PNAS RESEARCH ARTICLE | ECONOMIC SCIENCES OPEN ACCESS

Estimating perceptions of the relative COVID risk of different social-distancing behaviors from respondents' pairwise assessments

Ori Heffetz[a,b,c,1,2] and Matthew Rabin[d,e,1]

Contributed by Matthew Rabin; received November 23, 2022; accepted December 29, 2022; reviewed by Angus Deaton and Amy N. Finkelstein

How do people compare the effectiveness of different social-distancing behaviors in avoiding the spread of viral infection? During the COVID pandemic, we showed 676 online respondents in the United States, United Kingdom, and Israel 30 pairs of brief videos of acquaintances meeting. We asked respondents to indicate which video from each pair depicted greater risk of COVID infection. Their choices imply that on average, respondents considered talking 14 min longer to be as risky as standing 1 foot closer, being indoors as standing 3 feet closer, being exposed to coughs or sneezes as 3 to 4 ft closer, greeting with a hug as 7 ft closer, and with a handshake as 5 ft closer. Respondents considered properly masking as protecting the wearer and interlocutor equally, removing the mask entirely or only when talking as standing 4 to 5 ft closer but wearing it under the nose as only 1 to 2 ft closer. We provide weaker evidence on beliefs about the interaction effects of different behaviors. In a more limited, ex post analysis, we find little evidence of differences in beliefs across subpopulations.

risk perceptions | social distance | perceived health-risk tradeoffs | belief elicitation | pairwise choice

Significance

The COVID-19 pandemic led to a worldwide campaign to change our lives. It involved governmental and health-professional communications with explicit guidance, such as keeping social interactions short and outdoors, maintaining six-foot distance, avoiding handshakes, and properly wearing masks. What have people come to believe regarding the relative COVID-infection risk of different combinations of such social-distancing behaviors? We measure and estimate these relative-risk perceptions with surveys that we conducted in three countries during COVID. To the extent that people's perceived risk tradeoffs affect their behavior, our investigation helps focus health-risk communications on relative risks, potentially helping people make healthier choices.

Although its details varied by country and evolved over time, the COVID-19 pandemic led to a worldwide campaign (by persuasion and mandates) to change our lives. We have been told that we can reduce COVID-infection risk by not touching our faces, covering our coughs, properly washing our hands, disinfecting surfaces, staying at home when we feel ill, and (once available) getting tested frequently and getting vaccinated. The campaign also included guidelines for our private interactions with others: We should replace hugs and handshakes with elbow bumps, stand six feet apart, wear masks (properly), and keep our interactions brief and outdoors. These conversational guidelines were unusual not only in their scale and domain but also in their concreteness. Billions of people were told where and how far apart to be when conversing with others and whether and how to make physical contact.

There are nonetheless reasons to doubt that this communication provided people with a sharp sense of how important these different precautions are. It is impossible to imagine precise calibrations of the health costs of failing to follow different parts of these guidelines, in terms of decreasing life expectancy, or even "merely" providing the exact probability that a behavior will lead to infection. And as in many other domains, the risks vary massively by the age and health status of individuals. The presence of this variation was well conveyed, but it rendered any effort to convey the level of danger to a "typical" person meaningless. Likewise, advice was given as the experts and officials were racing to understand the dangers of this new virus, and doing so when the virus itself was evolving; precise statements might later backfire if they made changes to guidelines more salient. All said, conveying the absolute risks would therefore be a fanciful task. Even now, and even with a specific person in mind, we doubt experts would venture estimates of the percentage-point increase in the risk of infection or dying from one handshake.

By contrast, we assume almost all members of the public who accepted the reality of the pandemic would have a strong sense of expert advice on the direction of risk for any particular aspect of behavior. A long conversation is riskier than a short one, an unmasked one is riskier than a masked one, and an inside one is riskier than outside. While it would be useful to confirm (as we do) what people thought of as good and bad changes in particular behaviors, it would seem unlikely that we could learn much by eliciting such beliefs.*

*There are a few exceptions we can think of: We were a bit curious, for instance, about what people thought about hugs vs. handshakes.

Author contributions: O.H. and M.R. designed research; performed research; analyzed data; and wrote the paper.

Reviewers: A.D., Princeton University; and A.N.F., Massachusetts Institute of Technology.

The authors declare no competing interest.

[1]O.H. and M.R. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: oh33@cornell.edu.

This article contains supporting information online at http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2219599120/-/DCSupplemental.

Published February 7, 2023.

In between knowing which single behaviors are bad—as most of us do—and having precise beliefs about how bad each behavior is—as none of us does—is having beliefs about the relative risk of different behaviors. Although indications of the relative health risk of different activities seem conspicuously absent from most medical communication—and we hence do not know what experts' opinions on the relative risks are, nor how much expert consensus there is—quantifying and conveying risks relative to each other may be more achievable than precise calibrations. At the same time, it may be more helpful than simply identifying good and bad behavior because people often face tradeoffs: We cannot always improve particular behaviors in isolation but rather need to know which of two bundles of behaviors puts us at a greater risk of contracting COVID. Should we opt for longer work meetings outdoors—or shorter ones indoors? Should we keep greater distance in those meetings but then, to be heard, have to remove our masks when talking—or stand closer with masks always on? Should we go to a crowded indoor mask-enforced space where people queue in long lines—or to an outdoor market, with much shorter lines but unmasked people?

In this paper, we measure what people have come to believe regarding the relative dangers of infection of different conversational behaviors. As complicated as many aspects of the disease are, to our understanding, the risk of infection is an unusually concrete, objective, essentially one-dimensional outcome variable. It is also of nearly universal relevance and concern to everybody participating in our survey (unlike other proximate health outcomes, such as determinants of raising blood pressure, whose impact is neither concrete nor universally relevant). The likelihood of infection may vary across people a great deal, and the likelihood of illness from contracting the virus varies even more widely, but (without being experts) we are unaware of variation in the relative risk of different activities. What is riskier for a young healthy person is also riskier for the elderly or those with exacerbating conditions. Moreover, although advice changed over time, especially early in the pandemic, it is our impression that advice in this domain changed little over the period of our study.[†] And while the delta variant arose between our first and last rounds of survey (the omicron arose well after our last survey), we are also unfamiliar with variation in advice that depended on the Greek alphabet. Although we hope that techniques similar to those employed in this article could be used to study a broader range of health domains, all these factors provide an unusual (if unfortunate) advantage for studying relative-risk perceptions in the COVID context. In addition to the central importance of the domain itself, the ubiquity of its relevance and the relative uniformity of appropriate measures allow us to gather evidence without too much focus on targeting the right set of respondents.

To study perceptions, we showed a total of 676 respondents in the United States, United Kingdom, and Israel between May and September 2021, a sequence of pairs of 5-seconds videos played simultaneously, side by side, each showing the same acquaintances meeting and conversing. Respondents were asked to judge, for one of the two people designated, which of the two scenarios in each pair is riskier. From their responses to 30 pairs of such videos, we estimate their perceptions of how risks changed by the features of the conversation. We use videos—rather than written descriptions—for several reasons. It allows us to vary features of situations without suggesting the importance of these features and to make the salience of those features be as naturalistic as possible. We were curious, for example, not just whether people knew how important it is to cover the nose, when asked, but also whether that is something they attended to in conversations.[‡] Of course, witnessing naturalistic situations is still different from participating in them, and we speculate below how our design might have distorted perceptions relative to the case where people participate in conversations. There was one exception to our no-written-descriptions approach. Because we could not replace our 5-s clips with 300- to 3,600-s "clips," in treatments where it is varied, duration is therefore written below each video, as shown in Fig. 1 on page 3. We also had treatments that did not indicate conversation duration, and others where we provided identical duration for the two videos.[§]

We collected four convenience online samples from three countries over the course of 4 mo. We began by collecting data from 100 US participants in May 2021 and (after verifying that responses were not pure noise, and making no material modifications except for our randomization of videos into pairs, as explained below) from about 200 in August. To expand our data numerically and geographically, we collected samples of about 200 each in the United Kingdom in August and in Israel in September. Our samples are not meant to be nationally representative. While we discuss below a few seeming differences across the three countries, we find generally similar estimated tradeoffs. We had no ex ante hypotheses we were testing about qualitative features of perceptions—and did not preregister any such hypotheses. Nor can we contrast quantitative perceptions with expert opinion—which we do not know.[¶]

Overall (Table 1 on page 6), respondents perceive 14 extra minutes of social interaction to be as risky as standing one foot closer to the other person and interacting indoors as equivalent to standing 3 ft closer outdoors. Masks loom large to respondents: Wearing no mask by one of the interacting parties is perceived as equivalent to standing 4 to 5 ft closer (9 ft if both are maskless); wearing mask under one's nose as reducing protection by a third; and pulling one's mask down when talking as roughly equivalent to wearing no mask at all. Hugs and handshakes are seen as comparably risky, and as risky as standing 5 to 7 ft closer during the entire conversation, relative to no physical contact, and almost three to four times riskier than elbow bumps. We were surprised

[†]Dramatic shifts in understanding of the virus occurred early in the pandemic. Most notable was evolution in beliefs about how "airborne" the virus is, shifting the emphasis from keeping surfaces clean to ridding shared air of the virus.

[‡]Our design reflects an initial focus on the translation respondents might have had from ubiquitous messaging quantifying the appropriate distance into respondents' objective experience of distance. In our exit questionnaire, we asked respondents to estimate the distances in the videos. Although individuals' assessments varied and (per usual) respondents tended to provide round numbers in local units, the median respondent is generally well calibrated. We summarize these findings in *SI Appendix* .

[§] The closest existing studies we found investigate research questions different from ours and do not use videos. Svenson et al. (1) ask online US respondents: "Assume that two persons are in a face to face conversation … what percentage of the airborne viruses reaching a person at 2 ft will reach a person at 4 feet?" with a similar question for a decrease in distance. They find that relative to their reading of the existing evidence, respondents underestimate the effects on exposure of moving closer and away from another person. Their question and findings are orthogonal to ours because they looked at the perceived exposure effect of distance, while we compare the perceived relative effects of distance, time, location, maskiness, etc. Ref. 2 replicates Ref. 1 and adds time of a conversation but does not trade time against distance as we do. Luckman et al. (3) ask online UK respondents to place themselves at the closest distance they would keep from a stranger, on a diagram representing figures with or without masks. Their question too is different from ours, as they study, and find, stated intentions of risk-compensating behavior: Respondents, especially younger ones, place themselves closer when masks are in use. In addition to risk perceptions, such behavior could also be driven by, e.g., personal preferences, health condition, habits, and perceived norms of acceptable behavior among different ages.

[¶]That said, a few findings seem to us as likely misperceptions. For example, respondents did not perceive being outdoors as dramatically safer than wearing a mask indoors. In a few instances where we think that the ranking is clear, we were surprised that our participants saw essentially no difference. In part prompted by such findings, we speculate below how the high salience of certain behaviors—either in the real world or only in our videos—may have affected their perceived risk.

Fig. 1. Example risk-assessment snapshot. Each respondent makes thirty such risk assessments. Three design features are randomized at the respondent level (features shown in the example are underlined): a) "risk... for the person in red/blue;" b) "risk... is higher/lower;" c) "all videos depict conversations of the same duration" (in the instructions)/"interaction duration: ... minutes" (under each video).

to find that on average, respondents believed wearing a mask was as important for self as for the others.

Since we had no strong hypotheses on how perceptions would vary by country, we did not design the experiment to disentangle national variation from any time trends in beliefs.[#,‖] Nonetheless, we found some differences that seem noteworthy. One difference was in perceptions regarding how people wear their masks. Respondents in the United States and the United Kingdom recognized a substantial difference in risk from wearing the mask below the nose versus fully, whereas Israelis thought that this difference mattered little. Israelis also seemed to think that a person's own mask mattered more than the other person's mask. Finally, being indoors vs. outdoors may have loomed larger in the United Kingdom (and perhaps in Israel) than in the United States.

Our primary intended and implemented estimates treated each of the variables independently. This independence is probably implausible—e.g., presumably the risk from any initial physical contact would not depend on the distance and duration of the conversation, whereas the other factors would. It may, however, be plausible that participants perceived little such interaction, and we indeed present evidence consistent with this possibility. We are also able to estimate some perceived interdependencies as well as the perceived risk of combinations of behaviors. For example (Fig. 2 on page 4 and Table 2 on page 7), masks are perceived more effective relative to other factors indoors than they are outdoors, and the perceived change in risk from an additional minute of conversation or an additional foot of distance is smaller at longer conversations and longer distances, respectively (i.e., we find perceived decreasing marginal effects of duration and distance).

We hope that one use of this article is to edge debates on public health a little closer to more precise communication. To the best of health professionals' knowledge, is being indoors as risky as standing 3 ft closer outdoors? Is the risk of an extra 10 min of conversation smaller when the conversation is already 40 min long? Is wearing a mask as effective as having the other person wear one? The public constantly makes choices based on such perceptions. It would be good to know what the experts think. Shouldn't authorities find out and communicate it to the public?

Survey Design and Data

Relative-Risk Assessments. Respondents complete thirty rounds of relative-risk assessments. Fig. 1 provides an example snapshot from one round. Each round consists of a pair of silent videos, appearing on the screen side by side. The two videos play simultaneously, in repeated loops, until the respondent clicks on one of them, moving to the next round.

Each video depicts the first moments of a social encounter between two people. It starts with the two people entering the scene walking toward each other and greeting each other. Sometimes they stop at a distance they then maintain, and other times, they hug, shake hands, or touch elbows, before repositioning themselves at a certain distance. The rest of the video is a conversation, during which one of the two people may take off the mask when talking or may cough or sneeze. The scene is cut after 5 s, and the video is replayed.

At the top of the screen, above the two videos, a fixed instruction reads: "Click on the video in which the risk of infection for the person in [red/blue] is [higher/lower]." Respondents are uniformly split into: a) being asked about red or blue, b) being asked about higher or lower risk, and c) being told the specific duration of each interaction or being asked to assume that all interactions are of the same duration. For split c), "specific duration" respondents are shown a caption under each video that reads: "Interaction duration: [5/10/15/20/30/40/60] min"

[#] Our guess is that most of the national differences were not due to time trends or due to self-selection into our national samples, but we have no data from our study or externally in support of that hunch. Because the emphasis on the value of being outdoors seemed to increase over time, however, we think that this difference could plausibly be based on time trends rather than national differences.

[‖] Despite the contentious nature of some aspects of the pandemic, we also had little reason to suspect that beliefs about conversational behaviors would vary based on the political views of our participants and did not design the study to identify any such differences. We did collect such data, as well as demographic characteristics, after the main part of the survey, and found no striking (or unstriking) patterns. We report several such analyses in *SI Appendix* (and summarize them below), and our data, publicly available at https://osf.io/bvyj2/, include all collected variables.

(uniformly randomized at the single-video level); "same duration" respondents are asked, before starting the survey, to assume that "all videos depict conversations of the same duration" (with no captions under the videos).** For each respondent, these three randomized design features are held fixed throughout the survey.

Videos. Our production team shot 379 videos during a single shooting day, featuring the same two actors, wearing the same clothes and always each appearing on the same side of the scene.†† Our posted data package available at: https://osf.io/bvyj2/ provides full details on our video catalog (by video features). Here, we provide a brief summary. The videos differ along several dimensions. First, 32 baseline videos show the following 2 × 4 × 2 × 2 combinations: the scene is located indoors or outdoors; after greeting remotely, the actors position themselves 2, 4, 6, or 8 ft apart; and either person does or does not wear a surgical mask throughout the video. Second, subsets of the baseline videos show the following variations, one per video: In addition to greeting remotely, the encounter starts with a hug, handshake, or elbow (after which the persons reposition themselves at the above distances and continue the encounter as usual); the person in red, blue, or both wear a cotton mask (two different types) rather than a surgical mask; the person in red or blue (but not both) coughs, sneezes, takes the mask off (by pulling it down) when talking, or wears the mask below their nose.‡‡ The indoor videos are all shot in the same location; outdoor videos use two different locations.

Survey Flow. Prior to making the thirty relative-risk assessments, respondents are provided with detailed instructions, are walked through an example round, are provided (system-generated) feedback and clarification regarding their assessment in that round, and are given an opportunity to confirm their choice or to cancel and retry. Importantly, they are asked, and are repeatedly reminded throughout the instructions, to assume that both people are not vaccinated for COVID-19 and, if randomized into "same duration," to also assume that all videos depict conversations of the same duration. They are also asked to consider each video on its own, as if it were the only interaction they saw, ignoring the behavior of the same actors they observed in other videos. After completing the instructions and example, respondents can start the thirty assessment rounds.

To ease respondents in, and as an attention check, the first four rounds consist of (randomly selected) "easy" video pairs: Using the anticipated (and observed) near-universal agreement on what is safer or less safe on each dimension, the encounter in one video is strictly riskier than that in the other video. In one pair, for example, the two videos are identical except for the distance. The remainder 26 assessment rounds consist of "noneasy" pairs, all involving tradeoffs (US May sample), or any pairs, that may or may not involve tradeoffs (all other samples).§§

After completing the thirty rounds, respondents a) estimate interaction distances (in feet or meters, both from memory and in real time; *SI Appendix*); b) restate interaction duration (only from memory and only in the "specified same duration" treatment; see footnote ††); and c) list sources that informed their understanding of COVID and ways to reduce infection risk, including specific media outlets and governmental resources. They then proceed to an exit questionnaire, asking them about d) past infection, symptoms, and severity (separately regarding self, family, relatives, and friends); e) social-distancing behavior regarding mask wearing, keeping distance, and hugging (self); f) vaccination uptake/intentions, and their timing (self); g) perceptions regarding being high risk (self, family), concerns regarding getting infected (self, others), opinion about the local COVID situation and whether they, family, or friends work in healthcare; and h) personal demographics and political views. Our posted data package available at: https://osf.io/bvyj2/ provides screenshots, full survey text, and details about sample-specific modifications and adjustments. The survey was programmed on Otree (4). The study was approved by the Hebrew University Committee for the Use of Human Subjects in Research. Informed consent was elicited on the first page of the web survey.

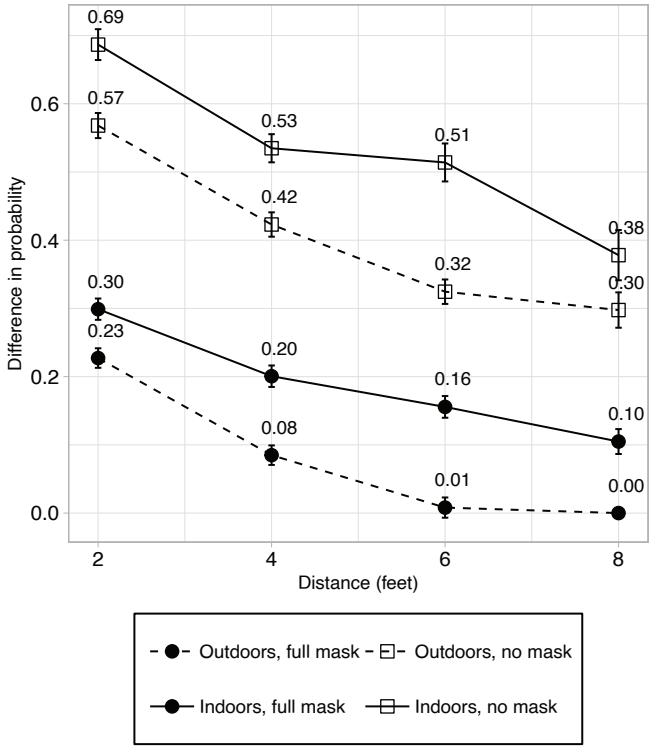

Fig. 2. Risk assessments: location, distance, and mask wearing. Source: authors' online surveys during 2021 on convenience samples in the United States, United Kingdom, and Israel. Average marginal effects (and SEs) from logit regression. Observations: $N = 13{,}689$ relative-risk assessments where masks are either worn in full or not at all. Dependent variable: 1 if the left video is assessed riskier; 0 otherwise. Independent variables: 31 main difference variables for the 32 baseline-feature combinations (2 in/outdoors × 4 distances × 2 blue mask full/none × 2 red mask full/none; omitted combination: outside, 8 ft apart, blue and red full masks), and controls for differences in duration and in 5 extra features (cough, sneeze, hug, handshake, and elbow). The 31 main independent variables: 1 if feature combination or extra feature appears only in the left video, −1 if only in the right video, and 0 if both/none. The figure shows the estimated effects on the 16 of the 32 baseline-feature combinations where either both blue and red wear full masks or none does (omitted combination coefficient = 0); *SI Appendix*, Fig. S1 for all 32 effects. Each depicted coefficient represents the estimated difference in probability that its combination is indicated as riskier when compared with the omitted combination (outside, 8 ft apart, full masks).

**Respondents in the "same duration" treatment are further randomized into two subtreatments: unspecified same duration and specified same duration (10 and 40% of all respondents, respectively). In the latter, the instruction prior to starting the survey explicitly specifies the duration of all interactions. Specifically, respondents are asked to assume that "all videos depict conversations of the same duration: [5/10/15/20/30/40/60] min" (uniformly randomized at the respondent level).

††Due to COVID-risk considerations, only pairs of domestic-partner actors were auditioned.

‡‡In order to avoid unnecessary ambiguity, respondents are never asked to assess the risk of infection for a person who is seen coughing or sneezing.

§§Across all respondents and first four rounds, the strictly riskier video is indeed indicated as riskier 90.6% of the time. To the extent that these rounds serve as attention check, our respondents seem attentive.

Data. Respondents in the United States and the United Kingdom were recruited using Prolific (https://www.prolific.co) and in Israel using Sekernet (https://sekernet.co.il). Median respondent age was around 30 in each US and UK sample, and 43 in Israel; overall, 50% of respondents were female, with a share of 47 to 52% in each sample; the median survey-completion time is 15 to 18 min in all four samples. *SI Appendix*, Table S1 panel A reports the exact statistics as well as respondents' reported infection and vaccination status; its panel *B* reports official local daily infection and death rates during each of the four survey samples.

Pooled across the four samples (total $N = 676$) and all rounds ($N = 30$), our main data consist of 20,280 pairwise relative-risk assessments.[¶¶]

Results

Semiparametric Relative-Risk Estimates. Fig. 2 shows the impact on relative-risk assessment of each of sixteen combinations of features: location (in/outdoors) × distance (2/4/6/8 ft) × full-mask wearing (none/both actors). By estimating the effect of each individual combination, rather than of its underlying features, we avoid imposing assumptions regarding how effects vary with distance or regarding how location, distance, and mask wearing interact. Each of the sixteen points on the graph shows the estimated difference in probability that respondents assess the relevant combination as riskier compared with the combination at the bottom right: outdoors, 8-ft apart, both wearing full masks. That combination is considered the safest (both by common wisdom and, as the figure shows, by our respondents). Its value of 0 (by construction) means that when both videos in a pair include this combination of features, respondents assess the one on the left as riskier with a 50-percent chance.

The estimates reported in the figure are the average marginal effects (and their SEs) from a logit regression based on the assessments made in the 13,689 cases where neither video involves partial mask wearing. The dependent variable is 1 if the video on the left is assessed riskier and 0 otherwise. The main independent variables represent feature-combination differences between the videos in a pair. To construct them, first, each single video is coded as a vector of indicator variables (0/1)—an indicator for each feature combination; second, each pairwise assessment—the unit of observation—is coded as a vector of differences (−1/0/+1) between the left- and right-video vectors. To make the graph readable, it reports only estimates from the above 16 coefficients; however, the regression has 31 such difference variables as main regressors, based on 31 0/1 indicators at the video level for the main 32 feature combinations that include the sixteen above plus the sixteen where one of the actors fully wears a mask while the other wears none (omitted category: the "safest" combination above).[##] In addition to the 31 main regressors—the nonparametric part of the specification—the regression also controls for any differences (across the videos in a pair) in duration and in these five extra features: cough, sneeze, hug, handshake, and elbow. (Hence, it is a semiparametric specification.)

The figure shows, first, that all three features matter: socializing indoors (solid lines) is assessed riskier than socializing outdoors (dashed lines); socializing with both persons maskless (hollow squares) is assessed riskier than with both fully masked (filled circles); and socializing at closer distances is increasingly riskier. The riskiest baseline combination at the top left—indoors, no masks, 2 ft apart—is 69% likelier to be assessed riskier than the safest combination at the bottom right—outdoors, full masks, 8 ft apart—suggesting that in a direct pairwise comparison, it is assessed riskier almost 85% of the time.

Second, interactions across the three features appear modest: The four lines move roughly together. Specifically, the vertical gap between the no-mask indoors and outdoors lines (the top two lines) is not dramatically larger than its counterpart gap between the two full-mask lines (the bottom two lines), meaning that the perceived effects of masks inside are not much larger than outside. Similarly, these two gaps are only somewhat larger at longer distances (6 to 8 ft) than at shorter ones (2 to 4 ft), meaning that the perceived effects of being indoors versus outdoors vary only modestly with distance. Finally, the top two (no-mask) lines are only modestly closer to the bottom two (full-mask) lines at longer than at shorter distances, suggesting that the perceived effects of masks decline only modestly with distance.

Third, the perceived effect of distance is only somewhat nonlinear: While the lines clearly slope downward more steeply at closer than at longer distances, none of the four lines is very far from linear.

This general absence of strong nonlinearities and interactions motivates the analysis in the rest of this paper. It suggests, first, that a simple linear regression model is a convenient way to organize our main findings. It also suggests that distance could serve as a convenient numeraire, allowing us to discuss the assessed-risk effects of changes in location, mask wearing, and other features in terms of the (average) change in distance, measured in feet, that would lead to an equivalent effect. We return to investigating nonlinearities and interactions below.

Average Tradeoff Estimates. Table 1 reports our main results. Each column is based on a single logit regression. Each pair of videos is an observation. The dependent variable is 1 if the left video is indicated as riskier and 0 otherwise. The independent variables are twelve 1/0/−1 variables indicating differences in features across the videos in a pair—coded 1 if a feature is present in the left but not right video, −1 if it is present in the right but not left video, and 0 if there is no difference across the videos—and two variables indicating distance differences (in feet) and duration differences (in minutes). Panel *A* reports the estimated coefficient on distance. In panel *B*, the estimated coefficients are normalized into distance equivalents (in feet) by dividing them by the distance coefficient.

The leftmost column pools all four samples and rounds. The Distance coefficient in panel A, −0.040 (SE 0.001), means that on average—over the entire set of respondents and video pairs—and holding all else constant, an additional foot of distance increases the probability that a video is chosen as riskier by 4%. This can be seen for instance in Fig. 2: The likelihood that a participant chooses a video as riskier is about 24% points higher when that video shows a two-foot distance than when it shows an eight-foot distance.

Moving to panel *B*, the Duration row, −0.69 (0.05), means that on average, respondents consider an additional ten minutes of interaction to be as risky as standing roughly 0.7 ft closer. The next row shows that being indoors rather than outdoors is on average perceived equivalent to standing 3 ft closer outside (As an indoors–outdoors falsification test, we also run a specification with two separate indicator variables for the two outdoors locations where videos were shot. We find essentially identical Indoors estimates: −2.97 (0.15) and −3.05 (0.14) ft.) Next,

[¶¶] Respondents went back to watch the previous pair of videos in 102 of those 20,280 assessment rounds and reversed their previous assessment when doing so in 53 of those 102 times.

[##] *SI Appendix*, Fig. S1 , a busier version of Fig. 2, reports all 32 feature-combination estimates. It conveys a similar picture.

Table 1. Relative-risk perceptions, by sample

	Pooled	US May	US Aug	UK Aug	Israel Sep
A. Regression coefficient					
Distance (feet)	−0.040	−0.040	−0.038	−0.042	−0.039
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
B. Ratio of coefficient to Distance coefficient					
Distance (feet; used as numeraire)	1.00	1.00	1.00	1.00	1.00
Duration (minutes×10)	−0.69	−0.71	−0.53	−0.84	−0.67
	(0.05)	(0.14)	(0.10)	(0.08)	(0.09)
Indoors	−3.02	−2.15	−1.69	−4.22	−3.39
	(0.13)	(0.31)	(0.24)	(0.25)	(0.23)
Self wears mask: fully	4.53	4.21	4.66	3.51	5.56
	(0.16)	(0.40)	(0.34)	(0.24)	(0.31)
under nose	2.92	2.53	1.89	1.61	5.28
	(0.24)	(0.64)	(0.50)	(0.40)	(0.46)
fully but removed when talking	0.15	−0.38	−0.59	−0.49	1.69
	(0.26)	(0.67)	(0.57)	(0.44)	(0.45)
Other wears mask: fully	4.27	4.95	4.20	3.80	4.44
	(0.15)	(0.42)	(0.32)	(0.25)	(0.28)
under nose	2.89	3.23	2.70	2.02	3.88
	(0.23)	(0.61)	(0.47)	(0.37)	(0.42)
fully but removed when talking	−0.33	0.47	−1.30	−0.17	0.02
	(0.27)	(0.77)	(0.62)	(0.43)	(0.47)
Other: coughs	−2.95	−2.96	−3.30	−3.18	−2.37
	(0.22)	(0.62)	(0.46)	(0.38)	(0.37)
sneezes	−4.36	−4.89	−4.37	−5.27	−3.25
	(0.23)	(0.64)	(0.48)	(0.44)	(0.38)
Greeting: hug	−7.42	−7.07	−6.16	−7.99	−8.14
	(0.27)	(0.70)	(0.52)	(0.47)	(0.51)
handshake	−5.40	−5.54	−4.57	−5.78	−5.73
	(0.25)	(0.63)	(0.50)	(0.43)	(0.46)
elbow	−1.98	−1.14	−1.39	−2.68	−2.11
	(0.21)	(0.54)	(0.44)	(0.37)	(0.38)
Observations	20,280	3,000	5,460	5,820	6,000
Respondents	676	100	182	194	200

Source: authors' online surveys during 2021 on convenience samples in the United States, United Kingdom, and Israel. Each column's estimates are from a single logit regression, based on only that column's sample. Dependent variable: response to pairwise assessment (across two videos): "Click on the video in which the risk of infection for the person in [red/blue] is [higher/lower]" (depending on treatment); coded 1 if left (right) video is clicked on in higher (lower) treatment; 0 otherwise. Independent variables: differences in features of social interaction (across the videos in a video pair); features are dichotomous (0/1), except for Distance and Duration. Self/Other: coded by [red/blue] treatment. Standard errors, clustered at the respondent level, are in parentheses.

fully wearing a mask or (three rows below it) having the other person fully wear a mask is perceived equivalent to being 4.5 and 4.3 ft farther, respectively, while wearing a mask (by self or other person) under one's nose is perceived equivalent to being only 2.9 ft farther, and pulling it down when talking is perceived roughly as risky as not wearing it at all.

The extra features at the lower rows all loom large. In increasing order: elbow bumps, the other person's coughs, the other person's sneezes, handshakes, and hugs are perceived equivalent to standing 2 to 7.4 ft closer. Representing a downside of our methodology, we suspect that these estimates may overstate naturalistically occurring perceptions: A quick hug is difficult to ignore in a repeatedly replayed 5-s clip but may hardly be remembered—perhaps rightly so—in an imagined full-length video depicting an entire 30-min interaction. Given that as a group, these features may appear disproportionately salient in our videos, we think that comparisons among themselves may be more meaningful than comparisons with other features. Thus, sneezes are perceived roughly 1.5 times riskier than coughs; hugs and handshakes are perceived roughly 3.7 and 2.7 times riskier than elbow bumps.

The four sample-specific columns are generally similar. The Distance coefficient in panel *A* is particularly stable and, given the reported standard errors in panel *B*, many of the (mostly small) cross-column differences could simply reflect sampling variation. This general similarity in perceived-risk tradeoffs across the samples is notable given that the samples differ not only in respondent populations but also in other features, including survey timing and local COVID conditions. Of the few possible exceptions, we mention two that stand out. First, being indoors rather than outdoors may be perceived equivalent to roughly twice as many feet of closeness in the UK sample as in the US samples (with the Israel sample somewhere in between). Second, wearing a mask under one's nose—apparently a common practice in Israel throughout the pandemic—is perceived there roughly as risk-reducing as wearing a mask in full; in contrast, in the US and UK samples, it is perceived as substantially less risk-reducing.

Table 2. Relative-risk perceptions, by location, duration, and distance

	Location		Duration (minutes)		Distance (feet)	
	Indoors	Outdoors	30–60	5–20	2–4	6–8
A. Regression coefficient						
Distance (feet)	−0.040	−0.045	−0.042	−0.039	−0.062	−0.027
	(0.002)	(0.001)	(0.002)	(0.001)	(0.003)	(0.005)
B. Ratio of coefficient to Distance coefficient						
Distance (feet)	1.00	1.00	1.00	1.00	1.00	1.00
Duration (minutes×10)	−0.78	−0.75	−0.53	−1.93	−0.48	−1.16
	(0.12)	(0.07)	(0.13)	(0.32)	(0.06)	(0.27)
Indoors			−2.73	−3.00	−1.76	−6.60
			(0.22)	(0.21)	(0.14)	(1.29)
Self wears mask: fully	5.46	3.85	4.00	4.86	3.32	6.90
	(0.41)	(0.21)	(0.27)	(0.26)	(0.22)	(1.35)
under nose	4.24	2.65	2.91	3.25	2.40	4.12
	(0.56)	(0.35)	(0.43)	(0.39)	(0.27)	(1.07)
only when not talking	1.12	0.03	0.06	−0.01	0.05	1.88
	(0.61)	(0.36)	(0.49)	(0.41)	(0.30)	(0.98)
Other wears mask: fully	4.97	3.82	3.88	4.48	3.11	7.23
	(0.39)	(0.21)	(0.26)	(0.25)	(0.20)	(1.42)
under nose	3.54	2.30	2.42	3.01	2.13	4.28
	(0.52)	(0.32)	(0.39)	(0.36)	(0.26)	(1.06)
only when not talking	0.39	−0.12	0.00	−0.02	−0.71	1.81
	(0.60)	(0.41)	(0.48)	(0.45)	(0.31)	(1.00)
Other: coughs	−3.18	−2.65	−2.60	−3.41	−2.18	−4.77
	(0.52)	(0.31)	(0.38)	(0.37)	(0.25)	(1.17)
sneezes	−5.21	−4.00	−4.06	−5.02	−3.16	−6.91
	(0.57)	(0.33)	(0.41)	(0.40)	(0.30)	(1.53)
Greeting: hug	−7.11	−7.28	−7.59	−7.67	−4.22	−13.42
	(0.69)	(0.37)	(0.49)	(0.44)	(0.31)	(2.60)
handshake	−5.42	−5.13	−5.10	−5.82	−2.93	−12.48
	(0.61)	(0.33)	(0.44)	(0.41)	(0.26)	(2.41)
elbow	−1.93	−2.30	−1.52	−2.23	−1.08	−6.04
	(0.49)	(0.29)	(0.38)	(0.34)	(0.22)	(1.40)
Observations	3,488	7,249	5,727	8,107	7,064	3,607
Respondents	672	676	451	483	676	676

Source: authors' online surveys during 2021 on convenience samples in the United States, United Kingdom, and Israel (pooled). Each column's estimates are from a single logit regression, based on only that column's subsample, which consists of all pairwise assessments in which both videos are within the indicated location/duration/distance range. Dependent variable: response to pairwise assessment (across two videos): "Click on the video in which the risk of infection for the person in [red/blue] is [higher/lower]" (depending on treatment); coded 1 if left (right) video is clicked on in higher (lower) treatment; 0 otherwise. Independent variables: differences in features of social interaction (across the videos in a video pair); features are dichotomous (0/1), except for Distance and Duration. Self/other: coded by [red/blue] treatment. Standard errors, clustered at the respondent level, are in parentheses.

Interactions and Nonlinearities. Table 2 reproduces the pooled column in table 1 for three data splits: by location (in- vs. outdoors), duration (30 to 60 vs. 5 to 20 min), and distance (2 to 4 vs. 6 to 8 ft). Each column is based solely on assessment rounds where both videos satisfy the condition in the column's title. For example, the first and second columns are each based on a regression that includes only rounds where both videos are shot indoors and outdoors, respectively. (Since in such rounds, the indoors indicator does not vary, the "Indoors" cells are empty.)

These two location columns do not differ much, in panel *A*, on the importance of distance, and are pretty similar, in panel *B*, on the negative effects of duration and the different greetings. However, both the protective effects of masks and the risky effects of coughs and sneezes are consistently larger indoors.

The two duration columns again do not differ much, in panel *A*, on the importance of distance. In panel *B*, while essentially everything tends to matter slightly more, in terms of distance, in shorter-than in longer-duration conversations, the most dramatic difference is in the marginal effect of duration itself. In longer conversations, a normalized duration coefficient of −0.53 (0.13) suggests that standing 1 foot closer is equivalent to talking for almost 19 additional minutes. But in shorter conversations, every minute matters much more: The coefficient almost quadruples to −1.93 (0.32), suggesting that 1 foot closer is equivalent to only slightly more than 5 min longer.

Finally, consistent with Fig. 2, the two distance columns show that the effect of distance is nonlinear too. In panel *A*, the average effect of an extra foot of distance in the 2 to 4-foot range is more than twice its effect in the 6 to 8-foot range (−0.062 vs. −0.027). As a result, while many of the other effects are similar in shorter and longer distances in absolute terms, once converted to feet, they become more than twice larger in the longer distances—not because they matter more, but because a foot matters less. That said, some effects in the 6- to 8-foot column are conspicuously larger than twice their 2 to 4-ft counterparts, including indoors, hugs, handshakes, and elbow bumps. We did not expect this difference, which could also result from our video methodology: It is possible that in a longer distance conversation, a zero-distance

greeting looms larger because it is more salient, and it takes up more time of the 5-s video.

Robustness and Heterogeneity. *SI Appendix*, Tables S2–S4 reproduce the pooled column in Table 1 for seventeen additional subsamples. The estimates appear generally similar to each other, with no particularly noteworthy exceptions. The subsamples include a) data splits by our three 50–50 between-subject randomizations (*SI Appendix*, Table S2): risk is for the person in red vs. blue, risk is higher vs. lower, and interactions are "same duration" vs. "interaction duration: … minutes" (we further split the same-duration sample into unspecified- vs. specified-same-duration subsamples (see footnote ††), and again find no systematic differences); b) only certain assessment rounds (*SI Appendix*, Table S3): excluding all rounds that include extra features or "zero-distance" features (hugs, handshakes, or elbow bumps), excluding the first four ("easy") rounds, or splitting the remaining rounds into ("early") rounds 5 to 17 and ("late") rounds 18 to 30; and c) data cuts by respondent characteristics (*SI Appendix*, Table S4): gender, political-party affiliation (in the US samples only), and vaccination status (at least one dose).

Our detailed exit survey allows for many additional splits. We hope that researchers with specific hypotheses of interest will explore them using our data, which are publicly available at https://osf.io/bvyj2/.

Discussion

As an overarching goal, we hope to contribute to an understanding of modes of health communication in general and, in the future, to investigate some worries we have about such communication in other important domains. In particular, we are intrigued by the dearth of communication about tradeoffs which, as economists, we believe is crucial for helping people make sound choices. We do not know what the experts think. To the extent that experts believe that the perceived risk tradeoffs we see in our data are wrong and to the extent that correcting them would actually improve outcomes—an admittedly big behavioral assumption—our results may indicate a colossal health-risk communication failure that could be costing many lives all around the world.

Our pairwise-assessment methodology—a video adaptation of pairwise-choice methods routinely used by empirical economists—could be applied to other health-risk domains. Potential domains include weight loss, cardiac health, and dental health. Like the COVID domain, in these domains, we are also asked by the experts to change our behavior in order to reduce health risk. And in these domains too, we cannot always—or simply do not always want to—improve particular behaviors in isolation but would rather like to know which of two bundles of behaviors is better for us. How should we trade off low-carb food items versus low-fat ones vs. low-calorie ones versus exercising more? What is better for our hearts: using the stairs at home four times a week or keeping using the elevator but spending time at the gym on weekends? And what is better for our teeth: replacing one soda a week with a glass of fruit juice or brushing 10 times a week rather than 8?

By investigating perceptions regarding such tradeoffs, we hope that future research, and the findings in this paper, will help focus health-risk communications on relative risks—potentially helping people make healthier choices.

Data, Materials, and Software Availability. Anonymized survey data have been deposited at https://osf.io/bvyj2/.

ACKNOWLEDGMENTS. We thank Amitabh Chandra, Katrina Ligett, and Neel Mukherjee for helpful conversations, and participants in the Cornell Behavioral Economics Reading Group (BERG), Jerusalem Economics, Society, and Culture (JESC) lab, fifth Coller Conference on Behavioral Economics (CCBE), 30th Annual Retreat of the Center for Rationality, and Princeton Center for Health and Wellbeing (CHW) Lunch Seminar for comments that greatly improved the paper. Shira Biblarsh, Ronny Gelman, Aharon Haver, Mattar Klein, Jonathan Peled, and especially Itamar Bellaiche and Tamar Yerushalmi provided excellent research assistance. Heffetz thanks NIH/NIA grant R01-AG065364 to Hebrew University; for financial support and hospitality while conducting research on this article, Heffetz thanks the School for Public and International Affairs (SPIA) at Princeton University and Rabin thanks the Centre for Experimental Research on Fairness, Inequality and Rationality (FAIR) at the Norwegian School of Economics (NHH).

Author affiliations: [a]Bogen Family Department of Economics, Federmann Center for the Study of Rationality, The Hebrew University of Jerusalem, Jerusalem 9190401, Israel; [b]S.C. Johnson Graduate School of Management, Cornell University, Ithaca, NY 14853; [c]National Bureau of Economic Research, Cambridge, MA 02138; [d]Department of Economics, Harvard University, Cambridge, MA 02138; and [e]Business School, Harvard University, Cambridge, MA 02138

1. O. Svenson, S. Appelbom, M. Mayorga, T. L. Öjmyr, Without a mask: Judgments of corona virus exposure as a function of inter personal distance. *Judgm. Decis. Making* **15**, 881–888 (2020).
2. O. Svenson, Perceived corona virus exposure as a function of interpersonal distance and time of a conversation. *Discover Soc. Sci. Health* **2**, 1–8 (2022).
3. A. Luckman *et al.*, Risk compensation during COVID-19: The impact of face mask usage on social distancing. *J. Exp. Psychol.: Appl.* **27**, 722–738 (2021).
4. D. L. Chen, M. Schonger, C. Wickens, oTree–An open-source platform for laboratory, online, and field experiments. *J. Behav. Exp. Finance* **9**, 88–97 (2016).