Appendix: Additional Analyses

A1. Within-Participant Analysis of Responses

Specifications 2 and 4 of Table 4 show that with participant fixed effects, perceptions of relative and absolute quality vary by health score, and Specifications 2 of Table 4 show the same for WTP. To get a sense of how much variation there is across participants in the variation by health score, we study how the spread between the highest and lowest response with each task varies across participants. Figures A1–A3 show that for all three tasks, there is a fair amount of dispersion in spreads across participants.

Figure A1. Histogram and kernel density plot of spread between highest and lowest WTP across health score ranges at the participant level

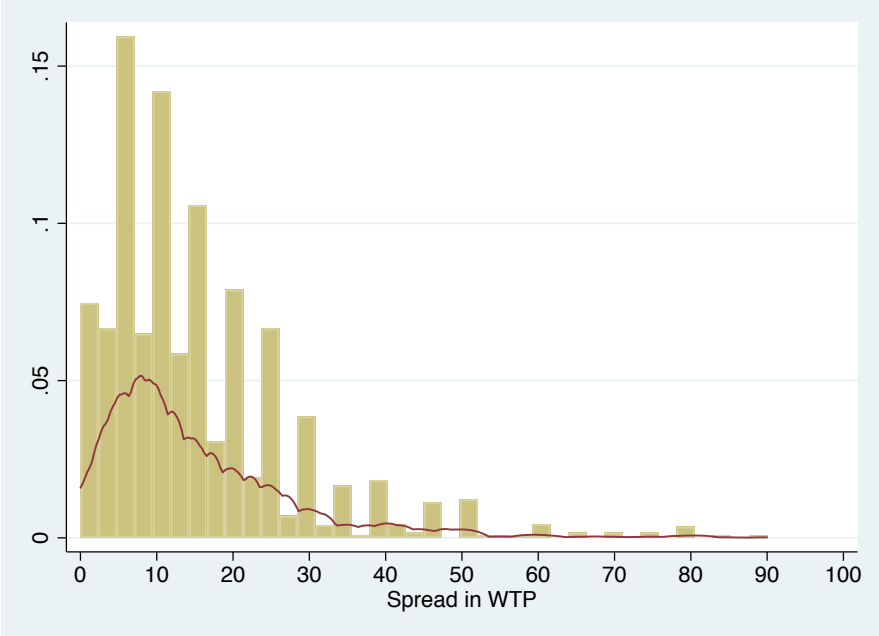


Figure A2. Histogram and kernel density plot of spread between highest and lowest guess in the Score Distribution task across health score ranges at the participant level

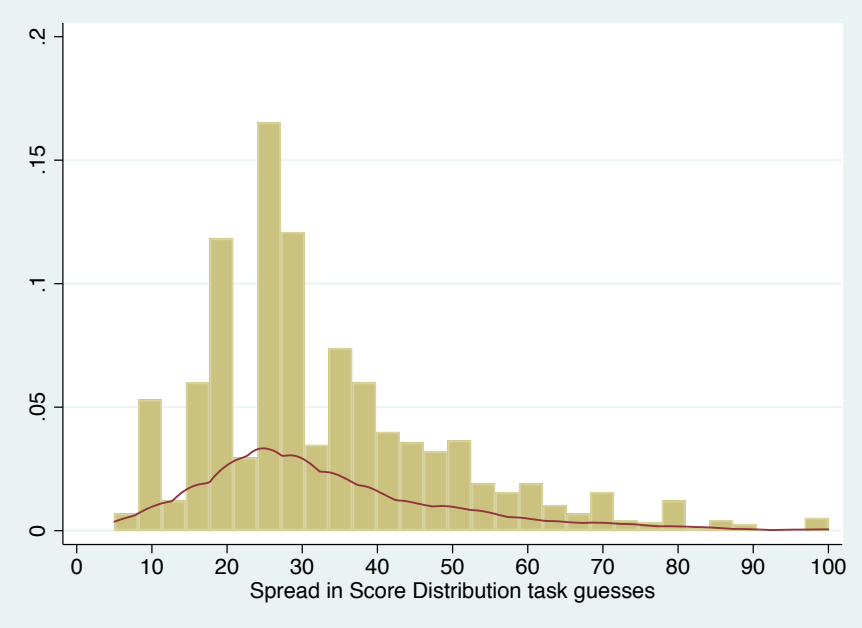
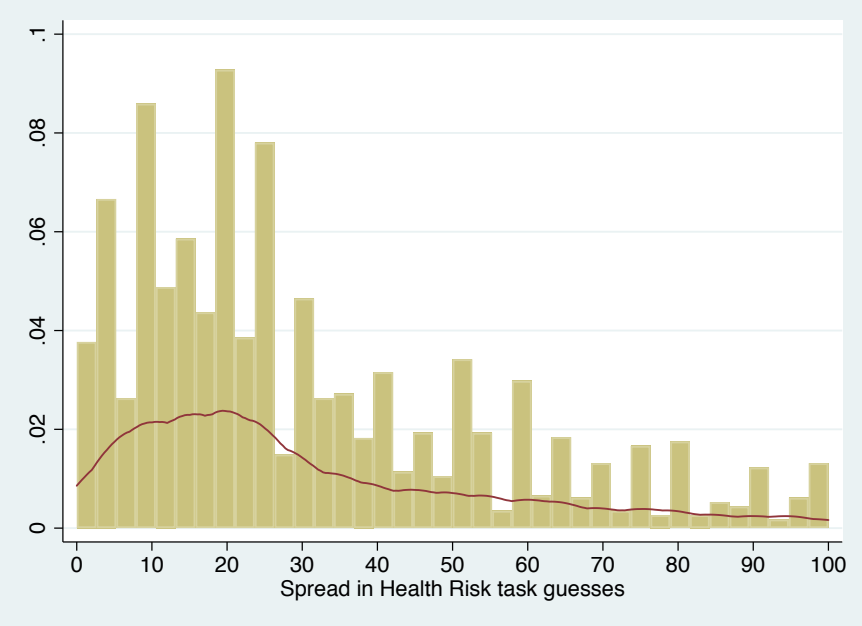


Figure A3. Histogram and kernel density plot of spread between highest and lowest guess in the Health Risk task across health score ranges at the participant level



A2. Attentional Robustness Checks

In this robustness check, which is based on an anonymous referee's suggestion, we excluded 2.1% of participants because they provided the same guess for all of the health score ranges in the Health Risk task. While it is possible that these participants believe health risks do not change with health score, it is also possible that these participants gave the same answer every time to avoid thinking about the problem.

We also excluded participants who exhibited monotonicity violations (see Table A1). In WTP elicitation, 25.4% of participants had a monotonicity violation with respect to increases in health score. In the Health Risk task, a larger fraction of participants (62.9%) had a monotonicity violation. There is some overlap between these groups, as 20.5% of participants had a monotonicity violation in both cases. Altogether, we excluded all 67.9% of participants because they exhibited a monotonicity violation.

We find that our primary regression results are robust to excluding participants who exhibit these patterns. First, when looking at our regression analysis of mistakes in the two guessing tasks (Tables 4 and A2), the signs and significance remains unchanged for all coefficients. Importantly, the effect sizes are similar as well, which shows that our findings about mistakes along both pathways are largely unchanged.

Second, when looking at our regression analysis of the difference in WTP with differences in beliefs in the control condition (Table 7 and Table A3), the size, signs, and significance of the coefficients remains largely unchanged. The only noticeable difference is a decrease in the statistical significance in the impact of beliefs about absolute health risk on WTP. The magnitude of this coefficient is largely unchanged, so the loss of statistical significance is likely due to a reduction in statistical power.

Finally, when looking at our regression of the difference in WTP across treatments (Table 9 and Table A4), we see the same general pattern, with a decrease at the lower end, an increase at the higher end with both treatments, and a stronger effect for the DD condition. However, none of the treatment coefficients are now statistically significant, and again we attribute this to a reduction in statistical power.

A3. Follow-up Experiment with Incentives for WTP

Method

We recruited all participants who participated in the primary experiment (reported in the main text) who were located in San Francisco or had a restaurant meal in San Francisco in the past six months, and assigned them to the same condition as their original condition assignment (control, D, or DD). Participants received $1.50 for completing this follow-up study, and they were further informed that they had a one-in-six chance of winning a restaurant gift card, the value of which could range between $0 and $50, depending on their choices.

All participants began by reading a description about restaurant hygiene scores. Specifically, they were informed: "Every restaurant in San Francisco (California) gets regularly inspected by the Health Department. During each inspection, they receive a Food Safety score, also known as a 'health score.' This score is calculated based on the type and the number of health code violations observed. 100 is the best possible score, and 0 is the worst possible score. During today's session, you will be asked to choose between gift cards of different amounts at restaurants with different health scores." They could not proceed until they answered two attention-check questions correctly.

Participants then were informed about the choices they would have to make throughout the experiment:

> "During today's session, you will be asked to make several choices about real restaurants with different health scores. In each case, you will be shown two different options called Restaurant A and Restaurant B. Both are actual restaurants located in the San Francisco area (we have masked their names for the purposes of this survey), feature the same cuisine, the same number of stars on Yelp, and the same number of dollar signs on Yelp. But they have two different health scores.
>
> You will be asked whether you would like a $25 gift card to Restaurant A or a gift card of another amount to Restaurant B. These gift cards are non-transferrable. For instance, if you choose a gift card for Restaurant B over a gift card for Restaurant A, you can only spend that gift card at Restaurant B.
>
> You will have a 1 in 6 chance of winning a gift card. If you are chosen to receive a gift card, one of the questions during today's session will be randomly selected and we will

send a link to the gift card that you chose on that one question to your MTurk account, so it is important to choose carefully on each question."

After answering one practice question, participants then were presented with nine pairs of restaurants in random order. For instance, in one scenario, they saw:

"Restaurant A and Restaurant B are both located in the San Francisco area, have the same cuisine type (New American), same consumer rating (4.0), and same dollar signs ($$$). However, they have different health scores: Restaurant A received 94 out of 100 while Restaurant B received 75 out of 100 during each of their recent restaurant health inspections.

We are going to ask you which gift card you would rather have: a $25 gift card to Restaurant A or a gift card to Restaurant B.

Would you rather have...

Question number	**A gift card to Restaurant A (Health score: 94) for**		**A gift card to Restaurant B (Health score: 75) for**
1	$25	or	$0
2	$25	or	$1
:	:	or	:
49	$25	or	$48
50	$25	or	$49
51	$25	or	$50

At which dollar value you would switch to picking the gift card for Restaurant B? If you prefer a $25 gift card to Restaurant A no matter what, you can select the last option.

*Remember, you are actually answering 51 questions by giving us this answer. We will randomly pick one question if you are selected to win a gift card and give you the option you chose on that one question. Each question is equally likely to be chosen for payment."

Results

Altogether, 144 participants completed this follow-up experiment. Because of the relatively small sample, we pool all three conditions when analyzing the data. Following our analysis of the primary experiment, we look at the difference in WTP for the difference in health scores. However, in the follow-up study, differences in WTP and difference in health scores are identified differently than in the primary

study. For a given pair of restaurants, the difference in WTP is given by the switch point on the price list. Differences in health scores were found by comparing the health scores of the two restaurants.

Figure A4 provides the differences in WTP for each range of differences in health score by study for participants in the follow-up study. Because there is not a clear reference restaurant in the follow-up study, there is not a clear sign on any of the differences. For example, if the switch point to the lower hygiene score is at $45, then the difference in WTP could be viewed as $25 − $45= −$20 or $45 − $25 = $20. Thus, to facilitate comparisons with the primary study, we includes each observation in the follow-up study twice in this figure: once with the restaurant with the lower health score as the reference restaurant and once with the restaurant with the higher health score as the reference restaurant.

If demand effects were the reason for variation we observed in the relationship between WTP and ratings in our primary study, we would expect that adding incentives in this follow-up study would substantially weaken the relationship between WTP and ratings. However, as can be seen in Figure A4, we find that the relationship between WTP and rating is even stronger when WTP is incentivized.

Figure A4. Difference in WTP for each range of differences in health score (relative to WTP and guess of health score for most frequent restaurant) for participants in the follow-up study

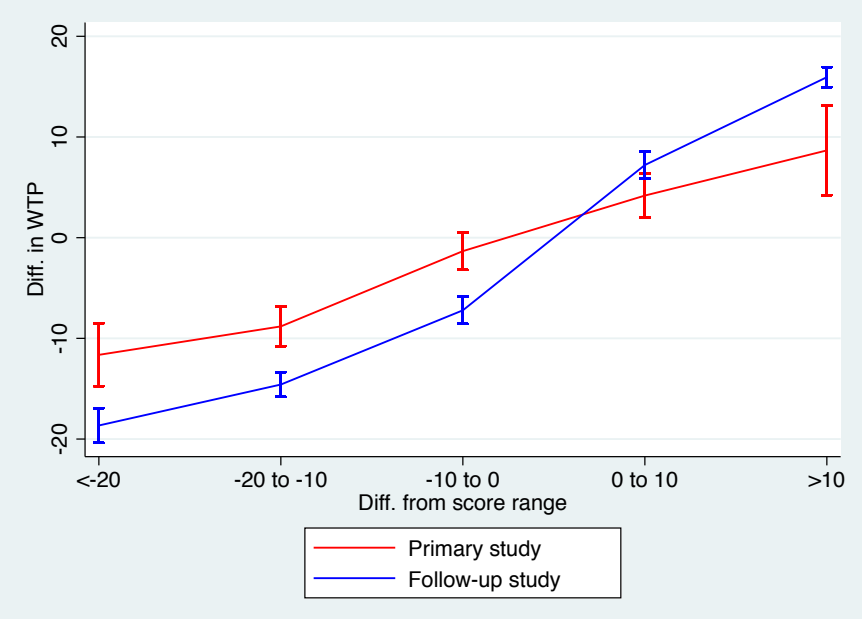


Table A1. Monotonicity violations by task

WTP task monotonicity violations	Health Risk task monotonicity violations						
	0	1	2	3	4	5	Total
0	369	259	140	54	12	23	857
1	51	72	65	22	9	4	223
2	5	11	24	7	3	2	52
3	0	3	8	1	1	0	13
4	1	0	2	0	1	0	4
Total	426	345	239	84	26	29	1,149

Table A2. Regressions of participant guesses (control condition)

VARIABLES	(1) Actual guess % w/ high risk	(2) Actual guess % w/ high risk	(3) Actual guess % at or above	(4) Actual guess % at or above
Health score in 71–75	46.891***	46.891***	−19.631***	−19.631***
	(1.901)	(2.075)	(1.363)	(1.489)
Health score in 76–80	42.255***	42.255***	−13.333***	−13.333***
	(1.261)	(1.376)	(1.132)	(1.236)
Health score in 81–85	45.241***	45.241***	−5.438***	−5.438***
	(0.887)	(0.969)	(0.674)	(0.736)
Health score in 91–95	−4.914***	−4.914***	−6.779***	−6.779***
	(0.977)	(1.066)	(0.689)	(0.753)
Health score in 96–100	−23.044***	−22.929***	−9.885***	−9.900***
	(0.916)	(0.994)	(1.339)	(1.448)
Constant	15.955***	10.795***	33.060***	29.994***
	(4.411)	(0.750)	(4.696)	(0.687)
Observations	1,174	1,174	1,174	1,174
R-squared	0.626	0.877	0.095	0.771
Demographic controls	Yes	No	Yes	No
Participant fixed effects	No	Yes	No	Yes

Notes: The dependent variable is the under-guess (actual guess) in the task. In parentheses are robust standard errors clustered by participant. *** p<0.01, ** p<0.05, * p<0.10. Demographic controls are dummies for age quartile, gender, whether located in San Francisco or Los Angeles, whether had more than one meal in a restaurant per month, and whether had a restaurant meal in San Francisco in the past six months.

Table A3. Regressions of Difference in WTP (control condition)

VARIABLES	(1) Diff WTP	(2) Diff WTP	(3) Diff WTP	(4) Diff WTP
Difference in health score <20	−9.179***	−10.246***		
	(1.323)	(0.997)		
Difference in health score −20 to −10	−5.715***	−5.690***		
	(0.646)	(0.583)		
Difference in health score 0 to 10	4.423***	4.748***		
	(0.482)	(0.558)		
Difference in health score >10	11.288***	8.597***		
	(2.637)	(1.113)		
Diff in guess of % w/ high risk			−0.039*	−0.033*
			(0.023)	(0.019)
Diff in guess of % at or above			−0.122***	−0.143***
			(0.016)	(0.013)
Constant	−0.842	−1.341***	0.024	−1.047***
	(1.528)	(0.277)	(1.445)	(0.076)
Observations	1,174	1,174	1,126	1,126
R-squared	0.267	0.843	0.213	0.817
Demographic controls	Yes	No	Yes	No
Participant fixed effects	No	Yes	No	Yes

Notes: The dependent variable is the difference in WTP at a given score range from the WTP at the most frequent restaurant. In parentheses are robust standard errors clustered by participant. *** p<0.01, ** p<0.05, * p<0.10. Demographic controls are dummies for age quartile, gender, whether located in San Francisco or Los Angeles, whether had more than 1 meal in a restaurant per month, and whether had a restaurant meal in San Francisco in the past 6 months.

Table A4. Regressions of Difference in WTP (all conditions)

VARIABLES	(1) Diff WTP	(2) Diff WTP
Difference in health score <20	−9.597***	−9.141***
	(0.828)	(1.362)
Difference in health score −20 to −10	−6.104***	−5.605***
	(0.375)	(0.656)
Difference in health score 0 to 10	4.680***	4.437***
	(0.331)	(0.494)
Difference in health score >10	13.158***	11.160***
	(2.233)	(2.605)
Condition D	−0.257	−0.189
	(1.011)	(0.850)
Difference in health score <20 *D		0.318
		(1.927)
Difference −20 to −10 *D		−0.780
		(0.961)
Difference 0 to 10 *D		−0.198
		(0.729)
Difference >10 *D		1.450
		(5.035)
Condition DD	−0.786	−1.168
	(1.088)	(1.026)
Difference in health score <20 *DD		−2.438
		(2.319)
Difference −20 to −10 *DD		−0.726
		(0.904)
Difference 0 to 10 *DD		0.866
		(0.842)
Difference >10 *DD		4.688
		(5.415)
Constant	1.634	1.732
	(1.532)	(1.544)
Observations	3,403	3,403
R-squared	0.278	0.281
Demographic controls	No	No
Participant fixed effects	Yes	Yes

Notes: *** p<0.01, ** p<0.05, * p<0.10. Demographic controls are dummies for age quartile, gender, whether located in San Francisco or Los Angeles, whether had more than one meal in a restaurant per month, and whether had a restaurant meal in San Francisco in the past six months.

Table 7. Regressions of Difference in WTP (control condition)

VARIABLES	(1) Diff WTP	(2) Diff WTP
Difference in health score <20	−9.612***	
	(0.708)	
Difference in health score −20 to −10	−5.450***	
	(0.414)	
Difference in health score 0 to 10	4.684***	
	(0.400)	
Difference in health score >10	9.402***	
	(0.948)	
Constant	−1.798***	−0.533
	(0.235)	(1.378)
Observations	2,258	2,092
R-squared	0.827	0.178
Demographic controls	No	Yes
Participant fixed effects	Yes	No

Notes: The dependent variable is the difference in WTP at a given score range from the WTP at the most frequent restaurant. In parentheses are robust standard errors clustered by participant. *** p<0.01, ** p<0.05, * p<0.10. Demographic controls are dummies for age quartile, gender, whether located in San Francisco or Los Angeles, whether had more than one meal in a restaurant per month, and whether had a restaurant meal in San Francisco in the past six months.