0.371	0.372	0.156	0.121
F-statistic on F.E.	1035.96	25.71	31.37	6.36
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. The outcome in this table is the number chosen from a 1-7 scale that asked how one felt towards the end of the experiment, with 1 being tired and 7 being alert. In comparison to the specifications in Section A, these specifications feature session fixed effects, with the F-statistic from the joint F-test on the fixed effects included at the bottom of the table. * $p < .05$, ** $p < .01$

G.9 Selected outcomes: accounting for outside weather

Table G.9.1: Precision task

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	0.518 (0.048)* [0.339]	0.931 (0.064) [0.538]	0.598 (0.075) [0.495]	0.450 (0.250) [0.722]
Male	1.921 (0.000)**	2.016 (0.000)**	1.845 (0.000)**	2.360 (0.000)**
Male × Heat		-0.194 (0.687) [1.000]		
Nairobi	-10.53 (0.000)**	-10.24 (0.000)**		
Nairobi × Heat		-0.581 (0.414) [0.950]		
Outside temperature (pooled)	-0.116 (0.237)	-0.150 (0.146)		
Outside temperature (pooled) × Heat	0.00612 (0.934)	0.0756 (0.450)		
Outside relative humidity (pooled)	-0.0345 (0.048)*	-0.0392 (0.032)*		
Outside relative humidity (pooled) × Heat	0.0153 (0.445)	0.0247 (0.262)		
Outside temperature (site)			-0.170 (0.129)	0.470 (0.138)
Outside temperature (site) × Heat			0.0735 (0.473)	0.507 (0.271)
Outside relative humidity (site)			-0.0331 (0.121)	-0.0356 (0.275)
Outside relative humidity (site) × Heat			0.0355 (0.133)	0.0168 (0.710)
Outcome control mean	17.95	17.95	23.63	12.99
R-squared	0.551	0.552	0.0779	0.0492
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Within Supplement Section G.9, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes, aside from Supplement Table G.9.11. Additionally, within Supplement Section G.9, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes, aside from Supplement Table G.9.11. The precision task is the slider task. The outcome is the number of correct sliders in 3 minutes. Earnings are based off either being weakly above (high) or below (low) the median within treatment cohort (median pair is randomly assigned to high or low). "Outside temperature (pooled)" captures the demeaned (relative to the pooled sample) average outside temperature in Celsius for the day of the session. "Outside relative humidity (pooled)" captures the demeaned (relative to the pooled sample) average outside relative humidity in percentage for the day of the session. "Outside temperature (site)" captures the demeaned (relative to the site sample) average site-specific outside temperature in Celsius for the day of the session. "Outside relative humidity (site)" captures the demeaned (relative to the site sample) average site-specific outside relative humidity in percentage for the day of the session. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.9.2: Fairness

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-0.0167 (0.171) [0.627]	-0.0276 (0.231) [1.000]	-0.0241 (0.120) [0.495]	-0.0105 (0.559) [1.000]
Male	-0.0427 (0.001)**	-0.0366 (0.038)*	-0.0811 (0.000)**	-0.00895 (0.597)
Male × Heat		-0.0119 (0.607) [1.000]		
Nairobi	0.0670 (0.000)**	0.0506 (0.042)*		
Nairobi × Heat		0.0324 (0.373) [0.950]		
Outside temperature (pooled)	-0.00651 (0.031)*	-0.00484 (0.153)		
Outside temperature (pooled) × Heat	0.000474 (0.874)	-0.00288 (0.501)		
Outside relative humidity (pooled)	-0.000907 (0.196)	-0.000658 (0.386)		
Outside relative humidity (pooled) × Heat	-0.0000836 (0.922)	-0.000566 (0.606)		
Outside temperature (site)			-0.00584 (0.079)	-0.00645 (0.723)
Outside temperature (site) × Heat			0.00130 (0.759)	-0.0227 (0.358)
Outside relative humidity (site)			-0.00134 (0.085)	0.000337 (0.843)
Outside relative humidity (site) × Heat			0.000998 (0.384)	-0.00441 (0.088)
Outcome control mean	0.318	0.318	0.309	0.326
R-squared	0.0136	0.0141	0.0383	0.00963
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Within Supple-

ment Section G.9, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes, aside from Supplement Table G.9.11. Additionally, within Supplement Section G.9, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes, aside from Supplement Table G.9.11. Fairness here refers to the real effort dictator game, where the endowment is determined by the number of correct sliders made in the precision task. The outcome is the share of joint earnings (2400 tokens for high, 1200 tokens for low) that each participant desires to give to the anonymous partner. "Outside temperature (pooled)" captures the demeaned (relative to pooled) average outside temperature (°C) for the session day. "Outside relative humidity (pooled)" captures the demeaned (relative to pooled) average outside relative humidity (%) for the session day. "Outside temperature (site)" and "outside relative humidity (site)" are site-specific variants. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.9.3: Risk-taking

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-1.551 (0.940) [1.000]	20.84 (0.601) [1.000]	-5.037 (0.870) [0.891]	1.411 (0.958) [1.000]
Male	159.8 (0.000)**	166.8 (0.000)**	312.2 (0.000)**	26.93 (0.344)
Male × Heat		-14.13 (0.750) [1.000]		
Nairobi	-95.17 (0.005)**	-81.27 (0.064)		
Nairobi × Heat		-27.91 (0.680) [1.000]		
Outside temperature (pooled)	-2.759 (0.673)	-4.451 (0.526)		
Outside temperature (pooled) × Heat	2.410 (0.693)	5.864 (0.525)		
Outside relative humidity (pooled)	-3.141 (0.031)*	-3.368 (0.030)*		
Outside relative humidity (pooled) × Heat	1.720 (0.314)	2.179 (0.273)		
Outside temperature (site)			-4.236 (0.549)	26.57 (0.282)
Outside temperature (site) × Heat			8.381 (0.404)	-49.64 (0.087)
Outside relative humidity (site)			-1.852 (0.291)	-4.214 (0.086)
Outside relative humidity (site) × Heat			1.753 (0.484)	1.957 (0.571)
Outcome control mean	366.3	366.3	409.0	329.1
R-squared	0.0445	0.0447	0.114	0.0107
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Within Supplement

Section G.9, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes, aside from Supplement Table G.9.11. Additionally, within Supplement Section G.9, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes, aside from Supplement Table G.9.11. The outcome is the variance of the coin toss from menu A, in tokens. Note that the expected value is not constant across each coin, so that the outcome does not capture the trade-off between expected value and variance, and that Coin 7 will be as good as Coin 5, even though Coin 5 strictly dominates Coin 7 with expected utility. "Outside temperature (pooled)" captures the demeaned (relative to the pooled sample) average outside temperature in Celsius for the day of the session. "Outside relative humidity (pooled)" captures the demeaned (relative to the pooled sample) average outside relative humidity in percentage for the day of the session. "Outside temperature (site)" captures the demeaned (relative to the site sample) average site-specific outside temperature in Celsius for the day of the session. "Outside relative humidity (site)" captures the demeaned (relative to the site sample) average site-specific outside relative humidity in percentage for the day of the session. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.9.4: Rational choice violation I

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-0.00433 (0.559) [1.000]	0.00505 (0.745) [1.000]	-0.00440 (0.695) [0.787]	-0.00422 (0.668) [1.000]
Male	0.00192 (0.795)	0.00323 (0.773)	0.00341 (0.749)	-0.00347 (0.747)
Male × Heat		-0.00272 (0.861) [1.000]		
Nairobi	0.0107 (0.279)	0.0182 (0.241)		
Nairobi × Heat		-0.0149 (0.499) [0.950]		
Outside temperature (pooled)	-0.00264 (0.258)	-0.00350 (0.221)		
Outside temperature (pooled) × Heat	0.00153 (0.536)	0.00326 (0.385)		
Outside relative humidity (pooled)	-0.000410 (0.398)	-0.000529 (0.338)		
Outside relative humidity (pooled) × Heat	0.0000446 (0.948)	0.000281 (0.729)		
Outside temperature (site)			-0.00318 (0.333)	-0.00809 (0.266)
Outside temperature (site) × Heat			0.00287 (0.504)	-0.00377 (0.738)
Outside relative humidity (site)			-0.000462 (0.516)	-0.000828 (0.314)
Outside relative humidity (site) × Heat			-0.000137 (0.892)	0.000990 (0.481)
Outcome control mean	0.0277	0.0277	0.0275	0.0279
R-squared	0.00144	0.00172	0.00421	0.00365
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Within Supplement

Section G.9, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes, aside from Supplement Table G.9.11. Within Supplement Section G.9, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes, aside from Supplement Table G.9.11. The outcome in this table is an indicator of transitivity violation using both menus A & B. A transitivity violation comes from choosing two coins in the interior region of the intersection of both menus, where it cannot be said that one preferring coin A to coin B and also preferring coin B to coin C implies that one prefers coin A to coin C. "Outside temperature (pooled)" captures the demeaned (relative to the pooled sample) average outside temperature in Celsius for the day of the session. "Outside relative humidity (pooled)" captures the demeaned (relative to the pooled sample) average outside relative humidity in percentage for the day of the session. "Outside temperature (site)" captures the demeaned (relative to the site sample) average site-specific outside temperature in Celsius for the day of the session. "Outside relative humidity (site)" captures the demeaned (relative to the site sample) average site-specific outside relative humidity in percentage for the day of the session. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.9.5: Patience

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-0.00005823 (0.854) [1.000]	-0.00005823 (0.854) [1.000]	-0.00060946 (0.093) [0.495]	0.00096101 (0.279) [0.722]
Heat (Male)		-0.0002442 (0.702) [1.000]		
Heat (Nairobi)		0.00157046 (0.102) [0.513]		
Outcome control mean	0.9938358	0.9938358	0.9937076	0.9939217
Observations	6612	6612	3200	3412

Note: Standard errors clustered at the individual level. Per-comparison p-values in parentheses. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. For (1) and (2), "Heat" refers to the difference between the aggregate delta statistics for the treatment and control groups, divided by the square root of the sum of squared standard errors. For (2), "Heat (Male)" refers to the difference between the aggregate delta statistics for male treatment and control groups, subtracting the difference between the aggregate delta statistics for female treatment and control groups, and then dividing by the square root of the sum of squared standard errors. "Heat (Nairobi)" refers to the difference between the aggregate delta statistics for Nairobi treatment and control groups, subtracting the difference between the aggregate delta statistics for California treatment and control groups, and then dividing by the square root of the sum of squared standard errors. (3) and (4) carry out a similar estimation to "Heat" for (1) and (2) but for the California sample and Nairobi sample, respectively. Within Supplement Section G.9, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes, aside from Supplement Table G.9.11. Additionally, within Supplement Section G.9, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes, aside from Supplement Table G.9.11. Given that the non-linear least squares specification used to measure patience does not incorporate interaction terms, we instead use p-values from analogous differences in this table for multiple testing adjustments. Also note that the non-linear least squares specification does not allow one to control for parental education status, and thus the results (aside from the q-values) are similar to that in Appendix Table D.1.5. The outcome in this table is the aggregate δ statistic from the non-linear least squares specification featured in Andreoni et al. (2015), carried out at the individual-choice level. δ is the daily discount factor between two future days. For comparability with regression results, individuals who did not respond "Female" or "Male" to the gender survey question were dropped (2% of the sample). Before estimation of aggregate parameters, individuals who never altered their decision from a specific corner solution in all convex time budgets were dropped (as they provided insufficient variation for the calculation of utility parameters) as were individuals who exhibited generalized axiom of revealed preference (GARP) violations. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.9.6: Time inconsistency

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	0.00048029	0.00048029	-0.00878209	0.0179143
	(0.961)	(0.961)	(0.489)	(0.487)
	[1.000]	[1.000]	[0.787]	[1.000]
Heat (Male)		-0.00824147		
		(0.673)		
		[1.000]		
Heat (Nairobi)		0.02669638		
		(0.353)		
		[0.950]		
Outcome control mean	1.001042	1.001042	0.9886241	1.05313
Observations	6612	6612	3200	3412

Note: Standard errors clustered at the individual level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. For (1) and (2), "Heat" refers to the difference between the aggregate beta statistics for the treatment and control groups, divided by the square root of the sum of squared standard errors. For (2), "Heat (Male)" refers to the difference between the aggregate beta statistics for male treatment and control groups, then subtracting the difference between the aggregate beta statistics for female treatment and control groups, and then dividing by the square root of the sum of squared standard errors. "Heat (Nairobi)" refers to the difference between the aggregate beta statistics for Nairobi treatment and control groups, subtracting the difference between the aggregate beta statistics for California treatment and control groups, and then dividing by the square root of the sum of squared standard errors. (3) and (4) carry out a similar estimation to "Heat" for (1) and (2) but for the California sample and Nairobi sample, respectively. Note that the effects presented above are multiplied by -1, so that a positive difference reflects more time inconsistency. Within Supplement Section G.9, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes, aside from Supplement Table G.9.11. Additionally, within Supplement Section G.9, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes, aside from Supplement Table G.9.11. Given that the non-linear least squares specification used to measure patience does not incorporate interaction terms, we instead use p-values from analogous differences in this table for multiple testing adjustments. Also note that the non-linear least squares specification does not allow one to control for parental education status, and thus the results (aside from the q-values) are similar to that in Appendix Table D.1.6. The outcome in this table is the aggregate β statistic from the non-linear least squares specification featured in Andreoni et al. (2015), carried out at the individual-choice level. β measures present bias, and values less than 1 denote time inconsistency. For comparability with regression results, individuals who did not respond "Female" or "Male" to the gender survey question were dropped (2% of the sample). Before estimation of aggregate parameters, individuals who never altered their decision from a specific corner solution in all convex time budgets were dropped (as they provided insufficient variation for the calculation of utility parameters) as were individuals who exhibited generalized axiom of revealed preference (GARP) violations. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.9.7: Trust

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	0.00762 (0.622) [1.000]	-0.0324 (0.377) [1.000]	-0.0210 (0.385) [0.787]	0.0333 (0.090) [0.428]
Male	0.0679 (0.000)**	0.0496 (0.021)*	0.0546 (0.029)*	0.0773 (0.000)**
Male × Heat		0.0367 (0.206) [1.000]		
Nairobi	-0.0994 (0.000)**	-0.118 (0.001)**		
Nairobi × Heat		0.0383 (0.505) [0.950]		
Outside temperature (pooled)	-0.0105 (0.018)*	-0.00799 (0.147)		
Outside temperature (pooled) × Heat	0.00576 (0.246)	0.000613 (0.945)		
Outside relative humidity (pooled)	-0.00132 (0.206)	-0.00100 (0.351)		
Outside relative humidity (pooled) × Heat	0.00141 (0.226)	0.000745 (0.597)		
Outside temperature (site)			-0.0108 (0.075)	0.00977 (0.620)
Outside temperature (site) × Heat			0.00494 (0.608)	-0.0400 (0.125)
Outside relative humidity (site)			-0.00213 (0.113)	0.00224 (0.183)
Outside relative humidity (site) × Heat			0.00201 (0.254)	-0.00373 (0.098)
Outcome control mean	0.417	0.417	0.495	0.348
R-squared	0.0506	0.0521	0.0151	0.0266
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Within Supplement

Section G.9, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes, aside from Supplement Table G.9.11. Additionally, within Supplement Section G.9, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes, aside from Supplement Table G.9.11. The outcome in this table is the share of endowed tokens (out of 600) entrusted to the other person in the first round of the trust game. "Outside temperature (pooled)" captures the demeaned (relative to the pooled sample) average outside temperature in Celsius for the day of the session. "Outside relative humidity (pooled)" captures the demeaned (relative to the pooled sample) average outside relative humidity in percentage for the day of the session. "Outside temperature (site)" captures the demeaned (relative to the site sample) average site-specific outside temperature in Celsius for the day of the session. "Outside relative humidity (site)" captures the demeaned (relative to the site sample) average site-specific outside relative humidity in percentage for the day of the session. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.9.8: Public contribution

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-9.000 (0.643) [1.000]	-30.70 (0.480) [1.000]	-36.56 (0.220) [0.566]	14.30 (0.580) [1.000]
Male	36.18 (0.087)	35.51 (0.189)	5.571 (0.855)	55.07 (0.042)*
Male × Heat		1.587 (0.966) [1.000]		
Nairobi	-159.7 (0.000)**	-179.4 (0.001)**		
Nairobi × Heat		39.27 (0.541) [0.950]		
Outside temperature (pooled)	-4.394 (0.543)	-2.204 (0.780)		
Outside temperature (pooled) × Heat	7.287 (0.176)	2.858 (0.734)		
Outside relative humidity (pooled)	-1.440 (0.295)	-1.130 (0.430)		
Outside relative humidity (pooled) × Heat	0.450 (0.714)	-0.163 (0.914)		
Outside temperature (site)			0.0000295 (1.000)	-34.24 (0.333)
Outside temperature (site) × Heat			5.395 (0.538)	-14.90 (0.662)
Outside relative humidity (site)			-0.941 (0.522)	-2.919 (0.426)
Outside relative humidity (site) × Heat			0.599 (0.722)	-2.128 (0.539)
Outcome control mean	529.7	529.7	629.6	442.5
R-squared	0.0404	0.0406	0.00393	0.0158
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Within Supplement

Section G.9, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes, aside from Supplement Table G.9.11. Within Supplement Section G.9, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes, aside from Supplement Table G.9.11. Public contribution here refers to the public goods game. The outcome is the amount of tokens (out of 1200) put into the shared fund. "Outside temperature (pooled)" captures the demeaned (relative to the pooled sample) average outside temperature in Celsius for the day of the session. "Outside relative humidity (pooled)" captures the demeaned (relative to the pooled sample) average outside relative humidity in percentage for the day of the session. "Outside temperature (site)" captures the demeaned (relative to the site sample) average site-specific outside temperature in Celsius for the day of the session. "Outside relative humidity (site)" captures the demeaned (relative to the site sample) average site-specific outside relative humidity in percentage for the day of the session. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.9.9: Fluid intelligence

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	0.0149 (0.062) [0.339]	-0.0123 (0.332) [1.000]	0.000612 (0.942) [0.891]	0.0273 (0.037)* [0.255]
Male	0.0181 (0.041)*	0.00482 (0.704)	0.0229 (0.006)**	0.0159 (0.300)
Male × Heat		0.0266 (0.104) [1.000]		
Nairobi	-0.136 (0.000)**	-0.148 (0.000)**		
Nairobi × Heat		0.0243 (0.229) [0.950]		
Outside temperature (pooled)	0.000256 (0.883)	0.00188 (0.314)		
Outside temperature (pooled) × Heat	0.00266 (0.210)	-0.000692 (0.786)		
Outside relative humidity (pooled)	0.000204 (0.637)	0.000410 (0.368)		
Outside relative humidity (pooled) × Heat	0.000200 (0.725)	-0.000227 (0.719)		
Outside temperature (site)			0.00211 (0.228)	0.00202 (0.852)
Outside temperature (site) × Heat			0.0000503 (0.983)	-0.00107 (0.943)
Outside relative humidity (site)			0.000486 (0.225)	0.000468 (0.704)
Outside relative humidity (site) × Heat			0.000238 (0.686)	-0.00143 (0.348)
Outcome control mean	0.862	0.862	0.935	0.799
R-squared	0.122	0.124	0.0135	0.00695
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Within Supplement

Section G.9, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes, aside from Supplement Table G.9.11. Additionally, within Supplement Section G.9, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes, aside from Supplement Table G.9.11. Fluid intelligence is measured through Raven's Progressive Matrices. The outcome is the correct share of 6 matrices. Each correct puzzle yields a 50 KSh Airtime Voucher ($1 Amazon Gift Card in California), providing the earnings for the next module. "Outside temperature (pooled)" captures the demeaned (relative to the pooled sample) average outside temperature in Celsius for the day of the session. "Outside relative humidity (pooled)" captures the demeaned (relative to the pooled sample) average outside relative humidity in percentage for the day of the session. "Outside temperature (site)" captures the demeaned (relative to the site sample) average site-specific outside temperature in Celsius for the day of the session. "Outside relative humidity (site)" captures the demeaned (relative to the site sample) average site-specific outside relative humidity in percentage for the day of the session. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.9.10: Joy of Destruction

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	0.0229	-0.0236	-0.0411	0.0792
	(0.063)	(0.231)	(0.004)**	(0.000)**
	[0.339]	[1.000]	[0.051]	[0.001]++
Male	0.0366	0.0637	0.0259	0.0464
	(0.001)**	(0.000)**	(0.107)	(0.007)**
Male × Heat		-0.0531		
		(0.015)*		
		[0.222]		
Nairobi	0.110	0.0383		
	(0.000)**	(0.130)		
Nairobi × Heat		0.141		
		(0.000)**		
		[0.002]++		
Outside temperature (pooled)	-0.00633	0.000978		
	(0.016)*	(0.742)		
Outside temperature (pooled) × Heat	0.0137	-0.000839		
	(0.000)**	(0.840)		
Outside relative humidity (pooled)	-0.000512	0.000566		
	(0.468)	(0.464)		
Outside relative humidity (pooled) × Heat	0.00142	-0.000660		
	(0.113)	(0.514)		
Outside temperature (site)			0.00114	-0.000951
			(0.709)	(0.943)
Outside temperature (site) × Heat			-0.00119	-0.00136
			(0.771)	(0.945)
Outside relative humidity (site)			0.000717	-0.000320
			(0.432)	(0.833)
Outside relative humidity (site) × Heat			-0.00109	0.000761
			(0.295)	(0.731)
Outcome control mean	0.110	0.110	0.0771	0.139
R-squared	0.0722	0.0809	0.0156	0.0269
Observations	1859	1859	864	995

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Within Supplement

Section G.9, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes, aside from Supplement Table G.9.11. Additionally, within Supplement Section G.9, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes, aside from Supplement Table G.9.11. The outcome in this table is the share of the anonymous partner's earned Airtime Vouchers (Amazon Gift Cards in the California sample) destroyed by the participant. The earned Airtime Vouchers and Amazon Gift Cards resulted from the number of Raven's Progressive Matrices answered correctly in the previous module, where one puzzle answered correctly yielded one Airtime Voucher or one Amazon Gift Card. Airtime Vouchers were worth 50 KSh each, while Amazon Gift Cards were worth 1 dollar each. "Outside temperature (pooled)" captures the demeaned (relative to the pooled sample) average outside temperature in Celsius for the day of the session. "Outside relative humidity (pooled)" captures the demeaned (relative to the pooled sample) average outside relative humidity in percentage for the day of the session. "Outside temperature (site)" captures the demeaned (relative to the site sample) average site-specific outside temperature in Celsius for the day of the session. "Outside relative humidity (site)" captures the demeaned (relative to the site sample) average site-specific outside relative humidity in percentage for the day of the session. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.9.11: Joy of Destruction (only outside temperature)

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	0.0230 (0.064)	-0.0176 (0.290)	-0.0410 (0.004)**	0.0792 (0.000)**
Male	0.0366 (0.001)**	0.0625 (0.000)**	0.0251 (0.109)	0.0466 (0.006)**
Male × Heat		-0.0517 (0.016)*		
Nairobi	0.113 (0.000)**	0.0491 (0.009)**		
Nairobi × Heat		0.128 (0.000)**		
Outside temperature (pooled)	-0.00607 (0.021)*	-0.000645 (0.815)		
Outside temperature (pooled) × Heat	0.0120 (0.000)**	0.00105 (0.775)		
Outside temperature (site)			-0.000780 (0.783)	0.000823 (0.946)
Outside temperature (site) × Heat			0.00174 (0.631)	-0.00563 (0.777)
Outcome control mean	0.110	0.110	0.0771	0.139
R-squared	0.0707	0.0806	0.0141	0.0268
Observations	1859	1859	864	995

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Results from this table are excluded from the multiple hypotheses testing adjustments. The outcome in this table is the share of the anonymous partner's earned Airtime Vouchers (Amazon Gift Cards in the California sample) destroyed by the participant. The earned Airtime Vouchers and Amazon Gift Cards resulted from the number of Raven's Progressive Matrices answered correctly in the previous module, where one puzzle answered correctly yielded one Airtime Voucher or one Amazon Gift Card. Airtime Vouchers were worth 50 KSh each, while Amazon Gift Cards were worth 1 dollar each. "Outside temperature (pooled)" captures the demeaned (relative to the pooled sample) average outside temperature in Celsius for the day of the session. "Outside temperature (site)" captures the demeaned (relative to the site sample) average site-specific outside temperature in Celsius for the day of the session. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.9.12: Cognitive reflection

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-0.00243 (0.822) [1.000]	-0.0119 (0.599) [1.000]	-0.0232 (0.203) [0.566]	0.0158 (0.208) [0.722]
Male	0.0839 (0.000)**	0.0776 (0.000)**	0.181 (0.000)**	0.00139 (0.919)
Male × Heat		0.0126 (0.616) [1.000]		
Nairobi	-0.237 (0.000)**	-0.240 (0.000)**		
Nairobi × Heat		0.00526 (0.894) [1.000]		
Outside temperature (pooled)	-0.00426 (0.277)	-0.00384 (0.387)		
Outside temperature (pooled) × Heat	0.00487 (0.197)	0.00399 (0.495)		
Outside relative humidity (pooled)	-0.00111 (0.214)	-0.00106 (0.284)		
Outside relative humidity (pooled) × Heat	0.00244 (0.007)**	0.00233 (0.058)		
Outside temperature (site)			-0.00357 (0.405)	0.0136 (0.397)
Outside temperature (site) × Heat			0.00550 (0.402)	-0.0184 (0.299)
Outside relative humidity (site)			-0.000152 (0.883)	-0.00140 (0.484)
Outside relative humidity (site) × Heat			0.00247 (0.117)	0.00152 (0.335)
Outcome control mean	0.325	0.325	0.455	0.212
R-squared	0.182	0.182	0.0843	0.00808
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Within Supplement

Section G.9, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes, aside from Supplement Table G.9.11. Additionally, within Supplement Section G.9, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes, aside from Supplement Table G.9.11. The outcome in this table is the share of questions (out of 5) from the Cognitive Reflection Test answered correctly. "Outside temperature (pooled)" captures the demeaned (relative to the pooled sample) average outside temperature in Celsius for the day of the session. "Outside relative humidity (pooled)" captures the demeaned (relative to the pooled sample) average outside relative humidity in percentage for the day of the session. "Outside temperature (site)" captures the demeaned (relative to the site sample) average site-specific outside temperature in Celsius for the day of the session. "Outside relative humidity (site)" captures the demeaned (relative to the site sample) average site-specific outside relative humidity in percentage for the day of the session. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.9.13: Charitable donation

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-4.963 (0.874) [1.000]	-136.5 (0.006)** [0.080]	-12.73 (0.663) [0.787]	-0.896 (0.986) [1.000]
Male	8.366 (0.793)	15.05 (0.733)	-56.54 (0.111)	71.31 (0.135)
Male × Heat		-10.36 (0.864) [1.000]		
Nairobi	325.5 (0.000)**	221.9 (0.002)**		
Nairobi × Heat		207.0 (0.018)* [0.108]		
Matched with ingroup charity	-9.866 (0.761)	-69.02 (0.132)	34.89 (0.363)	-68.84 (0.232)
Matched with ingroup charity × Heat		123.5 (0.085)		
Earnings in tokens	0.00419 (0.688)	0.00426 (0.684)	-0.0169 (0.150)	0.0279 (0.110)
Outside temperature (pooled)	3.899 (0.655)	13.85 (0.161)		
Outside temperature (pooled) × Heat	-10.87 (0.167)	-30.85 (0.005)**		
Outside relative humidity (pooled)	1.642 (0.383)	3.120 (0.136)		
Outside relative humidity (pooled) × Heat	-1.576 (0.499)	-4.331 (0.145)		
Outside temperature (site)			7.435 (0.429)	41.62 (0.394)
Outside temperature (site) × Heat			-15.87 (0.086)	-114.0 (0.094)
Outside relative humidity (site)			0.0764 (0.969)	9.951 (0.034)*
Outside relative humidity (site) × Heat			0.965 (0.708)	-18.34 (0.007)**
Outcome control mean	410.4	410.4	245.0	561.8
R-squared	0.0670	0.0707	0.0119	0.0184
Observations	1806	1806	861	945

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. Multiple testing adjusted False Discovery Rate (FDR) q-value significance level in square brackets. Within Supplement Section G.9, for each specification (1)-(4), multiple hypothesis testing adjustments are performed on the set of p-values associated with Heat, across primary outcomes, aside from Supplement Table G.9.11. Additionally, within Supplement Section G.9, in (2) multiple hypothesis testing adjustments are performed on the set of p-values associated with the interaction between Heat and Male, across primary outcomes, as well as on the set of p-values associated with the interaction between Heat and Nairobi, across primary outcomes, aside from Supplement Table G.9.11. The outcome in this table is the amount of tokens earned in the experiment that is donated to the randomly selected charity. In Nairobi, "Matched with ingroup" charity is an indicator taking on a value of one if a participant is matched to a charity associated with her ethnicity, and 0 otherwise. In California, "Matched with ingroup charity" is an indicator taking on a value of one if a participant has resided in the San Francisco Bay Area for five years or more and is matched with a charity in the San Francisco Bay Area. "Earnings in tokens" captures the amount of tokens earned in the experiment. "Outside temperature (pooled)" captures the demeaned (relative to the pooled sample) average outside temperature in Celsius for the day of the session. "Outside relative humidity (pooled)" captures the demeaned (relative to the pooled sample) average outside relative humidity in percentage for the day of the session. "Outside temperature (site)" captures the demeaned (relative to the site sample) average site-specific outside temperature in Celsius for the day of the session. "Outside relative humidity (site)" captures the demeaned (relative to the site sample) average site-specific outside relative humidity in percentage for the day of the session. * $p < .05$, ** $p < .01$; + $q < .05$, ++ $q < .01$

Table G.9.14: Happiness

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-0.0788 (0.207)	0.0925 (0.484)	0.0481 (0.587)	-0.190 (0.029)*
Male	0.178 (0.004)**	0.165 (0.078)	0.214 (0.019)*	0.115 (0.173)
Male × Heat		0.0221 (0.868)		
Nairobi	1.141 (0.000)**	1.314 (0.000)**		
Nairobi × Heat		-0.345 (0.143)		
Outside temperature (pooled)	0.00371 (0.851)	-0.0152 (0.541)		
Outside temperature (pooled) × Heat	-0.0215 (0.301)	0.0166 (0.637)		
Outside relative humidity (pooled)	0.00282 (0.471)	0.000118 (0.978)		
Outside relative humidity (pooled) × Heat	-0.000101 (0.984)	0.00521 (0.429)		
Outside temperature (site)			-0.0122 (0.648)	-0.0369 (0.614)
Outside temperature (site) × Heat			0.00666 (0.859)	-0.00804 (0.949)
Outside relative humidity (site)			0.00124 (0.818)	-0.00258 (0.694)
Outside relative humidity (site) × Heat			-0.000553 (0.942)	0.0164 (0.139)
Outcome control mean	5.313	5.313	4.625	5.914
R-squared	0.182	0.184	0.00745	0.0182
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. The outcome in this table is the number chosen from a 1-7 scale that asked how one felt towards the end of the experiment, with 1 being sad and 7 being happy. "Outside temperature (pooled)" captures the demeaned (relative to the pooled sample) average outside temperature in Celsius for the day of the session. "Outside relative humidity (pooled)" captures the demeaned (relative to the pooled sample) average outside relative humidity in percentage for the day of the session. "Outside temperature (site)" captures the demeaned (relative to the

site sample) average site-specific outside temperature in Celsius for the day of the session. "Outside relative humidity (site)" captures the demeaned (relative to the site sample) average site-specific outside relative humidity in percentage for the day of the session. * $p < .05$, ** $p < .01$

Table G.9.15: Alertness

	(1) Pooled	(2) Pooled	(3) California	(4) Nairobi
Heat	-0.420 (0.000)**	-0.667 (0.000)**	-0.531 (0.000)**	-0.324 (0.001)**
Male	0.386 (0.000)**	0.337 (0.002)**	0.734 (0.000)**	0.00881 (0.928)
Male × Heat		0.101 (0.500)		
Nairobi	1.883 (0.000)**	1.701 (0.000)**		
Nairobi × Heat		0.364 (0.137)		
Outside temperature (pooled)	-0.0186 (0.407)	0.00258 (0.928)		
Outside temperature (pooled) × Heat	0.00911 (0.685)	-0.0340 (0.346)		
Outside relative humidity (pooled)	-0.000881 (0.868)	0.00204 (0.720)		
Outside relative humidity (pooled) × Heat	0.000249 (0.967)	-0.00558 (0.418)		
Outside temperature (site)			0.00865 (0.778)	-0.0427 (0.589)
Outside temperature (site) × Heat			-0.0256 (0.492)	-0.306 (0.015)*
Outside relative humidity (site)			0.00511 (0.471)	-0.00184 (0.810)
Outside relative humidity (site) × Heat			-0.00898 (0.285)	-0.00644 (0.510)
Outcome control mean	5.247	5.247	4.281	6.090
R-squared	0.304	0.305	0.0755	0.0316
Observations	1878	1878	874	1004

Note: Standard errors clustered at the session level. Per-comparison p-values in parentheses. The outcome in this table is the number chosen from a 1-7 scale that asked how one felt towards the end of the experiment, with 1 being tired and 7 being alert. "Outside temperature (pooled)" captures the demeaned (relative to the pooled sample) average outside temperature in Celsius for the day of the session. "Outside relative humidity (pooled)" captures the demeaned (relative to the pooled sample) average outside relative humidity in percentage for the day of the session. "Outside temperature (site)" captures the demeaned (relative to the

site sample) average site-specific outside temperature in Celsius for the day of the session. "Outside relative humidity (site)" captures the demeaned (relative to the site sample) average site-specific outside relative humidity in percentage for the day of the session. * $p < .05$, ** $p < .01$

References

Andreoni, J., M. A. Kuhn, and C. Sprenger (2015). Measuring time preferences: A comparison of experimental methods. *Journal of Economic Behavior & Organization 116*, 451–464.