Erasmus School of Economics

BSc International Bachelor Economics and Business Economics

BACHELOR THESIS (FEB13100X)

Behavioural & Health Economics

Debiasing the mind: Langerian mindfulness and our probability judgment accuracy

Name student: Yiwen Chen

Student ID number: 498101

Supervisor: Chen Li

Second assessor: Merel van Hulsen

Date final version: 1/08/2021

Erasmus University Rotterdam
Make it happen.

ABSTRACT

This thesis explores the effect of Langerian mindfulness on improving our probability judgment accuracy. Specifically, the focus is on individuals' probability judgment of the test results of COVID-19. As false positives and negatives are common in the testing for COVID-19, interpretations of the test results are important. The purpose of this research is to test whether Langerian mindfulness can mitigate the cognitive biases we are prone to during probability judgments. In a survey study, respondents were randomly assigned into either the mindful or low-mindful treatment group. Using the Langer Mindfulness Survey (LMS14), the mindful group was found to be significantly more induced by the mindfulness exercises, than the low-mindful group. Incorporating the within-subject design for the individual improvement scores, respondents' estimations before and after the treatment were compared. Results found no significant results for the effect of Langerian mindfulness on respondents' improvement scores. However, with an interaction term, students in the mindful group were found to reduce probability judgment inaccuracies significantly. Policy implications for universities towards students in emotional distress due to the pandemic should advocate for (Langerian) mindfulness practices.

Table of Contents

Chapter 1: INTRODUCTION

According to Yablon (2004), the fundamental truth about probability judgment shows that individuals often struggle to use statistical probabilities to come to consistent and accurate judgments. Inaccurate probability judgments are especially problematic during the current Coronavirus disease 2019 (COVID-19) pandemic, as misperceptions concerning the chance of getting infected may affect people's wellbeing. For instance, implications of COVID-19 on our wellbeing include increased difficulty sleeping and eating due to anxiety (Panchal et al., 2021). To a great extent, anxiety is created by overlooking the inaccuracy of COVID-19 tests (McVean, 2021). Test results only become valuable and interpretable if deemed trustworthy, but is this the case for the COVID-19 tests? This thesis focusses on the interpretations of the COVID-19 test results.

Since the WHO declared COVID-19 a pandemic in March 2020, our society has been facing an unprecedented challenge. Specifically, uncertainties concerning a lot of aspects of our lives increased, giving people a harder time making decisions (Arora et al., 2020). Questions we may ask ourselves include how to deal with the conflicting advice on face masks, which additional personal practices we could employ to protect ourselves and others, whether it is necessary to get tested and in turn, to what extent should we rely on the results of the tests? Do we choose to follow the public policies, or would we rather make personal decisions? And finally, we may ask ourselves whether these choices are justified and rational. *In toto*, the pandemic has brought the ideal conditions for cognitive biases to thrive, leading to inaccurate judgments and beliefs, illogical decisions, and perceptual distortion (Landucci & Lamperti, 2020). Individuals hoarding food and toilet paper at the beginning of the pandemic serves as an example of illogical decision-making, as supply problems were absent. The repercussion that ensued from panicking customers rampaging the aisles left supermarkets with a shortage of toilet paper. This illustrates that such decision flaws led to pernicious consequences. Here, we can identify that the uncertainties and ambiguity in the context of the pandemic is the root of our psychological ramifications and flawed decisions.

Especially in situations where decision-making under uncertainty is especially difficult and tedious, people's descriptive behaviour (i.e., actual) often deviates from the normative standards (i.e., ideal behaviour) (Stanovich & West, 2000). This discrepancy is referred to as the "normative-descriptive" gap and can often be mitigated by behavioural interventions (Thaler and Sunstein, 2008; McDonald, 2008). In fact, in the field of psychological research on decision-making, evidence shows the role of our systematically flawed reasoning and lack of rationality in faulty decisions (Barton & Grüne-Yanoff, 2015). In terms of the consequences of decision-making, Huang (2018) illustrated that individuals display emotional distress, such as anxiety, stress, and negative emotions, leading to our susceptibility to cognitive biases. In the example of individuals hoarding toilet paper, potential stress and anxiety over the novel COVID-19 led to irrational behaviour. The crucial question to ask here is whether we can mitigate such cognitive biases from hindering our decision-making process by limiting our information processing capacity (Landucci & Lamperti, 2020). Therefore, behavioural interventions are necessary as rationality under normative economics is not a common phenomenon in human beings (Stanovich & West, 2000; Ariely, 2009; Kahneman, 2011).

The concept of "nudging" is known to improve people's decision-making skills (Thaler & Sunstein, 2008). According to Thaler and Sunstein, a nudge is described as a concept that indirectly influences an individual's behaviour to make better choices for themselves, without forbidding any actions. However, critics were quick to point out caveats in the theory, namely, that the assumption that "we are hardly educable" lacks sufficient evidence. Gigerenzer (2015) argued that teaching people to be "risk savvy" is a better alternative to nudging. To support this, empirical evidence has found that the theory of choice architecture (*id est* nudges) is unsuccessful at improving an individual's decision-making competency (Selinger & Whyte, 2010). A further suggested alternative, similar to nudges, is called "boosts".

Huang (2018) differentiated boosts from nudges as it aims to improve an individual's decision-making processes, in comparison with merely the outcomes and implicit actions. This is done by focusing explicitly on the analysis of effective decision-making processes, educating individuals to improve their cognitive processes by reducing cognitive biases. For example, an attempt to reduce stress through mindfulness practices, which could be in the form of yoga, guided meditation or breathing, acts as a boost.

This thesis attempts to reduce individual's inaccurate judgments following the approach of a boost. In particular, I will investigate whether increasing people's mindfulness level will lead to an improvement in their probability judgment accuracy.

I will use the Langerian mindfulness, also known as socio-cognitive mindfulness, which is defined as the act of noticing new things and making use of information relevant to a given situation (Langer et al., 1978). Previous studies have found a positive acute effect on learning (Langer et al., 1989), performance (Langer, 2009), problem-solving (Ostafin & Kassman, 2012), attention, and cognitive flexibility (Murphy et al., 2012). This thesis adopts this definition of mindfulness as it is claimed that the attainability of the state of Langerian mindfulness can be achieved instantly. As surveys are the only experimental method available during the current pandemic, the instantaneousness of Langerian mindfulness will be favourable.

Probability judgment accuracy will be studied in the context of the current COVID-19 pandemic. Specifically, I am interested in whether inaccurate probability judgment causes individuals to misperceive the test results from the rapid influenza diagnostic test (RIDT). If so, is it possible to utilize Langerian mindfulness to debias judgment? Thus, the following research question will be investigated:

RQ: "To what extent can Langerian mindfulness improve an individual's probability judgment accuracy?"

To answer this question, this paper will build upon the existing literature. As the pandemic remains prevalent, healthy people will soon have to prove that they are truthfully healthy, employing a rapid test. Already, RIDTs are part of the requirements to enter the Netherlands (Government of the Netherlands, 2021), and may soon become part of our daily routine.

Currently, the PCR tests are still facing many concerns regarding its low sensitivity of 70%, meaning that 30% of people with a positive test in fact are not infected (Good et al., 2020). This means the possibility of false negatives is high. False negatives are particularly concerning as an individual may perceive himself as non-contagious. It has become common knowledge that physicians and patients often place inaccurate confidence in the results of a test (Jha, 2020). The specificity of the PCR test is estimated to be 99.9%, meaning that 99.9% of people who get tested negative, do not have the virus (Good et al., 2020). In other words, the rate of false negatives is 0.1%. To get an appropriate estimation in a certain context, the Bayesian rule has to be applied, as it depends on the sensitivity, specificity, and pre-test probability.

For rapid tests, however, both the sensitivity and specificity are lower than the PCR test, leading to false positives being the main concern. On average, the sensitivity of the RIDTs is 58% for individuals who have symptoms, and 72% for those who are asymptomatic. As the prevalence of the coronavirus is gradually falling, the proportion of false positives increases. This may lead to unwarranted anxiety by wrongfully quarantining individuals and the crowd in which they came in contact with (Halliday, 2021). Furthermore, the U.S. Food and Drug Administration (FDA) expressed its concerns about the rapid test by issuing a letter to the clinical laboratory staff and health providers regarding the potential false positives (Centers for Disease Control and Prevention, 2020). Thus, it is also imperative to identify how the general population estimates their chances of receiving a false positive or negative.

To add to the existing literature, this thesis investigates whether mindfulness can improve an individual's probability judgment accuracy, specifically in rapid tests.

H1: The probability judgment of the mindful group is significantly more accurate than the low-mindful group

Currently, there is limited existing literature on the probability judgment accuracy for COVID-19's rapid test. By investigating the hypothesis, we can learn if mindfulness significantly impacts students' susceptibility to inaccurate probability judgments. According to Cao et al. (2020), the well-being of university students has been hit the severely due to the pandemic. Their data shows a positive correlation between anxiety levels due to the effects of the pandemic, making an individual more prone to cognitive biases. Furthermore, their results illustrate the negative correlation between social support and anxiety levels. From this, the authors highlight the importance of monitoring students' well-being during the pandemic by reducing their stress level. This could be in the form of mindfulness training to reduce cognitive biases, which leads to potential policy implications for universities with online tutoring platforms (OTPs) to reduce emotional distress amongst students. Therefore, it is of social relevance to discover methods to reduce cognitive biases, especially practices to mitigate the psychological consequences of the lockdown due to the current pandemic. To investigate whether being a student has a significant effect on their probability judgment improvement, the hypothesis below will also be tested.

H2: Improvements in probability judgment is significantly higher for students, compared to non-students

This paper is organized as follows. Chapter 2 explores the existing literature on the fundamental concepts of the research question, namely, probability judgment accuracy, Langerian mindfulness, and prior findings on the effect of Langerian mindfulness on cognitive biases. Then, in Chapter 3, the data and methodology are elucidated. Following this, Chapter 4 unpacks the data to form the results section. A discussion on the experiment will be in Chapter 5. Finally, the conclusion can be found in Chapter 6, along with the limitations and discussion, and policy implications of this research.

Chapter 2: LITERATURE REVIEW

2.1. Probability Judgment Accuracy

In the context of decision-making under risk and uncertainty, probability learning, and intuitive statistics, a plethora of theoretical literature on the normative procedure of dealing with risk probabilities is available (Kahneman & Tversky, 1972). However, we have to question whether this information is beneficial for our daily lives as individuals frequently show their inability to accurately make and understand statistical probability judgments (Korobkin & Ulen, 2000). Errors in our probability reasoning have been the focus of modern behavioural economics, demonstrating a caveat in the assumption of a rational agent. Korobkin and Ulen (2000) believe that it is necessary to substitute the rationality assumption with a "more nuanced description" of human behaviour. Indeed, a general conclusion made by Kahneman and Tversky (1972) inferred that people often deviate from the principles of probability theory in decision-making under risk. In turn, inaccurate probability judgments provoke suboptimal decisions.

2.1.1. Subjective Probability Judgment

During our cognitive decision process, subjective probabilities are a critical component (Whitcomb et al., 1995). Subjective probabilities are those derived from an individual's own judgment, typically made according to the most salient characteristic of the sample (Kahneman & Tversky, 1972). Conversely, objective probabilities are calculated according to the laws of probability.

Yablon (2004) argues that heuristics and biases used in an individual's subjective probability judgment do not necessarily lead to their inaccuracy. Instead, the concepts facilitate our decision-making under uncertainty and are typically worthwhile the shortcut. Given the fact that uncertainty provokes human responses such as fear, Yablon (2004) believes that cognitive biases may often be the best available response within the context of uncertainty.

2.1.2. Measuring Probability Judgment

To investigate the research question, the respondents will be given a hypothetical scenario to act as a thought experiment for probability judgment elicitation.

Consider the following thought experiment:

> "There is a type of cancer that afflicts 0.1% of the population. A screening test correctly identifies the cancer 90% of the time when it is truly there, and correctly reports there is no cancer 95% of the time when it is truly not there. You take the screening test, and

it reports that you have the cancer. What is the probability that you really do? (Maymin & Langer, 2021)"

When we try to solve such a problem, we are typically prone to anchoring on large numbers, such as 90%, to then draw erroneous conclusions that the probability of having cancer is relatively certain with a positive screening test outcome. Using the Bayes formula (see Equation (1)), the correct answer would be a 1.77% chance of having this cancer (see Appendix C for calculations). The high probabilities are commonly derived by ignoring the low base rate which, in this case, is that the prevalence of cancer in the population is merely 0.1%. The cognitive bias that is hindering a more accurate intuitive response, is known as the base rate neglect. Although the thought experiment is merely a hypothetical case, a study by Toft et al. (2019) found that false positives in colorectal cancer screening results led to negative psychosocial consequences, such as anxiety. Thus, it may well be that false positive COVID-19 tests could also lead to adverse outcomes of such kind.

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B|A) * P(A) + P(B|A') * P(A')} \tag{1}$$

$$with\ P(A') = 1 - P(A)$$

The hypothetical example given above will be used to elicit individual's probability judgment accuracy. As Eide (2011) argues, the problem does not lie in the failure to use Bayes' rule, but in our intuition when faced with decision-making under uncertainty. To come to a probability estimation, rational agents will take into consideration all the relevant statistics given. As aforementioned, individuals typically neglect the base rates, and do not incorporate these statistics into their estimations. Due to this, they are susceptible to the base rate neglect. In our daily lives, it is not necessary for individuals to be able to know or use Bayes' rule, as it can be used intuitively. Thus, base rate neglect will be represented by an individual's decision quality. Specifically, the respondent's decision quality will be equated with the predictive accuracy of their estimation (Koehler, 1996).

2.1.3. Accuracy of the COVID-19 Tests

To estimate the accuracy of a COVID-19 test, statistics on the prevalence (or pre-test probability) is needed. The prevalence represents the presence of COVID-19 in the given population at a given time (i.e. the number of people who can spread the virus). Imagine a test with the sensitivity and specificity of 70% and 99.9%, respectively. Good et al. (2020) estimated that, if the base rate of infection in the given population is 90% pre-test, then the probability of a false negative is 73%; whereas, if the base rate is 10%, then the chances of having a false negative decreases to 3.2%. False positives, on the other hand, are less common due to the high specificity. Specifically, with 10% and 90% pre-test probability, the PCR test correctly identifies a positive test 98.7% and 100% of the time, respectively.

For rapid influenza diagnostic tests (RIDTs), however, both the sensitivity and specificity are lower than the PCR test, leading to false positives being the main concern. On average, the RIDTs correctly identifies COVID-19 in 72% of individuals with symptoms and in 58% of the asymptomatic cases (Dinnes et al., 2021). The authors utilized summary results of a specific antigen test (SD Biosensor STANDARD Q) as an illustrative example. They estimated that in a population of 10,000 individuals with no symptoms, where the pre-test probability is 0.5%, out of the 125 people who test positive, 90 (72%) individuals would have a false positive, whereas 0.2% of the population have the possibility of a false-negative result.

2.2. Representative Heuristics

Research done by Tversky and Kahneman (1971) illustrate the commonalities between the systematic errors in the intuitive subjective judgments of experienced scientists, and representative heuristics. Representative heuristics are utilized during decision-making, or probability judgments under uncertainty, by using simple, rules-of-thumb to reduce cognitive effort. Then, they proposed a set of biases which belong to the group of representative heuristics (Kahneman & Tversky, 1972). The authors came up with two conditions to identify a representative heuristic. An individual (applying this heuristic) will evaluate the probability of an event happening by the degree to which it is:

1. Similar in essential properties to its parent population, and
2. Reflects the salient features of the process by which it is generated (p. 431)

The higher the degree to which the two conditions holds, the more representative it is deemed. The researchers hypothesized that an event A will be judged more probable than an event B, considering an individual perceives A to be more representative than B. Put differently, an individual's subjective probabilities will be based on the order of representativeness. To illustrate a simple example, consider the following sequences of coin tosses: HTTHTH and HHHHTH. Evidence has shown that, on average, respondents estimate the former sequence to be more probable than the latter, despite the equal likelihood of the sequences. This may be because the first sequence looks more representative of a random process (satisfying condition 2) and there is a balance of heads and tails (satisfying condition 1).

Deviations of subjective probabilities from objective ones can often be attributed to one (or more) of the following cognitive biases, base rate neglect, anchoring and adjustments, or confirmation bias. In this paper, I will focus on base rate neglect.

2.2.1. Base Rate Neglect

In the literature on errors in probabilistic reasoning, base rate neglect was the centre of attention (Tversky & Kahneman, 1973). To define base rate neglect, consider the following illustrative example from a study by Tversky and Kahneman (1973). In the study, participants were given a fictional personality sketch of a graduate student named "Tom W". The task inquired individuals to rank nine different areas of graduate studies, in order of the likelihood that Tom is pursuing that specific field. At that point in time, the most populated field of study was education and humanities. Nonetheless, the results revealed that 95% of the participants estimate a higher likelihood of Tom pursuing computer science compared to education/humanities. Individuals based their probability judgment of how well the description of Tom W fit with the stereotype of a certain industry, and the base rate was not taken into consideration or appreciated enough in this process. What the result suggests is that individuals fail to take the base rate information into account when making probability judgments. But why does this happen?

Bar-Hillel (1984) argues that it is due to relevance, or more specifically, the belief of a lack of relevance. In other words, when we make a judgment, we often disregard any base rate information as we deem it irrelevant for decision-making. What we hold to be relevant, depends on the specificity. When the given information is specific, such as individuating information, we automatically denote it as highly relevant. Conversely, base rate information is, by nature, very general, which may be why we categorize it as low relevance. Broadly speaking, Bar-Hillel (1984) explains the role of the representativeness heuristic in probability estimations under uncertainty. A heuristic, according to Kahneman and Tversky (1972), is a mental shortcut to reduce cognitive load and facilitate decision-making. Although the efficacy is low, it typically suffices for an approximation. Representative heuristics occurs when individuals assess the likelihood of a given scenario based on how it fits in with the characteristics of the population it was drawn from (Kahneman & Tversky, 1972). Consequently, individuals susceptible to base rate neglect also apply representative heuristics.

2.3.Availability Heuristics

As a sequel to their heuristics research, Tversky and Kahneman (1973) explored the availability heuristic – where an individual evaluates the probability of an event by the ease with which relevant information comes to one's memory or imagination. Similar to the representative heuristics, availability heuristics are used as mental shortcuts to come to a decision.

2.3.1. Media & the Availability Cascade

In a study on the judged frequency of lethal events, Lichtenstein et al. (1978) found two biases. The first bias is described by over-estimating few frequencies and underestimating the large ones, whereas the second exaggerates the frequency of specific causes while under-estimating other reasons. This relates to the prior findings by Tversky and Kahneman (1973), where an individual's judgment is fuelled by first-hand experiences but also by indirect exposures to events through the media, movies, television, and more. Kahneman (2011) used the media exposure of aeroplane crashes and celebrity divorces to illustrate that perceived frequency for these two events tend to be exaggerated.

It is interesting to understand the idea of an 'availability cascade' in the context availability heuristics and the media. Economists Sunstein and Kuran (1998) came up with the idea of a self-reinforcing cycle, serving as an explanation to collective belief formation. Let us use COVID-19 as an illustrative example. As of March 2021, the average global infection fatality rate (IFR) is approximated to be around 0.15% (Ioannidis, 2021). Subjective judgment may be influenced by the increasingly exaggerated attention-grabbing headlines, in turn leading to an emotional reaction from individuals, such as stress and anxiety. From this, the media may turn the emotional distress of the population to another headliner, causing individuals to be able to retrieve this information when being asked for it. When the low mortality is considered, the hysteria which media brings to the public seems to be unjustified (Mitzner, 2020).

2.4.Langerian Mindfulness

When it comes to mindfulness, it is necessary to differentiate Langerian mindfulness, which this paper focuses on, from the Buddhist traditions, meditative, and psychometric forms. The latter forms of mindfulness concentrate on both internal and external awareness through the means of focussing on the present moment, unjudgementally. The aforementioned definition of Langerian mindfulness stated that it is the "process of consciously making use of information relevant to the situation" (Langer et al., 1978). Although mindfulness is a multidimensional construct, the variations join consensus in certain qualities (Kabat-Zinn, 2003). Mindlessness, on the other hand, describes a rigid mindset that adheres and relies on one's automatic system, oblivious to other perspectives (Langer, 2009).

Due to its distinct definition, drawing novel distinctions as a result of Langerian mindfulness presents several positive outcomes. For instance, an individual will gain greater sensitivity towards their surrounding environment, along with an enhanced realization of considering multiple perspectives in decision making and problem solving (Langer & Moldoveanu, 2000). Furthermore, the practice invites an individual to improve their openness towards new information. Specifically, the aim is to improve the state of three notions: novelty-producing; novelty-seeking; and engagement.

Since the development of Langerian mindfulness in 1978, the concept has been used as a methodology for investigating improvements in well-being. To name a few, Geng et al. (2019) explored whether mindful learning could improve optimism of cancer patients' family caregivers. Specifically, four sessions of Langerian mindfulness training were provided to the treatment group, resulting in a significant difference in positive feelings relative to the control group. The scales used in this research were the Langer Mindfulness Scale and the Positive Aspects of Caregiving. Similarly, a positive association between an individual's mindfulness and well-being was found by utilizing Langerian mindfulness to improve clinician's perception of control in their clients (Pagnini et al., 2016). Perceived control, defined by the authors, refers to an individual's personal belief about their own capability of "exerting influence on internal states and behaviours, as well as one's external environment" (Langer, 1977). The reason that perception of control indirectly measures well-being is because prior research found that exertion of control lead to better immune responses, increased life satisfaction, and their overall psychological well-being.

Huang (2018) believes that mindfulness has the potential to reduce extensive information search, as the utilization of cognitive attention is enhanced. Furthermore, research has found that socio-cognitive mindfulness has a positive effect on different adult development outcomes, such as one's creativity and decision-making skills (Sternberg, 2000; Pirson et al., 2018). Three notions are used in the Langerian approach towards mindfulness: novelty-seeking, novelty-producing, and engagement (Bodner & Langer, 2001). Novelty-seeking addresses one's curiosity and openness, stimulating an individual's desire to explore. The latter two notions emphasize one's creativity, and a specific way of being involved in an activity, respectively. The purpose of such notions is to improve consistency in the research of mindfulness theory within adult development in social contexts (Pirson et al., 2018).

2.4.1. Langerian Mindfulness and Cognitive Biases

In the existing literature on the effect of Langerian mindfulness and decision-making, Maymin and Langer (2021) outline an arising paradox originating from the relationship between rationality and mindfulness. Rationality, in accordance with normative economics, suggests one correct answer, independent of the context (Oppenheimer, 2008). On the contrary, mindfulness advocates multiple answers depending on the perspective and context. To tackle the uncorroborated paradox, Langer and Newman (1979) argue that it is it is ultimately due to the absence of mindfulness that individual choices are influenced.

Behavioural and cognitive biases, defined by Kahneman (2011), describes systematic deviations from the established rationality in judgments and decision making. Thus, to test whether Langerian mindfulness is effective at mitigating or eliminating cognitive biases, Maymin and Langer (2021) conducted an experiment testing the 22 biases identified by Kahneman (2011). Twenty-two questions testing an individual's susceptibility to the cognitive bias were asked after inducing Langerian mindfulness. Along with, respondents were asked to complete the Langerian mindfulness survey (LMS14) assessing three notions – novelty-seeking, novelty-producing, and engagement. Using three experimental groups, namely a control group, low mindful, and mindful group, the researchers concluded that mindfulness is effective in improving decision-making abilities. In total, respondents in the mindful group were shown to be less biased for 19 out of the 22 biases, including base rate neglect.

Chapter 3: DATA & METHODOLOGY

3.1. Data

To collect data, an experimental survey is conducted using Qualtrics (an online survey platform). Due to the COVID-19 pandemic, measurements are currently restricting in-person meetings, automatically making an online survey the default. The survey is sent out to predominantly university students via social media platforms, but it is not restricted to this occupation only (for the complete survey, see Appendix A). Before the survey, the participants have to complete an informed consent form. In total, 139 observations are collected. However, respondents that either did not consent (*n=9*) or did not complete the survey (*n=27*), are removed. The remaining valid observations (*n=103*) are primarily composed of Dutch (39%) individuals, with the remaining individuals having either an other European (30%) or a non-European nationality (31%). Furthermore, the individuals in the dataset have an average age of 24, with the majority of respondents being female (62%). As mentioned, the survey was generally sent to students, making up 82% of the dataset.

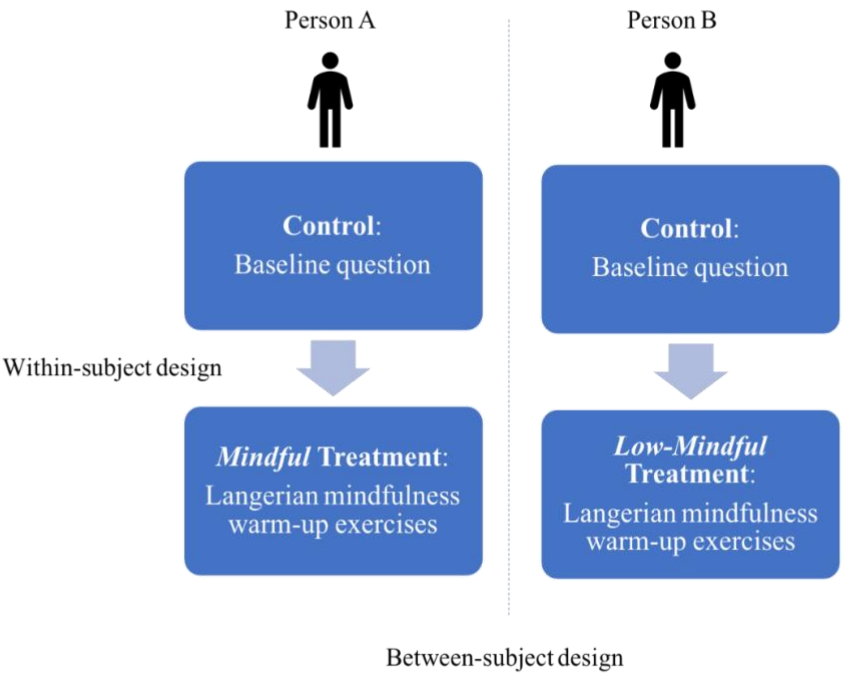
3.2. Methodology

To test the hypothesis, a mixed-subject (one within-subject and one between-subject) randomized controlled trial (RCT) design is employed. This experimental design (see Figure 1) will allow us to investigate within-individual variations over time, as well as potential differences between the two treatment groups (i.e. *Mindful and Low-Mindful*). The comparison between the two groups through repeated measures of probability judgment accuracy will allow for the probability judgment elicitation per individual.

The survey is divided into five parts, namely the consent, baseline question, treatment, manipulation check, and finally the demographic questions (full survey in Appendix A). After giving consent, the respondent is required to answer the baseline question on their probability judgment. A question unrelated to COVID-19 will be asked, namely the example using a screening test to detect a type of cancer (full question in the Chapter 2.1.2.). The purpose of the baseline question is to be able to compare individuals before, and after the treatment, for the within-subject design. Moreover, it allows for the computation of the improvements in accuracy variables at an individual level.

Figure 1

A graphical representation of the experimental design



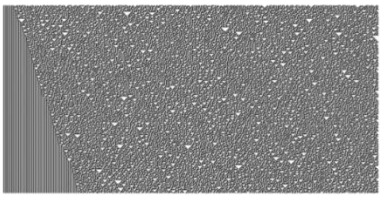
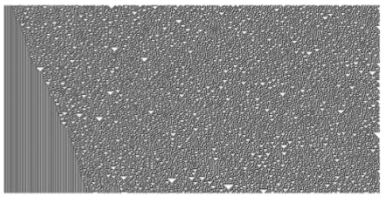
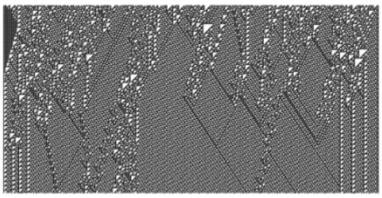
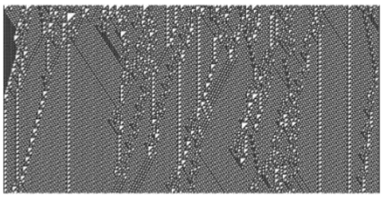
Notes: The figure illustrates the experimental design of this research. Individuals are exposed to both the within- and the between-subject design.

Then, using the randomizer, the individual will be assigned to either the *Mindful* or the *Low-Mindful* group. Both the treatment groups took part in four Langerian mindfulness-inducing exercises as a warm-up, forming the independent variable for the between-subject design (Maymin & Langer, 2021). This paper will follow the procedure of Maymin and Langer (2021) to induce mindfulness. The four exercises include computed images, spot-the-differences, visuals, and noticing new things, where the *Low Mindful* will receive an easier warm-up than the *Mindful* group, as they can, for example, choose to skip a question (see Figure 2, below).

As a manipulation check, the Langer Mindfulness Survey (LMS14) will be used for both the treatment groups to identify whether the warm-up exercises were successful. Pirson et al. (2018) developed the scale such that it is a reliable and valid measure. Similarly, prior studies have reported the robust validity of the scale, and a good internal consistency (Wang et al., 2016; Haigh et al., 2011). To test the internal consistency of the current data, the Cronbach's alpha score will be computed for the LMS14 for the scale as a whole, and the three notions separately.

Figure 2

Computed images for the two treatment groups, Low-Mindful and Mindful

Group	Question	Computed Images	
Low-Mindful	Which image do you prefer? You can skip the question if you like.		
Mindful	Which image do you prefer?		

Notes: The figure shows the first exercise of the Langerian mindfulness-inducing warm-up. The *low-mindful* computed image is generated using Rule 30, and for the *mindful*, Rule 110 is used (Acker, 2010).

The LMS14 scale consists of 14 items based on the three notions: including five novelty-seeking (NS) statements; five novelty-producing (NP); and four for engagement (E) (see Appendix B). The statements assess an individual's curious side, creativity, and activities, respectively (Langer & Moldoveanu, 2000). Participants are then asked to record their response on a 7-point Likert scale, ranging from strongly disagree (1) to strongly agree (7). Each individual will be evaluated based on their total mindfulness score (shown in Equation (1)), where it is notable that item 2 and 14 for the novelty-producing (NPN), and all four of the engagement statements, have the reverse effect, and is thus subtracted from the formula (Maymin & Langer, 2021). Haigh et al. (2011) identified the reverse-scored items and stated that a higher total score for Equation (1) indicates an increased propensity to mindfulness.

$$Mindfulness_i = \sum_{j=1}^{5} NS + \sum_{k=1}^{3} NP_{i,k} - \sum_{l=1}^{4} E_{i,l} - \sum_{m=1}^{2} NPN_{i,m} \qquad (1)$$

Subsequently, the outcome variable will be measured using two questions related to the respondent's probability judgment on COVID-19 rapid test accuracy. For the outcome variables, the first question focuses on false positives of the rapid tests, whereas the second question looks into potential false negatives. The latter question is deemed relevant for policymaking, as individuals with false-negative results should still be cautious not to spread COVID-19. Nine different probability intervals (between 0% - 100%) will be given to the respondents, from which they are required to select one of the options. It is expected that people typically use their subjective probability in surveys, rather than making the formal calculations. Furthermore, research has found that surveys which measure a respondent's probability judgment have traditionally asked for a response between 0 to 100 percent (Bruin & Carman, 2018). The two outcome variable questions are measured in the following two questions.

1. *"The prevalence of COVID-19 in a population is 1%. A rapid test correctly identifies the coronavirus 70% of the time when it is truly there, and correctly reports there is no coronavirus 98% of the time when it is truly not there. You take the rapid test and it reports that you have the coronavirus. What is the probability that you really do? Please pick the answer closest to your best guess of the probability."*

2. *"Now, consider the same statistics as above, but the rapid test reports that you do not have the coronavirus. What is the probability that you do have the coronavirus?"*

At the end of the survey, demographic questions are asked from the participants to perform a baseline test. If there are significant imbalances between the two experimental groups, the variable(s) will be included in the ordinary least squares regression (OLS) as a set of control variables (C) to reduce omitted variable bias. The collected variables will include gender, age, occupation, and nationality. Gender is a dummy variable that will take the value 1 if the individual is male, and 0 if female. Age is measured in years, and occupation takes the value 1 if student, 2 if employed, 3 if unemployed, and 4 if self-employed. Lastly, for nationality, the category is split into three, namely, Dutch, other European, and non-European. Using Dutch as the reference category, an individual's nationality variable will take the value 1 if they belong to one of the aforementioned categories, and 0 otherwise.

In this dataset, it is not guaranteed that randomisation was successful. As the sample size (*n=103*) is relatively small, the law of large numbers may not have taken its full effect. This means that the given sample may not reflect the behaviour of the true population. Ipso facto, a balance test will be conducted using the demographic variables, and the probability estimation of the baseline question, to ensure that the *Mindful* and *Low-Mindful* are, on average, similar before the intervention. Table 1 depicts the summary statistics of the mentioned main variables.

Table 1

Summary Statistics (N = 103)

Variable	Mean	Standard deviation	Min	Max
Baseline Estimation	59.4	37.9	0	95
LMS14 Total Score	24.3	10.2	-10	46
False Positives COVID-19 Estimation	47.6	27.7	0	95
False Negatives COVID-19 Estimation	66.0	38.0	0	100
Age (years)	24.1	7.0	18	60
Gender (male = 1)	0.38	0.49	0	1
Occupation	1.3	0.7	1	4
Nationality				
Dutch	0.39	0.49	0	1
Other European	0.30	0.46	0	1
Non-European	0.31	0.47	0	1

Notes: Baseline Estimation is defined by the mid-value of the interval in the survey. For example, the first option has a mid-value of 0%, the second option between 0%-10% has a mid-value of 5%, and so forth. The mean suggests that the average person estimated a probability of 59%. Mindful takes the value 1 if the individual is in the *Mindful* group, and 0 if in the *Low-Mindful* group. For the LMS14 Total Score, all values for each item were added, and the reverse-scored items were reversed accordingly. As the same choice options were given for the *False Positive COVID-19 Estimation* and *False Negative COVID-19 Estimation*, the summary statistics can be interpreted the same way as for the *Baseline Accuracy*. *Age* is defined in years, and *Gender* is a dummy that takes the value 1 if the respondent is a male, and 0 if female. For occupation, four options were given, namely, student, employed, unemployed, and self-employed, taking the values 1,2,3, and 4, respectively. As the mean is around 1, the main occupation of this sample is student. Nationality is classified into three different subgroup dummies, taking the value 1 if the individual belongs to a nationality within the subgroup, and 0 otherwise.

3.3. Data Analysis

Two preliminary analyses will be performed with the data. First, a baseline test will be conducted to test the success of randomization, using Equation (2) below. The treatment variable (T_i) takes the value 1 if the respondent belongs to the *Mindful* group and 0 for the *Low-Mindful* group.

$$Baseline\ Variable_{ji} = \alpha + \beta T_i + \varepsilon_{ji} \qquad (2)$$

Secondly, a manipulation check of the Langerian mindfulness warm-up will be analyzed using Equation (1) and (3). If the manipulation is successful, a t-test using

$\text{Mindfulness}_\text{i}$ should show a significant difference in comparison to the *Low-Mindful* group for the demographics and the baseline question. From this, we can make the assumption that the unobserved differences are also insignificant if the observed differences are insignificant. This means that that any differences in the outcomes could only have been caused by a difference in treatment, and a causal inference can be made. Furthermore, Cronbach's alpha of the items in the LMS14 scale will be calculated to test for internal consistency and reliability.

$$Mindfulness_i = \alpha + \beta T_i + \sum\nolimits_{j \epsilon C} \gamma_j Control\ Variable_{i,j} + \varepsilon_i \qquad (3)$$

To test *H1*, which states that the probability judgment of the mindful group is significantly more accurate than the low-mindful group, an OLS regression will be performed, depicted in Equation (4), below. A correct answer is defined by choosing the correct probability interval (20-30%), as the true probability of having the coronavirus when being tested positive, is merely 26% (using Bayes formula)). Each interval is converted into its corresponding mid-interval value, where the first option, "Definitely 0%", takes the value 0, the second option, "between 0% and 10%", takes value 5%, and so on. For the full calculations for the three probability judgment questions, see Appendix C. The mid-interval value of the correct interval is subtracted from the respondent's answer to generate their accuracy scores, and in turn improvement scores (see Equation (5), (6), and (7)). To interpret the accuracy score, an answer equal to zero is considered perfectly accurate. Positive values suggest an over-estimation, whereas a negative accuracy score shows an under-estimation by the individual. For the *Mindful* group, an accuracy score closer to zero is expected, relative to the *Low-Mindful* group.

$$Accuracy\_FalsePositive_i = \alpha + \beta T_i + \sum\nolimits_{j \epsilon C} \gamma_j Control\ Variable_{i,j} + \varepsilon_i \qquad (4)$$

For the accuracy variables, the signed integers of the values suggest the direction of inaccuracy presented by the participant. As mentioned, a positive value indicates an over-estimation, whereas a negative value shows an under-estimation of the individual's probability judgment. In this case, it is not ideal to use the absolute value as there will be a loss of information. In further research, the direction of the inaccuracy can be discussed and considered. For example, an overestimation can be explained by base rate neglect, as an individual ignores the base rate.

Then, to conduct the within-subject variable, which is comparing the accuracy score before and after the Langerian mindfulness training, several improvement variables are generated. In this paper, the focus will lie in measuring improvement defined by a reduction of the absolute error, suggesting that an individual becomes less inaccurate. Specifically, for the improvement variables, absolute values need to be taken to measure the magnitude of real numbers (i.e., the distance from zero). As the distance from zero cannot be negative, absolute values are necessary. This will also account for individuals who under-estimated their probability in the false-positive question and then over-estimated in the false-negative question, and vice versa. For example, if an individual scored an accuracy of -10 and 10, taking the absolute values will regard these two values as an equal distance from zero. Consequently, the individual's improvement score would be zero, as the absolute error stays the same. If absolute values are not taken, the individual would obtain an improvement score of 20, which is incorrect.

Having an improvement score of zero either means that the individual made an accurate probability judgment, or that no improvements were made. As the purpose of this paper is to identify the extent to which improvements can be made, the two interpretations do not affect the research. Deviations from zero into the positive integers indicate an improvement in the respondent's probability judgment. Similarly, deviations from zero into the negative integers suggests a deterioration in one's judgment accuracy. This is done for all three probability judgment questions. Subsequently, two improvement variables are created, representing the improvement in their estimation for the probability judgment questions (see Equation (5)). Then, an average improvement for the two questions is taken to analyse the overall progress.

$$Improvement\_FalsePositive_i = |BaselineAccuracy_i| - |Accuracy\_FalsePositive_i| \quad (5)$$

To estimate the average effect of the two levels of Langerian mindfulness exercises, the three improvement variables (generated from questions 1 and 2 separately, and the overall score) are regressed on the treatment dummy. As a reminder, T_i takes the value 1 if the individual is in the *Mindful* group, and 0 if in the *Low-Mindful.* The first regression is depicted in Equation (6), where the same is repeated for Question 2 and the overall score.

$$Improvement\_FalsePositive_i = \alpha + \beta T_i + \varepsilon_i \quad (6)$$

Finally, to test *H2*, further analysis can be conducted by focussing on the students in the dataset. This is done by creating a *Student* dummy, which takes the value 1 if the individual is a student, and 0 otherwise. By generating an interaction term between the new dummy and the *Mindful* group, we can potentially identify the effect which the mindfulness exercises had on students. As mentioned, the emphasis on students is to identify potential policy implications for universities to support their students in these times of crisis. The following Equation (7) depicts the regression described above.

$$Improvement\_FalsePositive_i = \alpha + \beta T_i + \delta Student_i + \gamma T_i * Student_i + \varepsilon_i \quad (7)$$

Chapter 4: RESULTS

4.1. Balance Test

The results of the balance test are displayed in Table 2, shown below. The test results show that there are no significant differences between the *Mindful* and the *Low-Mindful* group. This was measured uses the respondent's baseline estimation and the demographics. We can say that within the dataset, randomisation was successful. Furthermore, the null hypothesis that the sample (*n=103)* comes from the same population cannot be rejected. Due to this, none of the variables shown in Table 2 are used as confounding variables in the OLS regressions, and it can be said that the selection bias is eliminated. When comparing the outcome variables between the two groups (*Mindful* vs. *Low-Mindful*), we can deduce that the difference is attributable to the warm-up exercise level. In this study, the assumption of homogeneity across unobserved characteristics holds, as shown by the results of the balance test. Due to this assumption, we can make causal inferences from the following results.

Table 2

Balance Test

	(1)	(2)	(3)
Variable	Mindful Group	Low-Mindful Group	Difference
Baseline Estimation	56.86	52.02	-4.84
	(5.04)	(5.53)	(7.49)
False Positive COVID-19 Estimation	24.22	20.96	-3.25
	(4.05)	(3.70)	(5.48)
False Negative COVID-19 Estimation	27.65	34.33	6.68
	(5.41)	(5.18)	(7.49)
Age	24.41	23.75	-0.66
	(8.12)	(5.83)	(0.64)
Gender	0.35	0.40	0.05
	(0.483)	(0.495)	(0.599)
Occupation	1.28	1.31	0.03
	(0.64)	(0.76)	(0.81)

Nationality			
Dutch	0.35	0.42	0.07
	(0.48)	(0.50)	(0.47)
Other European	0.31	0.29	-0.03
	(0.47)	(0.46)	(0.78)
Non-European	0.33	0.29	-0.05
	(0.48)	(0.46)	(0.63)
Observations	51	52	103

Notes: Baseline Estimation is defined by the mid-value of the interval in the survey. For example, the first option has a mid-value of 0%, the second option between 0%-10% has a mid-value of 5%, and so forth. *False Positive COVID-19 Estimation* and *False Negative COVID-19 Estimation* are also measured in the same way. All the variables above are regressed on the *Mindful* treatment dummy. Age is measured in years, whereas gender is a dummy that takes the value 1 if the individual is a male, and 0 if female. For occupation, four options were given, namely, student, employed, unemployed, and self-employed, taking the values 1,2,3, and 4, respectively. As the mean is around 1, the main occupation of this sample is student. Nationality is classified into three different subgroup dummies, taking the value 1 if the individual belongs to a nationality within the subgroup, and 0 otherwise. Standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

4.2.Manipulation Check

The second preliminary test, the manipulation check, is presented in Table 3 and 4. This test aims to test the effectiveness of the warm-up exercises (for both the *Mindful* and *Low-Mindful* group), and whether it has a significant effect on the mindfulness level of the individual. As a reminder, the level of mindfulness is calculated using the LMS14 scale.

To interpret the coefficient in Table 3, the OLS regression illustrates that the *Mindful* group has a higher level of mindfulness relative to the *Low-Mindful* group, statistically significant at a five per cent level. Consequently, we can conclude that the warm-up exercises used for the *Mindful* group are more effective in improving an individual’s mindfulness than the simplified version used (for *Low-Mindful*). Specifically, the *Mindful* group had, on average, a higher estimated LMS14 total score of 5.06, in comparison to the *Low-Mindful* group, ceteris paribus. This is statistically significant at a five percent level.

Table 3

Manipulation Check by Regressing the LMS14 on the Mindful Treatment Group

	LMS14
Mindful	5.06**
	(1.96)
Constant	21.78***
	(1.45)
Observation	103
R-squared	.062

Notes: The total score for the LMS14 scale is calculated by summing the items and subtracting the sum of the reverse-scored items. The *Mindful* group takes the value 1 if the individual completed the *Mindful* warm-up exercises, and 0 if they completed the *Low-Mindful* exercises. The sample consists of 103 observations. Robust standard errors are in parentheses *** p<.01, ** p<.05, * p<.1

Furthermore, when the LMS14 scale is analysed separately using the three notions (novelty-seeking, novelty-producing, and engagement), the mindfulness training in the mindful group only had a positive effect on novelty-seeking relative to the low mindful group (at a 1% level). On average, the novelty-seeking items on the LMS14 induced the *Mindful* group more by 3.07 points, compared to the *Low-Mindful.* For the notions in Columns (2) and (3), no significant differences were identified between the two groups.

Table 4

Manipulation Check per Individual Notion

	(1) Novelty-Seeking	(2) Novelty-Producing	(3) Engagement
Mindful	3.07***	0.59	1.40
	(0.95)	(0.79)	(0.88)
Constant	25.26***	7.16***	-10.63***
	(0.76)	(0.49)	(0.68)
Observations	103	103	103
R-squared	.094	.006	.024

Notes: Splitting the items into the three individual notions, each notion is regressed on the *Mindful* group. Robust standard errors are in parentheses *** p<.01, ** p<.05, * p<.1

4.3.Cronbach's Alpha

For the current dataset, the computed Cronbach's alpha coefficient for the LMS14 is 0.793, demonstrating an acceptable level of internal consistency within the scale. Naturally, the items which have the opposite effect on one's mindfulness are reversed manually while calculating Cronbach's alpha. However, item 2 (I generate few novel ideas) is positively correlated with mindfulness despite it being a reverse-scored item. Both options, keeping item 2 as a positive factor or removing the item, results in an increase in the Cronbach's alpha, to 0.798 and 0.821, respectively. On this account, removing item 2 and rerunning the manipulation test results in an increase in the statistical significance level to 1% ($p=0.009$). Besides this, there are no other significant differences. For consistency, item 2 will remain on the scale as reverse-scored. Furthermore, individual Cronbach's alpha for NS, NP, and E was computed, arriving at 0.868, 0.480, and 0.675, respectively.

4.4. Testing the Hypotheses

To test *H1,* which states that the probability judgment of the *Mindful* group is significantly more accurate than the *Low-Mindful* group, the improvement scores are regressed on the *Mindful* group, shown in Table 5. Unfortunately, the three columns do not show any statistically significant results. This shows us that, on average, there are no significant differences in the improvements for the three questions between the *Low-Mindful* and *Mindful* group. The negative coefficient shown in the third column suggests that the *Mindful* group made an overall improvement that is lower, compared with the *Low-Mindful* group (on average, -4.86 lower, ceteris paribus). However, due to the insignificance of the coefficients, no clear interpretations can be made. Therefore, there is insufficient evidence to support *H1*.

Table 5

OLS Regressions on the Improvement of Accuracy for the Outcome Variables

	(1) Improvement False Positive	(2) Improvement False Negative	(3) Improvement
Mindful	1.61	-11.33	-4.86
	(7.36)	(10.41)	(7.86)
Constant	26.18***	29.22***	27.70***
	(4.61)	(7.59)	(5.54)
Observations	103	103	103
R-squared	0	.012	.004

Notes: The table above contains the results of three OLS regressions, on the variables that depict an individual's improvement in their accuracy for question 1, 2, and an average improvement for the two questions. Generating the dependent variables were done in accordance with Equation (5) and (6) in Chapter 3.3. Robust standard errors are in parentheses *** p<.01, ** p<.05, * p<.1

To extent the analysis, the same regressions are performed, but focussing on students only (depicted in Table 6). Adding a *Student* dummy to the regressions allows us to analyse the average effect being a student has on their improvements. In addition, an interaction effect between the *Mindful* group and *Student* is added to identify whether mindfulness has a larger effect on students.

Hypothesis 2 states that improvements in probability judgment is significantly higher for students, compared to non-students. From column (1), it can be identified that, on average, students made a statistically significant improvement from the baseline question, to the question on COVID-19 false positives, at a 5 percent level ($p=0.043$). What could also be interpreted from this, is that students, on average, improved their probability judgment accuracy by 21.37 percentage points, ceteris paribus. For the improvements in the false positive question and overall improvement, being in the *Mindful* group does not have a significant effect on an individual's improvement. Furthermore, the *Student* coefficients in Column (2) and (3) also show a positive coefficient, significant at 10 and 5 percent, respectively. Specifically, the third column, which represents the overall improvement within this experiment, shows that students on average improved significantly more in their accuracy, compared to the other occupations. From this, it can be concluded that there is sufficient evidence to accept *H2*.

Table 6

OLS Regressions on the Improvement of Accuracy (Focusing on Students)

	(1) Improvement False Positive	(2) Improvement False Negative	(3) Improvement
Student	21.37**	35.72**	28.54**
	(10.41)	(17.22)	(13.09)
Mindful	8.22	42.83**	25.53
	(17.32)	(18.38)	(15.50)
Student × Mindful	-8.59	-66.50***	-37.54**
	(19.11)	(21.82)	(17.83)
Constant	9	0.5	4.75
	(9.06)	(15.02)	(11.62)
Observations	103	103	103
R-squared	.035	.072	.048

Notes: The table above contains the results of three OLS regressions, on the variables that depict an individual's improvement in their accuracy for question 1, 2, and an average improvement for the two questions related to COVID-19. Generating the dependent variables were done in accordance with Equation (5) and (6) in Chapter 3.3. To focus the analysis on students, a Student dummy was generated, holding the value 1 if the individual is a student, and 0 otherwise. Robust standard errors are in parentheses. *** p<.01, ** p<.05, * p<.1

In the third row, it can be noted that the interaction effect between *Student* and *Mindful* is statistically significant for Column (2) and (3). Notably, it is important to account for the interaction effect when it is statistically significant. This means that the main effects of *Student* and *Mindful* cannot be independently interpreted in an accurate manner. Thus, from Column (2), the total effect of being a student in the *Mindful* group is equal to 12.06 percentage points from zero, suggesting an overall improvement. In Column (3), although the interaction term is statistically significant, the coefficient for the *Mindful* group is not. Due to this, inferences on the overall effect cannot be made. However, the interaction effect can be interpreted as an average reduction in the total effect by -37.541 percentage points, statistically significant at five percent. Unfortunately, Column (1) does not show a significant effect on one's improvement in their probability judgment accuracy.

Chapter 5: DISCUSSION

5.1. Experimental Design

As the baseline test differed from the outcome variable questions, it is important to elicit the purpose behind this. In the within-subject design, the participants first act as their own control, and then receive the treatment. This allows for a score to be obtained of the individual's initial state. This is important as Langerian mindfulness can be induced instantly (Langer, 1979). In order to avoid the learning effect, the test after the treatment should differ from the initial as the effect gradually increases over time. Furthermore, in psychology, repeated testing was found to affect the outcome of experiments (Cuncic, 2020).

5.2. Results

The results by itself may not substantiate the potential channels for inaccuracies used by the individuals. However, as it is not possible to identify which biases and heuristics the participants were prone to, I will refer to the literature when discussing the channels.

First, I believe that base rate neglect has the largest effect on an individual's inaccurate probability judgments. From the dataset, 78%, 62%, and 70% of the respondents overestimated the baseline question, false-positive, and false-negative question, respectively. With that being said, it is not necessary for an individual to know Bayes formula, but rather have an intuitive approach to an estimation. Simply, the measure which this research uses depicts an individual's predictive accuracy and decision quality. Humans naturally try to use intuition to solve problems. Research by Krynski and Tenenbaum (2007) identified the "causal-Bayes" for probability judgments, stating that overestimations occur when certain statistics do not directly fit into an intuitive causal model. In the first question on COVID-19, the false positive rate that is being estimated is not immediately clear, which often leads to an individual failing to incorporate the information into their calculations. Furthermore, individuals are less likely to disregard base rate information when the person is familiar or experienced with the presented sample (Hogarth & Soyer, 2011). It is typical of news media companies to present data on COVID-19 in the most attention-grabbing way. Although this may increase the number of clicks on the headline, it simultaneously conceals information such as the base rates. Linking this with mindfulness, the outburst of tragic COVID-19 headlines will obfuscate the other important statistics, eventually leading to collective fear. Here, mindfulness plays an important role in bringing our judgements and decisions closer to those of a rational agent. The impact of the media leads to the next discussion point: the availability heuristic.

Availability heuristics are important to consider in the context of COVID-19. The fact that quarantining, curfews, and lockdowns play an essential part in controlling the number of COVID-19 cases may be a reason why people are reading the news more frequently (Molla, 2020). Additionally, social media platforms are increasingly being used as a source of news (Shearer & Mitchell, 2021). As anyone can post anything on social media, the facts are not always accurate. In the Netherlands, a survey with two thousand residents found that 40 percent of the respondents were faced with misleading or inaccurate information regarding the coronavirus (DutchNews, 2021).

6. CONCLUSION

6.1.Conclusion

Even though the COVID-19 cases are currently declining in the Netherlands, the pandemic remains prevalent. As the vaccination coverage is growing (indicating an increase in the percentage of fully vaccinated individuals), rapid tests are heavily employed while attempting to reopen Dutch society. As of the first of July 2021, the Digital Corona Certificate (DCC) will be introduced to the population, giving individuals three options. The options include getting vaccinated, a rapid test that lasts for 40 hours, or a recovery statement (if an individual was infected up to six months prior) (Rijksoverheid, 2021). For the individuals who do not choose to get vaccinated, the rapid tests act as the next best alternative. However, this paper elucidates that the false-positive rates are relatively high, in comparison to false-negatives. Individuals should be aware of their probability judgment inaccuracies as it could lead to psychological distress. Thus, methods to mitigate stress from the pandemic are of substance. Aforementioned, the purpose of this paper was to investigate the following research question:

"To what extent can Langerian mindfulness improve an individual's probability judgment accuracy?"

From the analysis of the data, the two hypotheses were tested. *Hypothesis 1* states that the probability judgment of the mindful group is significantly more accurate than the low-mindful group. *Hypothesis 2* proposed that the improvements in probability judgment is significantly higher for students, compared to non-students. As a result, it was found that there was insufficient evidence to accept *H1*, but *H2* can be accepted. Although the *Mindful* group did not seem to be significantly different in their probability judgment accuracy, robust evidence was found for students. In particular, being a student has a statistically significant effect on an individual's improvement in their probability judgment accuracy. Furthermore, the interaction effect between *Mindful* and *Student* is also statistically significant for the false-negative question and the overall improvement. To answer the research question, Langerian mindfulness did not have a larger effect on the *Mindful* group than the *Low-Mindful* to improve probability judgment accuracy. However, the effect on students seemed to be particularly significant. In fact, interacting *Student* with *Mindful* reduces the extent of an individual's improvement in accuracy. As the coefficient remains positive, it is plausible to assume that students in the *Mindful* group are on average more accurate in their probability judgment. In addition, more than half of the respondents for each question overestimated the probability, suggesting the potential presence of base rate neglect.

In terms of the accuracy and rigorousness of this experiment, the success of the randomization method allowed for the high internal validity of this experiment. A balance test is used to make sure that the observed characteristics between the two groups are statistically indifferent prior to the experiment. This also ensures that demographic heterogeneity is being accounted for. Also, 27 individuals were removed from the data as they did not complete the survey. Attrition may bias the sample as the data consists of participants who chose not to leave the survey in advance.

Furthermore, as high internal validity is traded-off by high external validity, the majority of the sample belonging to the occupation of students limits the external validity of this experiment. Although the number of students in each treatment group is balanced, this data may not be representative of the general population. In order to improve the external validity, the experiment should be repeated with different samples.

6.2. Limitations

In the paper by Maymin and Langer (2021), the notion *Engagement* did not change across the groups. Indeed, this paper found that *Engagement,* along with *Novelty-Producing,* hardly differed between the *Mindful and Low-Mindful* individuals. This was concluded by regressing three notions on *Mindful* independently. Although the answers to the warm-up exercises are irrelevant for the analysis of the hypothesis, it is useful to identify where the problem arose. For instance, the find-the-difference exercise used in the experiment was unsuccessful. From the statistics, 60% of the *Mindful* group either did not see a difference or did not think that there was a difference. This information is useful for further research as the intention was to evoke novelty. The fact that the exercise was ineffective to the majority of the *Mindful* group, may consequently have led to the insignificant results. For the *Low-Mindful* group, a similar occurrence was identified, where over fifty percent of the sample group could not identify the difference between the two images. To compare this to the results of Maymin and Langer (2021), a similar exercise returned contrasting results for the authors. In fact, the majority of their sample were able to identify the difference successfully, where the rest were merely outliers.

Several limitations in the research design may have affected the results. First of all, the *Low-Mindful* group does not act as a perfect control group as the respondents are still induced with some Langerian mindfulness warm-up exercises. The groups merely vary in the extend of their effort that has to be put into answering the survey. To account for this, a within-subject design was adopted, which came with the advantage of a higher statistical power. However, it may have been the case that it led to fatigue effects amongst the participants, which declines the performance due to longer participation and the demanding tasks. The between-subject design of this experiment nominated participants either into the *Low-Mindful* group, or the *Mindful* group. This is in line with the experiment by Maymin and Langer (2021), except that the control group is dropped in this paper. The author found that it could mitigate potential sample selection bias, which could have contaminated the observed effect. However, individuals in the *Low-Mindful* group still received a warm-up treatment. What the effect of this was cannot be determined, as there is no counterfactual. Furthermore, due to the format of the survey being online, the warm-up exercise to notice new things also brought its limitations. For example, the results showed answers of unseriousness, where individuals either put 'nothing', or an inadmissible answer. Although the specific answers did not matter, the exercise should have been answered carefully, in order for it to stimulate the right notions.

6.3.Policy Implications

Determining the extent to which individuals could improve their probability judgment inaccuracy could be important for universities and policymakers. As mentioned in the introduction and literature review, utilizing 'boosts' to effectively educate individuals on their decision-making processes could potentially promote an adaptation of growth mindsets (Huang, 2018). This paper illustrates the benefits of using the Langerian mindfulness method as a boost.

Focussing on students, potential policies can be adopted as the current state of online tutoring may hinder decision-making, and in turn, mindfulness. Langerian mindfulness, in its natural state, is easy to immediately operationalize, which is why universities could adopt this method into their well-being programs for their students. This research shows that the method was especially effective on students, compared to non-students.

6.4.Further Research

Further research should focus on addressing the limitations aforementioned. Specifically, more attention should be paid to diving deeper into warm-up exercises which focus on engaging both the notions of *Novelty-Producing* and *Engagement* more. In order to resolve the problems caused by the online format of the survey, a field experiment should be conducted so that individuals are more motivated to complete the warm-up exercises with accuracy. This paper only speculates the causes and biases of probability judgment inaccuracies, more robust research should be done to expand our knowledge on the reasons for inaccuracies. In addition, this research cannot conclude anything associated with the long-term effects of the Langerian mindfulness.

BIBLIOGRAPHY

Acker, D. (2010). Elementary Cellular Automata. Retrieved from https://devinacker.github.io/celldemo/

Ariely, D. (2009). The end of rational economics. Harvard business review, 78-84.

Arora, A., Jha, A. K., Alat, P., & Das, S. S. (2020). Understanding coronaphobia. Asian Journal of Psychiatry, 54.

Bar-Hillel, M. (1984). Representativeness and fallacies of probability judgment. Acta Psychologica, 91-107.

Barton, A., & Grüne-Yanoff, T. (2015). From Libertarian Paternalism to Nudging—and Beyond. Review of Philosophy and Psychology, 341–359.

Bruin, W. B., & Carman, K. G. (2018). Measuring Subjective Probabilities: The Effect of Response Mode on the Use of Focal Responses, Validity, and Respondents' Evaluations. Risk Analysis, 2128.

Cao, J., Yang, T., Lai, I. K., & Wu, J. (2020). Is online education more welcomed during COVID-19? An empirical study of social impact theory on online tutoring platforms. International Journal of Electrical Engineering & Education, 1-12.

Centers for Disease Control and Prevention. (2020). Division of Laboratory Systems (DLS). Retrieved from Centers for Disease Control and Prevention (CDC): https://www.cdc.gov/csels/dls/locs/2020/fda_issues_safety_letter_about_false_positive_results_with_antigen_tests_for_rapid_detection_of_sars-cov-2.html

Cuncic, A. (2020). Understanding Internal and External Validity. Retrieved from Very Well Mind: https://www.verywellmind.com/internal-and-external-validity-4584479

Dinnes, J., JJ, D., S, B., M, T., A, A., C, D., . . . MDF. (2021, 3 24). How accurate are rapid tests for diagnosing COVID-19? Retrieved from Cochrane: https://www.cochrane.org/CD013705/INFECTN_how-accurate-are-rapid-tests-diagnosing-covid-19

DutchNews. (2021). More people are reading the news, and have more confidence in what they read. Retrieved from DutchNews.nl: dutchnews.nl/news/2021/06/more-people-are-reading-the-news-and-have-more-confidence-in-what-they-read/

Eide, E. (2011). Two tests of the base rate neglect among law students. Oslo: University of Oslo.

Geng, L., Wang, J., Cheng, L., Zhang, B., & Shen, H. (2019). Mindful Learning Improves Positive Feelings of Cancer Patients' Family Caregivers. International Journal of Environmental Research and Public Health.

Gigerenzer, G. (2015). On the Supposed Evidence for Libertarian Paternalism. Review of Philosophy and Psychology, 361–383.

Good, C. B., Hernandez, I., D, P., & Smith, K. (2020). Interpreting COVID-19 Test Results: a Bayesian Approach. J Gen Intern Med, 2490–2491.

Government of the Netherlands. (2021). Requirements for negative rapid test result. Retrieved from Government of the Netherlands: https://www.government.nl/topics/coronavirus-covid-19/visiting-the-netherlands-from-abroad/mandatory-negative-test-results-and-declaration/requirements-rapid-test-result

Haigh, E. A., Moore, M. T., Kashdan, T. B., & Fresco, D. M. (2011). Examination of the factor structure and concurrent validity of the Langer Mindfulness/Mindlessness Scale. Assessment, 11-26.

Halliday, J. (2021, 4). Rapid Covid testing in England may be scaled back over false positives. Retrieved from The Guardian: https://www.theguardian.com/world/2021/apr/15/rapid-covid-testing-in-england-may-be-scaled-back-over-false-positives

Hogarth, R. M., & Soyer, E. (2011). Sequentially simulated outcomes: Kind experience versus nontransparent description. Journal of Experimental Psychology, 434.

Huang, P. H. (2018). Boost: Improving Mindfulness, Thinking, and Diversity. the Colorado Law Faculty Scholarship.

Ioannidis, J. P. (2021). Reconciling estimates of global spread and infection fatality rates of COVID-19: An overview of systematic evaluations. European Journal of Clinical Investigation.

Jha, S. (2020, 3 31). False Negative: COVID-19 Testing's Catch-22 — And the consequences of being wrong. Retrieved from MedPage Today: https://www.medpagetoday.com/infectiousdisease/covid19/85717

Kabat-Zinn, J. (2003). Mindfulness-Based Interventions in Context: Past, Present, and Future. American Psychological Association.

Kahneman, D. (2011). Thinking, fast and slow. MacMillan.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. Cognitive Psychology, 430-454.

Koehler, J. J. (1996). The base rate fallacy reconsidered: Descriptive, normative, and methodological changes. Behavioral and brain sciences, 1-53.

Korobkin, R. B., & Ulen, T. S. (2000). Law and Behavioral Science: Removing the Rationality Assumption from Law and Economics. California Law Review, 1051-1144.

Krynski, T. R., & Tenenbaum, J. B. (2007). The role of causality in judgment under uncertainty. Journal of Experimental Psychology: General, 430.

Landucci, F., & Lamperti, M. (2020). A pandemic of cognitive bias. Intensive Care Medicine, 636–637.

Langer, E. (2009). Counter clockwise. New York: Ballantine Books.

Langer, E. J., & Moldoveanu, M. (2000). The Construct of Mindfulness. Journal of Social Issues.

Langer, E. J., & Newman, H. M. (1979). The role of mindlessness in a typical social psychological experiment. Personality and Social Psychology Bulletin, 295-298.

Langer, E., Blank, A. E., & Chanowitz, B. (1978). The mindlessness of ostensibly thoughtful action: The role of "placebic" information in interpersonal interaction. Journal of Personality and Social Psychology, 635-642.

Lichtenstein, S., Slovic, P., Fischhoff, B., Layman, M., & Combs, B. (1978). Judged Frequency of Lethal Events. Journal of Experimental Psychology: Human Learning and Memory, 551.

Maymin, P. Z., & Langer, E. J. (2021). Cognitive biases and mindfulness. Humanities and Social Sciences Communications.

McDonald, I. M. (2008). Behavioural Economics. The Australian Economic Review, 222-228.

McVean, A. (2021). When A Negative Doesn't Really Mean Negative: Why Quality Matters Just As Much As Quantity When Testing For COVID-19. The Skeptical Inquirer.

Mitzner, D. (2020, 10 19). Daniel Kahneman, COVID-19 and the cascade of fear. Retrieved from Jewish News Syndicate: https://www.jns.org/opinion/daniel-kahneman-covid-19-and-the-cascade-of-fear/

Molla, R. (2020). Everybody is reading the news more because of coronavirus. Retrieved from Vox: https://www.vox.com/recode/2020/3/17/21182770/news-consumption-coronavirus-traffic-views

Murphy, F. C., Michael, A. S., & B. J., A. (2012). Emotion modulates cognitive flexibility in patients with major depression. Psychological medicine.

Oppenheimer, J. A. (2008). Rational choice theory. The Sage Encyclopedia of Political Theory.

Ostafin, B. D., & Kassman, K. T. (2012). Stepping out of history: Mindfulness improves insight problem solving. Consciousness and Cognition.

Pagnini, F., Bercovitz, K., & Langer, E. (2016). Perceived control and mindfulness: Implications for clinical practice. Journal of Psychotherapy Integration, 91.

Panchal, N., Kamal, R., Cox, C., & Garfield, R. (2021). The Implications of COVID-19 for Mental Health and Substance Use. Coronavirus (COVID-19).

Pirson, M. A., Langer, E., & Zilcha, S. (2018). Enabling a Socio-cognitive Perspective of Mindfulness: The Development and Validation of the Langer Mindfulness Scale . Journal of Adult Development.

Rijksoverheid. (2021). Coronabewijs voor reizen en evenementen met CoronaCheck. Retrieved from Rijksoverheid: https://www.rijksoverheid.nl/onderwerpen/coronavirus-covid-19/coronabewijs

Selinger, E., & Whyte, K. P. (2010). Competence and Trust in Choice Architecture. Knowledge, Technology, & Policy, 461-482.

Shearer, E., & Mitchell, A. (2021). News Use Across Social Media Platforms in 2020. Retrieved from Pew Research Center: https://www.journalism.org/2021/01/12/news-use-across-social-media-platforms-in-2020/

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? Behavioral and brain sciences, 645-665.

Sternberg, R. J. (2000). Images of Mindfulness. Journal of Social Issues.

Sunstein, C. R., & Kuran, T. (1998). Availability Cascades and Risk Regulation. Stanford Law Review.

Thaler, R. H., & Sunstein, C. R. (2008). Nudge: Improving Decisions About Health, Wealth, and Happiness. Chicago: Yale University Press.

Toft, E. L., Kaae, S. E., Malmqvist, J., & Brodersen, J. (2019). Psychosocial consequences of receiving false-positive colorectal cancer screening results: a qualitative study. Scandinavian Journal of Primary Health Care, 145-154.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. Cognitive Psychology, 207-232 .

Wang, X., Geng, L., Zhou, K., Ye, L., Ma, Y., & Zhang, S. (2016). Mindful learning can promote connectedness to nature: Implicit and explicit evidence. Consciousness and Cognition, 1-7.

Whitcomb, K. M., Önkal, D., Curley, S. P., & Benson, P. G. (1995). Probability judgment accuracy for general knowledge. Cross-national differences and assessment methods. Journal of Behavioral Decision Making, 51-67.

Yablon, C. (2004). The Meaning of Probability Judgments: An Essay on the Use and Misuse of Behavioral Economics. U. Ill. L. Rev., 899.

APPENDICES

Appendix A: Survey

INTRODUCTION

Dear participant, welcome!
Thank you for participating in this survey.

This survey will be used for my bachelor thesis at the Erasmus School of Economics, to study the effect of mindfulness on decision making. Your response will be kept fully anonymous. I kindly ask you to answer as truthfully as possible. The survey will take you around 8-10 minutes.

If you have any questions, please contact me: 498101yc@student.eur.nl

CONSENT

I consent to participating in the research study as described above.

- Yes, I consent to participating in the research study
- No, I don't consent and I will not participate in the research study

BASELINE QUESTION

Please note that the following scenario is hypothetical.

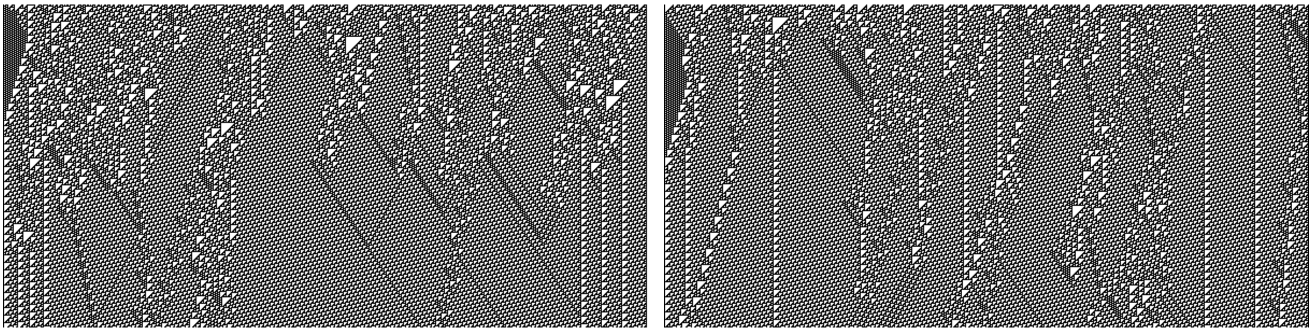
There is a type of cancer that afflicts 0.1% of the population. A screening test correctly identifies the cancer 90% of the time when it is truly there, and correctly reports that there is no cancer 95% of the time when it is truly not there. You take the screening test and it reports that you have the cancer. What is the probability that you really do?

Please pick the answer closest to your best guess of the probability.

- Definitely 0% you do not think there is any chance you have this type of cancer (0%)
- Between 0 and 10% (5%)
- Between 10 and 20% (15%)
- Between 20 and 30% (25%)
- Between 30 and 40% (35%)
- Between 40 and 50% (45%)
- Between 50 and 60% (55%)
- Between 60 and 70% (65%)
- Between 70 and 80% (75%)
- Between 80 and 90% (85%)
- Between 90 and 100% (95%)
- Definitely 100%—you are completely sure you have this type of cancer (100%)

TREATMENT GROUP: MINDFUL

1. COMPUTED IMAGE: Which image do you prefer?



2. SPOT-THE-DIFFERENCE: Can you spot the difference?

3. RUBIN'S VASE: What can you see here?

- o I see the vase on the left and the vase on the right, but no faces anywhere.
- o I see only the vase on the left, and the vase and faces on the right.
- o I see only the vase on the right, and the vase and faces on the left.
- o I see the vase and faces on the left and the vase and faces on the right.

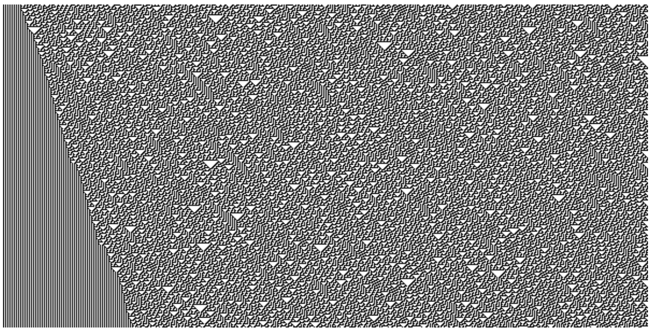
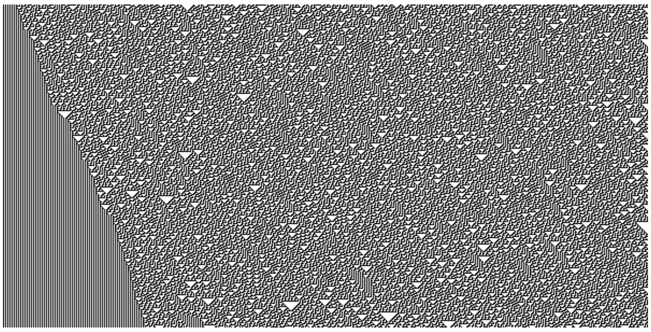
5. NEW THINGS: Look around where you are right now, and try to notice three new things that you have never noticed before.

What are those three things?

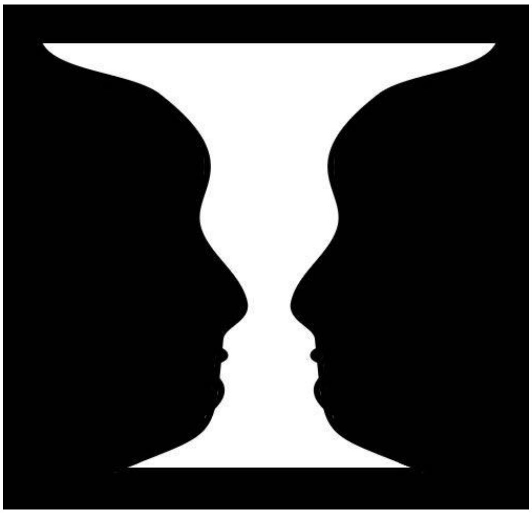
1. ____
2. ____
3. ____

TREATMENT GROUP: LOW MINDFUL

4. COMPUTED IMAGE: Which image do you prefer? You can skip the question if you like.

5. SPOT-THE-DIFFERENCE: Can you spot the difference?

3. RUBIN’S VASE: What can you see here?

○ I see the vase on the left and the vase on the right, but no faces anywhere.
○ I see only the vase on the left, and the vase and faces on the right.
○ I see only the vase on the right, and the vase and faces on the left.
○ I see the vase and faces on the left and the vase and faces on the right.

6. NEW THINGS: Look around where you are right now, and try to notice three new things that you may or may not have never noticed before.

What are those three things?

4. ______________________________
5. ______________________________
6. ______________________________

MANIPULATION CHECK:

LMS14: To the best of your knowledge, to what extent do you agree with the following statements?

7-point Likert scale, ranging from *strongly disagree (1)* to *strongly agree (7)*

	0 1 2 3 4 5 6 7
I like to investigate things	
I generate few novel ideas	
I make many novel contributions	
I seldom notice what other people are up to	
I avoid thought-provoking conversations	
I am very creative	
I am very curious	
I try to think of new ways of doing things	
I am rarely aware of changes	
I like to be challenged intellectually	
I find it easy to create new and effective ideas	
I am rarely alert to new developments	
I like to figure out how things work	
I am not an original thinker	

MEASURING OUTCOME VARIABLE:

COVID-19 False Positive: Please note that the following scenario is hypothetical.

The prevalence of COVID-19 in a population is 1%. A rapid test correctly identifies the coronavirus 70% of the time when it is truly there, and correctly reports there is no coronavirus 98% of the time when it is truly not there. You take the rapid test and it reports that you have the coronavirus.

What is the probability that you really do? Please pick the answer closest to your best guess of the probability

- Definitely 0% you do not think there is any chance you have this type of cancer (0%)
- Between 0 and 10% (5%)
- Between 10 and 20% (15%)
- Between 20 and 30% (25%)
- Between 30 and 40% (35%)
- Between 40 and 50% (45%)
- Between 50 and 60% (55%)
- Between 60 and 70% (65%)
- Between 70 and 80% (75%)
- Between 80 and 90% (85%)
- Between 90 and 100% (95%)
- Definitely 100%—you are completely sure you have this type of cancer (100%)

COVID-19 False Negative: Now, consider the same statistics as above, but the rapid test reports that you **do not** have the coronavirus.
What is the probability that you **do have** the coronavirus?

- Definitely 0% you do not think there is any chance you have this type of cancer (0%)
- Between 0 and 10% (5%)
- Between 10 and 20% (15%)
- Between 20 and 30% (25%)
- Between 30 and 40% (35%)
- Between 40 and 50% (45%)
- Between 50 and 60% (55%)
- Between 60 and 70% (65%)
- Between 70 and 80% (75%)
- Between 80 and 90% (85%)
- Between 90 and 100% (95%)
- Definitely 100%—you are completely sure you have this type of cancer (100%)

DEMOGRAPHICS:

AGE: Please indicate your age below:

____

GENDER: Please indicate your gender:

- Male (1)
- Female (2)
- Non-binary / third gender (3)
- Prefer not to say (4)

NATIONALITY: Please indicate your nationality:

▼ Afghanistan ... Zimbabwe

OCCUPATION: Please indicate your current occupation:

- Student (1)
- Employed (2)
- Unemployed (3)
- Self-employed (4)

Appendix B: The LMS14 Items

1. I like to investigate things (NS)
2. I generate few novel ideas (NP)*
3. I make many novel contributions (NP)
4. I seldom notice what other people are up to (E)*
5. I avoid thought-provoking conversations (E)*
6. I am very creative (NP)
7. I am very curious (NS)
8. I try to think of new ways of doing things (NS)
9. I am rarely aware of changes (E)*
10. I like to be challenged intellectually (NS)
11. I find it easy to create new and effective ideas (NP)
12. I am rarely alert to new developments (E)*
13. I like to figure out how things work (NS)
14. I am not an original thinker (NP)*

All the items marked with an asterisk are reverse-scored items.

Appendix C: Calculations

1. Baseline Question

$$P(A) = 0.001$$

$$P(B) = 0.999$$

$$P(B|A') = 0.050$$

$$P(B|A) = 0.900$$

$$P(Cancer|Positive) = \frac{0.9 * 0.001}{(0.9 * 0.001) + (0.05 * 0.999)}$$

$$= 1.77\%\ (midpoint\ of\ interval = 5\%)$$

2. COVID-19 False Positives

$$P(A) = 0.010$$

$$P(B) = 0.990$$

$$P(B|A') = 0.020$$

$$P(B|A) = 0.700$$

$$P(COVID19|Positive) = \frac{0.7 * 0.01}{(0.7 * 0.01) + (0.7 * 0.99)}$$

$$= 26.12\%\ (midpoint\ of\ interval = 25\%)$$

3. COVID-19 False Negatives

$$P(A) = 0.010$$

$$P(B) = 0.990$$

$$P(B|A') = 0.98$$

$$P(B|A) = 0.300$$

$$P(COVID19|Negative) = \frac{0.3 * 0.01}{(0.3 * 0.01) + (0.98 * 0.99)}$$

$$= 0.31\%\ (midpoint\ of\ interval = 35\%)$$