Journal of Experimental Psychology: Applied

# Is It Riskier to Meet 100 People Outdoors or 14 People Indoors? Comparing Public and Expert Perceptions of COVID-19 Risk

Shane Timmons, Cameron A. Belton, Deirdre A. Robertson, Martina Barjaková, Ciarán Lavin, Hannah Julienne, and Peter D. Lunn

Online First Publication, February 24, 2022. http://dx.doi.org/10.1037/xap0000399

Journal of Experimental Psychology: Applied


ISSN: 1076-898X
https://doi.org/10.1037/xap0000399

# Is It Riskier to Meet 100 People Outdoors or 14 People Indoors? Comparing Public and Expert Perceptions of COVID-19 Risk

Shane Timmons[1, 2], Cameron A. Belton[1], Deirdre A. Robertson[1, 2], Martina Barjaková[1], Ciarán Lavin[1], Hannah Julienne[1], and Peter D. Lunn[1, 3]
[1] Social Research Division, Economic and Social Research Institute
[2] School of Psychology, Trinity College Dublin
[3] Department of Economics, Trinity College Dublin

People have limited capacity to process and integrate multiple sources of information, so how do they integrate multiple contextual risk factors for Coronavirus disease (COVID-19) infection? In June 2020, we elicited risk perceptions from a nationally representative sample of the public ($N = 800$) using three psychologically-distinct tasks. Responses were compared to a sample of medical experts who completed the same tasks. Relative to experts, the public perceived lower risk associated with environmental factors (such as whether a gathering takes place indoors or outdoors) and were less inclined to treat risk factors as multiplicative. Our results are consistent with a heuristic simply to "avoid people" and with a coarse (e.g., "safe or unsafe") classification of social settings. A further task, completed only by the general public sample, generated novel evidence that when infection risk competes with risk in another domain (e.g., a different medical risk), people perceive a lower likelihood of contracting the virus. These results inform the policy response to the pandemic and have implications for understanding differences between expert and lay perception of risk.

***Public Significance Statement***

This study shows that, in Summer 2020, medical experts placed greater weight on environmental factors (such as being indoors or outdoors) than the general public when judging the risk of contracting COVID-19. The study also shows that the public perceived the risk of contracting COVID-19 to be lower when potential exposure to the virus was needed to avoid an alternative risk (e.g., using busy public transport to attend an urgent medical appointment).

*Keywords:* risk perception, COVID-19, affect heuristic, heuristics, policy making

*Supplemental materials:* https://doi.org/10.1037/xap0000399.supp



Throughout the Coronavirus disease (COVID-19) pandemic, humankind's main defense has been our behavior. When deciding what behaviors we are happy to undertake, we rely on our perception of risk (Brewer et al., 2007; Fischhoff et al., 1993; Slovic, 1987). Thus, the spread of infection partly depends on how accurately we can integrate multiple risk factors into everyday decisions. The accuracy of this integration process is the present focus.

Perceptions of the overall risk of COVID-19 infection affect compliance with public health advice (e.g., Dryhurst et al., 2020; Lohiniva et al., 2020; Wise et al., 2020). However, risk of infection varies by context and, therefore, it matters not only how much risk people perceive overall, but how they differentiate between high- and low-risk situations. For instance, Sarah might decide to attend a birthday dinner if it takes place outdoors on a restaurant terrace, but not if it takes place in her friend's dining room, even if more people would be at the restaurant dinner. Stephen might share the same overall perceived likelihood of infection as Sarah, but worry more about the number of people present than the location, and so decide the opposite. Moreover, multiple other aspects of a given social setting affect transmission risk, including the duration of the encounter and

Shane Timmons https://orcid.org/0000-0002-0200-5927
Deirdre A. Robertson https://orcid.org/0000-0002-2829-1722
Ciarán Lavin https://orcid.org/0000-0002-0184-0402
Hannah Julienne https://orcid.org/0000-0002-9601-5658
Peter D. Lunn https://orcid.org/0000-0002-7174-5320

Shane Timmons developed the study concept. All authors contributed to the study design. Shane Timmons, Cameron A. Belton, and Deirdre A. Robertson developed the materials. Cameron A. Belton programmed the study. Martina Barjaková and Ciarán Lavin coded the qualitative data, with supervision from Shane Timmons. Shane Timmons and Peter D. Lunn analyzed and interpreted the data. Shane Timmons drafted the manuscript and Peter D. Lunn provided critical revisions. All authors approved the final version of the manuscript for submission.

We thank the Behavioural Change Subgroup of the National Public Health Emergency Team (NPHET) for support, and various of its members for useful guidance and comments. We are also grateful to Helen Russell and Kieran Mohr for helpful comments on an initial draft of this manuscript.

Correspondence concerning this article should be addressed to Shane Timmons, Social Research Division, Economic and Social Research Institute, Whitaker Square, Sir John Rogerson's Quay, Dublin D02 K138, Ireland. Email: shane.timmons@esri.ie

mitigation behaviors (e.g., maintaining social distance, wearing a mask; Qian et al., 2020; Setti et al., 2020; Van Doremalen et al., 2020). All these factors must be integrated to assess risk accurately.

Decades of psychological research demonstrate that humans have limited ability to integrate multiple factors into judgments of absolute quantities, as is required to assess risk objectively and hence to make accurate decisions about infection risk. Comparison of just one cue to an internal scale is coarse (Miller, 1956). As the number of cues to be integrated increases, judgment performance against objective criteria declines (Karelaia & Hogarth, 2008), especially where multiple cues trade-off against each other (Lunn, Bohacek, et al., 2020). Relative weightings of cues are subject to multiple biases (Weber & Borcherding, 1993), for instance where judgments are driven by more salient cues, in contexts ranging from the social (Schkade & Kahneman, 1998) to the perceptual (Hunt et al., 2014). When thinking about COVID-19, multiple possible heuristics may be important, given the novelty of the situation, complexity of risk factors, and inherent uncertainty (Lipshitz & Strauss, 1997; Tversky & Kahneman, 1974).

The above (and many other) empirical results have led recent theories of judgment and decision-making to abandon the idea that humans even attempt to integrate all available cues into an overall assessment of the absolute value or risk of a given option in a decision (Vlaev et al., 2011). Rather, decisions may be based on the ease of generating reasons (Shafir et al., 1993), simplifying heuristics (Gigerenzer & Todd, 1999), mental models of sets of possibilities (Johnson-Laird & Byrne, 2002), the order of questions asked of memory (Johnson et al., 2007), gist rather than verbatim mental representations (Reyna, 2008), or relative comparisons to limited samples held in memory (Stewart & Simpson, 2008). The present study was not designed to test explicit hypotheses derived from these theories, but was motivated by them, because they imply that the complexity of risk judgments faced by members of the public during the COVID-19 pandemic is likely to lead to systematic misperceptions.

Our primary aim was hence exploratory: To measure variation in risk perceptions across social contexts. Ignoring, underweighting, or overweighting relevant factors implies systematic misperception of infection risk, with consequences for spread of the disease and implications for improving public communications. A further factor, which does not feature strongly in the theories cited above, is the possibility that individuals might fail to account for synergistic relationships between risk factors, as has been observed in other domains (Dawson et al., 2013)—a point we expand on below.

The contribution of the study is threefold. First, by measuring risk perceptions and illuminating potential underlying mechanisms, we provide empirical evidence to support efforts to reduce transmission. Second, the range of experimental tasks we devised demonstrates how techniques of psychological science can be used in an important applied setting—to inform interventions during a global pandemic. Third, our findings are of interest beyond the response to COVID-19. The pandemic offers a highly unusual opportunity to measure how well the public can absorb complex risk information. All citizens have been affected and the attention paid to the relevant public information is unprecedented in the modern media age.

The lack of veridical benchmarks for infection risk presents an obvious challenge. We chose to compare risk assessments generated by a representative sample of the public to those of a sample of medical professionals with expertise in public health, microbiology, and virology. This approach is imperfect. It assumes that aggregated judgments of an expert sample reflect "ground truth" better than those of the general public. We contend that this is highly likely. These professionals have access to better information about how the virus spreads (e.g., academic papers, preprints, researcher networks and newsletters, professional websites and blogs), training and background knowledge to understand and contextualize the information, a track-record that implies motivation to engage with it, and a contemporaneous incentive to pay attention to the issue given the amount of public attention that the pandemic brought to their area of expertise. This is not to imply that expert risk judgment is veridical, since there is ample evidence to the contrary (Adam & Reyna, 2005; Koehler et al., 2002; Wright et al., 2002). Nevertheless, expert judgments of risks differ systematically from lay judgments (e.g., Barke & Jenkins-Smith, 1993; Kraus et al., 1992), deploying a greater range of values to discriminate between risks, being less influenced by emotional responses such as dread, and matching objective estimates more closely (Slovic et al., 1979, 1985). Medical experts tend not to rely on simplified heuristics when evaluating risk in an expertise-relevant context (Fleming et al., 2012). Such studies comparing expert and lay risk perception are not uncontroversial, however. Rowe and Wright (2001) argue that expertise is confounded with demographic background characteristics, casting doubt on the inference that expertise itself leads to risk perceptions that are consistent with objective measures of risk. For present purposes, however, demonstrating causality is not our concern. Our contention is that it is desirable for policymakers and others seeking to reduce infection to know whether and how public risk perceptions depart systematically from expert perceptions, which are likely (albeit not certain) to be better informed and hence more accurate. Furthermore, in a rapidly evolving pandemic, policymakers do not have time to wait for definitive evidence about transmission to emerge; experts are all we have.

In addition to obtaining comparison benchmarks, a further challenge is to measure risk perceptions that are likely to generalize to real-world settings. We deployed three psychologically-distinct tasks: (a) an open-ended question to determine the cognitive availability of different risk factors; (b) a quantitative rating task to measure how factors are integrated; and (c) a ranking task to identify how factors are prioritized. Our logic was that where consistencies are observed across these different tasks, the cognitive mechanisms involved are likely to be relied on in everyday contexts. Lastly, since real-world situations sometimes require balancing COVID-19 risks against other risks (e.g., financial or social risks), a final task used vignettes to test whether the presence of such alternative risks affects the perceived risk of infection. In the next section, we detail each of these tasks after presenting details on the participants and the context in which the study was conducted.

## Method

The experiment proceeded over multiple stages and was programmed in Gorilla Experiment Builder (Anwyl-Irvine et al., 2020). Participants in the "Public" sample completed four stages: The Open-Ended Question followed by Risk Ratings, Risk Rankings, and Risk Vignettes tasks. The tasks were presented to participants sequentially before finishing with a section on background characteristics and an unrelated experiment to measure bias in survey estimates of compliance (reported in Timmons, McGinnity, et al., 2020). The "Expert" sample completed only the first three stages, with some small modifications. We report the design and results for each stage separately. We report how we determined our sample size, all data exclusions, all manipulations, and all measures. The preregistration, data and analysis code, and materials are available at https://osf.io/ptv2y/. The study was conducted in line with institutional ethics policy.

## Timing and Context

Data were collected in mid-June 2020 in Ireland, approximately 1 month after restrictions from a 6-week lockdown were lifted. Cases were at their lowest since the onset of the pandemic, having fallen sharply to a 7-day average of nine new infections per day (in a population of approximately 5 million). The public were encouraged to "stay local," social visits from different households had been recently permitted (indoors or outdoors) and nonessential shops and shopping centers had reopened in the preceding days. Restaurants, pubs serving food, hairdressers, and gyms did not reopen until the end of June. It was recommended (although not yet mandatory) to wear masks on public transport, which had begun operating at 20% capacity, and in other public places. Masks became mandatory on public transport in mid-July and in shops and shopping centers in early August. By September, approximately 90% of adults reported wearing masks in public places, with close to full compliance in shops and on public transport (Amárach Public Opinion Survey, 2021). Throughout the pandemic, Ireland's restrictions have been as or more stringent than most other European countries (Hale et al., 2021). The public broadly supported these restrictions; a majority consistently judged the Government reaction to be appropriate or insufficient throughout the pandemic (Amárach Public Opinion Survey, 2021).

## Participants

The Public sample consisted of 800 adults (421 men, 376 women, 3 other, aged 18–86 years) recruited from a market research agency's online panel to take part based on a sociodemographic quota. Timmons, Barjaková, et al. (2020) provide details on how recruitment from this panel compares to a probability sample. Table 1 shows the sociodemographics of the sample. The sample is well matched to Census figures, although there was a slight underrepresentation from the youngest age group. The results we report in this manuscript are robust to the inclusion of sociodemographic controls, including age. Participants were paid €6 for undertaking the 20-min online study. To determine the sample size, we identified the Risk Vignettes as the task that would require the greatest number of participants to be sufficiently powered. Each vignette in the task required just one response per participant, whereas we elicited multiple responses per participant for the risk ratings and the aims of the open text and ranking tasks were primarily descriptive. There were three versions of each vignette and the sample size was set to ensure a minimum of 250 responses per version.

The "Expert" sample consisted of 56 professionals with medical expertise in an area relevant for assessing risk of COVID-19 infection: Infectious diseases, clinical microbiology, virology, and public health. We recruited as many relevant experts as possible over the timeframe of the experiment, from the Expert Advisory Group of NPHET, the Irish Society of Clinical Microbiologists, the Infectious Diseases Society of Ireland, and senior research staff in seven specialist university labs in infectious disease, virology, and immunology. Our expert group was therefore more tightly defined than some previous studies of expert versus lay risk (e.g., in Slovic et al., 1985, a lawyer, an economist, and a geographer were classified as similarly "expert" in evaluating the risk of death from multiple activities). Previous research suggests that tasks used to measure risk perception need to match day-to-day tasks of experts (e.g., Rowe & Wright, 2001). Many of the Expert sample were indeed required to make judgments about risk factors associated with COVID-19 transmission in social settings, in order to inform the policy response to the pandemic. Moreover, the pandemic had required everyone—experts and the lay public—to think about and evaluate their own risk of contracting COVID-19 during the months prior to data collection. Experts, however, have more domain-relevant knowledge to inform these judgments. The majority of the Expert sample was senior medical professionals: Two had 10 years' or less experience, 20 had 11–20 years', 16 had 21–30 years', and 18 had over 30 years'. To ensure anonymity, we did not collect sociodemographic details other than area of expertise and years of experience. They completed the study voluntarily.

**Table 1**
*Public Sample Sociodemographics*

| Characteristic | Subgroup | *n* | % | Census (%) |
|---|---|---|---|---|
| Gender | Men | 306 | 52.6 | 49.6 |
| | Women | 267 | 47.0 | 50.4 |
| | Other/prefer not to say | 3 | 0.4 | |
| Age | 18–39 years | 250 | 31.3 | 38.3 |
| | 40–59 years | 329 | 41.1 | 36.3 |
| | 60+ years | 221 | 27.6 | 25.4 |
| Education | Degree or above | 371 | 43.3 | 42.0 |
| | Below degree | 485 | 56.7 | 58.0 |
| Employment | In labor force | 517 | 64.6 | 62.3 |
| | (of which, employed) | 455 | 88.0 | (83.3)[a] |
| | (of which, unemployed) | 62 | 12.0 | (16.7)[a] |
| | Not in labor force | 283 | 35.4 | 37.7 |
| Living area | Urban | 504 | 58.9 | 60.8 |
| | Rural | 352 | 41.1 | 39.1 |

[a] Our estimate of unemployment (12%) falls between the Census standard unemployment estimate (5.5%) and the COVID-adjusted estimate (16.7%) at the time of the study. This is possibly because our measurement of Employment is based on a self-report question and so some respondents may have been in receipt of the Pandemic Unemployment Payment or Temporary Wage Subsidy Scheme and selected "unemployed," whereas others may have selected "employed."

## Stage 1: Open-Ended Question

Participants were first asked to write three things they think about when deciding whether an activity might be risky or safe, considering the possibility of contracting the virus. The instructions specified that we were interested only in the risk of becoming infected and not in how bad it might be to contract the virus or to pass it on to someone else. Reponses to open-ended questions provide important information about participant attitudes, beliefs, and knowledge without imposition from researchers (Geer, 1988). They are often avoided due to resource constraints, as responses are relatively difficult to score and analyze (Reja et al., 2003). We nevertheless elicited perceived risk factors before participants were presented with any cues. By recording both the factors people listed and the order in which they listed them, we assessed the cognitive availability of different factors (Folkes, 1988; Schwarz et al., 1991). The approach is supported by "query theory"; factors reported first and most often are likely to be more heavily relied on when evaluating risk (e.g., Weber et al., 2007). Hence, our first research questions were:

> *Research Question 1a*: What risk factors for COVID-19 infection in social settings are most cognitively available to the public?

> *Research Question 1b*: Do the public and experts differ in the factors that are most cognitively available to them, when thinking about the risk of contracting COVID-19 in social settings?

### Results

Responses ($n$ = 2,568) were coded independently by two of the authors (M. Barjaková and C. Lavin), using a framework with 22 possible categories that was developed from a pilot study ($N$ = 40) and preregistered. Agreement on the full 22-category coding structure was "substantial" according to Landis and Koch (1977) criteria (81.7% agreement; $\kappa = .80$, $p < .001$). We extracted five broader categories: The number of people, location (i.e., indoors or outdoors, or whether the area is well-ventilated), duration, social distancing, and mask-wearing. We also extracted an additional category for references to hand hygiene as a sixth factor, given its coverage in public health advice. Agreement for these six categories was "almost perfect" (96.5%–99.7%; all $\kappa$s > 0.81, all $p$s < .001). Disagreements were solved through discussion, with input from a third author (S. Timmons) in two cases.

Some participants (13.9%) did not write any factors related to the risk of *contracting* COVID-19—for example, instead referring to the risk they would subsequently pass it to a family member—and were removed from the analyses. Frequencies are shown in Figure 1. Over half of the Public sample wrote about the number of other people and whether social distancing could be maintained, and almost one-third mentioned whether the activity took place indoors or outdoors. A similar pattern with higher proportions is observed for the Expert sample, although substantially more referenced location. Tests of proportions using a Bonferroni-corrected $\alpha$ of .008 for six comparisons show that, compared to the Public sample, Experts were more likely to mention location (30.9% vs. 63%; $z = 4.90$, $p < .001$) and duration (6.7% vs. 22.2%; $z = 4.16$, $p < .001$) as factors they consider when evaluating their risk of contracting COVID-19. Effect sizes are large: Experts were more than twice as likely to mention location and more than three times as likely to mention the amount of time spent in one place. No other comparisons were statistically significant (number of people: $z = 1.32$, $p = .188$; distancing: $z = 1.15$, $p = .249$; masks: $z = 0.85$, $p = .394$; hand hygiene: $z = 0.85$, $p = .397$). The equivalent analysis only for factors mentioned first shows that Experts were also almost three times more likely to mention location first (8.6% vs. 24.4%; $z = 3.82$, $p < .001$), although there was no difference for duration (0% of both groups).

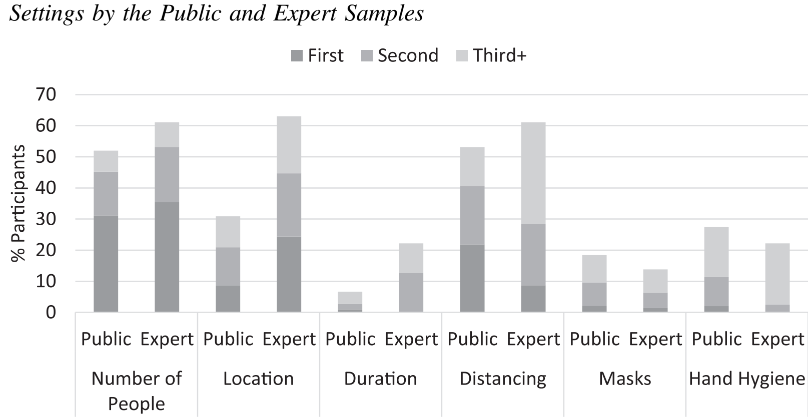
## Stage 2: Risk Ratings

The second task sought to determine the weight people give to specific risk factors when multiple factors are present and must be integrated to form a perception of risk. As described above, theory and empirical evidence suggest limitations in the ability to perform such integration. In addition, risk factors may be multiplicative. Few people are able to process anything beyond second-order interactions accurately (Halford et al., 2005) and when dealing with "synergistic risks" (i.e., risk factors that interact) people underestimate the risk arising from their combination, particularly if risks are unfamiliar (Dawson et al., 2013). In the present context, for example, people may underestimate the additional risk of meeting multiple people indoors, relative to meeting just a few people indoors or multiple people outdoors.

We presented participants with short descriptions of social situations that varied according to risk factors prominent in public health advice. We refer to these as "scenarios." The task was to rate each scenario for risk. In the controlled presentation of multiple factors and exposure to multiple scenarios, the task was similar to a conjoint experiment (Green & Srinivasan, 1978; Hainmueller et al., 2015). In a conjoint experiment, participants give a single evaluation of a stimulus that has multiple attributes. By varying the level of these attributes (e.g., the number of people present) and eliciting multiple responses, the relative influence of each attribute on evaluations can be determined. For example, if, over multiple judgments, variation in evaluations correlates strongly with the number of people present but less so with the location of the scenario, statistical models will show that the number of people has a stronger effect on overall

**Figure 1**

*Responses to the Open-Ended Question About COVID-19 Risk Factors in Social Settings by the Public and Expert Samples*



evaluations than the location. Regressing the level of the attribute onto the overall evaluation reveals the weighting assigned to each of the varied factors, even though responses are elicited on a single scale. Conjoint techniques are used in multiple domains, including health, to identify influence of different factors on judgments while avoiding self-report and social desirability biases (Horiuchi et al., 2020; Ryan & Farrar, 2000). This design allowed us to assess how people assigned relative weightings to different COVID-19 risk factors and how they processed interactions. Hence, our second set of research questions was:

> *Research Question 2a*: How do people weight specific risk factors for COVID-19 infection when multiple factors must be integrated?
>
> *Research Question 2b*: Are there differences between how the public and experts weight risk factors for COVID-19?
>
> *Research Question 2c*: Are there differences between how the public and experts process interactions between risk factors for COVID-19?

The scenarios were defined by four factors: How many people were present, whether it took place indoors or outdoors, how long it lasted and whether maintaining 2m distance from others was easy or difficult. These were chosen because it is well-established that they are factors that influence the spread of the virus and they had been covered widely in public health communications prior to the study (e.g., European Centre for Disease Prevention and Control, 2020; Health Information and Quality Authority, 2020). Participants' task was to rate the riskiness of each scenario on a scale from "Not At All Risky" to "Extremely Risky" (adapted from the domain-specific risk-taking [DOSPERT] Scale; Blais & Weber, 2006). The scale was un-numbered but contained 51 (0–50) possible responses.[1] We opted for a qualitative rather than quantitative risk perception scale for multiple reasons: (a) objective risk probabilities in different social settings are not yet known; (b) participants could hold similar estimates for the probability of infection while differing in how risky they judge those probabilities (e.g., two participants might agree that the probability of infection from an activity would be 1% but one could consider 1% to be high risk while the other judges it to be low risk); and (c) difficulties with numerical risk perception (e.g., probability neglect) are well-documented (e.g., Reyna & Brainerd, 2008). Because we do not use a numerical scale, the question we ask participants is to consider their *possibility* of infection, rather than the probability or likelihood, as "possibilities" are more appropriate when uncertainty is greater and numeric probabilities cannot be estimated (e.g., Delgado & Moral, 1987; Dubois et al., 1993). Moreover, many studies that question the validity of expert risk perceptions rely on exact probability forecasts, rather than qualitative estimates of risk from different factors (Rowe & Wright, 2001). We also elicited perceived risk using a first-person question (i.e., the possibility that the participant themselves could contract the virus). This perception is more strongly linked to behavior than perceived risk for others. In addition, our findings would be communicated to policymakers interested in how perceived risk of infection varied between socio-demographic groups. Note that, although the responses given may then be more optimistic than if participants were to judge risk for others (Weinstein, 1982; Wise et al., 2020), our focus here is the *relative* influence of different risk factors.

Each participant responded to 14 scenarios selected from a larger set of 24, which were constructed by orthogonally manipulating the above four factors based on the following levels: Number of other people (5, 14, 100); location (outdoors, indoors); duration (15–30 min, 2–4 hr); and distancing (easy, difficult). The levels were informed by policy decisions, although there are further possible nuances to each (e.g., indoors could be divided into indoor situations that are well-ventilated versus not; Liu et al., 2020). We constrained the selection of the 14 trials for each participant such that at least two were high-risk on three or more factors (i.e., 100 people, indoors, 2–4 hr, difficult to distance) and at least two were low-risk on three or more factors (i.e., five people, outdoors, 15–30 min, easy to distance). Other scenarios were selected at random and the order was randomized.

Scenarios were presented to participants with four per page. The first page included two further scenarios as controls (Figure 2). One described a scenario with an extremely high possibility of infection (close contact with a confirmed case for a prolonged period of time and no access to "peronal protective equipment"; PPE) and the other described a scenario with an extremely low possibility of infection (a video call). These scenarios were presented on the first page to calibrate participants to the levels of risk that would likely fall at either end of the response scale. They also served as comprehension/attention checks.

After completing four pages (14 trial scenarios, 2 controls), participants were presented with an additional four scenarios. These final four scenarios incorporated a fifth factor of interest: Mask-wearing (Chu et al., 2020). We tested for it separately because the recommendation to wear masks came much later than other public health advice and was less consistent. At the time of the study, masks were advised as a voluntary precaution on public transport or inside shops (before subsequently becoming mandatory). These final four scenarios were ones participants had rated previously, with information on mask-wearing added. For this additional factor, we varied between-participants whether only they wore a mask in the scenario, or whether everyone did. The manipulation was designed to check whether the public had absorbed the message that masks primarily protect others rather than the wearer.

Up to this point, the Experts completed the same task as the Public sample. However, the Experts were shown the mask scenarios twice, once when only they wore a mask and once when everyone wore a mask, with the order randomized. This was done to increase the accuracy of the benchmark estimates, given the smaller Expert sample.

To familiarize participants with assessing scenarios with multi-dimensional risks, before completing the above task, participants undertook practice trials involving everyday risk. For instance, one practice scenario involved not wearing a seatbelt while in a car, with information shown regarding the speed and journey duration. Other practice scenarios involved physical activity and gambling.

After completing all scenarios, participants in the Public sample were then asked to rate on the same risk scale the highest level of risk of contracting COVID-19 they judged to be acceptable to take and to rate their confidence in their ability to judge such risk on a Likert scale from 1 (*Not at all confident*) to 7 (*Extremely confident*).

[1] The reason for choosing 51 possible responses was entirely practical: We wanted to allow sufficient granularity in responses to detect differences within the spatial constraints of the software used to present the scales on screen.

**Figure 2**
*Example First Page on the Risk Ratings Task*

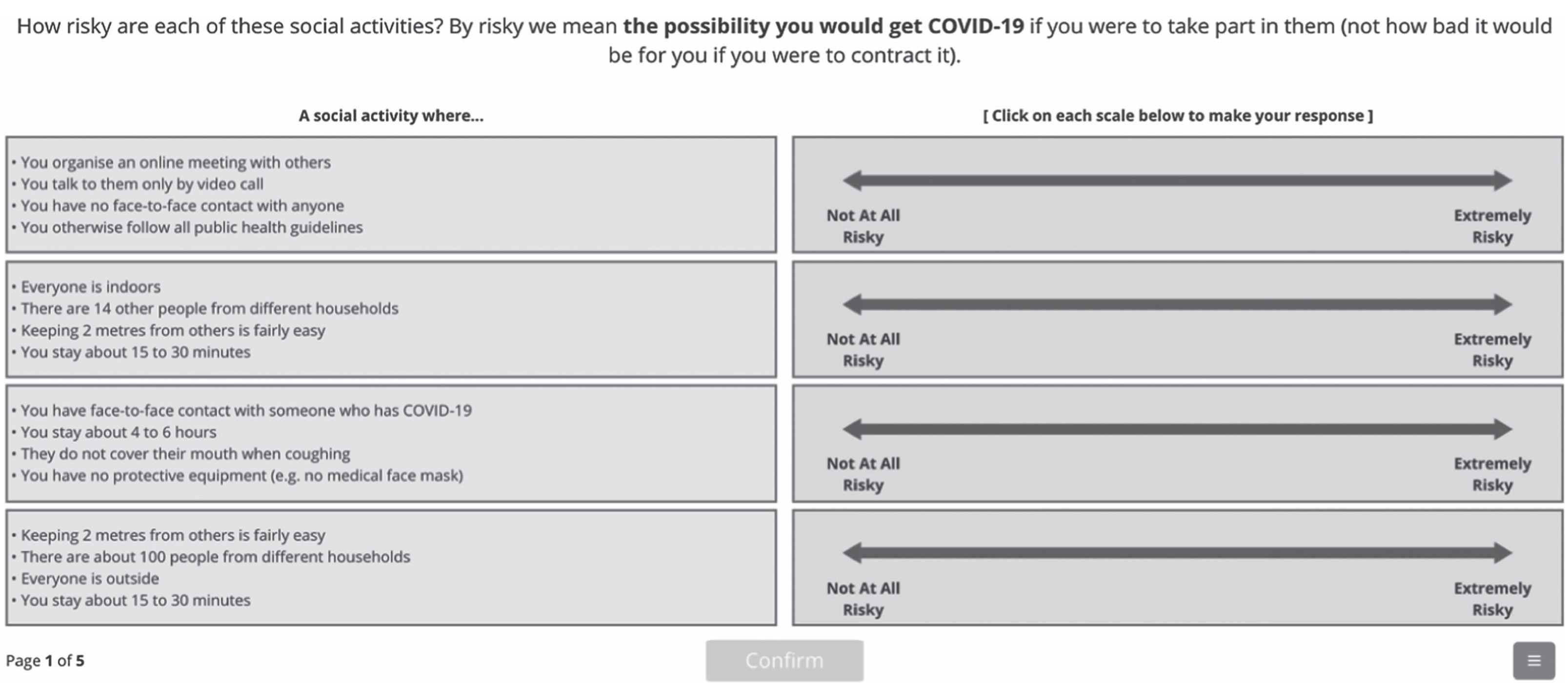
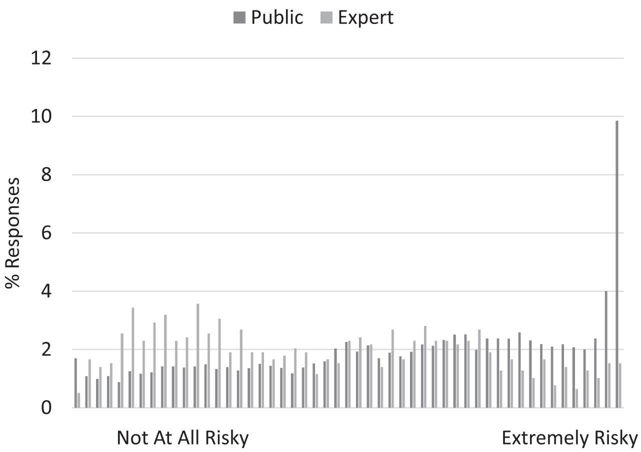


**Figure 3**
*Distributions of Responses to the 14 Trials by the Public and Experts*




### *Results*

Fifty-eight participants were removed from the Public sample following procedures outlined in the preregistration (mis-rating control activities, responding in the fastest 5% on every page, not varying their responses). Findings are not sensitive to these decisions.

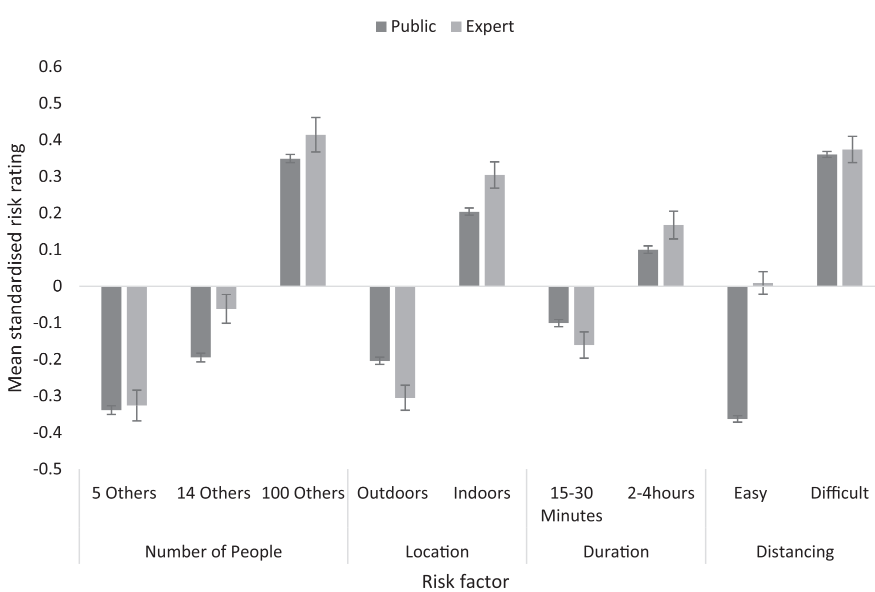
Figure 3 presents the distribution of responses to the 14 scenarios (i.e., excluding controls and the mask scenarios). Both the Public and Expert samples used the full length of the scale, although the Public had a greater tendency to give maximum responses. Taking each participant's average risk rating, Experts perceived risk to be lower on average, $M_{\text{Public}} = 30.39$, $SD = 9.64$; $M_{\text{Expert}} = 22.86$, $SD = 8.55$; $t(796) = 5.68$, $p < .001$; $d = 0.83$.

Individual standard deviations varied between 9.05 and 24.89 with a skewed distribution ($M = 14.30$, $SD = 2.58$). To avoid participants with larger standard deviations having undue influence on models, we standardize responses at the participant level and use OLS regression clustered standard errors with the four risk factors as predictors rather than using participant random effects. Results are closely similar with both approaches.

Figure 4 plots the mean standardized risk rating for each level of the risk factors, comparing the Public and Experts. Model 1, $F(5, 741) = 995.82$, $p < .001$, in Table 2 shows the main effects of the four risk factors for the Public sample. Judgments were sensitive to the levels of each of the four factors, with scenarios where there were 100 other people (compared to 5) and scenarios where it was difficult to keep 2 m from others (compared to easy) showing the largest and equivalent effect sizes. There is no change in coefficients when sociodemographic controls for gender, age, educational attainment, living area (urban/rural), and employment status are added. Model 2, $F(5, 55) = 130.68$, $p < .001$, shows a similar pattern for the Expert sample, however the Expert weighting of location (i.e., whether the scenario was described as taking place indoors or outdoors) was just as large as their risk judgments for meeting 100 others and meeting where it is difficult to socially distance.

Comparing Public weightings to Expert ones, the 95% confidence intervals show that Experts gave greater weighting to location and duration. For readers interested in $p$ values, we test for differences using $Z$-tests of the coefficients (Clogg et al., 1995). The final column in Table 2 shows that, compared to the Expert sample, the Public underweighted location ($p < .001$) and duration ($p = .010$), while there was no evidence for a difference on coefficients for number of people ($p_{14} = .482$, $p_{100} = .525$) or distancing ($p = .543$).

**Figure 4**
*Mean Standardized Risk Rating Assigned to Each Factor by the Public and Experts*




*Note*. Error bars are the standard error of the mean. For the Public, the average number of observations per factor is 4,617. For the Expert, it is 348.

**Table 2**
*Regression Models Predicting (Standardized) Risk Ratings by the Public and Experts*

| Risk factor | Model 1 (public) | Model 2 (expert) | Public versus Expert *Z*-test |
|---|---|---|---|
| Number of people: (*Ref*: 5 others) | | | |
| 14 others | 0.30*** | 0.27*** | −0.71 |
| | [0.28, 0.33] | [0.19, 0.35] | |
| 100 others | 0.66*** | 0.71*** | 0.80 |
| | [0.63, 0.70] | [0.60, 0.82] | |
| Location: Indoors (*Ref:* Outdoors) | 0.48*** | 0.68*** | 3.93*** |
| | [0.45, 0.51] | [0.58, 0.77] | |
| Duration: 2–4 hr (*Ref*: 15–30 min) | 0.16*** | 0.26*** | 2.59** |
| | [0.14, 0.18] | [0.19, 0.34] | |
| Distancing: Difficult (*Ref*: Easy) | 0.70*** | 0.67*** | −0.61 |
| | [0.67, 0.73] | [0.57, 0.77] | |
| Intercept | −1.00*** | −1.10*** | |
| | [−1.03, −0.97] | [−1.19, −1.01] | |
| Sociodemographic controls | No | N/A | |
| Observations | 10,388 | 784 | |
| $N$ | 742 | 56 | |
| $R^2$ | .48 | .49 | |

*Note.* Brackets contain 95% confidence intervals.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.



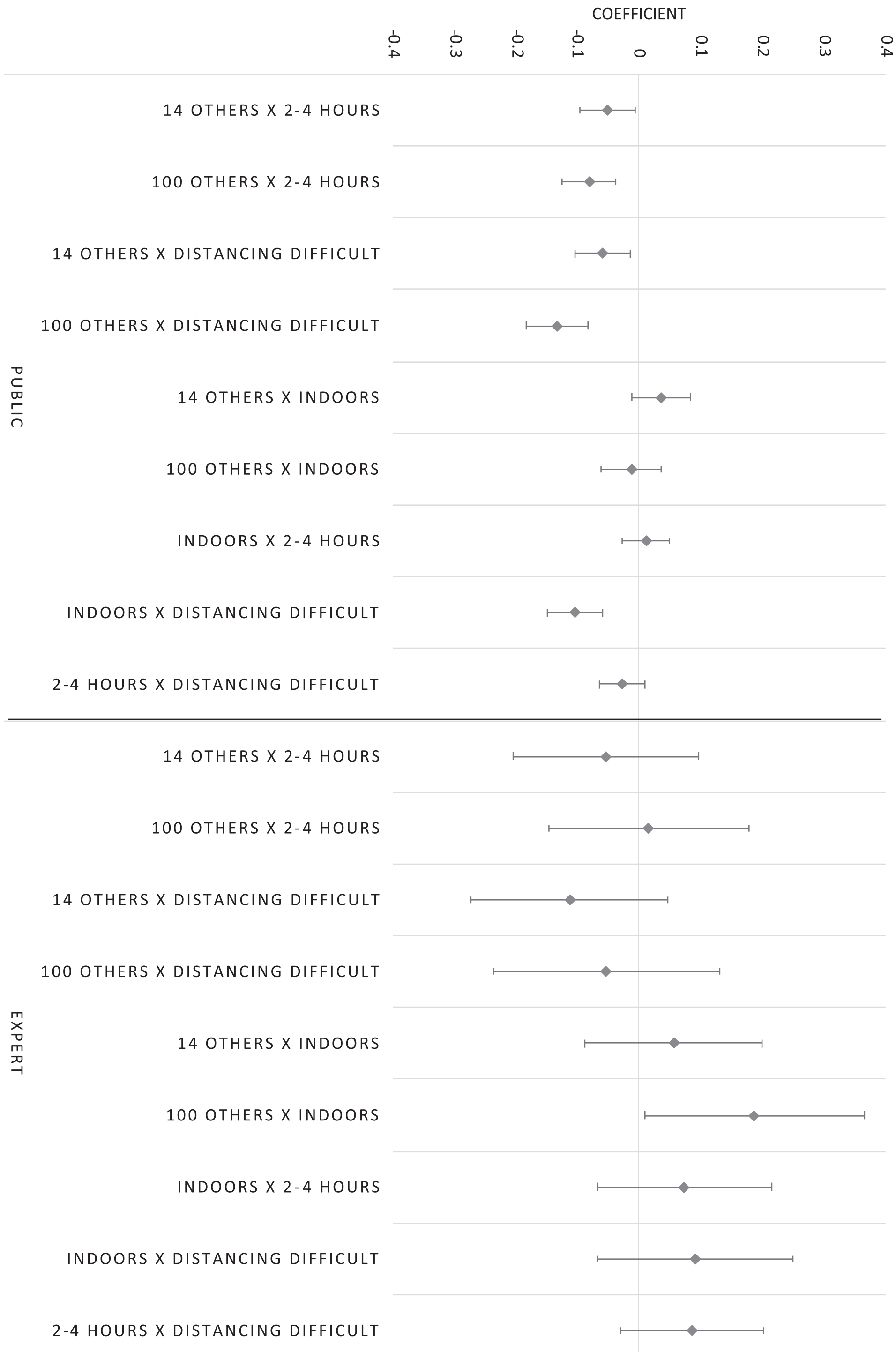
Multiple potential interactions between factors could be investigated. Since we had no confirmatory hypotheses and most people struggle to process interactions beyond second-order ones, we limit our exploratory analysis to two-way interactions and refrain from reporting $p$ values (Nosek & Lakens, 2014). We ran separate models for each possible two-way interaction (e.g., with the number of people interacted with location, then the number of people interacted with duration, and so on) and focus on the coefficient for the interaction terms. Figure 5 plots point estimates with 95% confidence intervals (confidence intervals give a range of values which is likely to include the true coefficient). Figure 6 plots their corresponding mean standardized risk rating. Interaction coefficients for the Public sample were mostly negative, while those for the Expert sample tended to be positive, particularly when location was one of the factors. In other words, when members of the general public integrated two risk factors, they did so subadditively (i.e., the whole was *less* than the sum of its parts). Experts, by contrast, tended to perceive more risk when the scenario described something high-risk (e.g., meeting many others) taking place indoors (i.e., Experts perceive the risk as a whole as *greater* than the sum of its parts). While these estimates for the Expert sample are imprecise with large confidence intervals, due to the smaller sample size, five of the nine coefficients for the Public sample have confidence intervals that do not contain zero.

Domain expertise is one possible explanation for the difference in how risk factors might be integrated. Another possibility is cognitive ability and educational attainment; the Expert sample was likely to have higher cognitive ability than the Public sample. We therefore compared the Experts to the participants in the Public sample with the highest attainment (those with at least a Master's degree, $n = 106$; Online Supplemental Materials). Again we refrain from null-hypothesis significance testing given the exploratory nature of this analysis and the reduced sample size. Descriptive statistics show that the Educated subsample had standardized means between the Public and Expert samples for all factors except for when distancing was easy. In this case, the Educated subsample gave responses much closer to the Public than Experts. Turning to the two-way interactions, again the means of the Educated subsample fell between the Public and Expert samples. We consider the comparisons on interaction coefficients in two ways: The difference between the coefficients (i.e., Are the coefficients for the Educated Public sample closer to Experts or the full Public sample?) and the valence of the coefficients. The interaction coefficients from the Educated subsample were more similar to the Public on five of the nine interaction models and more similar to Experts on four. On coefficient valence, the Public sample had positive coefficients on two interactions and Experts had positive coefficients on six. The Educated Public had positive coefficients on four. Hence, the Educated subsample showed responses in the middle ground between the full Public sample and the Expert sample. These findings imply that there may be a benefit of both higher cognitive ability *and* domain expertise in evaluating synergistic risks, although confirmatory research would be needed to explore this idea further.

**Mask-Wearing.** The Public sample decreased assessments of risk by 3.03 points ($SD = 11.67$) when the described scenario stated that only they wore a mask and by 8.44 ($SD = 12.32$) when everyone wore one. Expert perceptions of risk decreased by 4.26 ($SD = 9.50$) and 9.51 ($SD = 9.73$), respectively. We test for differences in this reduction using a regression model of change in risk (standardized at the participant level with clustered standard errors) predicted by participant group, mask condition, and their interaction. The model, $F(3, 979) = 49.32$, $p < .001$, $R^2 = .06$, shows that all participants' perceptions of risk reduced more when everyone wore a mask compared to just themselves (β = −0.39, 95% CI [−0.48, −0.31], $p < .001$). There was no evidence for a difference between Experts and the Public overall (β = −0.13, 95% CI [−0.28, 0.02], $p = .100$), nor for an interaction (β = 0.01, 95% CI [−0.12, 0.14], $p = .872$), implying that the Public had absorbed the message that masks have more of a protective effect on others than on the wearer.

**Figure 5**

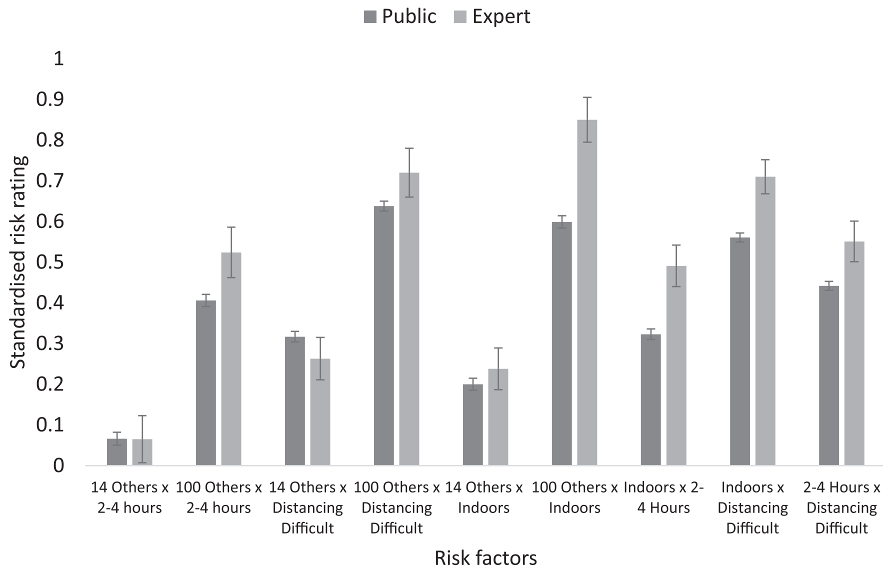
*Plot of Coefficient for Two-Way Interactions on the Risk Ratings*




*Note*. Error bars indicate the 95% confidence interval.

Models 3a–4b in Table 3 regress the change in risk perception on the four manipulated risk factors. We report separate models by participant group and mask condition. Negative coefficients imply that mask-wearing negated the risk due to the specific factor. The Public judged that wearing masks significantly reduced risk from all factors except duration. Expert results were broadly similar, except with respect to ease of distancing. The Public reduced their rating of risk when the scenario described that only they wore a

**Figure 6**
*Standardized Risk Ratings for Interaction Coefficients in Figure 5*




*Note*. Error bars are the standard error of the mean. Note that the average number of observations for the Public sample was 2,022 and 136 for the Expert sample.

mask and distancing was difficult compared to when distancing was difficult and they did not wear a mask. The reduction was nonsignificant when everyone was described as wearing a mask. Experts, on the other hand, only judged risk to reduce significantly when distancing was difficult if everyone wore a mask. Full models, which interact the risk factors with participant group and mask condition, support this pattern. Experts' perceptions of risk reduced less than the Public when everyone wore a mask but distancing was easy ($\beta = 0.16$, 95% CI [0.02, 0.30], $p = .024$) but their perception of risk reduced more than the Public when everyone wore a mask and distancing was hard ($\beta = -0.31$, 95% CI [−0.50, −0.13], $p = .001$).

When asked about the level of risk that was acceptable to take, the Public reported having low tolerance for risk, on average 12.75 out of 50 ($SD = 12.15$, $Mdn = 9$), with a strong skew. They also reported being highly confident in their ability to judge risk ($M = 5.65$, $SD = 1.13$, $Mdn = 6$), with 95.69% responding at the mid-point or above on the 1–7 scale. Findings reported in this section are the same when risk tolerance and confidence are added to the models as controls.

**Table 3**
*Regression Models of the Change in Risk Due to Mask-Wearing*

| Risk factor | Model 3a (public—only you) | Model 3b (public—everyone) | Model 4a (expert—only you) | Model 4b (expert—everyone) |
|---|---|---|---|---|
| Number of people (*Ref*: 5 others) | | | | |
| 14 others | $-0.16^{**}$ | $-0.23^{***}$ | $-0.27^{*}$ | −0.20 |
| | [−0.25, −0.07] | [−0.34, −0.13] | [−0.52, −0.03] | [−0.41, 0.01] |
| 100 others | $-0.15^{**}$ | $-0.30^{***}$ | $-0.24^{*}$ | $-0.29^{*}$ |
| | [−0.15, −0.06] | [−0.41, −0.18] | [−0.46, −0.01] | [−0.53, −0.05] |
| Indoors (*Ref:* Outdoors) | $-0.27^{***}$ | $-0.22^{***}$ | −0.17 | $-0.22^{*}$ |
| | [−0.35, −0.18] | [−0.31, −0.13] | [−0.37, 0.04] | [−0.41, 0.03] |
| 2–4 hr | −0.05 | −0.04 | −0.07 | −0.14 |
| (*Ref*: 15–30 min) | [−0.13, 0.02] | [−0.12, 0.05] | [−0.25, 0.11] | [−0.35, 0.06] |
| Difficult to distance (*Ref*: Easy) | $-0.19^{***}$ | −0.07 | −0.13 | $-0.32^{**}$ |
| | [−0.27, −0.11] | [−0.16, 0.01] | [−0.33, 0.07] | [−0.49, −0.14] |
| Intercept | $0.15^{**}$ | $-0.26^{***}$ | 0.02 | −0.21 |
| | [0.04, 0.25] | [−0.37, −0.15] | [−0.27, 0.30] | [−0.44, 0.01] |
| Observations | 1,520 | 1,448 | 224 | 224 |
| *N* | 380 | 362 | 56 | 56 |
| $R^2$ | .06 | .04 | .05 | .11 |

*Note*. Brackets contain 95% confidence intervals.
$^{*} p < .05$. $^{**} p < .01$. $^{***} p < .001$.

## Stage 3: Risk Rankings

Our third task asked participants to rank the riskiness of individual factors. Ranking tasks are deployed to elicit preferences because they require differentiation of options and thereby force stronger trade-offs than rating tasks (Krosnick, 1999). Although rankings do not quantify differences, they can shed light on which risk factors people place more importance on in isolation, when other contextual information is limited. The ranking task also allowed us to introduce non-COVID risks for comparison, such as driving without a seatbelt. Our third set of research questions was:

> *Research Question 3a*: How are individual risk factors prioritized in the absence of other contextual information?
>
> *Research Question 3b*: How are COVID-19 risks prioritized against non-COVID risks?
>
> *Research Question 3c*: Do the public and experts differ in how they prioritize COVID-19 and non-COVID risks?

Participants ranked eight factors—the five COVID-19 risk factors from Stage 2 and the three non-COVID factors used for the practice trials (gambling, driving without a seatbelt, and risky sporting activities)—in order of how risky each one would be for them, using the interface shown in Figure 7. Each participant saw the activities presented in a randomized order. This task required participants to prioritize specific risk factors over others when information about the context was limited to just one factor. Non-COVID risks were included in this task to provide insight for policymakers into how the public thought about specific COVID risks compared to everyday risks.

### Results

The fastest 5% of participants, who spent less than 32s on the task ($n = 37$), were excluded from the following analyses (although this does not alter the findings). Figure 8 charts the mean rank assigned to each factor by the Public and Expert samples, with the *X*-axis ordered by weightings assigned by the Public in Stage 2. The chart shows consistency among Experts between the rating task in Stage 2 and this task: Maintaining social distance, meeting a large group of people, and meeting indoors had the largest coefficients in Stage 2 and were ranked as most risky in Stage 3. The Public rankings, however, differed from the weightings estimated in the Rating Task. Although the number of other people and maintaining distance were the most heavily weighted risk factors in Stage 2, in the ranking task meeting indoors and meeting for a long time were judged to be more important than distancing, (Wilcoxon Signed Rank, $Z_{\text{Duration}} = 5.97$, $p < .001$; $Z_{\text{Location}} = 4.28$, $p < .001$).

Comparing judgments of COVID risks to non-COVID ones, the Public ranked not wearing a seatbelt similarly to meeting with a large group, and they ranked gambling and risky sporting activities as less risky than the other COVID risks (except for going where not many others wear a mask). By contrast, Experts ranked not wearing a seatbelt and engaging in a risky sporting activity as riskier than COVID risks.

## Stage 4: Risk Vignettes (Public Only)

In addition to comparing public and expert perceptions of risk factors, we tested whether perceived risk of COVID-19 infection is altered by other risks. Until a vaccine is widely available, everyday situations pit the potential for infection against other needs, such as going to work, attending medical appointments for other issues, or visiting friends and family. The affective response to different kinds of risk can bias perceptions in specific directions (i.e., the affect heuristic; Slovic et al., 2007). For example, if the anticipated thrill of a sky dive elicits a stronger affective response than the worry of injury, a prospective sky-diver is likely to take the risk (Finucane et al., 2000). Risk homeostasis theory posits that individuals attempt to balance tolerable levels of risk across domains (e.g., Wilde, 1998), however, we could find no empirical studies that directly examine whether *perceptions* of risk in one domain are biased by the presence of risk in another. A potential extension of the affect heuristic is to hypothesize that an affective response induced by an alternative everyday risk may lead people to downplay the perceived risk of infection from COVID-19, perhaps especially since the latter is relatively novel. For example, working in a busy factory, Paul might perceive his risk of contracting COVID-19 at work to be lower if his income is vital for meeting his mortgage repayments than if his income is less important. One can distinguish between whether Paul judges it to be more *reasonable* to take the risk when facing a serious financial risk, in line with risk homeostasis theory, from whether the second risk alters his perception of the likelihood of infection. Hence, the task determined whether people perceive the risk of COVID-19 infection independently of the presence of everyday risks. Our final research question was:

> *Research Question 4*: Does the presence of an alternative risk diminish the perceived risk of infection from COVID-19?

Participants in the Public sample saw a series of six vignettes, presented in random order. The experts did not complete this task. Three described a situation in which an individual must decide between engaging in a potentially risky COVID-19 behavior (e.g., using busy public transport, working in a crowded factory) or facing an alternative risk (financial, medical or psychosocial). Three further vignettes described factors of interest for policy (whether cases were increasing or decreasing, familiarity with others, whether Government and public health officials agreed) and are reported separately. Participants were asked two questions about each vignette. First, they were asked to judge the riskiness of the COVID-19 behavior, again considering only the possibility of infection. Second, they were asked how reasonable it would be for the individual to engage in the behavior. They gave both responses on the same scale used in Stage 2, from Not At All Risky (or Reasonable) to Extremely Risky (or Reasonable). For this task, the question was posed in a third-person format (i.e., the risk for the protagonist in the vignette) rather than first-person, thereby removing over-optimism bias from these judgments (e.g., Wise et al., 2020). This was important as we needed to "anchor" all participants to the same situation (King & Wand, 2007). Removing over-optimism bias also means the manipulations faced a more stringent test of whether they diminish perceptions of risk than if they were applied to the participants' own perception. The logic here is that, if justifications are

**Figure 7**
*Example Ranking Task*

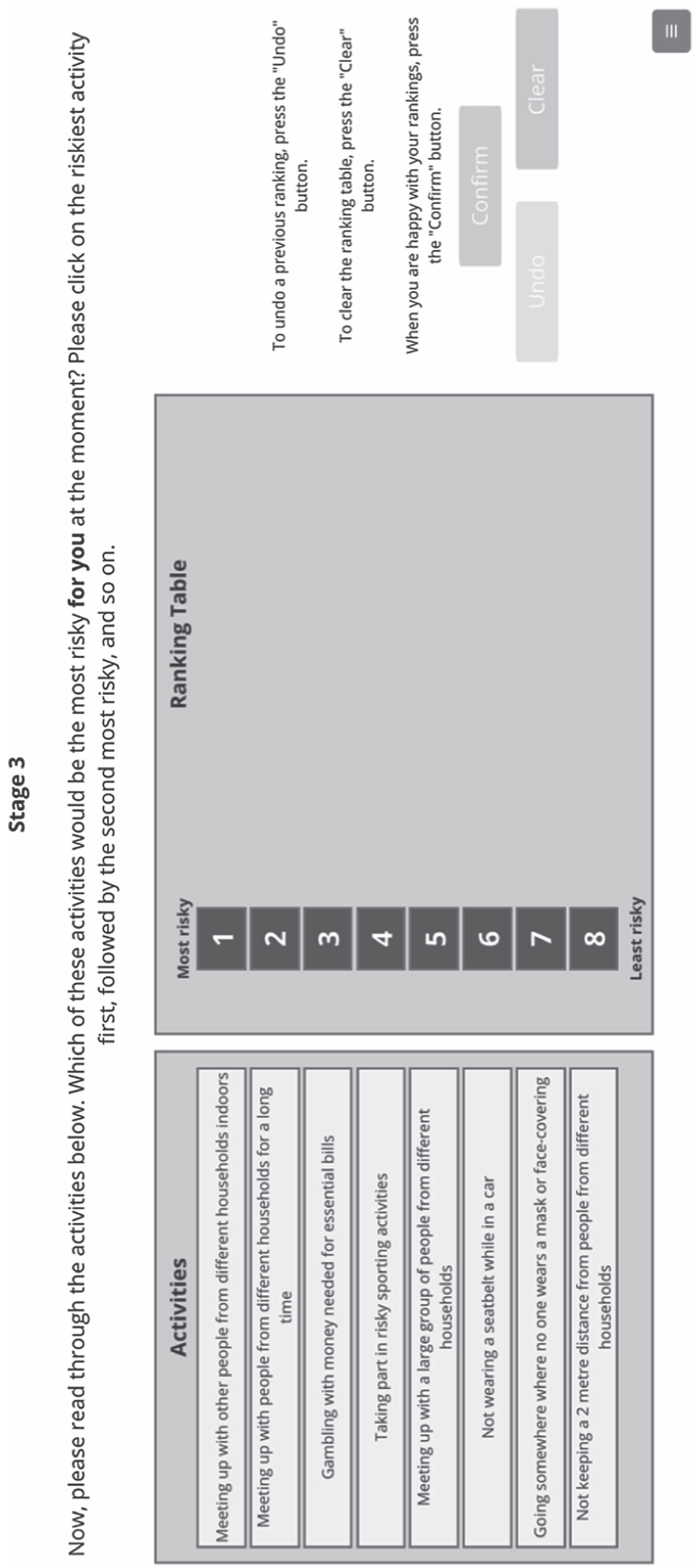
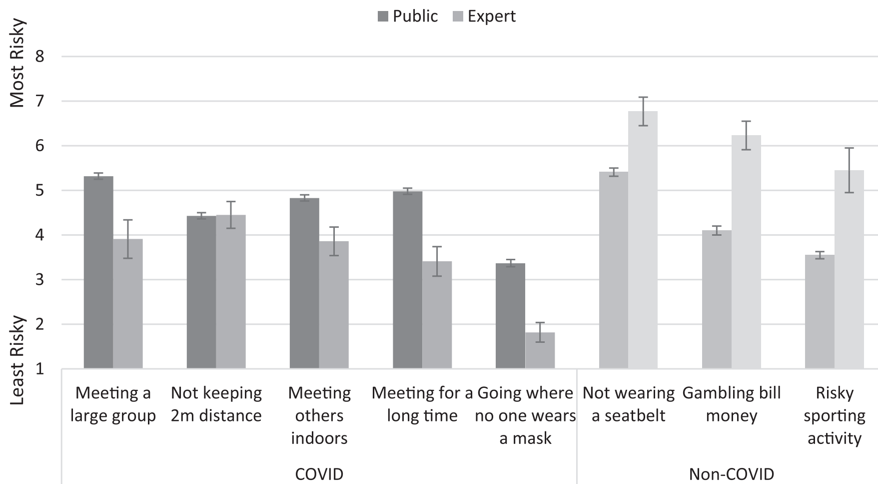


**Figure 8**

*Average Rankings Assigned to Risks in the Ranking Task by the Public and Experts*




*Note*. Error bars indicate the standard error.

sufficiently strong to diminish perceived risk for hypothetical others, they are also likely to influence over-optimistic perception of own-risk.

Each vignette described a situation in which there was likely to be many other people and distancing would be difficult to maintain. As demonstrated in the ratings task, these are features of social situations that people weight heavily when considering risk. Details about other factors (e.g., masks, ventilation) were deliberately not specified, in order to prevent ceiling or floor effects and to ensure the task did not become overly complex. Importantly, we were interested in the relative differences across different versions of vignettes and all COVID-19 risk factors were held constant. For each of the vignettes we report here, we created three isomorphs to vary the level of the alternative risk (i.e., low, moderate, high), as shown in Table 4. The alternative risks (not meeting mortgage repayments, missing an important medical appointment, being unable to have social contact for a prolonged period of time) were informed by a pretest in which a small sample of participants ($n = 22$) judged them to be equally worrying. The design was 3 (risk: financial, medical, psychosocial) × 3 (risk level: low, moderate, high) between-participants. Participants read one vignette from each type of risk and one vignette from each risk level. The pairing between risk type and risk level was counterbalanced across the sample.

## Results

The fastest 5% of participants read the vignettes and responded to the first question within 7.5s and are excluded from the following analyses. Again, their inclusion does not qualitatively alter results. The average response time otherwise was 36.5s (*Mdn* = 31.2s).

Figures 9 and 10 show average risk perceptions and reasonableness judgments to each vignette. Note that a one-way ANOVA showed no evidence that the Low Risk versions of each vignette elicited different perceptions of risk, $F(2, 739) = 0.70$, $p = .498$, $\eta^2 < .01$. We analyze each risk type (financial, medical, psychosocial) separately, meaning each participant has one score per question for each vignette and the primary comparison is between-groups for each risk level (low, medium, high). We report OLS regressions to test for differences between versions but results are closely similar if transformed or ordinal response variables are used (Table 5). All models include sociodemographic controls for gender, age, education, socioeconomic grade, employment status, and living area. Results are similar if a control for being in a high-risk group is added and if responses are standardized using each participant's mean and standard deviation from Stage 2. We preregistered directional hypotheses but do not adjust the alpha because we run two separate models on each vignette.

There was no evidence that facing a moderate or severe financial risk altered perceived risk of infection compared to a low financial risk, Model 5a, $F(2, 757) = 0.51$, $p = .600$, nor was there a difference between moderate and severe risk, $F(1, 757) = 0.01$, $p = .942$. However, Model 5b, $F(2, 757) = 9.83$, $p < .001$, shows that participants reported that taking such a risk was more reasonable when the financial risk was severe compared to low ($p < .001$) and moderate, $F(1, 757) = 6.82$, $p = .009$. The difference between low and moderate risk was in the predicted direction but not significant ($p = .072$).

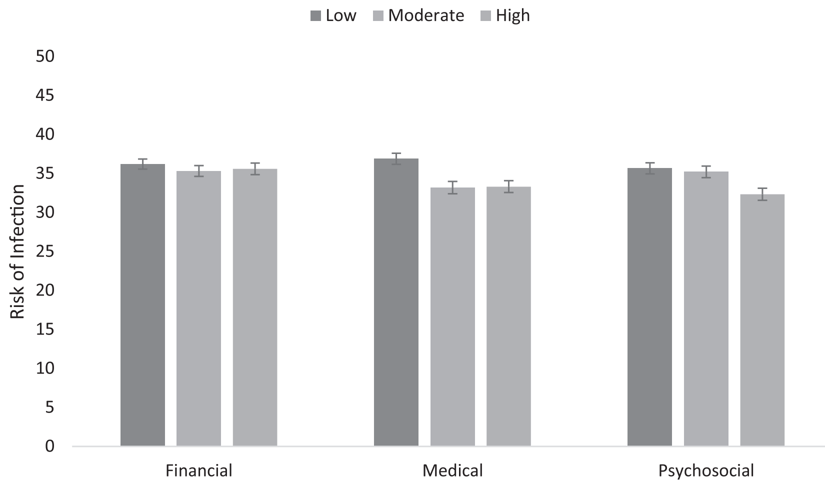
Regarding the medical vignette, participants judged the risk of contracting COVID-19 from traveling on public transport to be lower when the alternative medical risk was moderate or high compared to when it was low ($p < .001$, $p = .001$, respectively). There was no difference between moderate and high medical risk, $F(1, 758) = 0.01$, $p = .907$. Participants also judged that it was more reasonable to take the risk of contracting COVID-19 when there was moderate ($p < .001$) or severe ($p < .001$) medical risk scenarios, but there was no difference between moderate and severe risk, $F(1, 758) = 2.08$, $p = .151$.

Responding to the psychosocial vignette, participants judged the risk of contracting COVID-19 from traveling on public transport to be lower when the psychosocial risk (of loneliness) was high compared to when it was low ($p = .002$) or moderate, $F(1, 759) = 7.41$, $p = .007$, but there was no difference between low and

**Table 4**
*Risk Trade-Off Vignettes*

| Vignette structure | Financial | Medical | Psychosocial |
|---|---|---|---|
| Introduction | Paul does maintenance on machines and basic IT systems. He's been offered a day's work helping a factory to reopen. Paul knows the factory. There will be 40–50 people working fairly close together on the factory floor. The building is quite old and it's all indoors. | Mary has a doctor's appointment tomorrow. She's had increasing abdominal pain for several weeks, but has put off going during lockdown. Mary can't afford a taxi and the bus takes 45 min each way. Given the time of the appointment, she thinks the bus may be quite busy. With how long each journey will take, she'll probably need to use the public toilets near the bus stop. | Jim and Tony have been best friends for years, but live far apart and haven't seen each other since February. Jim is going to visit Tony at home, and both are looking forward to it. |
| Low risk | Paul has kept some work going from home during lockdown. He's paying his bills and feels like he's coping financially. The extra money from this job would be a boost to his regular income. | Mary's doctor has offered an online video consultation and thinks Mary's complaint is probably minor. The doctor can always arrange for someone to collect Mary to bring her to the surgery if it's needed. | Neither has been particularly lonely, as they have both seen family. They have good internet connections and have spoken using video calls, but it's just not the same. But Jim's car won't start. He could get a bus to the station and take a train, but the station has become busy again and he's worried the train could be busy too, especially when he comes back later on. Jim could wait a day or so to fix the car instead. |
| Moderate risk | Paul's income is down. He has some savings and got the Government welfare payment. He's managing to pay his bills but things are quite tight. The extra money would help. | Mary's doctor has offered a consultation over a video call, but given the nature of Mary's complaint it would be better to see her in person, just in case. | Jim has been a bit lonely, as he has no family or friends who could visit. His internet is not great, so he's only managed a few short video calls with other friends. He's really looking forward to seeing Tony in person. But Jim's car won't start. He could get a bus to the station and take a train, but the station has become busy again and he's worried the train could be busy too, especially when he comes back later on. Jim could wait a day or so to fix the car instead. |
| High risk | Paul's income is down and he has no savings. The Government welfare payment has not been enough to cover the bills and he's worrying about the next mortgage payment. Some extra money would really help. | Mary's doctor can't do an online consultation, and has said it's important to see her in person given the nature of Mary's complaint. It could be something serious. | Jim has been really lonely during lockdown, as he has no family or friends who could visit. His mobile phone coverage is quite poor and he's only received the occasional text message. But Jim's car won't start. He could get a bus to the station and take a train, but the station has become busy again and he's worried the train could be busy too, especially when he comes back later on. He's not sure how long it will take to fix his car and it could be a long time before he has any other social contact. |

**Figure 9**
*Average Risk Assigned to Each Vignette*



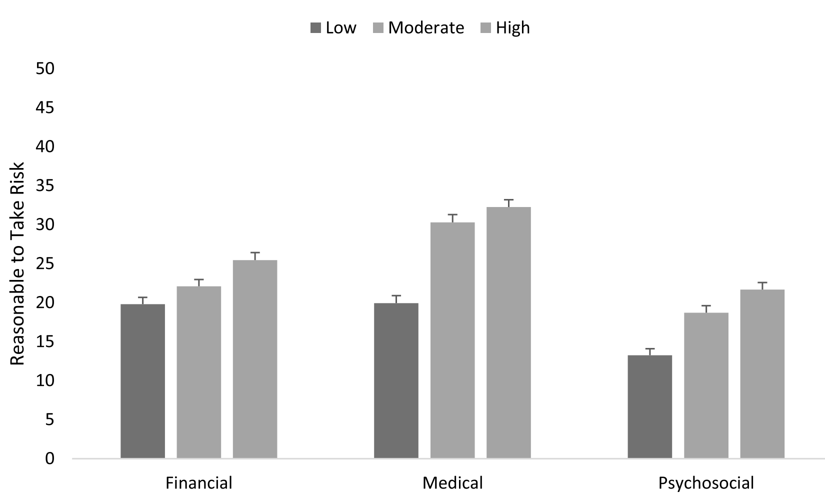

*Note*. Error bars indicate the standard error.

moderate risk ($p = .671$). However, participants judged that it was more reasonable to take the risk when the loneliness risk was high compared to low ($p < .001$) and moderate, $F(1, 759) = 5.60$, $p = .018$, and when it was moderate compared to low ($p < .001$).

## Deviations From Preregistration

We deviated from the preregistered analysis plan in the following ways. First, the possibility to collect data from the Expert sample arose after the research questions for the Public sample were preregistered. An additional preregistration was uploaded for the Expert sample. Second, in the risk ratings task, we standardized ratings at the participant level to account for large differences in individual standard deviations, rather than using mixed-effects models or transforming the response scale to an ordinal scale. We also opted not to exclude participants who classified themselves as high-risk. This decision was made prior to any analysis—we had not anticipated the proportion of participants who would fall into this category (26.6%), which on reflection is in line with population estimates released after the study (Clark et al., 2020). The participants closely match Census estimates on all demographic questions. We also preregistered checks for sociodemographic differences but do not report them in this paper as these were primarily to inform potential targeting of health communications for policy. Finally, on the Risk Vignettes task, we retained the raw response scale rather than transforming to ordinal scales for ease of interpretation; again this choice does not affect the results.

**Figure 10**
*Average Judgment That the Risk Was Reasonable to Take for Each Vignette*




*Note*. Error bars indicate the standard error.

**Table 5**
*Regression Models of Risk Perception and Reasonability Judgments to Each Vignette Type*

| Vignette condition (*Ref.* Low risk) | Financial | | Medical | | Psychosocial | |
|---|---|---|---|---|---|---|
| | Model 5a (risk perception) | Model 5b (reasonability judgment) | Model 6a (risk perception) | Model 6b (reasonability judgment) | Model 7a (risk perception) | Model 7b (reasonability judgment) |
| Moderate risk | −0.53 | 1.57 | −3.85*** | 10.41*** | −0.89 | 5.86*** |
| | [−2.43, 1.37] | [−0.89, 4.04] | [−5.89, −1.81] | [7.85, 12.97] | [−2.92, 1.14] | [3.42, 8.31] |
| High risk | −0.41 | 4.94*** | −3.90*** | 12.55*** | −3.52** | 8.51*** |
| | [−2.31, −1.50] | [2.47, 7.42] | [−5.95, −1.85] | [9.98, 15.12] | [−5.54, −1.50] | [6.08, 10.93] |
| Intercept | 35.38*** | 22.86*** | 37.36*** | 23.84*** | 35.65*** | 13.94*** |
| | [32.81, 37.95] | [19.52, 26.20] | [34.61, 40.12] | [20.38, 27.30] | [32.91, 38.40] | [10.65, 17.24] |
| $N$ | 760 | 760 | 761 | 761 | 762 | 762 |
| $R^2$ | .04 | .08 | .06 | .19 | .07 | .10 |

*$p < .05$. **$p < .01$. ***$p < .001$.

## Discussion

This study set out to elicit lay perceptions of everyday COVID-19 risk, benchmarked against perceptions of relevant medical experts. Results from the multistage experiment suggest that the public had absorbed information about some main risk factors well. Public and expert samples broadly agreed on the risks involved when meeting in large groups, not maintaining social distance and not wearing a mask. However, experts perceived substantially greater risk associated with being indoors, spending long periods with others, and being exposed to multiple simultaneous risk factors. Table 6 summarizes our research questions and findings. Looking across tasks, the results are suggestive of a heuristic approach underlying public evaluations of risk. We outline this here, followed by a discussion of study limitations.

Differences between public and expert responses were not consistent across tasks, although there were commonalities. Experts were more likely to mention location and duration in their open text responses and to weight both (especially location) more heavily in the Risk Rating task. In these two tasks, differences between experts and the public did not arise regarding the number of people present, maintaining social distance, or wearing a mask. However, in the Risk Ranking task, when risks were considered in isolation, the public placed greater weight on location and duration as risk factors than distancing. This pattern is consistent with public reliance on an "avoid people" heuristic when multiple situation attributes need to be integrated. That is, the number of close interactions with others, one of whom might be infectious, is cognitively available and dominates complex judgments once four or more risk factors must be juggled simultaneously. Yet, when making less cognitively demanding binary comparisons to generate a ranking, important environmental factors receive more weight. One possibility is that when contextual information is limited or uncertain—such as judging whether it is safe to go to a gathering where this an unknown number of people or where ability to maintain social distance is unclear—the public prioritize the gathering's location and how long they plan to be there. Otherwise, however, it may be easier to rely on the simple heuristic to avoid people. This finding is potentially important, given that real-world judgments tend to be multidimensional and there is growing evidence that the efficacy of social distancing depends on the environment (Jones et al., 2020).

The Ratings Task also revealed that, relative to an expert sample, the public neglects the synergistic nature of risk. Interaction coefficients implied that factors were combined subadditively, as with risk perceptions in other domains (Dawson et al., 2013). Conversely, equivalent interactions for experts had positive signs, particularly when the scenario described a high-risk encounter (e.g., meeting a large group of people) taking place indoors. Moreover, experts and the public perceived different interactions between mitigation behaviors. The public judged masks to reduce risk significantly more when social distancing was maintained, while experts judged masks to reduce risk significantly more when distancing was not maintained (i.e., when risk was higher). These findings may indicate a less granular perception of risk, perhaps including a degree of binary categorization by the public, whereby social settings are primarily classed as "safe" or "not safe," with only limited further differentiation. Future studies could explore the scope for teaching the public to integrate information as experts

**Table 6**
*Research Questions and Main Findings*

| RQ | Question | Finding |
|---|---|---|
| 1a | What risk factors for COVID-19 infection in social settings are most cognitively available to the public? | Meeting with large groups of people and not maintaining social distancing were the two most cognitively available risk factors for the public. Less than one-third mentioned location, and few people mentioned duration of interactions or mask-wearing. |
| 1b | Do the public and experts differ in the factors that are most cognitively available to them, when thinking about the risk of contracting COVID-19 in social settings? | Experts and the public showed similar awareness of risks, although experts were significantly more likely to think about whether they meet others indoors or outdoors and, to a lesser extent, the duration of a social gathering. |
| 2a | How do the public weight specific risk factors for COVID-19 infection when multiple ones must be integrated? | The public weighted the number of other people and the ability to maintain social distance most heavily when evaluating risk, followed by the location and lastly the duration. |
| 2b | Are there differences between how the public and experts weight risk factors for COVID-19? | Experts judged location and duration to be more important than the public did. They judged location to be as important as the number of people and distancing. |
| 2c | Are there differences between how the public and experts process interactions between risk factors for COVID-19? | The public tended to combine risks subadditively, whereas experts combined risks multiplicatively when a high-risk factor occurred indoors. |
| | | The public judged masks to reduce risk significantly more when distancing is maintained, whereas experts judged masks to diminish risk more so when distancing is not maintained. |
| 3a | How are individual risk factors prioritized in the absence of other contextual information? | The public judged meeting with a large group of people to be the most important risk factor, followed by duration, location, distancing, and mask-wearing. |
| 3b | How are COVID-19 risks prioritized against non-COVID risks? | The public judged COVID-19 infection to be riskier than some other everyday risks (such as gambling and risky sporting activities). They judged meeting with a large group of other people to be as risky as driving without a seatbelt. |
| 3c | Do public and experts differ in how they prioritize COVID-19 and non-COVID risks? | Experts didn't differentiate between most of the COVID-19 risk factors, except for mask-wearing (which they judged as the least important) and judged COVID-19 infection to be less risky than other everyday risks. |
| 4 | Does the presence of an alternative risk diminish the perceived risk of infection from COVID-19? | Facing a moderate medical risk and high psychosocial risk decreased the perceived risk of infection of COVID-19, but there was no evidence for an effect of financial risk. |

*Note.* COVID-19 = Coronavirus disease.



do (e.g., Attari et al., 2010; Fleming et al., 2012), as well as whether differences are due to cognitive ability, education, or domain-relevant expertise.

The Risk Vignettes, completed only by the public sample, showed that factors independent of COVID-19 risk can decrease the perceived possibility of infection. Financial, medical, and psychosocial risk all increased how reasonable participants judged COVID-19 risks to be to take, but high psychosocial risk and even moderate medical risk led participants to judge that the possibility of infection itself was lower than when the alternative risk was quite low. The findings therefore suggest that facing alternative risks is likely to make people more vulnerable during the pandemic. The findings also have implications for psychological understanding of risk. One possibility is that the results provide novel evidence that the affect heuristic extends to situations in which two sources of dread compete (e.g., Finucane et al., 2000), with the perceived level of one source of risk diminishing. However, this interpretation would benefit from more explicit testing. Moreover, why financial risk did not "compete" with COVID-19 risk in the same way as medical and psychosocial risk is unclear from our findings. One possibility for future research would be to test whether the relatedness of competing risk domains matters, as COVID-19, medical and psychosocial risks all relate to health risks, or whether certain domains are more heavily imbued moral connotations (O'Connor et al., 2021).

## Limitations

The findings from each stage have implications for public health interventions and psychological understanding of risk, but there are caveats worth noting. First, while we were interested exclusively in the perceived risk of becoming infected, there are multiple downstream components to COVID-19 risk to consider when generalizing the findings to everyday activity. These include the likelihood of removing the virus if it is picked-up (e.g., through hand hygiene), the risk of spreading the virus to others, the severity of symptoms and the likelihood of mortality. Second, our focus was infection through immediate social interaction; we ignored infection through other means. This included infection via face-touching after touching a fomite, which can be mitigated by observing good hand hygiene. Among the expert medical community, there is growing evidence that infection is driven primarily by airborne transmission (i.e., aerosols and droplets) rather than by fomites (Goldman, 2020; Mondelli et al., 2021). Third, this study was commissioned by policymakers to identify gaps in public comprehension of COVID-19 during the summer of 2020 in Ireland. While there is no reason to believe that the implied psychological mechanisms are specific to this context, there is no guarantee that the findings extend to other nations and times. Public health communications in Ireland have not departed notably from international norms and have been based heavily on

World Health Organization (WHO) advice. Nevertheless, the method we used offers a way to conduct diagnostic studies elsewhere, or to test further for an "avoid others" heuristic or subadditivity in multifactor COVID-19 risk judgments.

The main findings have relatively straightforward implications for policy communications. Medical experts placed greater weight on location when evaluating risk of COVID-19 transmission than the lay public. In the absence of other means for prioritizing messages, public communication that emphasizes the importance of location and ventilation seems sensible. However, the implications for psychological theory and for understanding expert versus lay risk perceptions are less straightforward. The Expert and Public samples were not matched on sociodemographic factors. An exploratory comparison found that the difference in risk perceptions between the experts and a subsample with postgraduate education was less extreme than the difference between the experts and public generally. One possibility is that cognitive ability (proxied by education) simply aids synergistic risk perception. Another is that the high-attainment subsample possessed greater capacity to educate themselves about the risks of transmission, given the broad impact of the pandemic. In other words, they may have developed some domain-specific expertise. If so, domain-relevant expertise may be more important for risk perception than the literature suggests (e.g., Rowe & Wright, 2001). However, further research to assess expertise in samples matched by sociodemographics would be required to address this question.

## Conclusion

To the extent that judgments of relevant medical experts can be treated as accurate benchmarks for how COVID-19 transmits, this study provides evidence that the public struggle to integrate environmental risks factors when evaluating the risk of becoming infected with COVID-19 in social settings. In particular, relative to medical experts, the public are likely to underestimate the benefits of interacting outdoors rather than indoors and focus more on how many people they come close to. This difficulty, coupled with the novel finding that perceived risk can be diminished by independent factors (such as other psychological needs), implies that people are likely to place themselves in environments with higher risk of infection unknowingly, thereby potentially contributing to the spread of the virus. Controlled diagnostic experiments can help to inform public health communications by identifying departures from medical advice and highlighting heuristics people use to evaluate risk, as well as advancing our understanding of the psychology of large-scale risks (Lunn, Belton, et al., 2020).

## References

Adam, M. B., & Reyna, V. F. (2005). Coherence and correspondence criteria for rationality: Experts' estimation of risks of sexually transmitted infections. *Journal of Behavioral Decision Making*, *18*(3), 169–186. https://doi.org/10.1002/bdm.493

Amárach Public Opinion Survey. (2021, July 13). *Public opinion tracking research*. https://www.gov.ie/en/collection/6b4401-view-the-amarach-public-opinion-survey/

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*(1), 388–407. https://doi.org/10.3758/s13428-019-01237-x

Attari, S. Z., DeKay, M. L., Davidson, C. I., & Bruine de Bruin, W. (2010). Public perceptions of energy consumption and savings. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(37), 16054–16059. https://doi.org/10.1073/pnas.1001509107

Barke, R. P., & Jenkins-Smith, H. C. (1993). Politics and scientific expertise: Scientists, risk perception, and nuclear waste policy. *Risk Analysis*, *13*(4), 425–439. https://doi.org/10.1111/j.1539-6924.1993.tb00743.x

Blais, A. R., & Weber, E. U. (2006). A domain-specific risk-taking (DOSPERT) scale for adult populations. *Judgment and Decision Making*, *1*(1), 33–47. https://sjdm.org/journal/vol1.1.htm

Brewer, N. T., Chapman, G. B., Gibbons, F. X., Gerrard, M., McCaul, K. D., & Weinstein, N. D. (2007). Meta-analysis of the relationship between risk perception and health behavior: The example of vaccination. *Health Psychology*, *26*(2), 136–145. https://doi.org/10.1037/0278-6133.26.2.136

Chu, D. K., Akl, E. A., Duda, S., Solo, K., Yaacoub, S., Schünemann, H. J., Hajizadeh, A., & the COVID-19 Systematic Urgent Review Group Effort (SURGE) study authors. (2020). Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: A systematic review and meta-analysis. *Lancet*, *395*(10242), 1973–1987. https://doi.org/10.1016/S0140-6736(20)31142-9

Clark, A., Jit, M., Warren-Gash, C., Guthrie, B., Wang, H. H. X., Mercer, S. W., Sanderson, C., Mckee, M., Troeger, C., Ong, K. L., Checchi, F., Perel, P., Joseph, S., Gibbs, H. P., Banerjee, A., Eggo, R. M., & the Centre for the Mathematical Modelling of Infectious Diseases COVID-19 working group. (2020). Global, regional, and national estimates of the population at increased risk of severe COVID-19 due to underlying health conditions in 2020: A modelling study. *The Lancet. Global Health*, *8*(8), e1003–e1017. https://doi.org/10.1016/S2214-109X(20)30264-3

Clogg, C. C., Petkova, E., & Haritou, A. (1995). Statistical methods for comparing regression coefficients between models. *American Journal of Sociology*, *100*(5), 1261–1293. https://doi.org/10.1086/230638

Dawson, I. G., Johnson, J. E., & Luke, M. A. (2013). Helping individuals to understand synergistic risks: An assessment of message contents depicting mechanistic and probabilistic concepts. *Risk Analysis*, *33*(5), 851–865. https://doi.org/10.1111/j.1539-6924.2012.01878.x

Delgado, M., & Moral, S. (1987). On the concept of possibility-probability consistency. *Fuzzy Sets and Systems*, *21*(3), 311–318. https://doi.org/10.1016/0165-0114(87)90132-1

Dryhurst, S., Schneider, C. R., Kerr, J., Freeman, A. L., Recchia, G., Van Der Bles, A. M., Spiegelhalter, D., & Van Der Linden, S. (2020). Risk perceptions of COVID-19 around the world. *Journal of Risk Research*, *23*(7–8), 994–1006. https://doi.org/10.1080/13669877.2020.1758193

Dubois, D., Prade, H., & Sandri, S. (1993). On possibility/probability transformations. In R. Lowen & M. Roubens (Eds.), *uzzy logic* (pp. 103–112). Springer. https://doi.org/10.1007/978-94-011-2014-2_10

European Centre for Disease Prevention and Control. (2020, March). *Considerations relating to social distancing measures in response to COVID-19—second update*. www.ecdc.europa.eu

Finucane, M. L., Alhakami, A., Slovic, P., & Johnson, S. M. (2000). The affect heuristic in judgments of risks and benefits. *Journal of Behavioral Decision Making*, *13*(1), 1–17. https://doi.org/10.1002/(SICI)1099-0771(200001/03)13:1<1::AID-BDM333>3.0.CO;2-S

Fischhoff, B., Bostrom, A., & Quadrel, M. J. (1993). Risk perception and communication. *Annual Review of Public Health*, *14*(1), 183–203. https://doi.org/10.1146/annurev.pu.14.050193.001151

Fleming, P., Townsend, E., Van Hilten, J. A., Spence, A., & Ferguson, E. (2012). Expert relevance and the use of context-driven heuristic processes in risk perception. *Journal of Risk Research*, *15*(7), 857–873. https://doi.org/10.1080/13669877.2012.666759

Folkes, V. S. (1988). The availability heuristic and perceived risk. *The Journal of Consumer Research*, *15*(1), 13–23. https://doi.org/10.1086/209141

Geer, J. G. (1988). What do open-ended questions measure? *Public Opinion Quarterly*, *52*(3), 365–367. https://doi.org/10.1086/269113

Gigerenzer, G., & Todd, P. M. (1999). Fast and frugal heuristics: The adaptive toolbox. In G. Gigerenzer, P. M. Todd, & the ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp. 3–34). Oxford University Press.

Goldman, E. (2020). Exaggerated risk of transmission of COVID-19 by fomites. *The Lancet. Infectious Diseases*, *20*(8), 892–893. https://doi.org/10.1016/S1473-3099(20)30561-2

Green, P. E., & Srinivasan, V. (1978). Conjoint analysis in consumer research: Issues and outlook. *The Journal of Consumer Research*, *5*(2), 103–123. https://doi.org/10.1086/208721

Hainmueller, J., Hangartner, D., & Yamamoto, T. (2015). Validating vignette and conjoint survey experiments against real-world behavior. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(8), 2395–2400. https://doi.org/10.1073/pnas.1416587112

Hale, T., Angrist, N., Goldszmidt, R., Kira, B., Petherick, A., Phillips, T., & Tatlow, H. (2021). A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nature Human Behaviour*, *5*, 529–538. https://doi.org/10.1038/s41562-021-01079-8

Halford, G. S., Baker, R., McCredden, J. E., & Bain, J. D. (2005). How many variables can humans process? *Psychological Science*, *16*(1), 70–76. https://doi.org/10.1111/j.0956-7976.2005.00782.x

Health Information and Quality Authority. (2020, June 8). *Restrictive public policy measures to limit COVID-19*. www.hiqa.ie/reports-and-publications

Horiuchi, Y., Markovich, Z. D., & Yamamoto, T. (2020). *Does conjoint analysis mitigate social desirability bias?* MIT Political Science Department Research Paper No. 2018-15 Political Analysis. https://doi.org/10.2139/ssrn.3219323

Hunt, L. T., Dolan, R. J., & Behrens, T. E. (2014). Hierarchical competitions subserving multi-attribute choice. *Nature Neuroscience*, *17*(11), 1613–1622. https://doi.org/10.1038/nn.3836

Johnson, E. J., Häubl, G., & Keinan, A. (2007). Aspects of endowment: A query theory of value construction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(3), 461–474. https://doi.org/10.1037/0278-7393.33.3.461

Johnson-Laird, P. N., & Byrne, R. M. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review*, *109*(4), 646–678. https://doi.org/10.1037/0033-295X.109.4.646

Jones, N. R., Qureshi, Z. U., Temple, R. J., Larwood, J. P. J., Greenhalgh, T., & Bourouiba, L. (2020). Two metres or one: What is the evidence for physical distancing in covid-19? *BMJ*, *370*, Article m3223. https://doi.org/10.1136/bmj.m3223

Karelaia, N., & Hogarth, R. M. (2008). Determinants of linear judgment: A meta-analysis of lens model studies. *Psychological Bulletin*, *134*(3), 404–426. https://doi.org/10.1037/0033-2909.134.3.404

King, G., & Wand, J. (2007). Comparing incomparable survey responses: Evaluating and selecting anchoring vignettes. *Political Analysis*, *15*(1), 46–66. https://doi.org/10.1093/pan/mpl011

Koehler, D. J., Brenner, L., & Griffin, D. (2002). The calibration of expert judgment: Heuristics and biases beyond the laboratory. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 686–715). Cambridge University Press. https://doi.org/10.1017/CBO9780511808098.041

Kraus, N., Malmfors, T., & Slovic, P. (1992). Intuitive toxicology: Expert and lay judgments of chemical risks. *Risk Analysis*, *12*(2), 215–232. https://doi.org/10.1111/j.1539-6924.1992.tb00669.x

Krosnick, J. A. (1999). Survey research. *Annual Review of Psychology*, *50*(1), 537–567. https://doi.org/10.1146/annurev.psych.50.1.537

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310

Lipshitz, R., & Strauss, O. (1997). Coping with uncertainty: A naturalistic decision-making analysis. *Organizational Behavior and Human Decision Processes*, *69*(2), 149–163. https://doi.org/10.1006/obhd.1997.2679

Liu, Y., Ning, Z., Chen, Y., Guo, M., Liu, Y., Gali, N. K., Sun, L., Duan, Y., Cai, J., Westerdahl, D., Liu, X., Xu, K., Ho, K.-F., Kan, H., Fu, Q., & Lan, K. (2020). Aerodynamic analysis of SARS-CoV-2 in two Wuhan hospitals. *Nature*, *582*(7813), 557–560. https://doi.org/10.1038/s41586-020-2271-3

Lohiniva, A. L., Sane, J., Sibenberg, K., Puumalainen, T., & Salminen, M. (2020). Understanding coronavirus disease (COVID-19) risk perceptions among the public to enhance risk communication efforts: A practical approach for outbreaks, Finland, February 2020. *Eurosurveillance*, *25*(13), Article 2000317. https://doi.org/10.2807/1560-7917.ES.2020.25.13.2000317

Lunn, P. D., Belton, C. A., Lavin, C., McGowan, F. P., Timmons, S., & Robertson, D. A. (2020). Using behavioral science to help fight the coronavirus. *Journal of Behavioral Public Administration*, *3*(1), 1–15. https://doi.org/10.30636/jbpa.31.147

Lunn, P. D., Bohacek, M., McGowan, F. P., & Ní Choisdealbha, Á. (2020). The surplus identification task and limits to multiattribute consumer choice. *Journal of Experimental Psychology: Applied*, *26*(2), 312–338. https://doi.org/10.1037/xap0000252

Miller, G. A. (1956). The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81–97. https://doi.org/10.1037/h0043158

Mondelli, M. U., Colaneri, M., Seminari, E. M., Baldanti, F., & Bruno, R. (2021). Low risk of SARS-CoV-2 transmission by fomites in real-life conditions. *The Lancet. Infectious Diseases*, *21*(5), Article E112. https://doi.org/10.1016/S1473-3099(20)30678-2

Nosek, B. A., & Lakens, D. (2014). A method to increase the credibility of published results. *Social Psychology*, *45*(3), 137–141. https://doi.org/10.1027/1864-9335/a000192

O'Connor, C., Relihan, D. P., Thomas, A. J., Ditto, P. H., Stanford, K., & Weatherall, J. O. (2021). Moral judgments impact perceived risks from COVID-19 exposure. *PsyArxiv*. https://doi.org/10.31222/osf.io/d64a8

Qian, G., Yang, N., Ma, A. H. Y., Wang, L., Li, G., Chen, X., & Chen, X. (2020). COVID-19 transmission within a family cluster by presymptomatic carriers in China. *Clinical Infectious Diseases*, *71*(15), 861–862. https://doi.org/10.1093/cid/ciaa316

Reja, U., Manfreda, K. L., Hlebec, V., & Vehovar, V. (2003). Open-ended vs. close-ended questions in web questionnaires. *Developments in Applied Statistics*, *19*(1), 159–177.

Reyna, V. F. (2008). A theory of medical decision making and health: Fuzzy trace theory. *Medical Decision Making*, *28*(6), 850–865. https://doi.org/10.1177/0272989X08327066

Reyna, V. F., & Brainerd, C. J. (2008). Numeracy, ratio bias, and denominator neglect in judgments of risk and probability. *Learning and Individual Differences*, *18*(1), 89–107. https://doi.org/10.1016/j.lindif.2007.03.011

Rowe, G., & Wright, G. (2001). Differences in expert and lay judgments of risk: Myth or reality? *Risk Analysis*, *21*(2), 341–356. https://doi.org/10.1111/0272-4332.212116

Ryan, M., & Farrar, S. (2000). Using conjoint analysis to elicit preferences for health care. *BMJ*, *320*(7248), 1530–1533. https://doi.org/10.1136/bmj.320.7248.1530

Schkade, D. A., & Kahneman, D. (1998). Does living in California make people happy? A focusing illusion in judgments of life satisfaction. *Psychological Science*, *9*(5), 340–346. https://doi.org/10.1111/1467-9280.00066

Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H., & Simons, A. (1991). Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality and Social Psychology*, *61*(2), 195–202. https://doi.org/10.1037/0022-3514.61.2.195

Setti, L., Passarini, F., De Gennaro, G., Barbieri, P., Perrone, M. G., Borelli, M., Palmisani, J., Di Gilio, A., Piscitelli, P., & Miani, A. (2020). Airborne transmission route of COVID-19: Why 2 meters/6 feet of inter-personal distance could not Be enough. *International Journal of Environmental Research and Public Health*, *17*(8), Article 2932. https://doi.org/10.3390/ijerph17082932

Shafir, E., Simonson, I., & Tversky, A. (1993). Reason-based choice. *Cognition*, *49*(1–2), 11–36. https://doi.org/10.1016/0010-0277(93)90034-S

Slovic, P. (1987). Perception of risk. *Science*, *236*(4799), 280–285. https://doi.org/10.1126/science.3563507

Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2007). The affect heuristic. *European Journal of Operational Research*, *177*(3), 1333–1352. https://doi.org/10.1016/j.ejor.2005.04.006

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1979). Rating the risks. *Environment*, *21*(3), 14–39. https://doi.org/10.1080/00139157.1979.9933091

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1985). Rating the risks: The structure of expert and lay perceptions. In V. T. Covello, J. L. Mumpower, P. J. M. Stallen, & V. R. R. Uppuluri (Eds.), *Environmental impact assessment, technology assessment, and risk analysis* (pp. 131–156). Springer. https://doi.org/10.1007/978-3-642-70634-9_7

Stewart, N., & Simpson, K. (2008). A decision-by-sampling account of decision under risk. In N. Chater & M. Oaksford (Eds.), *The probabilistic mind: Prospects for Bayesian cognitive science* (pp. 261–276). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199216093.003.0012

Timmons, S., Barjaková, M., Robertson, D., Belton, C., & Lunn, P. (2020). Public understanding and perceptions of the COVID-19 test-and-trace system. *Economic and Social Research Institute (ESRI) Research Series*. https://doi.org/10.26504/sustat96

Timmons, S., McGinnity, F., Belton, C., Barjaková, M., & Lunn, P. (2020). It depends on how you ask: Measuring bias in population surveys of compliance with COVID-19 public health guidance. *Journal of Epidemiology and Community Health*. Advance online publication. https://doi.org/10.1136/jech-2020-215256

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

Van Doremalen, N., Bushmaker, T., Morris, D. H., Holbrook, M. G., Gamble, A., Williamson, B. N., Tamin, A., Harcourt, J. L., Thornburg, N. J., Gerber, S. I., Lloyd-Smith, J. O., De Wit, E., & Munster, V. J. (2020). Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1. *The New England Journal of Medicine*, *382*(16), 1564–1567. https://doi.org/10.1056/NEJMc2004973

Vlaev, I., Chater, N., Stewart, N., & Brown, G. D. (2011). Does the brain calculate value? *Trends in Cognitive Sciences*, *15*(11), 546–554. https://doi.org/10.1016/j.tics.2011.09.008

Weber, E. U., Johnson, E. J., Milch, K. F., Chang, H., Brodscholl, J. C., & Goldstein, D. G. (2007). Asymmetric discounting in intertemporal choice: A query-theory account. *Psychological Science*, *18*(6), 516–523. https://doi.org/10.1111/j.1467-9280.2007.01932.x

Weber, M., & Borcherding, K. (1993). Behavioral influences on weight judgments in multiattribute decision making. *European Journal of Operational Research*, *67*(1), 1–12. https://doi.org/10.1016/0377-2217(93)90318-H

Weinstein, N. D. (1982). Unrealistic optimism about susceptibility to health problems. *Journal of Behavioral Medicine*, *5*(4), 441–460. https://doi.org/10.1007/BF00845372

Wilde, G. J. (1998). Risk homeostasis theory: An overview. *Injury Prevention*, *4*(2), 89–91. https://doi.org/10.1136/ip.4.2.89

Wise, T., Zbozinek, T. D., Michelini, G., Hagan, C. C., & Mobbs, D. (2020). Changes in risk perception and self-reported protective behaviour during the first week of the COVID-19 pandemic in the United States. *Royal Society Open Science*, *7*(9), Article 200742. https://doi.org/10.1098/rsos.200742

Wright, G., Bolger, F., & Rowe, G. (2002). An empirical test of the relative validity of expert and lay judgments of risk. *Risk Analysis*, *22*(6), 1107–1122. https://doi.org/10.1111/1539-6924.00276

Received March 12, 2021
Revision received August 10, 2021
Accepted August 19, 2021 ■