

The Role of Statistical Literacy in Risk Assessment and Behavior – Evidence from the COVID-19 Pandemic*

Daniel Metzger[†], Mikael Paaso[‡] and Vesa Pursiainen[§]

April 2022

Abstract

We study the role of statistical literacy in people's assessment of risk by analyzing risk expectations and behavior in the context of the COVID-19 pandemic. Using a large-scale panel survey of about 4,000 Dutch households, we find that more statistically literate people adjust their risk estimates more strongly to the actual infection risk, measured by the current number of confirmed COVID-19 cases in their home province. Consistently, they are more likely to socially distance, especially when local cases are high. We also show that statistically more literate people evaluate the risks and benefits of a vaccination against COVID-19 differently and are more likely to get vaccinated against COVID-19.

JEL classification: D84, I12, I18, I20

Keywords: statistical literacy, COVID-19, risk estimates, social distancing, vaccination

*We are grateful to CoViD-19 Impact Lab for sharing data with us and Deniz Okat for working with us on the survey. We appreciate the helpful comments from Ben Charoenwong, Johannes Haushofer, Hans-Martin von Gaudecker, Stefan Jaspersen, Alan Kwan, Xin Liu, Sami Torstila and Tom Zimmermann as well as from seminar participants at the Rotterdam School of Management.

[†]Erasmus University Rotterdam. E-mail: metzger@rsm.nl

[‡]Erasmus University Rotterdam. E-mail: paaso@rsm.nl

[§]University of St. Gallen and Swiss Finance Institute. E-mail: vesa.pursiainen@unisg.ch

1 Introduction

Laboratory and real-world experiments often find a considerable amount of unexplained variation in people's choices under uncertainty, beyond that explained by traditional measures such as preferences or risk aversion. Part of this may be explained by cognitive constraints – decision-making under uncertainty can be very difficult – or by limited knowledge of concepts relevant for such decision-making. For example, people may not know or be able to apply Bayes' rule, they may succumb to computational errors, or struggle to integrate relevant information. For this reason, cognitive abilities can be an important factor for processing available information and forming beliefs, as well as influencing preferences (Dohmen et al., 2010).

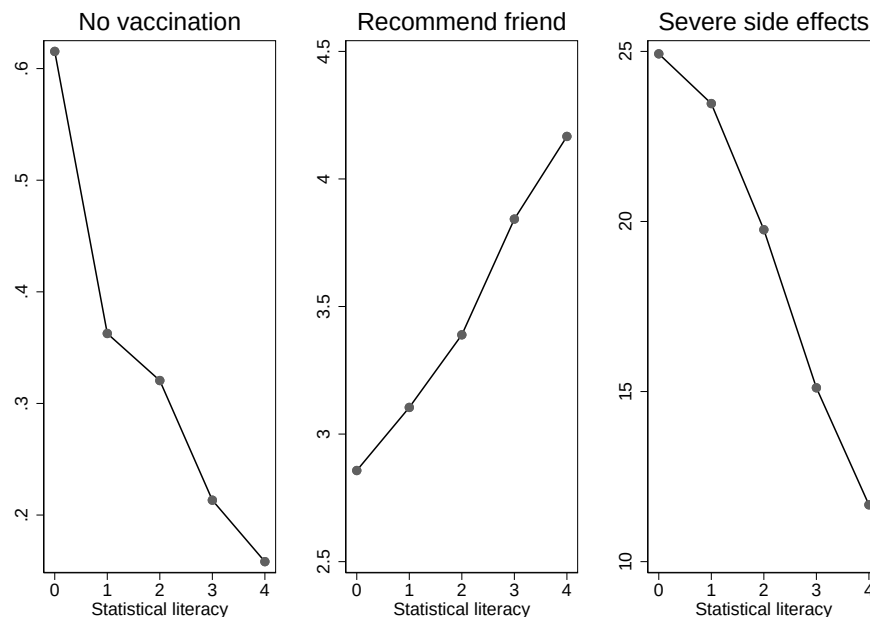
In this paper, we conjecture that *statistical literacy*, the ability to understand and reason with statistics and data, is another important factor in understanding decision-making under uncertainty. For example, without understanding the concepts of statistical measures such as mean and standard deviation, knowing the distributions of different alternatives is likely not helpful in the decision-making process. Such statistical concepts are important across a range of domains, from investing to assessing health-related risks.

We use the COVID-19 pandemic as a setting to study the role of statistical literacy in risk assessments. When it comes to assessing risk and behavior such as social distancing or vaccination against COVID-19, much of the discussion is driven by data and statistics, such as infection and mortality rates or undetected cases. This data might be difficult both to understand as well as to estimate accurately (Wu et al., 2020). For instance, most countries did not mandate a vaccination against COVID-19, meaning that people had to make their own assessments about the benefits of the vaccine relative to its potential harms. Our survey evidence, shown in Figure 1, shows that in April/May 2021, individuals in the US judge the costs and benefits of such vaccinations differently depending on their statistical literacy. More statistically literate people are more likely to obtain a vaccination and recommend a vaccination to friends. They also assign a lower likelihood to severe side effects from the

vaccination.¹ These differences are large. While more than 60% of the individuals with the lowest level of statistical literacy claim that they will not get a vaccination, this number goes down to less than 15% for the individuals with the highest level of statistical literacy. Similarly, while individuals with the lowest statistical literacy score estimate the likelihood of developing severe side effects after a COVID-19 vaccination to be about 25%, this estimate drops by more than half to about 12% for the most statistically literate individuals.

Figure 1: Statistical Literacy and Vaccination Against COVID-19

This figure shows the percentage of US respondents unwilling to take a vaccination against COVID-19, recommending vaccination to a friend and the expected probability of experiencing severe side effects as a result of a COVID-19 vaccine against a statistical literacy measure. The sample consists of Amazon MTurk workers in the US and the survey was conducted in April-May 2021. Details on the Statistical Reasoning measure and the survey can be found in the "Data & Methodology" section of the paper as well as in the Appendix.



In line with the documented correlations between statistical literacy and attitudes towards vaccination against COVID-19, we hypothesize that more statistically literate people are likely to better understand the threats posed by COVID-19 and more likely to adjust

¹The data come from an online survey conducted in April-May 2021. Details on the construction of the statistical literacy measure can be found in the internet appendix of this paper.

their behavior according to the actual risk of an infection, based on the available information. We construct a measure of statistical literacy using questions from the Statistical Reasoning Assessment test of Garfield (1998, 2003) and administer it in August 2020 as a separate wave of LISS, an extensive panel survey with a representative sample of about 4,000 households in the Netherlands. We combine our measure of statistical literacy with a rich set of information about the same person's COVID-19-related assessments of risk and social distancing behaviors, as well as comprehensive demographic background information. The COVID-related survey waves were conducted during the period between March and June 2020 and allow us to observe differences in views and behaviors, both cross-sectionally and over time.

Our main outcome of interest is the estimated probability that an acquaintance gets infected with COVID-19. In our baseline tests, we study whether more statistically literate individuals are more sensitive to less/more severe threats of getting infected, estimated by the current number of coronavirus cases in the region they live in. We find that statistical literacy is positively associated with significantly higher sensitivity of risk estimates to the actual infection levels. The baseline sensitivity of estimates of an acquaintance being infected to new cases is such that a 10% increase in cases per capita increases the probability estimate by 1.18 percentage points (compared to an average estimate across all waves and areas of 35.9 percent). For every unit of statistical literacy, this sensitivity increases by 1.74 percentage points, so that the most statistically literate people in our sample update the probability of an acquaintance being infected by 8.15 percentage points in response to a 10% increase in cases per capita.

More statistically literate people are also generally more likely to socially distance, including maintaining distance to others and avoiding public or busy places. Similar to risk estimates, the social distancing behavior of more statistically literate people is more sensitive to the local COVID-19 situation. Across various types of social distancing, the correlation between the level of social distancing and the number of local new cases increases with

statistical literacy.

One possible concern is that statistical literacy is merely a proxy for general intelligence or educational attainment. While we do not observe exogenous variation in statistical literacy, we attempt to exploit the richness of the LISS data to rule out alternative explanations. We run regressions with a wide range of controls, accounting for the most plausible alternative explanations, and show that even controlling for the level and field of education, risk aversion, health status, political views and other relevant characteristics, variation in statistical literacy is associated with more accurate updating of expectations. While these do not completely rule out the possibility of an omitted variable driving our results, they do mitigate this concern by eliminating the most obvious alternative explanations. We also use the "post-double-selection" algorithm of Belloni et al. (2014) to select variables from a large set of potential controls to further mitigate the concern of an omitted variables bias.

Our findings contribute to the literature on the determinants of social distancing behaviors and compliance with non-pharmaceutical interventions. Prior literature shows that compliance with restrictions and public health recommendations can be affected by, e.g., political views (Painter and Qiu, 2021; Barrios and Hochberg, 2021; Grossman et al., 2020), civic capital (Barrios et al., 2021; Durante et al., 2021) and trust in the government (Fancourt et al., 2020). Similarly, several studies suggest that differences in information (and misinformation) are important determinants of compliance (e.g., Charoenwong et al., 2020).

Risk perceptions can be an important channel through which these determinants of the effectiveness of policy interventions operate. Graso (2022) finds that mis-estimation of COVID-19-related risks is associated with higher support for policy interventions, while differences in risk perceptions have also been found to affect other health decisions like smoking (e.g., Viscusi, 1991; Lundborg and Andersson, 2008). Lammers et al. (2020) find that understanding exponential growth increases support for social distancing. Our findings suggest that general statistical literacy can be an important factor affecting risk perceptions hence the effectiveness of health policy.

Taken more broadly, individual risk perceptions can also have important economic consequences, including risk taking in personal investments (e.g., Choi and Robertson, 2020; Giglio et al., 2021) and even macroeconomic fluctuations (e.g., Pflueger et al., 2020). To the extent statistical literacy affects risk perceptions in other domains as well, our findings may have broader significance outside the scope of health behaviors only.

We also contribute to the literature on the impact of statistical literacy on expectations and behaviors. Prior literature has documented the impact of statistical literacy and numeracy on expectations in settings such as risky investments (e.g., Jasper et al., 2013; Almenberg and Widmark, 2011). Similarly, cognitive abilities are associated with better financial decision making Agarwal and Mazumder (2013); Benjamin et al. (2013). In this paper, we find that statistical literacy is strongly associated with more "correct" expectations even in a context where news about infection rates is widely available and publicized.

2 Statistical Literacy Measure and Data Description

The dataset we use for our main analyses comes from the LISS (Longitudinal Internet Studies for the Social Sciences) panel, run by CentER Data at Tilburg University. This panel covers a representative sample of almost 5,000 Dutch households (of whom about 4,000 responded to our questions). The panel consists of 12 rotating annual or biennial standardized waves (covering topics such as health, values and economic conditions) as well as irregular questionnaires added by researchers.

To measure statistical literacy and attitudes toward statistics, we add a separate set of questions to the August 2020 wave of the LISS panel. Various researchers (including the CoViD-19 Impact Lab) added five custom waves between March and June 2020 with questions on behaviours and beliefs related to the COVID-19 pandemic.

We describe the data in more detail in the next sections. All variables are defined in Appendix Table A.1.

2.1 Measuring Statistical Literacy

Our main measure of statistical literacy, *Statistical Reasoning*, is based on a subset of four different questions we adapted from the Statistical Reasoning Assessment (SRA) of Garfield (1998, 2003). This measure assesses the correct reasoning of different statistical concepts (e.g., representativeness of data, correlations vs. causation, or independence). The Statistical Reasoning Assessment consists of around 20 questions, but we are unable to field all of them in our survey due to time constraints. Instead, we ask 4 questions on various concepts, aiming to cover as wide a range of concepts as possible.

Our measure can take values from 0 to 4 (0 = correct reasoning for none of the concepts, 4 = correct reasoning for all concepts). Figure 2 shows the distribution of our statistical literacy measures. Statistical literacy (measured by the number of correct answers) generally increases with education, but even among university-educated people there is substantial variation in the number of correct answers. This is interesting and important to separate effects of statistical literacy and general educational attainment on risk perceptions.

While our survey tool on *Statistical reasoning* is admittedly relatively short, we believe that it is able to capture meaningful variation of individuals' understanding and the ability to interpret statistics. Our measure goes beyond pure probability numeracy or general numeracy measures² and aims to capture the understanding of more general concepts such as representativeness of data, the difference between correlation and causation, or the interpretation of probabilities - concepts that are likely to be relevant in the course of the pandemic.

To provide validation for our measure, we compare it to an existing measure of *Probability numeracy* from an earlier survey wave performed in September 2008 and used in prior studies (Bruine de Bruin and Carman, 2011). While this measure focuses on probability numeracy, we try to cover a more broad range of statistical understanding. There is, however, a bit of overlap between the two in terms of concepts (but not questions). The probability numeracy

²Sophisticated calculations are not required when answering.

survey contains questions which required a basic understanding of percentages and probabilities.³ Figure A.1 uses a bin-scatterplot and shows a strong positive correlation between our measure of statistical reasoning and the external measure or probability numeracy. This gives us further evidence that our measure is capturing relevant variation across individuals related to statistical literacy.

Of course, another concern that naturally arises is the "opposite" concern of this, namely that our measure is simply a proxy for general intelligence and numeracy. We attempt to address this in the discussion of the results by controlling for a set of education-related fixed effects and proxies for intelligence and find that our results are robust to controlling for a very tight set of fixed effects. While this is not conclusive because we do not have random variation in statistical literacy, it suggests that statistical literacy is important on its own.

2.2 Questions about COVID-19

We obtain data on COVID-19-related risk assessments, behaviours, and attitudes, from several waves of the LISS panel conducted during the pandemic. We mostly focus on the multi-wave study titled "Effects of the Outbreak of Covid-19", developed by the COVID-19 Impact Lab, a group of researchers at the University of Bonn, University of Lausanne and Tilburg University. All variables are described in detail in the Appendix. The measures we use comprise the following categories.

- ***Probability estimates.*** The first survey wave (March 2020) includes likelihoods of infection, an acquaintance being infected and other measures on a 1-7 scale as well as expectations about the effectiveness of various NPIs to control the virus. Three of the survey waves (April, May and June 2020) include several questions about the respondent's subjective estimates of the probability of being infected, the probability

³The constructed probability numeracy scale runs from 0 (all questions were answered incorrectly) till 13 (all questions were answered correctly). On average subjects scored 10.2 correct answers. The median score was 11. Somewhat over 20% reached the maximum score of 13 and about 12% answered more than half of the questions incorrectly.

of an acquaintance being infected and the probability of being hospitalised if infected, etc..

- ***Social distancing behaviours.*** The March 2020 wave includes measures of self-reported levels of compliance with various recommendations and other behaviours aimed at limiting the spread of COVID-19. These include social distancing, avoiding cafes or public places, wearing masks, etc. We use these to measure social distancing behavior.
- ***COVID-19 case numbers.*** We obtain the locations of the LISS survey participants at the province level. We collect data on local COVID-19 infections and deaths from RIVM (Netherlands National Institute for Public Health and the Environment) and province population data from Statistics Netherlands. We use these to calculate weekly case numbers (adjusted for population) to control for the local severity of the pandemic.

2.3 Background Information

We collect a number of background variables from standardized questions that are repeated periodically and from non-standard questions that were added by various scholars in some previous waves. Our set of controls includes detailed information on age, gender, marital status, education, income, and the number of children.

We also make use of individuals preferences attitudes' towards risk and ambiguity collected from the LISS "Higher Order Risk Attitudes"-wave in 2009 as well as and altruism from the "Unaffiliated spirituality and social engagement" wave of 2009. Unfortunately, the fact that the data were collected in 2009 means that, because of panel attrition, our sample size drops considerably when including these controls. There is also the risk that risk preferences have changed since the data were collected.

3 Statistical Literacy and Risk Assessment

3.1 Statistical Literacy and Infection Rates - Baseline Results

We start by analyzing whether statistical reasoning skills are systematically related to respondents' estimates of infection and mortality risks, as well in their updating of these estimates. We then address main alternative explanations such as education.

Figure 3 shows the average likelihood estimates in March 2020, early on in the pandemic, using a seven-point scale, with seven being the highest likelihood ("Certain"). There is a clear positive correlation between statistical literacy and the estimated likelihood of getting infected in the next two months. More statistically literate people generally estimate the probability of infection to be higher, with those who answered all questions correctly answering on average 4.2 and those who only answered one question correctly answering on average 3.7 on a 1-7 scale. They also assess the health risks of the virus differently, assigning lower likelihoods to be hospitalized after being infected or to die from COVID-19 for otherwise healthy 30-years old individuals. In contrast, they estimate the difference in mortality risk between 75-year-olds and 30-year-olds to be substantially larger. More statistically literate people generally estimate the likelihood of the 30-year-old dying to be lower than less statistically literate people, but the difference in the likelihood of the 75-year-old vs. the 30-year-old dying to be clearly larger. It is worth noting that almost everyone in our sample overestimates the probability of infection and death (for 30-year olds and 75-year olds), which is consistent with people generally overestimating small frequencies (Lichtenstein et al., 1978). However, our results suggest that more statistically literate people understand that older people were much more vulnerable.

We then move on to another dimension of risk estimation, the sensitivity of risk estimates to the actual situation regarding infections. Figure 4 (left) shows the development of the probability estimates (the probability of an acquaintance being infected) over time from April to June for which we have probability estimates, separately for each level of statistical

literacy. The number of wave denotes the month during which the survey wave was performed. As the COVID-19 situation develops, perceived infection risk generally decreases from April to June, with the most statistically literate respondents (those with four correct answers) lowering their estimate of an acquaintance being infected from about 50% in April to 35% in June, while those who answered only one question correctly having initially lower estimates (of 35%) and revising them only to 28%. This is consistent with the decline in the number of new cases, and hence, with the actual infection risk. It is also consistent with adjustments based on new scientific evidence, such as transmission routes, that was discovered during that period. More importantly, the intensity of those adjustments in perceived risks differs by the level of statistical literacy.

To separate effects originating from changes in infection rates from other changes over that time period, we also exploit cross-sectional (regional) differences of our data. In Figure 4 (right), we show that the risk estimates are sensitive to the local number of cases. The figure shows the infection probability estimates depending on the actual infection risk at the province where the respondent resides, measured by the number of new cases per capita in the last week. For more statistically literate people, the estimates of infection risk are clearly more sensitive to the current number of new cases in their home province. The most statistically literate people adjust their estimates of an acquaintance being infected from about 30% in the bottom quartile of provinces by cases per capita to about 50% in the top quartile whereas the least statistically literate people revise their estimates from about 28% in low-infection provinces to 35% in high infection provinces. It is also striking that in low-infection-rate provinces, people have similar estimates of an acquaintance being infected. Taken together, these charts suggest that there are stark differences in either the COVID-19-related information or the interpretation of it, between people of different statistical literacy levels.

To better understand the adjustment of risk estimates to actual risk, and to adjust for potential confounding variables (such as education), we run regressions of the self-estimated

infection risk on the respondent's statistical literacy, the number of new cases, and their interaction. We set up this analysis according to the equation:

$$\begin{aligned}
 P(Infected)_{i,j,t} = & \alpha_0 + \alpha_1 \times Statistical\ reasoning_i \times \ln(New\ cases)_{j,t} \\
 & + \alpha_2 \times Statistical\ reasoning_i + \alpha_3 \times \ln(New\ cases)_{j,t} \\
 & + \alpha_4 \times X_{i,j,t} + \varepsilon_{i,j,t},
 \end{aligned} \tag{1}$$

where i indexes an individual, j a province, and t a the survey time. $P(Infected)$ is the respondent's self-estimated likelihood of getting infected with COVID-19 within the next two months. *Statistical reasoning* is the number of statistical reasoning questions answered correctly, ranging from zero to four, and $\ln(New\ cases)$ is the natural logarithm of the number of new cases per 100,000 in the respondent's home province during the week before the survey response (for all except the first COVID-19 wave, we have the day the respondent started to respond to the question). X is a vector of controls including gender, age and detailed categories of marital status, education, income, and number of children. We describe the data in detail in Table A.1 in the Appendix. We also control for *Wave fixed effects* or *Province $\times$ Wave fixed effects*, capturing any effects of timing and location, as well as *individual fixed effects*, capturing any time-invariant personal characteristics of the respondent.

This regression allows us to study both cross-sectional as well as time differences in risk estimates. The main coefficient of interest is α_1 , which measures the sensitivity of probability estimates to the actual COVID-19 case numbers in the home province of the respondent. A larger positive value suggests a stronger positive relationships between actual infection risk and the perceived risk. We use the same regression specification for other risk estimates.

The results are shown in Table 2. Higher statistical literacy is associated with significantly higher sensitivity of risk estimates to the actual infection levels, as indicated by the positive and statistically significant coefficient on the interaction term between *Statistical Reasoning* and $\ln(New\ cases\ per\ 100k)$. This relationship holds for the estimates of the risk of an

acquaintance being infected as well as risk of the respondent being infected. The baseline sensitivity of estimates of an acquaintance being infected to new cases is such that a 10% increase in cases per capita increases the probability estimate by 1.18% (compared to an average estimate across all waves and areas of 37.7%). For every unit of statistical literacy, this sensitivity increases by 1.74 percentage points, so that the most statistically literate people in our sample update the probability of an acquaintance being infected by 8.15% in response to a 10% increase in cases per capita.

We also run two quasi-placebo tests in columns (5) and (6). Conditional on being infected, the local case numbers should not affect the likelihood of becoming hospitalized. Indeed, we do not find significant difference in the estimated likelihood of hospitalization in infected between people of different statistical literacy levels. On the contrary, people who are more statistically literate estimate a higher likelihood of hospitalization conditional on being infected if they are older. Overall, those results show that more statistically literate individuals adjust their risk assessments in meaningful ways.

Discussion We have shown that statistical literacy is associated with significantly higher sensitivity of risk estimates to actual infection risk. However, statistical reasoning is likely to be correlated with many factors that affect behavior (such as education). To mitigate this issue, we include a rich set of detailed individual-level as well as geographic controls. At the same time, our setup suffers less from concerns related to generalizability that are very common to many experimental settings, which are often based on small (and not necessarily representative) samples and observe outcomes only once. A threat to the internal validity of our analysis would be the presence of omitted variables that are correlated with both statistical reasoning and COVID-19 outcomes. In particular, statistical reasoning might be correlated with educational attainment, general cognitive abilities, or other domain-specific knowledge such as health literacy.

There are several arguments for statistical literacy being important in its own right,

over and above these confounding factors. First, statistical literacy might not directly be explained by certain educational achievements levels. Indeed, most of the questions (see appendix A.3.1)) do not require a specific level of education or sophisticated mathematical knowledge. Moreover, we also observe significant variation in statistical reasoning within different educational achievement levels (as we will discuss below). Second, a strength of the data is the availability of very detailed demographic information. This allows us to include a large set of control variables including detailed information on education and income levels. As we will discuss in more detail below, we also show that the effect of statistical literacy is relatively stable across different educational groups. For robustness, we also perform the analysis excluding individuals with a science or healthcare-related background and our results still hold. Third, some of the outcomes we observe at different points in time and we also know the location of survey participants. This allows us to control for *Wave fixed effects* or *Province $\times$ Wave fixed effects*, capturing any effects of timing and location, as well as *individual fixed effects*, capturing time-invariant personal characteristics of the respondent and therefore highlight differences within an individual's own behavior through the first months of the pandemic.

In the next sections, we also include additional control variables addressing explicitly these potential concerns. We also make use of the availability of the high-dimensionality of control variables and use double/debiased machine learning ("post-double-selection" method) to estimate the treatment effect.

3.2 Educational Achievements

One obvious concern is that education, being correlated with statistical reasoning abilities, might be driving our findings. Note that in our previous analyses, we already controlled for educational attainment. Moreover, while statistical reasoning is correlated with educational achievements, there is still substantial variation of our statistical reasoning measure within these educational groups (see Figure 2). As discussed before, most concepts of sta-

tistical reasoning do not require a specific level of education or sophisticated mathematical knowledge. We can also test directly whether statistical literacy is irrelevant for individuals with a high degree of education, for instance. We, therefore, perform a regression analysis of risk assessments, including separate interaction terms for each of three education levels: secondary, vocational, and polytechnic/university.

The results are shown in Table 3. We see that the interaction term between statistical reasoning and the number of new cases is significantly positive for all three education levels *and* of similar magnitudes. This suggests that, within each education group, statistical literacy is associated with probability estimates being significantly more sensitive to the actual number of cases in the respondent's home province. This shows that the variation captured by statistical literacy does not simply reflect higher educational levels. At the same time, as can be seen from the estimated coefficients for the interaction terms between education categories and the number of new cases, higher education is also associated with probability estimates being more sensitive to the local number of cases.

3.3 Risk Aversion and Ambiguity Aversion

Theories of decision making under risk and uncertainty predict that decisions depend on preferences such as risk aversion or ambiguity aversion. While in theory, there is no clear relationship between those preferences and statistical literacy, there is the concern that, in practice, they might be systematically related. For instance, risk averse individuals might have been more (or less) interested in accumulating human capital related to statistics. In that case, the previous results might be biased by attitudes towards risk and / or ambiguity.

In order to address this concern, we control for risk aversion of Noussair et al. (2014) as well as for an ambiguity aversion index of Dimmock et al. (2016). The results are presented in Table 4. They are generally unchanged and we find that neither the risk nor ambiguity measures seem to strongly predict sensitivity of infection risk to cases. Risk aversion has been measured in a variety of studies and domains in the LISS panel. We focus on a risk

aversion index as well as self-reported risk aversion in the domain of sport which is related to health and, hence, in a similar domain as COVID-19. Results remain unchanged if we consider risk aversion in other domains such as sports (self-reported risk-aversion in the sports domain, as elicited by a 2010 wave in the LISS panel).

3.4 Science and Healthcare workers, Health Status, Political Leaning and Altruism

Besides the concern that our statistical reasoning measure might be capturing simply education, there are some other possible confounding factors that could affect our results. We perform several additional analyses (in Table 5) with the goal of testing whether our results are driven by people working in healthcare or science, or affected by the inclusion of controls from the prior literature on determinants of COVID-spread. As US studies (Painter and Qiu, 2021; Barrios and Hochberg, 2021; Grossman et al., 2020) have found that political leanings affect attitudes and behaviors during the pandemic, we control for political leaning. We also control for a measure of altruism.

We start our analysis by excluding individuals who work in the science or healthcare sector. Those individuals are likely to be more statistically literate and, at the same time, might judge infection risk more accurately because of their domain-specific knowledge. Excluding those workers leave results largely unchanged though. In columns (3) to (6), we use the full sample and include measures of health status, political leaning, and altruism as additional controls, also interacted with the local case numbers. In columns (7) and (8), we control for a measure of altruism, i.e., whether somebody feels or feels not responsible for the well-being of the society. In the last two columns (9) and (10), we include all controls simultaneously. We see that these additional controls do not substantially change our estimates of the effect of statistical literacy and that our results are robust to adding a wide set of controls that have been suggested by the literature to be relevant for different behavior during the pandemic.

3.5 Treatment Effect after Selection among High-Dimensional Controls

One strength of our data lies in the availability of a very large set of potential control variables. As explained in more detail in Section 2, the LISS Survey data consist of a large set of standardized survey waves as well as irregular questionnaires to many diverse topics added by researchers. In the previous two subsections, we included control variables addressing the main concerns identified by us or by previous research. In this section, we are agnostic about the set of control variables, and employ the "post-double-selection" algorithm of Belloni et al. (2014) to select variables from a high-dimensional set of potential controls to further mitigate the concern of an omitted variables bias. Basically, the algorithm finds controls whose omission would lead to a "large" omitted variable bias, and includes them in the regression. The algorithm can be summarized as follows:

1. In the first step, we select a set of control variables that are useful for predicting the treatment d_i . This step helps to insure validity of post-model-selection-inference by finding control variables that are strongly related to the treatment and thus potentially important confounding factors. The set of control variables is called X_1 .
2. In the second step, we select additional variables by selecting control variables that predict y_i . This step helps to insure that we have captured important elements in the equation of interest, ideally helping keep the residual variance small, as well as providing an additional chance to find important confounds. The set of control variables is called X_2 .
3. In the final step, we estimate the treatment effect a_0 of interest by the linear regression of y_i on the treatment d_i and the union of the set of variables selected in the two variable selection steps, i.e., $X_1 \cup X_2$.

Given that our main variable of interest is the interaction term between *Statistical reasoning* and the *New cases per 100k*, we use different specifications in which we interact / do

not interact *Statistical reasoning* as well as all available control variables with the number of *New cases per 100k*. Each of the specifications selects different set of control variables. We report results for each set of control variables as well as for a specification that uses the union of all control variables selected by the algorithm. Our initial set of available control variables includes 180 different candidate controls (categorical variables are dummied out). Please note that we focused on variables related to education including fields, demographic information, income, and migration status (among others) that are collected in the main waves of LISS and, hence, are available for the whole cross-section of individuals.⁴ When we interact those with the number of *New cases per 100k* the set of candidate controls increases to 361.

Table 6 shows the results. Panel a shows the results of a regression of the probability that an acquaintance becomes infected on statistical reasoning, the number of weekly new case, the interaction between the two (our main variable of interest) as well as a set of control variables identified in the selection stage. Please note that we only report the coefficient of main interest because there is no meaningful interpretation of the selected set of control variables. Panel B shows the inputs and outputs of the selection stage. In specifications (1) to (4), we use different specification of interactions of treatment and potential control variables with the number of new Covid cases. In specification (5), we pool control variables selected in specifications (1) to (4). Controls that are picked up throughout the different specifications are information on education (degree and field), occupation, income and wealth, gender, age, marital status, and migration status, mostly interacted with the number of new Corona cases.

For that reason, we compare the estimates of the post-double-selection estimates with the results of our analysis in Table 3 that reports specifications that control for education, income, and demographics as well. In that specifications, depending on the level of education, the

⁴Additional variables including those of smaller, specialized surveys can be added to the set of candidate controls (possibly at the expense of sample size).

estimates of the interaction between statistical reasoning and the number of new weekly cases ranges between 1.0254 and 1.3921. The estimates of that interaction term using the post-double-selection algorithm are between 1.056 - 1.876, and, hence, remain comparable and statistically significant. Overall, those tests increase our confidence of a causal interpretation of our findings.

4 Statistical Literacy and Risk Behaviour

Having studied the relationship between statistical literacy and risk perceptions, we then move to an analysis of whether different assessments of risk also lead to different risk behaviors. We have already shown some evidence (see Figure 1) that individuals who are more statistical literate are more likely to obtain a COVID-19 vaccination, are more likely to recommend a vaccination to friends and assign a lower likelihood to severe side effects from the vaccination using US survey evidence from April-May 2021. We now also perform a regression analysis of social distancing behaviors on statistical literacy in our LISS sample. We do not have time-series data on social distancing, so this regression is purely cross-sectional. We first test the general relationship between behaviors and statistical literacy, using the following specification:

$$\begin{aligned} \text{Maintain distance}_{i,j} = & \alpha_0 + \alpha_1 \times \text{Statistical reasoning}_i \\ & + \alpha_2 \times \ln(\text{New cases})_j + \alpha_3 \times X_{i,j} + \varepsilon_{i,j}, \end{aligned} \quad (2)$$

where the coefficient α_1 measures the general correlation between statistical literacy and likelihood to engage in social distancing, and α_2 measures the correlation between the current COVID-19 case numbers in the respondent's home province and social distancing.

We also test for the relationship between statistical literacy and the sensitivity of these behaviors to the the local COVID-19 situation, using a specification similar to Equation 1. In this specification, we cannot include individual fixed-effects and hence cannot observe

within-person variation. However, the coefficient estimate for α_1 tells us to what extent statistical literacy is associated with the sensitivity of social distancing behaviors to the local COVID-19 case numbers cross-sectionally between different provinces.

We perform these same regressions for three different measures of social distancing: *Maintain distance*, *Avoid public places*, and *Avoid busy places*. We also test whether the propensity of *not* changing the behaviour at all is related to the measure of statistical literacy.

The results are shown in Table 7. As shown in columns 1, 3 and 5, we find that statistical literacy is associated with a significantly higher likelihood of maintaining distance as well as avoiding public and busy places. In columns 2, 4 and 6, we include an interaction term between statistical literacy and the current number of new cases at the respondent's home province. This shows that higher statistical literacy is associated with significantly higher sensitivity of social distancing behavior to the local COVID-19 situation, as indicated by the positive estimated coefficient on the interaction term. In the multivariate setting, every point of statistical literacy is associated with a 1.1% increase in the proportion of people keeping a distance, a 1.24% increase in the proportion of people avoiding public places and a small decrease in the proportion of people reporting no changes to behavior. Columns 2, 4 and 6 show that the sensitivity of these behaviors to the number of new cases per capita also increases. A 10% increase in cases per capita increases the proportion of people keeping a distance by 1.5% among the least statistically literate people (one correct answer) but by 6% among the most statistically literate. Finally, in columns 7 and 8 we examine the relationship between not taking any avoidance behaviors and statistical literacy. Statistically literate people are *less* likely to report not changing behavior, but there is no association between this relationship and the local infection situation.

Relatedly, we had already shown in Figure 1 that more statistical literate individuals are more likely to report willingness to be vaccinated against COVID-19 and recommend a vaccine to friends using a sample of MTurk workers in April-May 2021. We also show those results in a regression analysis in Panel B of Table A.2 in the Appendix, controlling

for gender, age and education categories. The regression results confirm the non-parametric analysis of before.

Summarizing our analysis, we find that more statistically literate people are also more likely to socially distance and to vaccinate, and, consistent with the estimates of infection probabilities, these behaviors are more sensitive to the local situation.

5 Conclusion

Using a large-scale panel survey of about 4,000 Dutch households as well as an online survey with U.S. individuals, we study the role of statistical literacy in people's assessment of risk by analyzing risk expectations and behavior in the context of the COVID-19 pandemic. *Statistical literacy*, the ability to understand and reason with statistics and data, is measured using questions from the Statistical Reasoning Assessment test of Garfield (1998, 2003).

We find that individuals' estimates of COVID-19 infection risk and their social distancing behaviors systematically differ depending on their statistical literacy. More statistically literate people are more likely to socially distance, including maintaining distance to others and avoiding public or busy places. Importantly, we also show that more statistical literate people are more positive towards a vaccination against COVID-19, one of the central and heavily discussed measures in the pandemic.⁵

Studying the sensitivity of those risk estimates to local threats of getting infected (estimated by the current number of coronavirus cases in the region they live in), we show that more statistical literate people adjust both their estimates of risk and their corresponding social distancing behavior more strongly to new information. Those findings suggest that statistical literacy is associated with a better ability to follow changes in the current situation and to incorporate new information in the formation of expectations more accurately. It also seems that better understanding of the risk translates into a more acute sense of protecting

⁵In unreported analyses, we also find that more statistical literate people consider flu shots more important.

others from infection.

Statistical reasoning could potentially be correlated with other factors affecting distancing. The richness of the data, however, allows us to control for a large set of control variables, including detailed information on educational attainment and income, addressing leading alternative explanations. In additional robustness tests, we also control for other potentially relevant factors, including those that have been suggested by recent research. This list includes proxies for science related domain knowledge, risk aversion, political leaning, and general altruism. While we believe that this list is relatively exhaustive, we cannot formally rule out that there are remaining confounding factors. To further address this concerns we also make use of the richness of our data and use methods from machine learning ("post-double-selection" algorithm) to select from high-dimensional controls. Our estimates remain robust.

There are several mechanisms through which statistical literacy might affect risk estimates and behavior. For instance, it may directly affect the ability to interpret available data and statistics. Importantly, we have shown that the possibility to acquire statistical literacy as well its effectiveness for assessing risks seem largely be independent of general educational attainment (see Figure 2 and Table 3). Those findings may carry important policy implications as they suggest that teaching statistical literacy could be an effective way to improve risk assessing capabilities of the general population and may lead to better decision making under risk and uncertainty. Further (experimental) research is needed to study efficient ways to teach statistical literacy in different formal (and informal) settings. Moreover, we strongly believe that statistical literacy is important in many other domains that are unrelated to COVID-19. For instance, household finance decisions are usually decisions under risk, potentially affected by large amounts of data and statistics. Overall, a better understanding of the importance of statistical literacy for decision making across different seems a fruitful area for future research.

References

- Agarwal, Sumit, and Bhashkar Mazumder, 2013, Cognitive abilities and household financial decision making, *American Economic Journal: Applied Economics* 5, 193–207.
- Almenberg, Johan, and Olof Widmark, 2011, Numeracy, financial literacy and participation in asset markets, *Working paper* .
- Barrios, John, and Yael V Hochberg, 2021, Risk perceptions and politics: Evidence from the covid-19 pandemic, *Journal of Financial Economics* 142, 862–879.
- Barrios, John M., Efraim Benmelech, Yael V. Hochberg, Paola Sapienza, and Luigi Zingales, 2021, Civic capital and social distancing during the covid-19 pandemic, *Journal of Public Economics* 193, 1–11.
- Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen, 2014, Inference on treatment effects after selection among high-dimensional controls, *The Review of Economic Studies* 81, 608–650.
- Benjamin, Daniel J., Sebastian A. Brown, and Jesse M. Shapiro, 2013, Who is behavioral? cognitive ability and anomalous preferences, *Journal of the European Economic Association* 11, 1231–1255.
- Bruine de Bruin, Wändi, and Katherine G. Carman, 2011, Measuring risk perceptions: What does the excessive use of 50% mean?, *Medical Decision Making* 32, 232–236.
- Charoenwong, Ben, Alan Kwan, and Vesa Pursiainen, 2020, Social connections with COVID-19-affected areas increase compliance with mobility restrictions, *Science Advances* October 2020.
- Choi, James J, and Adriana Z Robertson, 2020, What matters to individual investors? evidence from the horse’s mouth, *Journal of Finance* 75, 1965–2020.

- Dimmock, Stephen G, Roy Kouwenberg, and Peter P Wakker, 2016, Ambiguity attitudes in a large representative sample, *Management Science* 62, 1363–1380.
- Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde, 2010, Are risk aversion and impatience related to cognitive ability?, *American Economic Review* 100, 1238–60.
- Durante, Ruben, Luigi Guiso, and Giorgio Gulino, 2021, Asocial capital: Civic culture and social distancing during COVID-19, *Journal of Public Economics* 194.
- Fancourt, Daisy, Andrew Steptoe, and Liam Wright, 2020, The Cummings effect: politics, trust, and behaviours during the COVID-19 pandemic, *Lancet* 396, 464–465.
- Garfield, J. B., 1998, The statistical reasoning assessment: development and validation of a research tool, *Proceedings of the 5th International Conference on Teaching Statistics* .
- Garfield, Joan B., 2003, Assessing statistical reasoning, *Statistics Education Research Journal* 2, 22–38.
- Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Stephen Utkus, 2021, The joint dynamics of investor beliefs and trading during the covid-19 crash, *Proceedings of the National Academy of Sciences* 118.
- Graso, Maja, 2022, The new normal: Covid-19 risk perceptions and support for continuing restrictions past vaccinations, *PLoS ONE* 17.
- Grossman, Guy, Soojong Kim, Jonah M. Rexer, and Harsha Thirumurthy, 2020, Political partisanship influences behavioral responses to governors recommendations for covid-19 prevention in the united states, *Proceedings of the National Academy of Sciences* .
- Jasper, John D., Chandrima Bhattacharya, Irwin P. Levin, Lance Jones, and Elaine Bossard, 2013, Numeracy as a predictor of adaptive risky decision making, *Behavioral Decision Making* 26, 164–173.

- Lammers, Joris, Jan Crusius, and Anne Gast, 2020, Correcting misperceptions of exponential coronavirus growth increases support for social distancing, *Proceedings of the National Academy of Sciences* 117, 16264–16266.
- Lichtenstein, Sarah, Paul Slovic, Baruch Fischhoff, Mark Layman, and Barbara Combs, 1978, Judged frequency of lethal events., *Journal of Experimental Psychology: Human Learning and Memory* 4, 551–578.
- Litman, Leib, Jonathan Robinson, and Tzvi Abberbock, 2016, Turkprime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences., *Behavior Research Methods* 1–10.
- Lundborg, Petter, and Henrik Andersson, 2008, Gender, risk perceptions, and smoking behavior, *Journal of Health Economics* 27, 1299–1311.
- Noussair, Charles N., Stefan T. Trautmann, and Gijs van de Kuilen, 2014, Higher order risk attitudes, demographics, and financial decisions, *Review of Economic Studies* 81, 325–355.
- Painter, Marcus O., and Tian Qiu, 2021, Political beliefs affect compliance with government mandates, *Journal of Economic Behavior and Organization* 185, 688–701.
- Pflueger, Carolin, Emil Siriwardane, and Adi Sunderam, 2020, Quarterly journal of economics, *Financial Market Risk Perceptions and the Macroeconomy* 135, 1443–1491.
- Viscusi, W. Kip, 1991, Age variations in risk perceptions and smoking decisions, *Review of Economics and Statistics* 73, 577–588.
- Wu, Sean L., Andrew N. Mertens, Yoshika S. Crider, Anna Nguyen, Nolan N. Pokpongkiat, Stephanie Djajadi, Anmol Seth, Michelle S. Hsiang, John M. Colford Jr., Art Reingold, Benjamin F. Arnold, Alan Hubbard, and Jade Benjamin-Chung, 2020, Substantial underestimation of sars-cov-2 infection in the united states, *Nature Communications* 11.

Tables and Figures

Figure 2: Distribution of Statistical Literacy

These graphs show the distribution of our statistical literacy measure. The left graph shows the distribution of *Statistical reasoning*, calculated as the number of correct responses to our four questions from the Statistical Reasoning Assessment (SRA) Garfield (2003). The right graph shows the distributions of this measure by education categories (secondary education, vocational training, and polytechnic/university education).

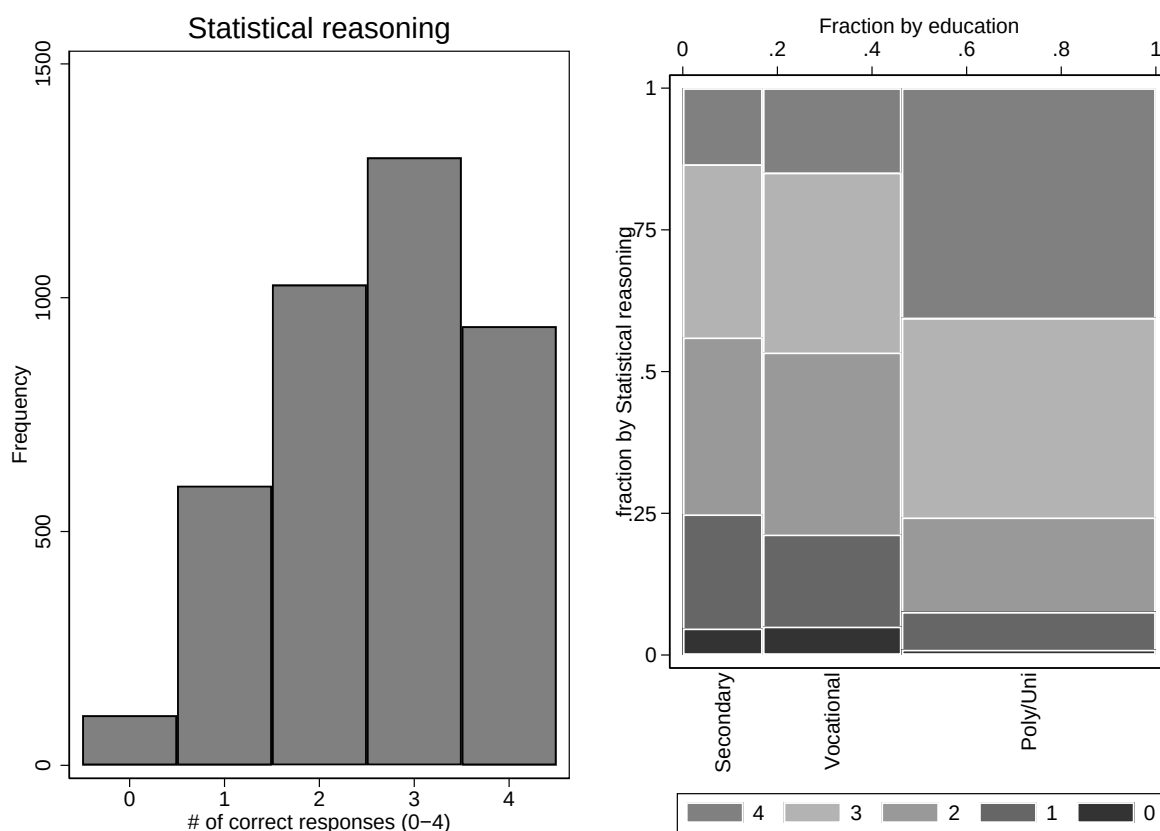


Figure 3: Perceptions of Risk – March 2020

These graphs show binned scatterplots of respondents' expectations of being infected and dying from the Corona virus for different age groups. The left graph presents the subjective likelihood of an infection in March (on a 1-7 scale) by # of correct answers to the 4 SRA questions. The graphs in the middle and to the right present respondents' expectations of the likelihood that an otherwise healthy 30-years-old is dieing from Corona (middle) and the difference in those expectations when compared to a 75-years-old one (right).

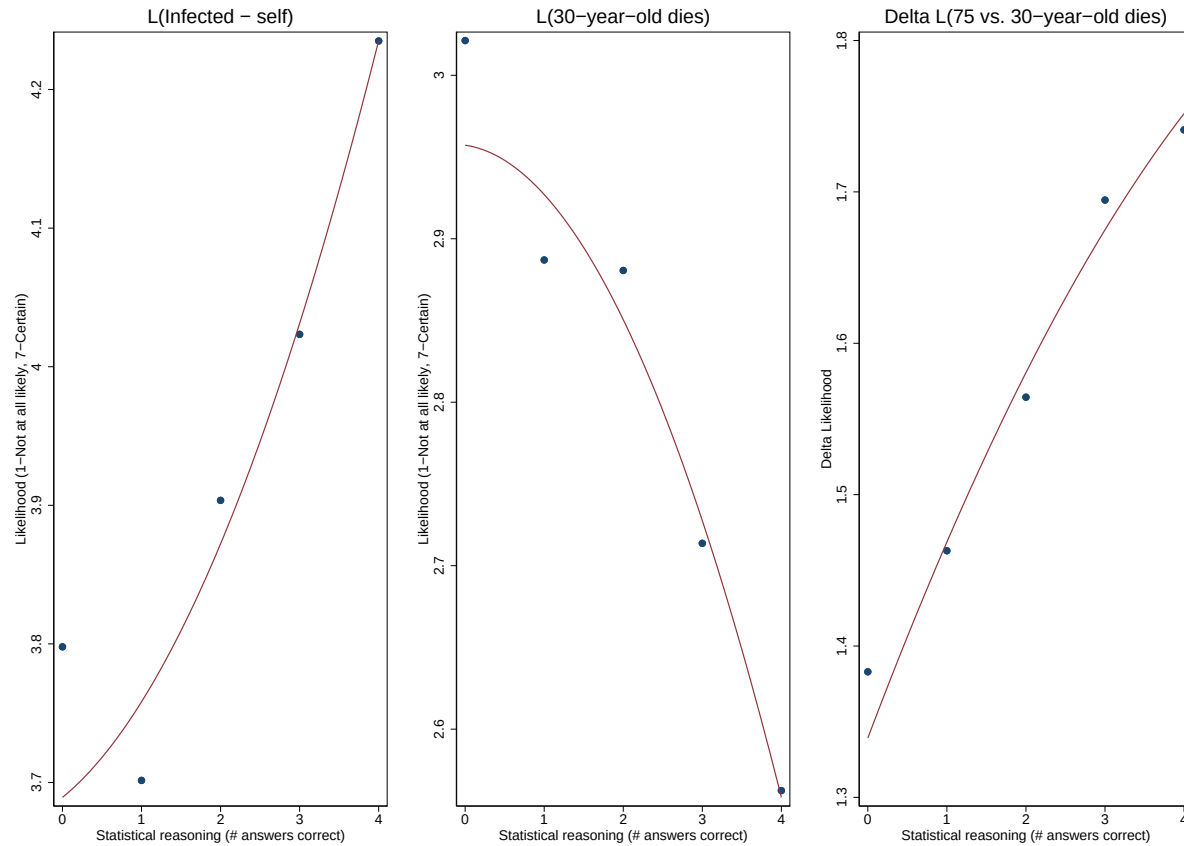


Figure 4: Perceptions of Risk – April-June 2020

These graphs show estimates of the respondent's expectations of the probability being infected and acquaintances being infected, over different survey waves (left) and across regions with different intensity of new cases (right). The number of the wave denotes the month (in 2020) when the wave was performed. $\ln(\text{New cases per capita})$ denotes the number of newly reported Corona infections over the last seven days before the survey was taken, in the region of the participant. The graphs plot different the corresponding estimated probabilities by # of correct answers to the 4 SRA questions.

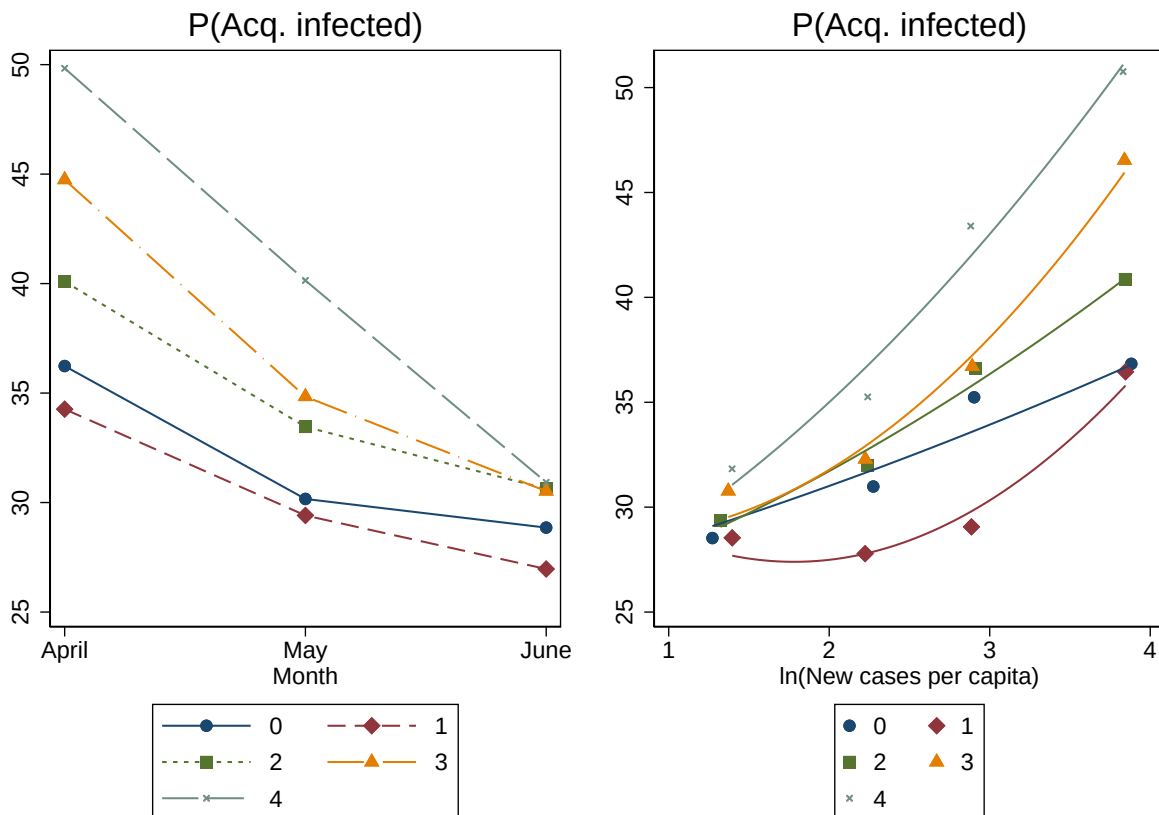


Table 1
Summary statistics

Summary statistics for the LISS survey sample. The unit of observation for COVID-19 panel data is a response (i.e. respondent-wave level). The unit of observation in the cross-sectional data is a respondent - sample sizes differ depending on the availability of the variable in question.

Panel A: COVID-19 panel data

	Mean	Std	p10	p50	p90	N
Estimates						
P(Acq. infected)	35.922	25.103	5.000	30.000	70.000	11,119
P(Infected)	28.785	22.398	2.000	25.000	50.000	11,119
P(Hosp. if infected)	28.950	27.090	0.000	20.000	70.000	11,119
Social distancing						
Maintain distance	0.925	0.263	1.000	1.000	1.000	3,651
Avoid public places	0.804	0.397	0.000	1.000	1.000	3,651
Avoid busy places	0.949	0.219	1.000	1.000	1.000	3,651
No changes	0.009	0.093	0.000	0.000	0.000	3,651
COVID-19 situation						
New cases per 100k	18.946	15.207	4.409	13.698	42.529	17,755
Hospitalisations per 100k	4.510	6.311	0.174	2.361	11.540	17,755
New deaths per 100k	2.017	2.656	0.054	0.774	5.258	17,755

Panel B: Cross-sectional variables

	Mean	Std	p10	p50	p90	N
Statistical literacy						
Statistical reasoning	2.596	1.085	1.000	3.000	4.000	3,972
Demographics						
Age	54.727	17.773	28.000	58.000	75.000	3,972
Male	0.476	0.499	0.000	0.000	1.000	3,972
Married	0.544	0.498	0.000	1.000	1.000	3,972
Number of children at home	0.650	1.037	0.000	0.000	2.000	3,972
Secondary education	0.285	0.451	0.000	0.000	1.000	3,812
Vocational education	0.241	0.428	0.000	0.000	1.000	3,812
Poly/Uni education	0.474	0.499	0.000	0.000	1.000	3,812
Net income ('000)	1.637	2.384	0.000	1.600	2.900	3,966

Table 2

Statistical Literacy and Sensitivity of Beliefs to Actual Infection Risk

The outcome variables in these tables are the self-reported expectations of the probability that an acquaintance gets infected with COVID-19 (columns (1) and (2)), that the respondent gets infected (columns (3) and (4)), and the probability of being hospitalized after a potential infection (columns (5) - (7)). All probabilities are reported in percent (0–100) and collected over three waves between March and June 2020. Statistical reasoning is based on the correct responses to questions from the Statistical Reasoning Assessment (SRA). $\ln(\text{New cases per capita})$ denotes the number of newly reported Corona infections over the last seven days when the survey was taken, in the region of the participant. The specifications include a large set of fixed effects which we describe in detail in Table A.1 in the Appendix.

	P(Acq. inf.)		P(Infected)		P(Hosp. if inf.)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Stat. r. x $\ln(\text{New cases p.c.})$	1.744*** (0.229)	1.696*** (0.213)	1.002*** (0.210)	0.937*** (0.182)	0.003 (0.257)	-0.038 (0.217)	
Statistical reasoning	-2.689*** (0.618)		-2.771*** (0.570)		-0.539 (0.718)		-12.666*** (2.347)
$\ln(\text{New cases per 100k})$	1.185* (0.642)		1.210** (0.603)		1.271* (0.750)		
Stat. r. x $\ln(\text{Age})$							3.108*** (0.601)
$\ln(\text{Age})$							2.327 (1.953)
Gender	Yes	No	Yes	No	Yes	No	Yes
Age cat.	Yes	No	Yes	No	Yes	No	No
Marital status cat.	Yes	No	Yes	No	Yes	No	Yes
Education cat.	Yes	No	Yes	No	Yes	No	Yes
Income cat.	Yes	No	Yes	No	Yes	No	Yes
Children cat.	Yes	No	Yes	No	Yes	No	Yes
Province x Day FE	No	Yes	No	Yes	No	Yes	No
Individual FE	No	Yes	No	Yes	No	Yes	No
N	10,234	10,518	10,234	10,518	10,234	10,518	10,514
R^2	0.125	0.742	0.075	0.747	0.059	0.760	0.058

Table 3
Education

The outcome variables in this table are the self-reported expectations of the probability that an acquaintance gets infected with COVID-19 (columns (1) and (2)), that the respondent gets infected (columns (3) and (4)), and the probability of being hospitalized after a potential infection (columns (5) and (6)). All probabilities are reported in percent (0–100) and collected over three waves between March and June 2020. We include interactions with educational achievement dummies. *Poly/Uni* is a dummy that is equal to one if the respondent has graduated from a polytechnic school or university, *Vocational* is a dummy if the highest degree is a vocational education. *Secondary* is one for the remaining cases. Statistical reasoning is based on the correct responses to questions from the Statistical Reasoning Assessment (SRA). $\ln(\text{New cases per capita})$ denotes the number of newly reported Corona infections over the last seven days. when the survey was taken, in the region of the participant. The specifications include a large set of fixed effects which we describe in detail in Table A.1 in the Appendix.

	P(Acq. inf.)		P(Infected)		P(Hosp. if inf.)	
	(1)	(2)	(3)	(4)	(5)	(6)
Stat. r. x Secondary x $\ln(\text{New cases})$	1.2630*** (0.3140)	0.8892** (0.4003)	1.0695** (0.4223)	0.3318 (0.3507)	0.8806 (0.5356)	0.2835 (0.4508)
Stat. r. x Vocational x $\ln(\text{New cases})$	1.0254*** (0.3230)	0.8746** (0.4207)	0.4154 (0.4161)	0.5146 (0.3818)	0.3621 (0.5163)	0.0307 (0.4375)
Stat. r. x Poly/Uni x $\ln(\text{New cases})$	1.3921*** (0.2615)	1.9164*** (0.3557)	0.7117* (0.3711)	1.0010*** (0.2963)	-0.8107** (0.4136)	-0.7610** (0.3555)
Stat. r. x Secondary	-0.2144 (0.8191)		-1.8642 (1.1785)		-1.9451 (1.5173)	
Stat. r. x Vocational	-0.7627 (0.7787)		-1.5833 (1.0877)		-0.6932 (1.3969)	
Stat. r. x Poly/Uni	-2.3017*** (0.6292)		-2.6161*** (0.9895)		0.3185 (1.1481)	
Secondary x $\ln(\text{New cases})$	1.2842* (0.7063)		-0.1313 (1.0500)		-1.3203 (1.3679)	
Vocational x $\ln(\text{New cases})$	2.6725*** (0.7873)	1.3736 (1.4731)	3.1042*** (1.1389)	1.2389 (1.3222)	0.7012 (1.3920)	1.1823 (1.6610)
Poly/Uni x $\ln(\text{New cases})$	3.3005*** (0.7657)	-0.1261 (1.4800)	2.8003** (1.1902)	0.1205 (1.2761)	3.8575*** (1.3384)	3.4699** (1.6600)
Gender	Yes	No	Yes	No	Yes	No
Age categories	Yes	No	Yes	No	Yes	No
Marital status categories	Yes	No	Yes	No	Yes	No
Education categories	No	No	Yes	No	Yes	No
Income categories	No	No	Yes	No	Yes	No
Children categories	Yes	No	Yes	No	Yes	No
Province x Day FE	No	Yes	No	Yes	No	Yes
Individual FE	No	Yes	No	Yes	No	Yes
N	10,370	10,090	9,824	10,090	9,824	10,090
R ²	0.122	0.747	0.080	0.749	0.061	0.764

Table 4
Risk Aversion and Ambiguity Aversion

The outcome variable in this table is the self-reported expectation of the probability that an acquaintance gets infected with COVID-19. Columns (1) and (2) show estimates for a subsample that excludes individuals with a background in science or health. Columns (3) and (4) include additional controls that have been found relevant in previous studies. All probabilities are reported in percent (0–100) and collected over three waves between March and June 2020. Statistical reasoning is based on the correct responses to questions from the Statistical Reasoning Assessment (SRA). $\ln(\text{New cases per capita})$ denotes the number of newly reported Corona infections over the last seven days when the survey was taken, in the region of the participant. The specifications include a large set of fixed effects which we describe in detail in Table A.1 in the Appendix.

	P(Acq. inf.)					
	(1)	(2)	(3)	(4)	(5)	(6)
Stat. r. x $\ln(\text{New cases p.c.})$	1.651*** (0.392)	2.204*** (0.431)	1.257*** (0.447)	1.228** (0.482)	1.366** (0.542)	1.245* (0.641)
Statistical reasoning	-2.858*** (1.072)		-2.309* (1.237)		-2.638* (1.504)	
$\ln(\text{New cases per 100k})$	-0.007 (1.351)		0.373 (1.652)		0.630 (1.453)	
$\ln(\text{New cases})$ x Risk a.	0.212 (0.237)	0.108 (0.246)				
Risk aversion index	-0.888 (0.637)					
$\ln(\text{New cases})$ x Risk t. (sport)			0.166 (0.208)	0.112 (0.209)		
Risk taking (sport)			-0.576 (0.567)			
$\ln(\text{New cases})$ x Amb. a.					1.818 (1.393)	3.730** (1.544)
Ambiguity Aversion					-4.587 (3.802)	
Gender	Yes	No	Yes	No	Yes	No
Age categories	Yes	No	Yes	No	Yes	No
Marital status categories	Yes	No	Yes	No	Yes	No
Education categories	Yes	No	Yes	No	Yes	No
Income categories	Yes	No	Yes	No	Yes	No
Children categories	Yes	No	Yes	No	Yes	No
Province x Day FE	No	Yes	No	Yes	No	Yes
Individual FE	No	Yes	No	Yes	No	Yes
N	3,376	3,258	2,756	2,556	1,875	1,644
R^2	0.095	0.764	0.111	0.785	0.089	0.797

Table 5

Science and Health Literacy, Political Leaning and Altruism

merge the rows - they are all the same outcome var The outcome variable in this table is the self-reported expectation of the probability that an acquaintance gets infected with COVID-19. Columns (1) and (2) show estimates for a subsample that excludes individuals with a background in science or health (working in these fields). The other columns include additional controls that have been found relevant in previous studies, such as political preferences, social responsibility and self-assessed health status. All probabilities are reported in percent (0–100) and collected over three waves between March and June 2020. Statistical reasoning is based on the correct responses to questions from the Statistical Reasoning Assessment (SRA). $\ln(\text{New cases per capita})$ denotes the number of newly reported Corona infections over the last seven days when the survey was taken, in the region of the participant. The specifications include a large set of fixed effects which we describe in detail in Table A.1 in the Appendix.

	P(Acq. inf.)		P(Acq. inf.)		P(Acq. inf.)		P(Acq. inf.)		P(Acq. inf.)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Stat. r. x $\ln(\text{New cases p.c.})$	1.877*** (0.268)	1.823*** (0.251)	1.753*** (0.235)	1.710*** (0.218)	1.822*** (0.242)	1.882*** (0.225)	0.766* (0.442)	1.654*** (0.494)	0.957** (0.464)	1.833*** (0.527)
Statistical reasoning	-3.096*** (0.723)		-2.733*** (0.634)		-3.085*** (0.659)		-1.153 (1.196)		-1.735 (1.255)	
$\ln(\text{New cases per 100k})$	0.591 (0.742)		-1.199 (1.128)		1.033 (0.684)		5.319*** (1.814)		1.230 (2.724)	
$\ln(\text{New cases}) \times \text{Health}$			0.734** (0.323)	0.653** (0.293)					1.039 (0.634)	0.100 (0.729)
Health			-4.267*** (0.861)						-4.718*** (1.755)	
$\ln(\text{New cases}) \times \text{Pol. L-R}$					0.100 (0.116)	0.246** (0.107)			-0.006 (0.230)	0.259 (0.266)
Pol. L-R					-0.016 (0.305)				0.344 (0.611)	
$\ln(\text{New cases}) \times \text{Not resp. soc.}$							-1.118** (0.528)	-1.160* (0.606)	-0.958* (0.556)	-1.057 (0.646)
Not resp. soc.							2.121 (1.459)	1.873 (1.538)		
Gender	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Age cat.	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Marital status cat.	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Education cat.	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Income cat.	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Children cat.	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Province x Day FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Individual FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
N	7,511	7,714	9,865	10,155	9,307	9,529	2,537	2,323	2,354	2,122
R ²	0.119	0.752	0.125	0.740	0.125	0.743	0.116	0.781	0.121	0.783

Table 6
Post-Double Selection

The outcome variable in this table is the self-reported expectation of the probability that an acquaintance gets infected with COVID-19. All probabilities are reported in percent (0–100) and collected over three waves between March and June 2020. Statistical reasoning is based on the correct responses to questions from the Statistical Reasoning Assessment (SRA). $\ln(\text{New cases per capita})$ denotes the number of newly reported Corona infections over the last seven days when the survey was taken, in the region of the participant. The specifications include a large set of fixed effects as in our main specification (Table (2)). Panel A shows the treatment effect of the interaction term between statistical reasoning and the number of weekly cases. The regressions include the main effects of statistical reasoning and of the number of weekly cases as well as the control variables selected in the selection stage. The details of the selection stage are displayed in Panel B. Treatment is the variable that is predicted in the first step of the algorithm when selecting the first set of control variables. On the second step, variables are selected that predict the outcome of interest (the probability that an acquaintance gets infected). Considered controls provides the number of potential control variables that are considered in the first and second step of the algorithm.

	P(Acq. inf.)				
	(1)	(2)	(3)	(4)	(5)
Panel A: Treatment Effect					
Stat. r. x $\ln(\text{New cases p.c.})$	1.876*** (0.260)	1.056*** (0.280)	1.236*** (0.271)	1.387*** (0.261)	1.113*** (0.275)
Employed controls	27	27 ## $\ln(\text{New cases p.c.})$	33	32	(3) $\cup$ (4) = 38
Fixed effects	Yes	Yes	Yes	Yes	Yes
Panel B: Selection Stage					
Treatment	Stat. r.	Stat. r.	Stat. r.	Stat. r. x $\ln(\text{New cases p.c.})$	N/A
Considered controls	180	180	180 ## $\ln(\text{New cases p.c.})$ = 361	180 ## $\ln(\text{New cases p.c.})$ = 361	N/A
Selected controls	27	27	33	32	38 N/A

Table 7
Social Distancing

The outcome variables in this table are self-reported COVID-19-related behaviors, including *Keeping distance (Yes-No)*, *Avoiding busy places (Yes-No)*, and *Avoiding public places (Yes-No)*. These responses are from a survey performed in March 2020. The last two columns show respondents' views on the *Importance of flu shot* from a survey from 2009, on a 1 (not important) -10 (very important) scale. Statistical reasoning is based on the correct responses to questions from the Statistical Reasoning Assessment (SRA). $\ln(\text{New cases per capita})$ denotes the number of newly reported Corona infections over the last seven days before the survey was taken, in the region of the participant. The specifications include a large set of fixed effects which we describe in detail in Table A.1 in the Appendix.

	Maintain distance		Avoid public places		Avoid busy places		No changes	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Statistical reasoning	0.0110** (0.0050)	-0.0308 (0.0233)	0.0124* (0.0070)	-0.0688** (0.0339)	0.0064 (0.0044)	-0.0283 (0.0180)	-0.0039* (0.0021)	0.0012 (0.0098)
$\ln(\text{New cases per 100k})$	0.0122 (0.0077)		0.0325*** (0.0117)		0.0175*** (0.0062)		-0.0047* (0.0029)	
Stat. r. x $\ln(\text{New cases p.c.})$		0.0152** (0.0077)		0.0274** (0.0111)		0.0116** (0.0059)		-0.0018 (0.0033)
Gender	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age categories	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Marital status categories	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Education categories	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Income categories	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Children categories	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province-Day FE	No	Yes	No	Yes	No	Yes	No	Yes
N	3,416	3,409	3,416	3,409	3,416	3,409	3,416	3,409
R^2	0.054	0.087	0.030	0.074	0.020	0.055	0.020	0.089

Appendix

A.1 Data Appendix

A.1.1 LISS Data - Variable Definitions

Table A.1
Description of Variables

Name	Definition	Source
Panel A: Statistical literacy		
Statistical reasoning	Counter variable that takes the value of 1 for every correct reasoning concept (max 1 per answer). It can therefore have a value from 0 to 4.	LISS Trust in Finance and Financial Technology Wave (September 2020)
Understand statistics	1-5 Scale - "Do you agree with the following statements? - I understand almost all the statistical terms I come across in newspapers or on television."	LISS Trust in Finance and Financial Technology Wave (September 2020)
Often use statistics	1-5 Scale - "Do you agree with the following statements? - I often use statistical information when I form my opinion or make decisions"	LISS Trust in Finance and Financial Technology Wave (September 2020)
Distrust in statistics	1-5 Scale - "Do you agree with the following statements? - Because it is easy to lie with statistics, I don't trust them at all"	LISS Trust in Finance and Financial Technology Wave (September 2020)
Statistics at university	Dummy - 1 if "Attended statistics course at university"	LISS Trust in Finance and Financial Technology Wave (September 2020)
Probability numeracy	1-13 Scale - Constructed based on number of correct responses to numeracy questions	LISS Disease Prevention Wave (September 2008)

Name	Definition	Source
Panel B: Demographics		
Age	Age in years	LISS Background Data
Gender	Dummy (1=Male, 2=Female)	LISS Background Data
Number of children at home	Number of living-at-home children in the household, children of the household head or his/her partner	LISS Background Data
Net income ('000)	Personal net monthly income in euros	LISS Background Data
Political leaning (-5=left 5 = right)	1-10 scale, Converted to -5 to 5. Answer to "In politics, a distinction is often made between "the left" and "the right". Where would you place yourself on the scale below, where 0 means left and 10 means right?"	LISS Politics and Values (January - March 2020)
Not responsible society	1-5 scale - Answer to "I do not feel responsible for the well-being of society (1-disagree entirely to 5-agree entirely)"	LISS Unaffiliated spirituality and social engagement (October 2009)
Self-reported health	1-5 scale - Answer to "How would you describe your health, generally speaking (1-Poor to 5-Excellent)"	LISS Health (November - December 2018)
Risk aversion index	Number of risky choices made (0-5), based on Noursair et al. (2014)	LISS Higher Order Risk Attitudes Wave (December 2009)
Ambiguity aversion index	First-round choices in ambiguity game, based on Dimmock et al. (2016)	LISS Ambiguity Aversion and Household Portfolio Choice (January 2010)
Risk aversion in sport-domains	Willingness to take risk (0-10 scale) in "leisure and sport"	LISS Commercial Opportunities Wave (December 2010)
Marital status categories	Civil status (5 categories)	LISS Background Data
Education categories	Highest level of education (7 categories)	LISS Background Data
Income categories	Income in 500 euro increments (from 500 or less to 7500 or more)	LISS Background Data
Children categories	Number of children living at home (from 0 to 6 or more)	LISS Background Data

Name	Definition	Source
Panel D: Corona		
Maintain distance	Dummy - 1 if "Maintain Distance" is chosen in response to "Which of these recommendations did you follow during the past week? Please check all that apply"	Impact Lab (March)
Avoid public places	Dummy - 1 if "Avoid Public Places"	Impact Lab (March)
Avoid busy places	Dummy - 1 if "Avoid Crowded Places"	Impact Lab (March)
P(Infected)	Percentage - Answer to "On a scale of 0 to 100 percent, how likely do you think that the following events will happen in the the next three months? - You will be infected"	Impact Lab (April-June)
P(Acq. infected)	Percentage - "Someone in your direct environment (family, friends, colleagues) will be infected"	Impact Lab (April-June)
P(Hosp. if infected)	Percentage - "You will have to go to the hospital if you get infected"	Impact Lab (April-June)
Likelihood 30yo dies	1-7 scale - Answer to "How big do you think the chances are that the following people infected with the coronavirus will have to go to the hospital or will die in the next two months because of the coronavirus? - an otherwise healthy 30 year old will die"	Impact Lab (March)
Likelihood 75yo dies	1-7 scale - "A 75 year old with diabetes and chronic cardiovascular disease will die"	Impact Lab (March)
Delta 75yo-30yo dies	Difference (subtracted) in likelihood of 75yo dying and likelihood of 30yo dying	Calculated
New cases per 100k	New weekly cases - Cases on day of survey response minus cases 7 days before response. Survey response date is set to March 20th for the first LISS wave on corona.	RIVM
Hospitalisations per 100k	New weekly hospital admissions - Cumulative admissions on day of survey response minus admissions 7 days before response	RIVM
Deaths per 100k	New weekly deaths - Cumulative deaths on day of survey response minus deaths 7 days before response	RIVM

A.1.2 MTurk Data

We also collect data on Amazon's MTurk platform. MTurk is a platform for Human Intelligence Tasks (HITs), where researchers (and others) pay to have workers complete various tasks. We made use of the CloudResearch (Litman et al. (2016)) platform for administration and participant screening and administered our survey using Qualtrics.

Using MTurk, we surveyed US citizens between the ages of 18-40⁶ who were on CloudResearch's "inclusion list" of workers who had provided consistent answers to surveys in the past. We excluded all MTurkers with rejection rates over 2% as well as the top 2% of survey takers on the platform (in an attempt to exclude people who complete many surveys without paying much attention). Our survey was conducted between 28/04-13/05/21 in three waves.

Our survey aimed to collect demographic information (age, education, state, political views), vaccination status / intention, risk perceptions relating to the virus and the vaccine as well as compliance, with various social distancing recommendations (keeping a distance and wearing a mask indoors). We included the same statistical reasoning test as in the LISS panel ((Garfield, 1998, 2003)). In addition, we included a module where we provided randomized information on the risks of getting vaccinated, varying the level of complexity in the information. Our survey took on average 7-8 minutes to complete. We did not find any differences in the interpretation of information by statistical literacy, which may be because it does not have an effect or because our information treatments were poorly designed. Hence, we do not include the results of any analyses using this data in our main paper (they are in the appendix) and use the MTurk results on social distancing as external validation for our LISS analyses.

We provided a payment of 1.5-2 USD plus a discretionary bonus to participants. The criteria for the bonus were not made explicit but it was suggested that it was conditional on effort, with the idea of inducing people to attempt to answer the statistical reasoning questions as well as possible but without committing us to any specific bonus structure.

⁶Our aim was to target age groups where vaccination rates were not yet high

Table A.2
Mturk Sample Summary Statistics

Panel A shows summary statistics for our MTurk sample. In Panel B, the dependent variable is shown above each column. $P(\text{Infected next 6 months})$ is the estimated probability of getting infected in the next 6 months, $P(\text{Severe side effects})$ is the estimated probability of experiencing severe side effects from the vaccine, Friend vaccine is the likelihood to recommend friends to get vaccinated, and No vaccine is a dummy taking the value one if the respondent does not intend to get vaccinated. Robust standard errors are reported in parentheses.

Panel A: Summary statistics

	Mean	Std	p25	p50	p75	N
Outcomes						
No vaccine	0.250	0.433	0.000	0.000	0.000	1,691
No vaccine (ex. vaccinated)	0.480	0.500	0.000	0.000	1.000	879
Friend vaccine	3.704	1.802	3.000	5.000	5.000	1,736
Probabilities						
P(Infected next 6m)	15.508	18.703	3.000	10.000	20.000	1,720
P(Severe side effects)	16.576	21.039	1.000	5.000	25.000	1,723
Stat. literacy						
Stat. literacy	2.714	0.987	2.000	3.000	3.000	1,737
Demographics						
Male	0.407	0.491	0.000	0.000	1.000	1,733
Age 18-25	0.198	0.399	0.000	0.000	0.000	1,737
Age 26-35	0.586	0.493	0.000	1.000	1.000	1,737
Age 36-45	0.215	0.411	0.000	0.000	0.000	1,737
N	1,737					

Panel B: Risk perceptions

	(1) P(Inf.)	(2) P(Ser. side e.)	(3) Friend vaccine	(4) No vaccine
Stat. literacy	-2.3729*** (0.4818)	-3.2635*** (0.5192)	0.2636*** (0.0435)	-0.0509*** (0.0107)
Male	-4.0597*** (0.8932)	-7.1626*** (0.9540)	0.2729*** (0.0823)	-0.0876*** (0.0199)
Age category FE	Yes	Yes	Yes	Yes
Education category FE	Yes	Yes	Yes	Yes
N	1,715	1,718	1,731	1,686
R^2	0.048	0.088	0.113	0.120

A.2 Robustness of Statistical Reasoning Measure

We also include questions of the *Survey of the Attitudes Toward Statistics (SATS)* to the LISS survey to measure self-assessed attitudes, understanding, and usage of statistics in the daily life.

In Table A.3 we run our base analysis on infection risk using those alternative measures of statistical literacy and attitudes towards statistics. Consistent with the main measure of statistical reasoning, we find that individuals with a higher level of probability literacy, who attended a statistics course at university, who understand statistics and who often use statistics in their daily life are *more* responsive to the number of local infections in their beliefs. At the same time, individuals who distrust statistics more, are *less* likely responsive.

Figure A.1: Correlation Between Statistical Reasoning and Probability Numeracy

These graphs show the distribution of our statistical literacy measures. The left graph shows the distribution of *Statistical reasoning*, calculated as the number of correct responses to our four questions from the Statistical Reasoning Assessment (SRA) of Garfield (2003). The right graph shows the distributions of this measure by education categories.

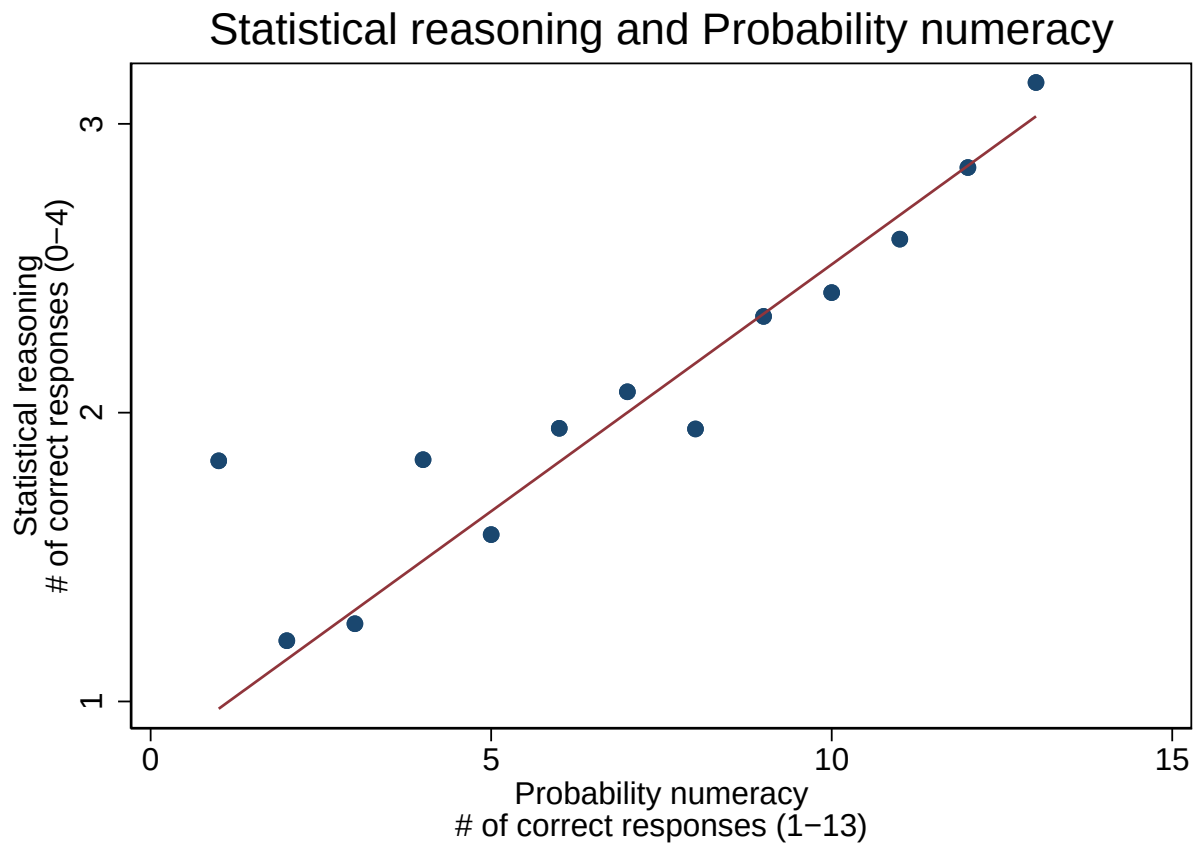


Table A.3
Probability Numeracy and Attitudes Toward Statistics

GET RID OF COLUMNS (7-10) This table shows the self-reported expectations on the probability that an acquaintance gets infected with COVID-19. All probabilities are reported in percent (0–100) and collected over three waves between March and June 2020. We include measures of probability numeracy as well as measures for the attitudes towards statistics, education in statistics, understanding and usage of statistics in daily life as well as distrust in statistics. $\ln(\text{New cases per capita})$ denotes the number of newly reported Corona infections over the last seven days. when the survey was taken, in the region of the participant. The specifications include a large set of fixed effects which we describe in detail in Table A.1 in the Appendix.

	P(Acq. inf.)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Prob. num. x $\ln(\text{New cases})$	0.3636*** (0.0727)	0.9086*** (0.1265)								
Probability numeracy	-0.4801** (0.2297)									
$\ln(\text{New cases per 100k})$	1.8784*** (0.7245)		5.3015*** (0.2589)		3.8689*** (0.5539)		4.6673*** (0.5428)		6.7089*** (0.5872)	
Stat. at uni. x $\ln(\text{New cases})$			4.4028*** (0.8662)	4.5771*** (0.7895)						
Statistics at uni			-10.7720*** (2.5516)							
Understand stat. x $\ln(\text{New cases})$					0.6942*** (0.1874)	1.1113*** (0.2638)				
Understand statistics					-2.3228*** (0.5607)					
Often use stat. x $\ln(\text{New cases})$							0.4376** (0.2003)	0.5746** (0.2717)		
Often use statistics							-0.6482 (0.5915)			
Distrust in stat. x $\ln(\text{New cases})$									-0.3772* (0.2037)	-1.2236*** (0.3165)
Distrust in statistics									0.4732 (0.6252)	
Gender	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Age categories	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Marital status categories	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Education categories	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Income categories	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Children categories	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Province x Day FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Individual FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
N	5,461	5,443	10,227	10,513	10,227	10,513	10,227	10,513	10,227	10,513
R ²	0.099	0.751	0.117	0.741	0.116	0.740	0.116	0.739	0.115	0.740

A.3 Statistical Reasoning

A.3.1 Survey Questions on Statistical Reasoning

Table A.4
Questions From the SRA Test

Question 1: *Which of the following sequences is most likely to result from flipping a fair coin 5 times?*

Answer option	Concept tested	Percent
a) Head, Head, Head, Tail, Tail		1.9%
b) Tail, Head, Head, Tail, Head		3.7%
c) Tail, Head, Tail, Tail, Tail		3.4%
d) Head, Tail, Head, Tail, Head		6.0%
e) All four sequences are equally likely	Correct reasoning: Understanding of independence	85.0%

Question 2: *A small object was weighed on the same scale separately by nine students in a science class. The weights (in grams) recorded by each student are shown below:*

6.2 grams, 6.0 grams, 6.0 grams, 15.3 grams, 6.1 grams, 6.3 grams, 6.2 grams, 6.15 grams, 5.2 grams. The students want to determine as accurately as they can the actual weight of this object. Of the following methods, which would you recommend they use?

Note: After the survey was conducted, the LISS team noticed a mistake. The final measurement, 5.2 grams, was initially meant to be 6.2 grams but was changed due to an unfortunate data-entry error

Answer option	Concept tested	Percent
a) Use the most common number, which is 6.2.		12.9%
b) Use the 6.15 since it is the most precise number.		6.9%
c) Add up the nine numbers and divide by nine.		20.0%
d) Throw out the 15.3, add up the other eight numbers and divide by eight.	Correct reasoning: Understands how to select an appropriate average Correct reasoning: Correctly interprets two-way tables.	60.2%

Question 3: *For one month, 500 elementary students kept a daily record of the hours they spent watching television. The average number of hours per week spent watching television was 28. The researchers conducting the study also obtained report cards for each of the students. They found that the students who did well in school spent less time watching television than those students who did poorly.*

Listed below are several possible statements concerning the results of this research. Please check every statement that you agree with.

Multiple answers are possible.

Answer option	Concept tested	Percent
a) Doing the research with more students would make the results more reliable.		38.4%
b) If a student decreased the amount of time spent watching television, his or her performance in school would improve.		20.1%
c) Even though students who did well watched less television, this doesn't necessarily mean that watching television hurts school performance.	Correct reasoning: distinguishes between correlation and causation	53.3%
d) One month is not a long enough period of time to estimate how many hours the students really spend watching television.		27.9%
e) The research demonstrates that watching television causes poorer performance in school.		8.9%
f) I don't agree with any of these statements.		13.8%

Question 4: *The following message is printed on a bottle of prescription medication: "WARNING: For applications to skin areas there is a 15% chance of developing a rash. If a rash develops, consult your physician."*

Which of the following is the best interpretation of this warning?

Answer option	Concept tested	Percent
a) Don't use the medication on your skin, there is a good chance of developing a rash.		5.6%
b) For application to the skin, apply only 15% of the recommended dose.		1.3%
c) If a rash develops, it will probably involve only 15% of the skin.		1.3%
d) About 15 of 100 people who use this medication develop a rash.	Correct reasoning: Correctly interprets probabilities	61.7%
e) There is hardly a chance of getting a rash using this medication.		