Active Unpacking of the Financial Consequences of Health Outcomes Improves Preventive Decisions

Zhihan Cui[†] Lu Liu[‡]

Current Version: September 2023

Abstract

Traditional health economic models assume that decision-makers (DMs) incorporate all dimensions of information regarding potential health outcomes when making preventive decisions. However, behavioral sciences suggest that individuals might deviate from this assumption in two key aspects. The first is opportunity cost neglect due to limited attention, whereby DMs may overlook some indirect or non-salient dimensions of loss from illness. The second is evaluability issues, whereby DMs may fail to comprehend the actual relevance of numerical health information to their circumstances. To increase people's attention to financial losses and make health outcomes more evaluable, we designed an intervention entitled "Active Unpacking with Money" (AUM), which directs DMs to *actively calculate the monetary losses from experiencing a negative health condition over a specified time period.* Through a series of six consecutive online experiments, we demonstrate that: (1) AUM amplifies people's perceived severity of health risks and their willingness-to-pay (WTP) for a guaranteed preventive measure; (2) AUM heightens people's sensitivity to the length of a disease; and (3) AUM bolsters sensitivity to probabilistic information. We clarify how AUM assists in mitigating both salience and evaluability challenges within these contexts, and why AUM is likely welfare-improving. Finally, we discuss its practical implications in health communication.

JEL Codes: C91, D14, I10, I12

*We are extraordinarily grateful for the guidance, support and encouragement of my advisors Geoffrey Heal, Jeffrey Shrader, Daniel Benjamin and Kathleen Nganouge without whom we can by no means complete the research. This project also gets important help and suggestions from the conversations with Ilana Brody, Eugene Caruso, Hengchen Dai, Mark Dean, David Dolifka, Evan Friedman, Craig Fox, Hal Hirshleifer, Xilin Li, RC Lim, Adriana Lleras-Muney, Jackson Lu, Yuhang Pan, Xuezheng Qin, Rui Sun, Yiwen Shen, Fanqi Shi, Junze Sun, Xu Tan, Megan Weber, Jueyu Wu, Mu Zhang, and Castiel Chen Zhuang.

[†]UCLA Anderson School of Management and SIPA, Columbia University; zhihan.cui@anderson.ucla.edu

[‡]New York University Shanghai; ll4748@nyu.edu

1 Introduction

It is well acknowledged that in many scenarios, people may fail to take substantial preventive measures despite their crucial role and cost-effectiveness in protecting health Banerjee and Duflo (2011); Olsen et al. (2010). Different disciplines have various ways to explain this phenomenon, and one important explanation is the behavioral biases[1] in judging potential or hypothetical health outcomes Thaler and Benartzi (2004); Banerjee and Duflo (2011). We argue that consumers may overlook financial losses from getting sick when these losses are non-salient and neglect time duration information of health outcomes when the evaluability of time information is low. We posit that explicitly mentioning money loss will help mitigate these problems. Accordingly, we introduce a novel intervention package, Active Unpacking with Money (AUM), which is designed to **direct people to actively calculate the potential financial losses from illnesses**. This campaigns against these biases and improves the quality of judgment about potential health outcomes.

Recent years have witnessed a consensus in both behavioral economics and psychology research, indicating that salience — the quality of a dimension, attribute, or piece of information being particularly noticeable or important — shapes consumer judgment and choices (Hoffman and Singh, 1997; Itti, 2007; Bordalo et al., 2012, 2013). When evaluating hypothetical health outcomes, multiple strands of literature predict that the financial impact, though an important part of the opportunity cost of illnesses, is less salient, given its non-physical, indirect/implicit (Frederick et al., 2009; Spiller, 2011), and relatively affect-poor (evaluated without involving rich emotions compared to physical pain) nature (Hsee and Kunreuther, 2000; Rottenstreich and Hsee, 2001). As a result, individuals may pay limited attention to or even neglect this loss when making decisions, leading to choices that might potentially deviate from their objectives. For instance, people may miss their opportunities to prevent specific diseases (Wettstein et al., 2012; Kimball et al., 2020) and regret contracting the disease. This ignorance of opportunity costs, usually caused by lack of salience implies insufficient risk perception, which may lead to sub-optimal decisions and welfare loss. This justifies using nudge for mitigation (Thaler and Sunstein, 2009; Sunstein, 2014).

Another important issue, possibly less familiar to economists, is that people might

[1] such as opportunity cost neglect, limited attention, probability weighting, etc. Details see Section 2.

find it difficult to evaluate numerical information (such as time duration and probability) regarding health outcomes even if they do pay significant attention or the information is salient (Hsee, 1996; Rottenstreich and Hsee, 2001; Hsee and Zhang, 2010). This is called evaluability problems in decision psychology, with which people show insensitivity to numerical variance in certain dimensions because people are not familiar with how these numbers affect their welfare (Morewedge et al., 2009; Hsee and Zhang, 2010). This insufficient reaction may be another barrier to optimizing their prevention behaviors (Cui et al., 2023). The insensitivity of numerical information may cause preference reversals (Hsee et al., 1999; Sunstein, 2018) and imply *inadequate* or *excessive* risk perception and prevention in different scenarios. Both may indicate welfare loss and justify proper intervention.

Building on existing literature in economics and psychology, we introduce the *Active Unpacking with Money (AUM)* process, a novel intervention designed specifically to counteract the behavioral biases identified. The term "Unpacking" means that the participant decomposes a packed health outcome description into separate components (Van Boven and Epley, 2003); "Unpacking with Money" indicates that we focus on decomposing the financial losses with a monetary scale; and "Active" suggests that the participants should conduct calculations actively on their own. In other words, AUM prompts participants to calculate the financial losses of a health outcome on their own. To our knowledge, no other research has employed such a targeted and comprehensive approach to address these biases in health decisions.

First, calculating financial losses increases salience (Spiller, 2011; Bordalo et al., 2022), bringing attention to the previously under-noticed or neglected opportunity costs of falling ill. Particularly, describing monetary losses unpacks the health outcome (Van Boven and Epley, 2003; Savitsky et al., 2005). Studies on unpacking suggest that people tend to underestimate the magnitude of a problem when judging it holistically, while our unpacking procedure allows them to decompose the problem, make the implicit opportunity costs explicit, and spur their risk perception (Wolff et al., 2019).

Additionally, money amounts provide a universal reference for scaling the severity of a health outcome, thus alleviating the evaluability problems. Prior literature has shown that a common and familiar scale provides an effective reference point for thinking and makes people more sensitive to magnitudes, (Hsee et al., 2009; Hsee and Zhang, 2010),

such as time duration (Morewedge et al., 2009) and associated probabilities (McGraw et al., 2010; Pachur et al., 2014).

Finally, it may be more effective if consumers conduct the unpacking procedure themselves, as it may increase their level of agency and engagement (Muncy and Hunt, 1984; Michaelidou and Dibb, 2008) and give them a personalized information profile (Kaufmann et al., 2018) that can further help them make decisions in the health problem, allowing the decision-maker to delve into the problem in a more effective way. These psychological mechanisms improve judgment quality and promote effective precautions and could be applied in various real-world scenarios, justifying the potential effectiveness of AUM.

To empirically showcase the effect of AUM, this paper presents the results of six online experiments that test four sequential hypotheses. Experiment 1, using a small-scale exploratory and demonstrative setup, reveals that people often overlook the financial implications of health conditions unless explicitly reminded, as seen when they view financial losses in an option list. By prompting subjects to provide open-ended responses about experiencing severe back pain, we found that they rarely consider financial losses without a reminder. However, when asked, nearly three times as many subjects highlighted the "impact on finances" as the most crucial dimension among unselected categories. This indicates that financial concerns related to health outcomes remain vital for many decision-makers, even if they initially overlook them.

Experiments 2 and 3 demonstrate that by increasing the salience of financial losses, AUM can significantly mitigate opportunity cost neglect and increase risk perception. In both experiments, we investigated people's perceived severity of a 3-month-long lung problem which is likely to lead to sick leave, and asked people's willingness-to-pay (WTP) for guaranteed prevention. Experiment 2 demonstrated that, compared to merely providing information about financial loss without allowing users to manually calculate monetary loss (Unpacking without Directly Mentioning Money), the AUM process significantly increased perceived severity and Willingness-to-Pay for prevention. Experiment 3 served as a supplement to Experiment 2, conducting a detailed exploration of the AUM mechanism through a five-arm design. It revealed that both the salience effect (explicitly reminding of the existence of monetary loss) and the evaluability effect (providing specific values) had significant positive roles, while the role of agency (active engagement) required further validation. Specifically, the salience effect impacts more on perceived

severity, and the evaluability effect impacts more on the WTP. Overall, these experiments show that AUM promotes people to think of money loss and thus take the disease more seriously. We detect that the possibility of pure demand effect and over-debiasing is thin, implying that the treatment effects are likely welfare-increasing.

Experiments 4 and 5 first showcase the prevalence of duration neglect (insensitivity to time duration information) in hypothetical health outcome judgments. Then, it shows that through increasing evaluability, AUM is effective in alleviating these biases. In Experiment 4, we employed a 3×2 factorial design. We found that when participants were merely provided with information about disease symptoms and duration (Control Group) or given the same information but primed for numerical thinking through a calculation task (Calculation Group), their sensitivity to disease duration information was notably lower. In contrast, the AUM process significantly enhanced this sensitivity. In the first two groups, people were entirely insensitive to whether the same severity of angina lasted for 8 or 16 weeks. However, in the AUM group, the perceived severity and WTP for prevention clearly increased when the duration was 16 weeks. With a 2×4 factorial design, Experiment 5 conducted a robustness check for Experiment 4 both at the within-subject and between-subject levels. It found that the conclusions at the between-subject level remained stable, and at the within-subject level (the same person evaluating the same disease multiple times but at different times), AUM could also increase sensitivity to duration. The combination of these two experiments validated the role of AUM in enhancing sensitivity to time, which is potentially welfare-increasing in many scenarios.

In Experiment 6, we explored the issue of probability insensitivity in hypothetical health outcomes and prevention measure judgments (McGraw et al., 2010; Pachur et al., 2014), and show that AUM effectively mitigates this problem by increasing evaluability. Employing a 2x4 factorial design, we had participants in both control and AUM groups see the same 3-month disease and income loss description, which had an 80% chance of occurrence. However, only the AUM group did an income loss calculation, and any incorrect answers were corrected. Following this, participants were introduced to four preventive measures, each reducing the disease contraction probability by 20pp, 30pp, 40pp, and 60pp, respectively. Our analyses, both between-subject and within-subject, revealed that AUM heightened individuals' sensitivity to probability changes. Echoing findings from Experiment 4, participants in the control group showed no varied responses to changes in

probability, whereas those in the AUM group did. This experiment emphasizes AUM's effectiveness, indicating its potential to enhance judgment and decision-making in intricate health contexts.

Based on all the experiments discussed, the AUM process clearly has a strong causal effect: it enhances individuals' risk perception of hypothetical health issues, heightens sensitivity to disease-related time durations, and bolsters understanding and judgments about prevention-related probabilities. As such, we assert that AUM is an intervention with significant practical potential. Furthermore, we've detailed its psychological underpinnings and offered practical recommendations.

This paper provides clear methodological insights, both theoretical and practical. On the theoretical front, our study not only integrates existing concepts but also introduces the numerical evaluability theory to health decisions, marking an important supplement to the health economics literature. We've introduced the numerical evaluability theory to health decisions, marking a pioneering contribution in this field. The empirical significance of our findings underscores the need for future research in health economics to consider these biases. On a practical note, we present an easily implementable intervention for health communicators, including governmental organizations, hospitals, and insurance companies. This intervention aids audiences in enhancing their health decision-making and has proven effective in various scenarios. Consider the recent hesitancy observed in certain populations toward COVID-19 vaccination. If individuals were made aware of the potential financial burdens of hospitalization or prolonged illness (Long COVID, which may lead to long sick leaves (Cutler, 2022)), they might be more motivated to get vaccinated. Furthermore, we establish a framework for future researchers to explore deeper into the impact of monetary mentions on promoting healthier behaviors.

The remainder of this paper proceeds as follows. Section 2 gives background information regarding behavioral biases in health judgments and sets up the theoretical foundations of AUM. Section 3 gives a sketch of the experimental design and offers a preview of the results. Section 4 shows the main results and primary mechanisms of our experimental findings. Section 5 delves into the implications and limitations of this paper. Section 6 concludes.

2 Theoretical Foundations

2.1 Background and Motivation

Despite the undeniable importance of preventive behaviors in maintaining health (Cohen et al., 1988; Kenkel, 1994), many individuals often overlook crucial preventive measures. These include purchasing medical insurance (Hsee and Kunreuther, 2000; Kunreuther et al., 2006; Baicker et al., 2012; Loewenstein et al., 2013), ensuring consistent consumption in preparation for potential health shocks (Gertler and Gruber, 2002; Dercon and Hoddinott, 2004), getting vaccinated (Dubé et al., 2013; MacDonald et al., 2015; Dror et al., 2020), and applying other non-pharmaceutical disease prevention strategies (Banerjee and Duflo, 2011; Linnemayr, 2015; Soofi et al., 2020). The failure to adopt these preventive behaviors can result in the spread of preventable illnesses, escalating healthcare costs, and, in severe cases, loss of life, underscoring the urgency of addressing this gap.

These failures can be partially attributed to behavioral biases in health decision-making, which may be multi-dimensional and complicated and, thus prone to cognitive biases. Myopic judgments (Thaler and Benartzi, 2004; Kan, 2007; Wang and Sloan, 2018), health illiteracy and misinformation (Thaler and Benartzi, 2004; Kan, 2007; Wang and Sloan, 2018; Southwell et al., 2019; Krishna and Thompson, 2021), and biased risk perception (Brnstrm and Brandberg, 2010; Wolff et al., 2019) are all decision biases that might drive people away from the optimal choice path and lead to insufficient prevention, thereby harming their health and well-being. Specifically, as mentioned in the book "Poor Economics" by Nobel Laureates Banerjee and Duflo (2011), people often overlook and invest insufficiently in the "low-hanging fruit" in preventive work - these are simple measures that do not cost much but have large potential preventive benefits, such as mosquito nets, vaccines, masks, and healthy diet.

At the same time, however, people sometimes worry excessively about certain health outcomes and devote themselves to over-prevention. A typical example is Side Effect Aversion (Waters et al., 2007b,a), where people are afraid to afraid to take vaccines that are proven safe by science due to fear of side effects (Rief, 2021), or even fear of all medication use (referred to as pharmacophobia, De las Cuevas et al. (2018)). Another example is over-insurance of modest risks (Sydnor, 2010), where people might invest too

much in health insurance and services due to concerns about health problems. These behavioral biases, whether leading to under-prevention or over-prevention, underscore the complexities of health decision-making and highlight the need for interventions that guide individuals toward optimal health choices. Before we engage in addressing these biases, we should first turn to the theories that explain why they emerge.

Exploring the theory of behavioral decision-making related to prevention holds considerable theoretical and practical significance for understanding these biases and designing interventions. Neoclassical health economics, decision psychology, and public health all employ different models for us to view and interpret the multifaceted nature of health decision-making. These fields incorporate theories such as cost-benefit analysis (Johannesson and Jönsson, 1991), expected utility theory (Meltzer, 2001), the life-cycle model (Bleichrodt and Quiggin, 1999), the theory of planned behavior (Ajzen, 1991) (a theory positing that intention, attitude, and perceived behavioral control influence behavior) the health belief model (Janz and Becker, 1984; Gielen and Sleet, 2003) (a model that focuses on individual perceptions of health threats and the benefits of avoiding them) and the protection motivation theory (Rogers and Prentice-Dunn, 1997; Norman et al., 2015) (a theory that examines how people are motivated to protect themselves from harmful events or situations).

Investigating the common grounds of these theories, we find that optimal preventive behavior necessitates a foundation built on the premise that "people have comprehensive, consistent preferences for different health outcomes." Economic theories suggest this depends on individuals conducting thorough evaluations and considerations of every dimension of a health outcome. They should also be able to correctly incorporate and evaluate this multi-dimensional information to form consistent preferences over different potential outcomes. In other words, rational individuals clearly understand the loss a disease causes them on multiple dimensions, and these losses correspond to cumulative disutility when summed. Then, they make choices and decisions based on the expected utility.

Decision theory in microeconomics and the behavioral economics literature suggests that, in reality, individuals may exhibit "bounded rationality," meaning their rationality is limited, leading to systematic deviations from the classical utility function (Kahneman and Tversky, 1984; Conlisk, 1996). Such deviations from perfect rationality can cause

individuals to either ignore, underestimate, or overestimate certain aspects of real-world decision-making. As a result, their judgments might not align with what is most beneficial for them. Based on the literature in behavioral and health economics, these deviations can be categorized into three distinct types, each underpinned by different economic theories.

(1) When making evaluations, people tend to discard dimensions or attributes that they deem "less important." This corresponds to one of the most classic theories in behavioral economics: Rational Inattention (Sims, 2003; Sims et al., 2015; Brown and Jeon, 2021). Its core idea is that people's attention and bandwidth are limited. As a result, they rationally allocate their attention, discarding those dimensions with a lower impact on their welfare. Another well-known theory under this framework is Sparsity-based Evaluation. This theory, first proposed by Gabaix (2014), suggests that real-world economic decisions and evaluations are influenced by numerous dimensions. People often ignore variables with less degree of change, minimal impact on deviation from optimal choices, and minor effects on the final utility function, focusing only on a few crucial variables. For instance, in health decisions, the price changes of some low-cost drugs or vaccines might not significantly impact people, even if they are related to the decision. Under Gabaix's model, such factors would be neglected when making health insurance choices.

(2) Individuals may overlook those less salient dimensions. Salience, first derived from a cognitive psychology concept, is defined as "the distinct subjective perceptual quality which makes some items in the world stand out from their neighbors and immediately grab our attention" (Itti, 2007; Hoffman and Singh, 1997). Salience differs from the rational attention problem mentioned in (1) as non-salient dimensions may indeed be important. However, due to contextual factors, the decision-maker may not sufficiently notice these dimensions or even overlook them altogether. For instance, when people think of smoking, immediate pleasure or stress relief might be salient, giving people a high valuation of cigarettes. However, in some countries, cigarette packs come with graphic images of the harmful effects of smoking. These images aim to make the long-term health risks of smoking more salient, countering the immediate gratification. The definition of salience has long been context-based and case-by-case until Bordalo et al. (2013, 2020, 2022) quantitatively described it in the theory of Salience and Consumer Choice, indicating

that salience issues could lead to various decision-making biases during consumer choices, such as reference-dependent preferences and inconsistent time preferences.

(3) Several other behavioral economics and decision science theories are also closely related to this topic. For instance, the theory of opportunity cost neglect suggests that non-salient opportunity costs—sometimes equated with implicit costs or encompassing both implicit and explicit costs—are frequently underestimated or ignored by individuals (Frederick et al., 2009; Spiller, 2011). However, this cognitive bias can be mitigated by enhancing the salience of the existence of opportunity costs (as in the cigarette example above). In addition, research by Wason (1968) has highlighted imperfections in people's deductive reasoning. These imperfections can lead them to overlook outcomes that, while indirect or implicit, are both logically sound and practically significant. Less fluent reasoning processes, for example, are more susceptible to biases (Trippas et al., 2014). The query theory posits that under limited cognitive resources, the factors people consider during health decision-making are often those that first come to mind or are most salient (Johnson et al., 2007; Nielsen and Phillips, 2008). On the contrary, those attributes that comes late are more prone to negligence.

When we synthesize these theories, a picture emerges suggesting that in the realm of health economics, individuals might often overlook less obvious yet crucial aspects of potential health outcomes. For instance, they might not fully consider the economic ramifications of sick leave and medical expenses or the long-term effects a disease might have on interpersonal relationships.

Even for dimensions with both genuine importance and salience, people may still exhibit insufficient sensitivity to information about these dimensions, especially when the information is numerically represented. This could result from two sources: numeracy and evaluability. Numeracy (Fagerlin et al., 2007; Brooks and Pui, 2010) is a relatively stable trait that depicts the propensity and ability of individuals to understand things numerically. It is not easily influenced by interventions. On the other hand, evaluability, which is less discussed in the economic literature, is primarily a context-based condition. However, this context often ties closely to the decision-makers' (DM's) experience. Hsee (1996); Hsee et al. (1999); Hsee and Zhang (2010) investigated the determinants of evaluability and established the significant General Evaluability Theory. In this theory, evaluability is influenced by factors such as the mode (whether the alternatives are

evaluated jointly or separately, for example, comparing two health insurance plans side by side versus evaluating them individually), the DM's familiarity with the evaluated attribute/alternative, and the inherent nature of the evaluated attribute (the fundamental characteristics that make an attribute more or less evaluable).

Firstly, familiarity with the dimension is crucial. Morewedge et al. (2009) suggested that individuals are more numerically sensitive to experiences they are familiar with or have encountered before and less sensitive to hypothetical scenarios they have not experienced. This could be particularly important in health decision-making, especially in preventive measures, as individuals often lack prior experience with the diseases they need to prevent rather than cure.

Secondly, Mode (evaluation mode) refers to a phenomenon known as the JE-SE (joint evaluation vs. separate evaluation) paradox (Hsee et al., 1999). Individuals are more sensitive to unfamiliar numerical indices under joint evaluation than under separate evaluation. This could have implications in many medical decision, as it might suggest a preference reversal between situations of no comparison (SE) and comparison (JE). For instance, Zikmund-Fisher et al. (2004) showed an interesting example as follows:

Table 1: An Example of JE-SE Preference Reversal in Health Decisions

Doctor Name	MD Institution	Eye Surgery Performance
Dr.Bettereyes	Harvard	80 successes/year
Dr.Seebetter	U.Iowa	300 successes/year

The existence of JE-SE preference reversal has been revealed in health psychology literature but not in economics yet. In the choice structure above, when the two doctors were evaluated jointly, the average rating for Dr.Bettereyes and Dr.Seebetter were 5.7 and 8.4 on a 0-10 scale, with the latter being strongly preferred. However, single evaluation generated the opposite: the scores were 6.4 and 5.9. The reason for this reversal is that numbers of 80/300 are not easily evaluated without a comparison as people are not familiar with them. But in joint evaluation, it is easy to find that 80 is "not sufficiently large." For more examples, see Zikmund-Fisher et al. (2004, 2010); Zikmund-Fisher (2019)).

Lastly, Nature (the nature of the item being evaluated) matters as well. Certain

indices that are instinctively encoded in human genetics can elicit a higher degree of sensitivity (Shen et al., 2012), such as knowing intuitively that water at 53 degrees Celsius is too hot for a bath, but 40 degrees is tolerable. However, it's more challenging to assess the value difference between a 0.53-carat and a 0.4-carat diamond. Affect-richness (Rottenstreich and Hsee, 2001) is also a part of the Nature category. Affect-rich indicators often imply poorer evaluability and numerical sensitivity (Pachur et al., 2014; Suter et al., 2016).

To sum up, the evaluability of individual preventive health assessments and decisions is frequently affected by these issues, making them a significant potential challenge to rational health decision-making.

While traditional behavioral economic models often group these biases together, differentiating them has considerable significance. For instance, categories (2) and (3) are often not distinctive in salience literature. However, they exhibit significant differences in psychological mechanisms. This has important theoretical and practical implications, as the interventions applicable to each might differ. For instance, if a vaccine has an efficacy rate of 83%, even if the center for Disease Control and Prevention (CDC) prominently displays this efficacy information, people might still be unclear about whether 83% truly represents a highly effective vaccine, making it challenging for them to make the best choice for themselves. By assessing the three types of biases, (1), (2), and (3), we can see that in realistic health decisions, bias (1) yields relatively smaller welfare losses since it relates to dimensions deemed 'less important'. However, both (2) and (3) have the potential to result in significant welfare losses. Given that numeracy is a relatively stable trait and not easily changed, our discussion on designing intervention methods in this paper mainly focuses on the issues of salience and evaluability.

2.2 Theoretical Foundation for AUM

In the preceding discussion, we identified several significant biases that individuals may exhibit when facing potential health outcomes. For instance, when not made salient, people may overlook the opportunity costs of illness, such as financial losses from sick leave. Such biases can lead to insufficient risk perception, subsequently causing individuals to fail to take adequate preventive measures. Furthermore, when the evaluability of health-related information, especially numerical data, is poor, people may be insensitive

to the variances in this information, like the duration of illness, incidence rate, or the effectiveness of a vaccine, or even completely ignore it. The second type of cognitive bias implies preference reversals, and this bias can result in either excessive or inadequate risk perception and preventive behavior. In general, these biases may amplify fears of minor risks while diminishing the perceived threats of major ones. Both the existing literature and empirical cases presented in this article demonstrate that these two types of biases are likely to be pervasive.

Salience-based nudges (Thaler and Sunstein, 2009; Sunstein, 2017b) are a common strategy that researchers and policymakers use to tackle the salience issue. These nudges work without employing manipulative constraints or imposing significant incentives (Sunstein, 2017a). They promote individuals to make certain choices by enhancing the salience of easily overlooked dimensions or options. Intervention designers often increase individuals' emphasis on a specific dimension or certain options. They change attitudes and behaviors by prominently displaying an option, priming a certain way of thinking, or reminding decision-makers of a dimension's existence or importance. These strategies aim to guide individuals towards the desired choice.

Salience-based nudges are effectively and widely applied in health interventions (Vallgårda, 2012; Blumenthal-Barby and Burroughs, 2012). They have been successfully applied in various scenarios related to prevention such as promoting healthy food choice (Wilson et al., 2016; Kroese et al., 2016; Bucher et al., 2016; Hoenink et al., 2020), health insurance take-ups (Wright et al., 2017), increasing vaccination (Milkman et al., 2021; Dai et al., 2021; Reñosa et al., 2021), and encouraging non-pharmaceutical preventive behaviors against diseases (Blackwell et al., 2017; Weijers and de Koning, 2021; Shiraly et al., 2022). These studies provide a robust foundation for utilizing salience-based nudging strategies to address the often-neglected financial implications of health issues. In our experimental design, a key manipulation involves emphasizing the potential financial losses associated with illness.

Another method that may serve as a helpful supplement is unpacking (Van Boven and Epley, 2003; Savitsky et al., 2005). This approach involves breaking down a general or vague description into more detailed components, providing a more clarified decomposition for the implicit attributes. An unpacked outcome description offers information on both explicit and implicit outcomes and costs, addressing the salience problem and

increases risk perception. Research has supported this theory, such as in areas like task workload (Kruger and Evans, 2004), group-work allocation (Savitsky et al., 2005), and risk assessment (Johnson et al., 1993; Tversky and Koehler, 1994; Wolff et al., 2019). For instance, subjects perceive the likelihood of death from specific causes (e.g., cancer or heart attack) to be higher than from general causes like "natural causes" (Tversky and Koehler, 1994). Apparently, these effects are highly related to our intervention design.

It's crucial to understand that while salience-based strategies typically *enhance* risk perception and behavior, from a neoclassical economist's viewpoint, this enhancement might not always be beneficial. Effective interventions shouldn't excessively amplify risk assessment, as this can lead to over-prevention behaviors like pharmacophobia[2] and vaccine hesitancy. The ultimate goal is to ensure that perceived risk aligns closely with the actual, objective risk.

Based on our research and findings, we posit that a high-quality risk assessment should exhibit ample sensitivity to fluctuations in risk dimensions, such as disease duration, incidence rate, and the efficacy of preventive measures. The evaluability challenge, which we touched upon earlier, directly pertains to this concern. In certain situations, individuals may not only show reduced sensitivity to these numerical indicators but might be entirely oblivious to their significance.

Numerical evaluability for hypothetical health outcomes is typically low. This is because these outcomes are hypothetical, and since people haven't experienced them, they tend to be unfamiliar. This unfamiliarity results in a lack of intuitive understanding and sensitivity to aspects like the severity and duration of illness. Diseases inherently have an affect-rich nature and don't have a clear quantitative mental scale, complicating quantitative evaluations. Existing research and the observation that people often overlook financial loss suggest that using financial loss as a metric to measure disease could be effective. By quantifying financial loss, we can address the challenges of unfamiliarity and the affect-rich nature of diseases. Economically independent individuals frequently encounter financial decisions, making monetary scenarios more familiar to them. Additionally, money, in contrast to pain, is less emotionally charged and has a clear quantitative scale, simplifying its evaluation (Rottenstreich and Hsee, 2001; McGraw et al., 2010). Consequently, biases such as probability weighting (Baillon et al., 2022) and even

[2]According to the NIH, pharmacophobia is a fear of medication and a negative attitude toward drugs in general

probability insensitivity and minimax heuristics (Pachur et al., 2014) can be mitigated.

According to the theories above, designing a financial loss reminder that specifies *an explicit money amount* may give the intervention the dual advantage of addressing both types of biases. This can lead to marked improvements in risk perception and numerical sensitivity. We term this intervention strategy *Unpacking Health Outcomes with Money* (**UM**). Arguably, this approach can significantly enhance the quality of decision-making.

However, the pure *Unpacking with Money* approach may have crucial limitations. The primary concern is the lack of personalization in the UM method. Economic losses from diseases correlate with factors like income levels, health insurance coverage, and the availability of paid sick leave. These factors can vary significantly among individuals. While developed countries might boast high-level health insurance and free medical care, variations in income levels and sick leave policies still influence potential income losses from illnesses. Medical information platforms, such as government websites and hospital portals, typically don't have direct access to individual income or sick leave details. Consequently, they can't provide precise health loss estimates for every user. Even for family doctors, obtaining this information presents technical and ethical challenges. In practical terms, UM might have to resort to presenting an "average" estimate to all users, which could limit its effectiveness. This realization drives our motivation to refine our intervention strategy.

So, how to make the UM (Unpacking with Money) approach more feasible and effectively align it with users' real-life experiences? First and foremost, information personalization is a direction for optimization. Several studies have validated its effectiveness in enhancing user experience and decision-making. Our theory suggests that when we provide users with economic loss information tailored to their circumstances, they find it highly familiar, making the information more evaluable. As a result, users are more inclined to make high-quality decisions. Additionally, user agency and engagement are pivotal. Research in social psychology, including cognitive dissonance theory and self-perception theory, indicates that users are more accepting of conclusions they derive through active involvement. With this in mind, our goal is to offer users a standardized, user-friendly calculation method, empowering them to compute potential economic losses from diseases independently. This not only addresses the operational challenges discussed earlier but also potentially amplifies intervention effects through personalization

and active engagement. We term this strategy *Active Unpacking with Money* (AUM).

While many health interventions have explored salience nudging, personalization and user engagement as key factors for success, our AUM strategy offers a nuanced approach that combines these strategies and conducts a novel improvement. By integrating financial metrics into health outcomes, we aim to simultaneously cope with salience and evaluability problems, systematically improve the quality of preventive decisions.

3 Experimental Design: A Roadmap

In this section, we outline a comprehensive roadmap detailing the experimental design schemes associated with AUM. Alongside this, we offer a concise rationale behind each experimental design to ensure clarity in our approach.

Our journey begins with Experiment 1, which probes how often individuals consider the financial ramifications of health conditions without any intervention. Using open-ended queries about severe back pain, we aimed to gauge if financial concerns naturally spring to mind for individuals contemplating health scenarios. Additionally, we introduced subtle reminders to ascertain their importance in this context. Experiments 2 and 3 delve into the primary effects of AUM, offering preliminary insights into its cognitive mechanisms. These studies evaluate if actively calculating potential financial losses can bolster risk assessments and the willingness-to-pay (WTP) for preventive actions. Experiment 3 further unpacks the intricacies of the AUM approach, shedding light on the influence of salience, evaluability, and agency. Experiments 4 and 5 confront the issue of duration neglect, where decision-makers might disregard the potential duration of an illness. By integrating the AUM methodology, we assess its efficacy in heightening sensitivity to disease duration variations, aiming to refine risk perception quality. Concluding with Experiment 6, we address the bias of probability insensitivity, focusing on perceptions surrounding health outcomes and preventive measure efficacy. By varying disease prevention probabilities, our goal is to discern if AUM can amplify individuals' adjustment to such probability shifts.

The details of the experimental design and the results will be elaborated in Section 4 and the appendix. For convenience, Table 2 demonstrates the logical structures and preview of the results of all six experiments.

Table 2: A Roadmap of All Six Experiments and Preview of Results

No.	Obs.	Experiment Design	Target of the Experiment (Group) Design	Main Procedures	Preview of Results
1	146	Free association	Reveal that, compared to direct physical harm, individuals tend to **overlook** the financial losses associated with illnesses. However, with small cues, they might become aware of such losses, showing that "opportunity cost neglect" may exist in real-world health judgments.	Engage in free association using statements on a symptom, self-code the statements from this free association into different categories of impacts. At last, they choose, at most, one important category that they consider having missed from the unselected categories.	Suggestive vidence for opportunity cost neglect. Less than 15% mentioned money loss during free association, yet approximately 30% chose money loss as their last-call selection among the unselected categories.
2	635	3-arm online RCT	Demonstrate the **first-order effect** of AUM: AUM mitigates the bias of opportunity cost neglect (of financial loss) and increases risk perception **Pairwise Comparisons:** Arm 3 vs Arm 2: Pure AUM effect without any potential information difference Arm 2 vs Arm 1: information difference Arm 3 vs Arm 1: A potential upper bound of AUM effect in real-world scenarios		Clear first-order effect detected even without introducing new information
	198	Arm 1 (Control)		WHO disease description + Irrelevant calculation	For both WTP and severity rating: **Arm 3 > Arm 2** ($p<0.01$), median WTP doubled; **Arm 2 > Arm 1** ($p<0.01$), median WTP doubled; **Arm 3 >> Arm 1** ($p<0.0001$), median WTP quadrupled.
	205	Arm 2 (Info)		WHO disease description + Income loss + Irrelevant calculation	
	210	Arm 3 (AUM)		WHO disease description + Income loss + Income loss calculation. No correction to avoid strong anchoring	

3	671(490)	5-arm online RCT	Examine the mechanisms of AUM, investigating the relative importance of salience, evaluability, and agency. Notably, there is no difference in information across the five groups. **Important Pairwise Comparisons:** Arm 1 vs Arm 2: Net effects of "Agency - Authority" Arm 1 vs Arm 3: Effects of giving correction Arm 3 vs. Arm 5: Net effect of salience in calculation + evaluability Arm 5 vs. Arm 4: Net effect of salience in the description		Salience impacts perceived severity more, while evaluability impacts WTP for prevention more.
	131(102)	Arm 1 (Full AUM)		Full AUM (with correction after a wrong answer)	For WTP: Arm 1 vs Arm 2 **ns**; **Arm 1 > Arm 3** ($p<0.05$); **Arm 3 > Arm 5** ($p<0.01$); Arm 4 vs Arm 5, **ns**. For severity rating: Arm 1 vs Arm 2 **ns**; Arm 1 vs Arm 3 **ns**; Arm 3 vs Arm 5 **ns**; **Arm 4 < Arm 5** ($p<0.05$).
	148(103)	Arm 2 (UM)		UM (directly showing the income loss amount). Act as a "maximum benchmark", impractical in real life	
	131(97)	Arm 3 (AUM)		AUM w/o correction, same as Experiment 2	
	137(99)	Arm 4 (Control)		Control (money loss not salient), claiming unpaid sick leave, not explicitly mentioning "income"	
	124(89)	Arm 5 (Salience)		Control (money loss salient), same as Arm 3 but no calculation task	
4	909(792)	3*2-arm online RCT	Show the second-order effect of AUM. AUM mitigates **duration neglect**: individuals become more attuned to the disease course, specifically the duration of symptoms, when calculating income loss. In this experiment, we permit variations in information, given that our key focus is on this second-order outcome. (1) Test the within-arm difference of 8w vs. 16w. Showing the existence of duration negelct. (2) Differences-in-differences, Arm 3 vs. Arm 1: AUM reduces duration neglect. (3) Differences-in-differences, Arm 3 vs. Arm 2: AUM reduces duration neglect, after teasing out the effect of "numerical priming" (4) Differences-in-differences, Arm 2 vs. Arm 1: pure numerical priming does not reduce duration negelct		Control group displays insensitivity to time, while AUM enhances their sensitivity to time. The interaction term is **significant in 3 of 4** arm-wise comparisons that included Arm 3. This significance is not observed only in Arm 2 vs Arm 3 for perceived severity. Not any interaction effect is observed across Arm 2 and Arm 1.
	327(255)	Arm 1 (Control)		No-money health description; Symptoms last for 8 weeks or 16 weeks	No rating/WTP difference ($p>0.3$) across time horizons of 8 weeks and 16 weeks
	303(273)	Arm 2 (Calculation)		No-money health description + Irrelevant calculation; Symptoms last for 8 weeks or 16 weeks	No rating/WTP difference ($p>0.3$) across time horizons of 8 weeks and 16 weeks
	279(264)	Arm 3 (AUM)		With-money health description + Income loss calculation; Symptoms last for 8 weeks or 16 weeks	Significant rating&WTP difference ($p<0.01$) across time horizons of 8 weeks and 16 weeks;

5	1200(867)	2*4-arm RCT	Provides further evidence of the effect of AUM on improving time sensitivity. We offer within-subject evidence and conduct robustness checks for between-subject findings by presenting an extended sequence of length values. (1) Run the within-arm regression of changing the time duration from 5, 7, 10 and 15 weeks for both groups (2) Detect the interaction effect between-subject: only count the first response (randomly assigned) as there is no reference/comparison for that response (3) Detect the interaction effect within-subject: count all four responses and run a fixed effect regression	Within-subject: the subject sees all four time lengths (5, 7, 10 and 15 weeks) in random order Between-subject: no additional experiments, but only count the first response (randomly assigned) as there is no references for that response	Within-subject: **very significant** ($p<0.001$) interaction effect; AUM increases within-subject probability sesnitivtiy Between-subject: **acceptably significant** ($0.01<p<0.05$) interaction effect in LogWTP, considerable, observable contrasts across arms. AUM increases between-subject probability sesnitivtiy
	648(439)	Arm 1 (Control)		No-money health description	**Very limited** sensitivity to time length between-subject, and medium sensitivity within-subject
	552(428)	Arm 2 (AUM)		With-money health description + Income loss calculation	**Medium** sensitivity between-subject, and high sensitivity within-subject
6	957(850)	2*4-arm RCT	Showcasing another important second order effect: probability sensitivity within non-extreme probabilities (20%-80%). A more real-life-resembling scenario with information of "effectiveness of preventive measure". (1) Run the within-arm regression of changing the effectiveness from 20, 30, 40, 60 pp of disease chance reduction for both groups, showcasing probability insensitivity in groups with no AUM (2) Detect the interaction effect between-subject: only count the first response (randomly assigned) as there is no reference/comparison for that response (3) Detect the interaction effect within-subject: count all four responses and run a fixed effect regression	Logic similar to Experiment 5. Between-subject: only look at the first response as there is no references or comparisons Within-subject: the subject sees all four probabilistic changes (20, 30, 40, 60 pp)	Within-subject: **significant interaction effect** ($p<0.05$) Between-subject: **highly significant interaction effect** ($p<0.001$), considerable, observable contrasts across arms AUM increases both between-subject and within-subject sensitivity to probability
	422	Arm 1 (Salience)		With-money health description, implying unpaid sick leave and income loss. No calculation	**NO sensitivity** ($p=0.9$) to probability between-subject, and medium-small sensitivity within-subject
	428	Arm 2 (AUM)		With-money health description + Income loss calculation + Correction after a wrong calculation	**Medium-small (significant)** sensitivity to probability between-subject, and medium sensitivity within-subject

4 Experimental Procedures and Results

In this part, we provide more detailed reports of the experimental procedures and main findings from Experiments 1-6.

4.1 Experiment 1: People Rarely Consider Financial Impacts of Health Conditions Unless Reminded

In Experiment 1, we employed a semi-quantitative approach to investigate the cognitive patterns that emerge when individuals contemplate a health outcome. Our methodology rests upon the theories of opportunity cost neglect and the query theory (Johnson et al., 2007). When individuals evaluate an outcome and make decisions based on this evaluation, they tend to rely on readily available information in their minds.

Experiment 1 was conducted on Amazon Mechanical Turk (MTurk), a popular platform for online decision experiments. We used the CloudResearch platform as an intermediary. The experiment took place on February 5th, 2023, and we obtained a total of N = 145 effective responses.[3] on February 5th, 2023, obtaining a total of N = 145 effective responses. Participants imagined experiencing severe back pain by reading a description provided by the World Health Organization, "**Severe back pain for three weeks, which would cause difficulty dressing, sitting, standing, walking, and lifting things. Meanwhile, you sleep poorly and feel worried.**"

Following this exercise, they were asked to generate sequential open-ended statements that reflected their thoughts on experiencing these symptoms. All respondents provided at least three responses. After typing "Done," they were asked to classify each of their responses into exactly one of the following categories: individual physical impact, individual emotional impact, impact on family, impact on social life, impact on finances, prevention, treatment, and "none of the above." It's important to note that this was the first time participants were exposed to the alternative list, ensuring they weren't primed or reminded of these options during their free association tasks. They were allowed to choose the same category for multiple statements. Finally, from the unchosen options, participants could select one that they also deemed important. This last step was designed to offer suggestive evidence of opportunity cost neglect.

[3]Eligibility criteria for all experiments included: Human Intelligence Task (HIT) approval rate of 95% or above, a minimum of 100 approved HITs, and being CloudResearch-approved participants.

This experiment's analysis is exploratory. We documented how often respondents selected "impact on finances" for each query in their open-ended statements and for the concluding supplementary question. For a robustness check, we hired a research assistant to independently code the open-ended statements into the same categories.

Figure 1: Frequency of Items Mentioned in Free Association

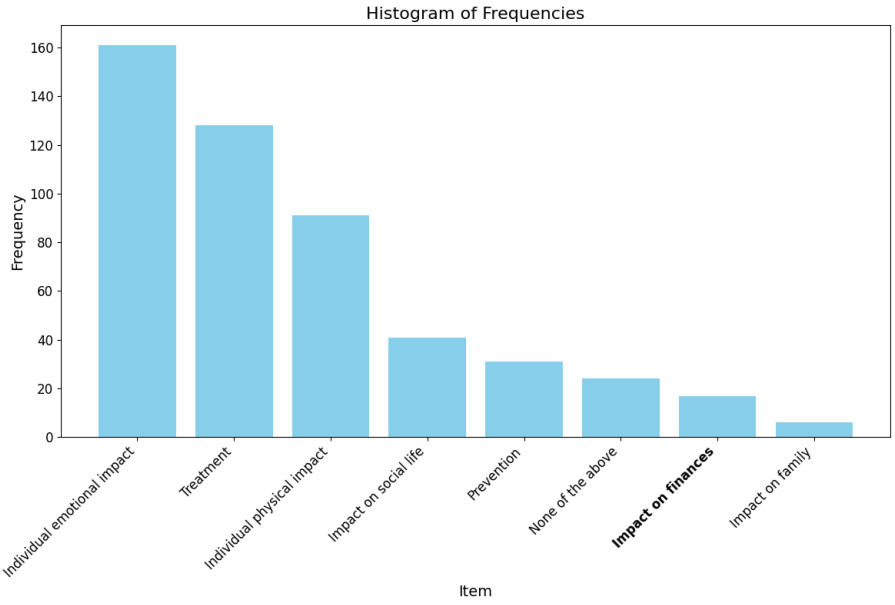
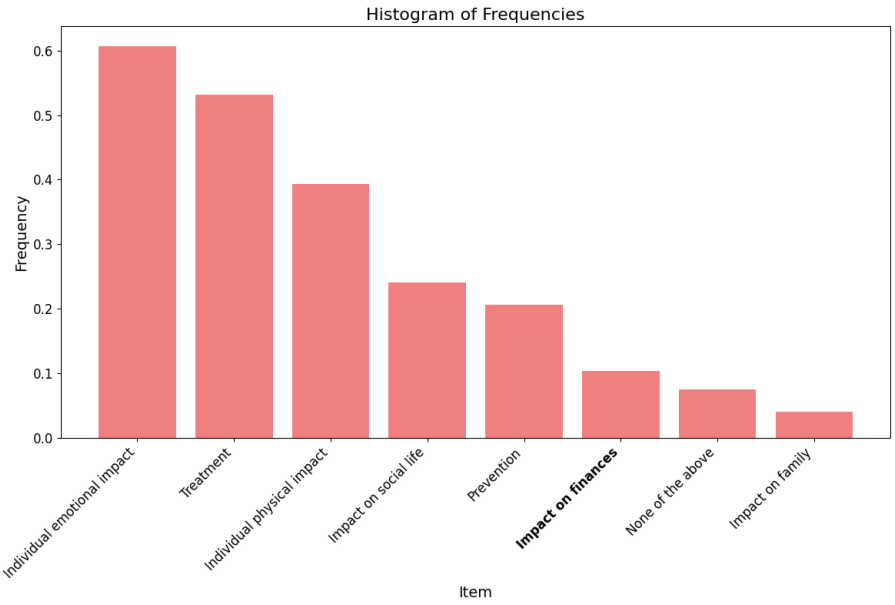


Figure 2: Proportion of Participants Mentioning the Item at Least Once

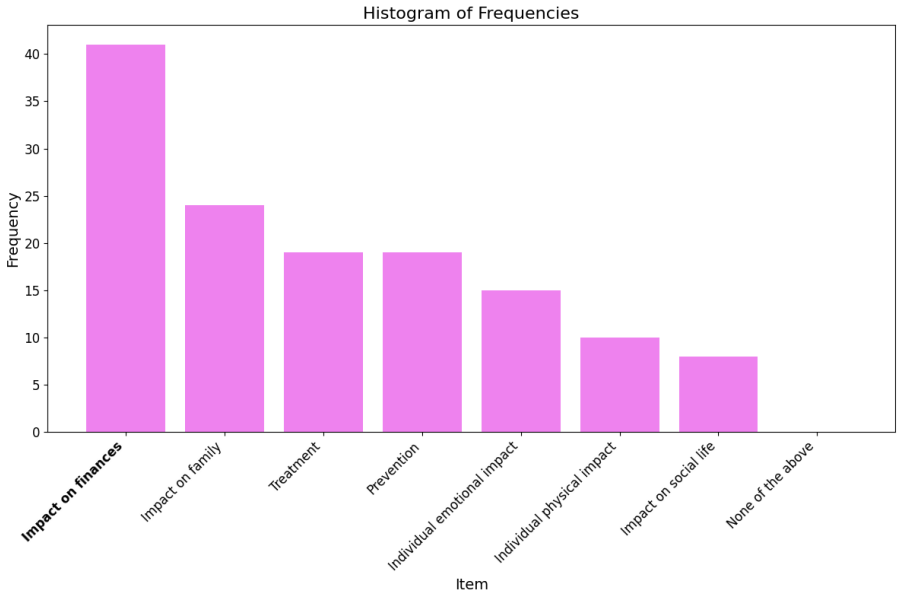


At a glance, our results highlight a consistent trend: participants rarely factored in financial consequences in their open-ended responses, a pattern we'll explore in depth

below. Out of 145 subjects, the self-reported classification showed that only 6, 7, 4, and 7 mentioned financial impacts in their first, second, third, and subsequent statements, respectively, totaling n=24 (16%). The research assistants' coding yielded even lower numbers: 3 for the first, 3 for the second, 4 for the third, and 7 for subsequent statements, totaling n=17 (12%), which ranked among the least concerned categories. If we do not include double-counting and look at the proportion of participants mentioning the item at least once, the numbers were even lower, respectively 19 (13%) and 15 (10%) for original response and RA-coded responses.

On the contrary, a considerable number of participants (n=41, 28%) reported that "impact on finances" was the most important dimension among the unselected categories. It ranks among the *highest* of all categories in the last call. While the graph under-represents emotional and physical impacts, the surge in "Impact on finances" queries underscores its significance.

Figure 3: Frequency of Participants Mentioning the Item at the Last Call



Qualitatively, there's a stark contrast between spontaneously considering financial implications and selecting "money" from previously overlooked categories. This provides preliminary evidence that financial concerns related to health outcomes are critical to many decision-makers, but they may be ignored if not explicitly prompted. Our data also reveals a significant increase in concerns about the 'Impacts on family,' representing another implicit opportunity cost of illness. Collectively, these findings suggest that

while implicit impacts might not be top-of-mind, they gain prominence when adequately highlighted. This experiment's results are consistent with theories on salience-based decisions, query theory, and opportunity cost neglect, setting the foundation for our following interventions: cognitive biases do exist in these scenarios, and there is potential that behavioral nudges can significantly change the

4.2 Experiment 2: AUM Enhances Health Risk Perception

We argued that due to salience and evaluability problems, people may overlook financial losses from negative health outcomes and thus have an insufficient risk perception, and that AUM is an effective debiasing tool. In Experiments 2 and 3, we showcase the effectiveness and psychological mechanisms of AUM through randomized online experiments.

4.2.1 Experimental Procedures

Experiment 2 is designed to empirically test the net effect of AUM on health risk perception. Experiment 1 suggests that just giving some "hints" about money may prompt people to think of financial losses for getting sick. However, this effect may still be peripheral, and many people still did not choose financial losses as a major concern. This finding and our theoretical predictions motivate us to introduce active participation in this experiment.

Experiment 2 was a preregistered experiment (AsPredicted #126197) conducted on MTurk through CloudResearch on Mar 23rd, 2023. 635 participants on MTurk took the experiment. After informed consent, the participants were randomly assigned to one of the three groups (control group, Active Unpacking with Money group (AUM) and Pure Information (INFO)). 613 effective responses [4] were collected.

All participants were presented with a description adapted from a WHO document. The control group merely saw symptom and duration information: "**The person has a persistent cough and fever, is short of breath, feels weak, has lost a lot of weight. This lasts for three months.**" In contrast, the AUM and INFO groups saw the same description with an additional sentence indicating financial loss information, "**...and has to take unpaid leave and miss three months' income.**" It is important to clarify two things here:

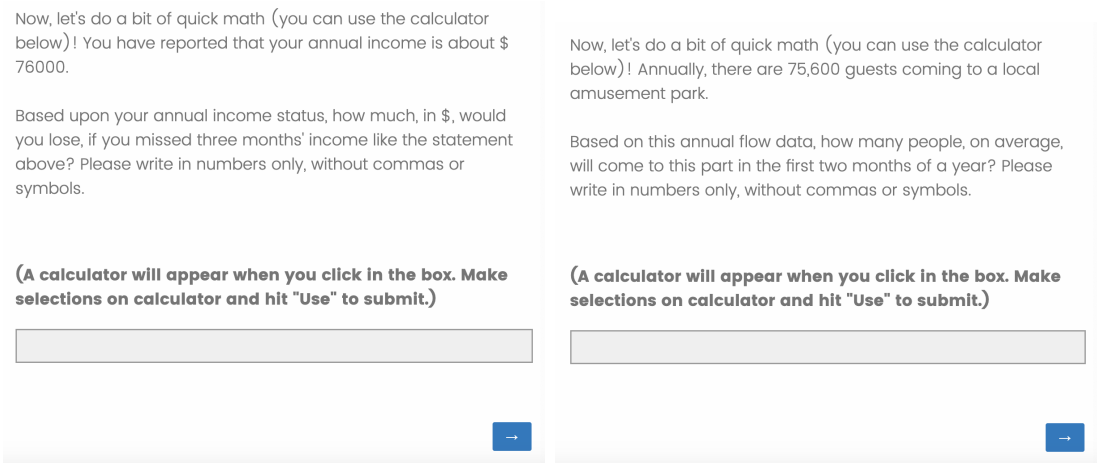
[4] participants who give non-zero WTP responses.

(1)In this setup and all the following ones, we only do the unpacking with missed income of the sick leave. Although this approach doesn't cover all financial costs (notably, it doesn't unpack medical expenses), we still believe it's the optimal design in an online environment. This is because medical expenses are greatly influenced by medical conditions and insurance status. Firstly, there's a significant variation among participants that's hard to measure, and secondly, estimating these expenses is challenging. In contrast, financial loss can be approximated relatively accurately just by knowing monthly income and the duration of sick leave. This setup not only minimizes distractions but is also easier for participants to understand. Experiments 2-6 will all adopt the AUM procedure focusing on sick leave.

(2)In this experiment, there may be informational differences between the control group and the treatment groups because getting the aforementioned symptoms does not guarantee sick leave (although very likely). However, there is no informational difference between the two treatment groups.

Following this, the AUM group was asked to calculate the hypothetical financial loss resulting from losing three months of personal income based on their annual income reported at the beginning of the experiment. The INFO and control groups, however, were asked to perform an unrelated calculation task that was similarly difficult and of comparable numerical scope. This task was designed to prevent any confounding effects of anchoring and adjustment (Epley and Gilovich, 2006) on the results. Accordingly, the AUM and INFO groups only differ in the nature of the calculation task (financial losses vs. a non-pecuniary computation), and the INFO and control groups only differ in the information on unpaid sick leave.

Figure 4: A Comparison of the User Interface of the AUM and Irrelevant Calculation



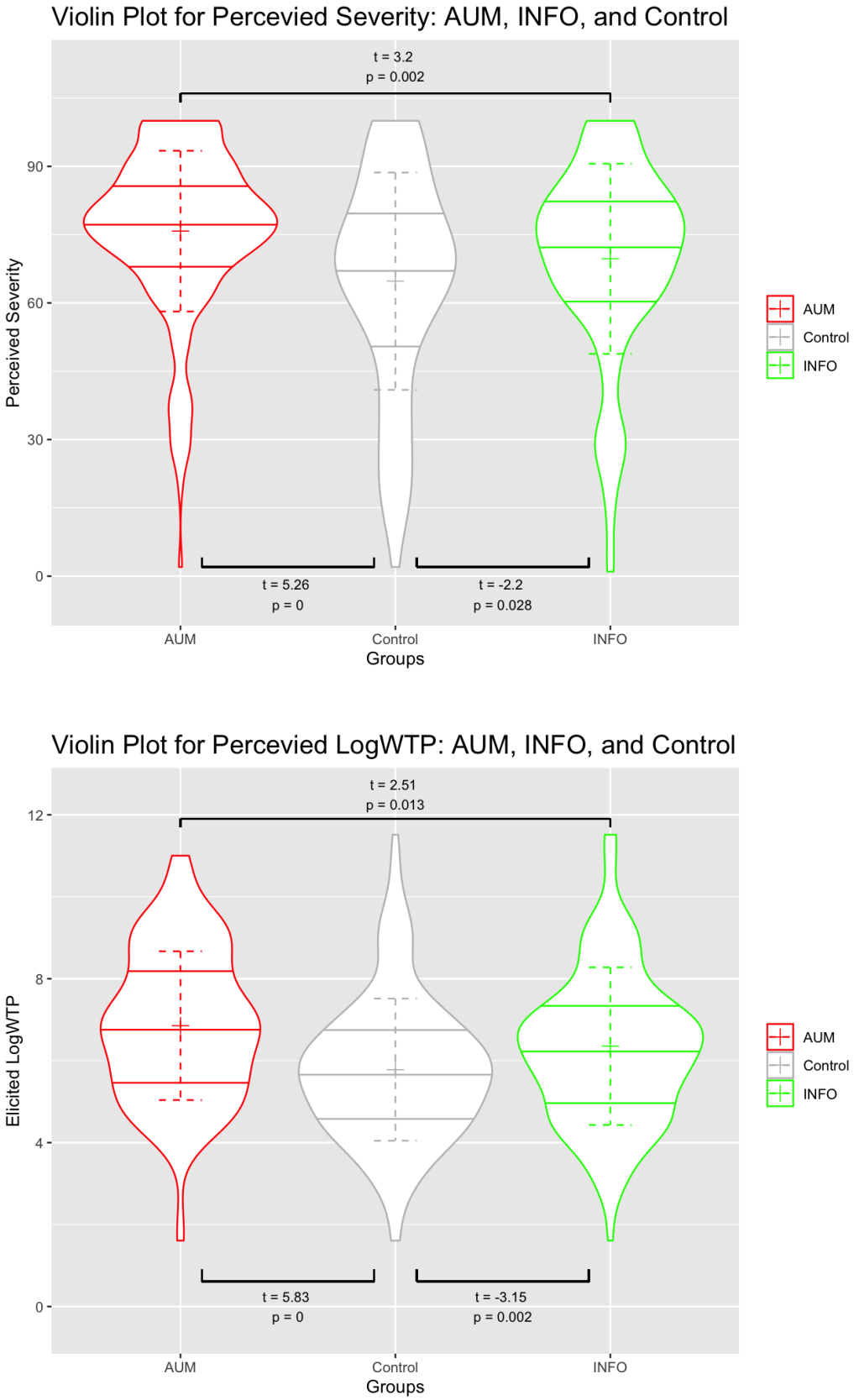
Afterward, participants were instructed to imagine that the health outcome described in the previous section would occur to them. They then answered questions about the severity rating of the health outcome using a 0-100 slider scale. Subsequently, participants were asked about their willingness-to-pay (WTP), in dollars, for a hypothetical prevention. This prevention would be guaranteed to eliminate the possibility of experiencing the health outcome, which would otherwise surely occur without any preventive measures. The design of this phase aimed to create a conceptual environment where participants could evaluate the negative health outcome without the influence of risk and uncertainty.

4.2.2 Results

After removing extreme outliers, we present two violin plots. These plots depict the responses of individuals in three different experimental conditions, focusing on their WTP (in log form) and Perceived Severity. In this experiment, we mainly focus on logarithms because of their higher robustness against extreme values. We did the balanced tests for the three groups, and random allocation is showcased[5].

[5]In all following experiments, balance tests are assumed to be passed and only abnormal results (if any) will be reported.

Figure 5: Perceived Severity and Log-WTP for a Prevention in Three Groups



The base of the logarithm is e. The violin plot indicates the density function at each value on the Y-axis. The solid horizontal lines are 25%, 50% (median) and 75% quantiles. The crossings in center positions are means. The dashed horizontal lines are +1 and -1 standard deviations.

Compared with the control, the INFO group has a significantly higher severity rating of the health outcome ($t(401) = 2.20$, $p=0.03$, Cohen's $d = 0.219$), and a significantly higher willingness-to-pay (in log form) for the hypothetical prevention ($t(401) = 3.14$, $p=0.002$, Cohen's $d = 0.313$).

Compared with the control, the AUM group has a significantly higher severity rating of the health outcome ($t(406) = 5.31$, $p<0.001$, Cohen's $d = 0.526$) and a significantly higher willingness-to-pay (in log form) for the hypothetical prevention ($t(406) =5.82$, $p<0.001$, Cohen's $d = 0.576$).

Finally, the AUM group outperformed the INFO group in increasing people's severity rating of the health outcome ($t(413) = 3.20$, $p=0.001$, Cohen's $d = 0.314$) and people's willingness-to-pay (in log form) for the hypothetical prevention ($t(413) = 2.51$, $p = 0.01$, Cohen's $d = 0.246$). All of our major findings are robust to adding common control variables, such as age, income, and completion time.

The findings in Experiment 2 demonstrate that the AUM procedure significantly enhances the level of the risk perception. Indeed, the median response within the AUM group ($1000) was four times as high as the Control group ($250), and twice as high as the INFO group ($500). It is important to note that these numbers remain significantly lower than a three-month wage loss for the majority of participants. This suggests two key insights: (1) participants did not merely replicate their earlier calculations, and (2) there was approximately no over-reporting. Even with the AUM procedure, participants still underestimated the true magnitude of the potential loss.

4.2.3 Exploring Heterogeneous Effects: Do We Have to Calculate Incorrectly?

In this experiment, an intriguing observation was that only 129 out of 210 (61.4%) participants correctly completed the AUM task. This might be due to a lack of attention or lower numeracy skills. However, what we are more interested in is whether the AUM still has a significant effect on those who made mistakes in the AUM calculation task. Therefore, in this section, we conducted two exploratory heterogeneity analyses. The first is a direct analysis where we compared the LogWTP of the 129 correct participants with the 81 who made errors. We found a statistically significant difference: those who were correct (mean=7.16, sd=1.68) had a notably higher WTP than those who were incor-

rect (mean=6.28, sd=2.03) with t=3.40, p<0.01. This seems to suggest that calculating correctly leads to a more pronounced effect.

However, this analysis did not account for the inherent differences between those who answered correctly and those who did not, which limits its interpretability. In this experiment, the INFO group had the same information set and performed a numerical task similar to the AUM group, which can help us delve deeper into the story behind the heterogeneity. We first conducted a balance test and found no significant difference in the accuracy rates between the two groups (with 132 correct participants in the INFO group, an accuracy rate of 64.3%, and the pairwise test p-value > 0.1). This suggests that regardless of the group, there might be participants with high or low numeracy and/or attention, and they can be matched correspondingly. Therefore, we can use a matching method to examine the effects of AUM in these two distinct populations.

In the INFO group, participants who answered correctly had a mean LogWTP of 6.64 (sd=2.05), while those who answered incorrectly had a mean LogWTP of 5.83 (sd=1.55). Within AUM and INFO groups, a joint Chow-test (with the interaction term being AUM and correctness) indicates no statistically significant differences (t=0.17) between participants who answered correctly and those who did not. A similar lack of significance is observed for perceived severity.

These findings suggest that, for now, what we have observed is that AUM is effective for both the high numeracy/attention correct group and the low numeracy/attention incorrect group, with little difference between them. This means that even for those who may not have sufficient numeracy or did not pay enough attention, AUM might still enhance their risk perception. However, the current results have not delved into whether correcting the answers for the low numeracy/attention incorrect group to match the correct group would be significantly beneficial. We explore this issue in more depth in the next experiment.

4.3 Experiment 3: Salience and Evaluability as Major Mechanisms of Risk Perception Change

4.3.1 Experimental Procedures

Experiment 3 was designed as a natural extension of Experiment 2. While Experiment 2 introduced us to the net increasing effects of the AUM process on risk perception, Experiment 3 aimed to dissect these effects in depth. Experiment 3 reveals the individual contributions of salience, evaluability, and active engagement in shaping risk perceptions. Employing a five-arm randomized design, we differentiate the effects of increasing salience, improving evaluability, and giving agency in enhancing people's risk perception.

We adopted the same stimuli (a non-lethal lung disease lasting for three months) and dependent variables (perceived severity and willingness to pay for prevention) as in the previous experiment. Also, unlike the last experiment, we carefully manipulated the expression of information in each group to ensure that participants in the five groups received identical information.

In the *Non-salient control* group (Group 1), monetary loss information is logically guaranteed, but its salience was kept minimal. The descriptive text was '**The person has a persistent cough and fever, is short of breath, feels weak, has lost a lot of weight during the past three months, and has to take an unpaid leave meanwhile.**" Since the sick leave is unpaid, a three-month loss of wages is inevitable.

In the *Salient control* group (Group 2), the description was modified to "... **and has to take an unpaid leave and miss three months' income**." This explicit wording enhanced the salience of income loss. However, it did not quantify the loss, thus not significantly increasing its evaluability.

In Group 3, the *AUM* group, we followed the same approach as in Experiment 2. We used the same description as Group 2 but asked the participants to calculate it after seeing the description. It, therefore, accounted for both salience and evaluability. However, the AUM group does not guarantee correct decisions by individuals. Thus, evaluability might be insufficient for those who make incorrect computations, or even offer a wrong benchmark for WTP elicitation.

To cope with this effect and to answer the remaining question in Section 4.2.3, we added Group 4, "the *Full-AUM* Group," which provided an error correction and offered

the right way of calculation and the correct answer for those who made a mistake in the AUM calculation. This provides the maximally effective treatment of AUM and offers unique information on the upper-bound treatment effect.

Finally, to separately observe whether agency alone constitutes a part of the mechanism, we included the fifth group. This group also ensured both salience and evaluability. The calculated result was directly displayed on the user interface, eliminating the need for users to manually calculate. The information presented to the users was "... **and has to take unpaid leave and miss three months' income, which is about $XXXXX according to your annual income.**" This treatment may act as another potential upper-bound benchmark, as it offers the strongest anchoring effect and an authoritative conclusion, which may or may not offset the Active engagement effect.

Experiment 3 used the same dependent variables (Perceived Severity on a 0-100 scale and WTP for guaranteed prevention) as Experiment 2. Demographics, self-reported numeracy (Fagerlin et al., 2007; Zikmund-Fisher et al., 2007), and economic resilience information were also collected[6]. This experiment was conducted on MTurk through CloudResearch between July 5th and July 14th, 2023, with 671 subjects completing their experiments. To minimize the noise from insufficient attention and extreme outliers, the main analysis drops the subjects with abnormally low income[7], extremely values of willingness-to-pay elicitation[8], insufficient total response time[9], or potentially inattentive response in numeracy scales (all reporting 1 or 6). These exclusion criteria led to 490 core responses. In a robustness check, where the exclusion criteria are stricter,[10], we obtained 636 effective responses.

4.3.2 Analysis

We analyze the results from three perspectives. Firstly, we employ different model specifications and observation inclusion criteria. Our aim is to compare the most comprehensive *Full-AUM process* with the average severity ratings and LogWTP of the other four groups. We also conduct specific pairwise comparisons to determine the relative importance of the

[6] We asked whether the participant would have enough cash, or money in your checking/saving account to cover a $500 expense if the participant encountered an emergency expense. If the participants answers "Yes" the same question will be asked about a $3000 expense.

[7] $\leq$\$10,000 a year

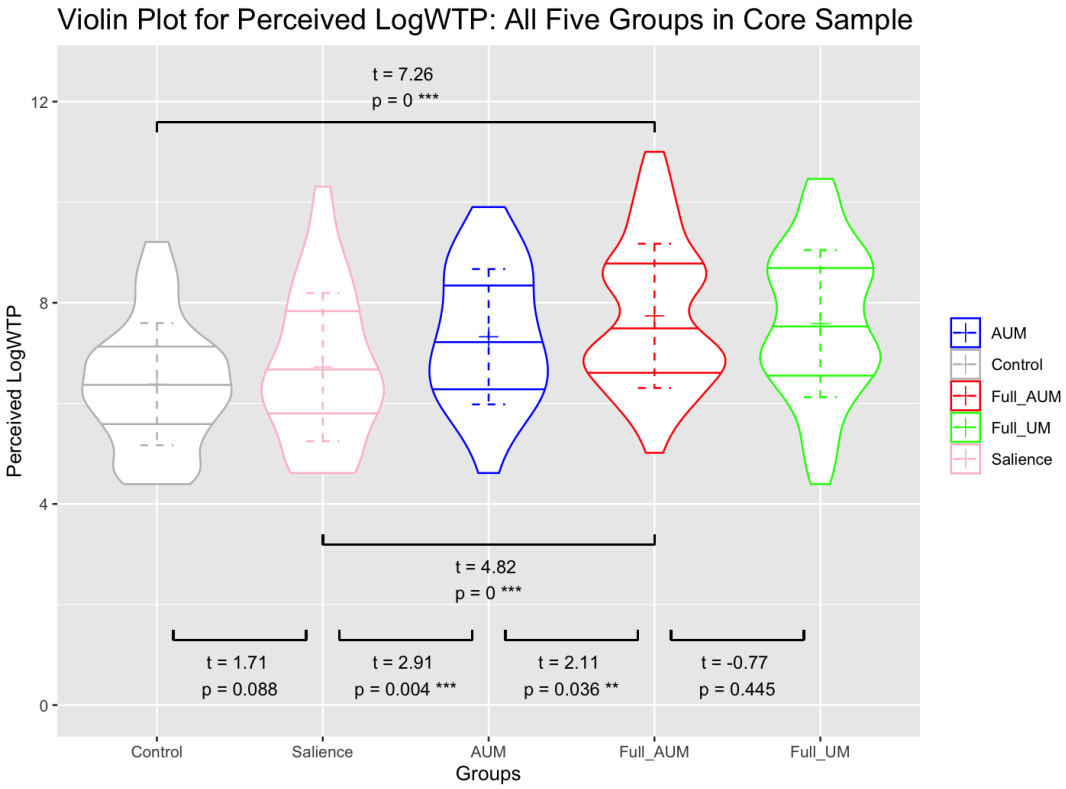
[8] $\leq$\$50 or $\geq$\$1M

[9] $<$120s

[10] WTP $\leq$\$50 or $\geq$\$1M, Income $\leq$ \$5000, response time $<$90s

three different mechanisms. Consequently, we perform exploratory heterogeneity analyses. Our aim is to examine whether the proposed mechanisms have stronger or weaker effects in specific subgroups with different socioeconomic backgrounds or psychological characteristics. This analysis aims to understand the potential variations in the effects of the mechanisms across different populations. Finally, we explore alternative mechanisms and conduct robustness checks to demonstrate why they are unlikely to significantly impact our results. This evidence underscores the robustness of our findings. It supports the argument that these alternative mechanisms are unlikely to compromise the validity of our conclusions.

Figure 6: Perceived Severity and Log-WTP for a Prevention in Five Groups



Note: The violin plot indicates the density function at each value on the Y-axis. The solid horizontal lines are 25%, 50% (median) and 75% quantiles. The crossings in center positions are means. The dashed horizontal lines are +1 and -1 standard deviations.

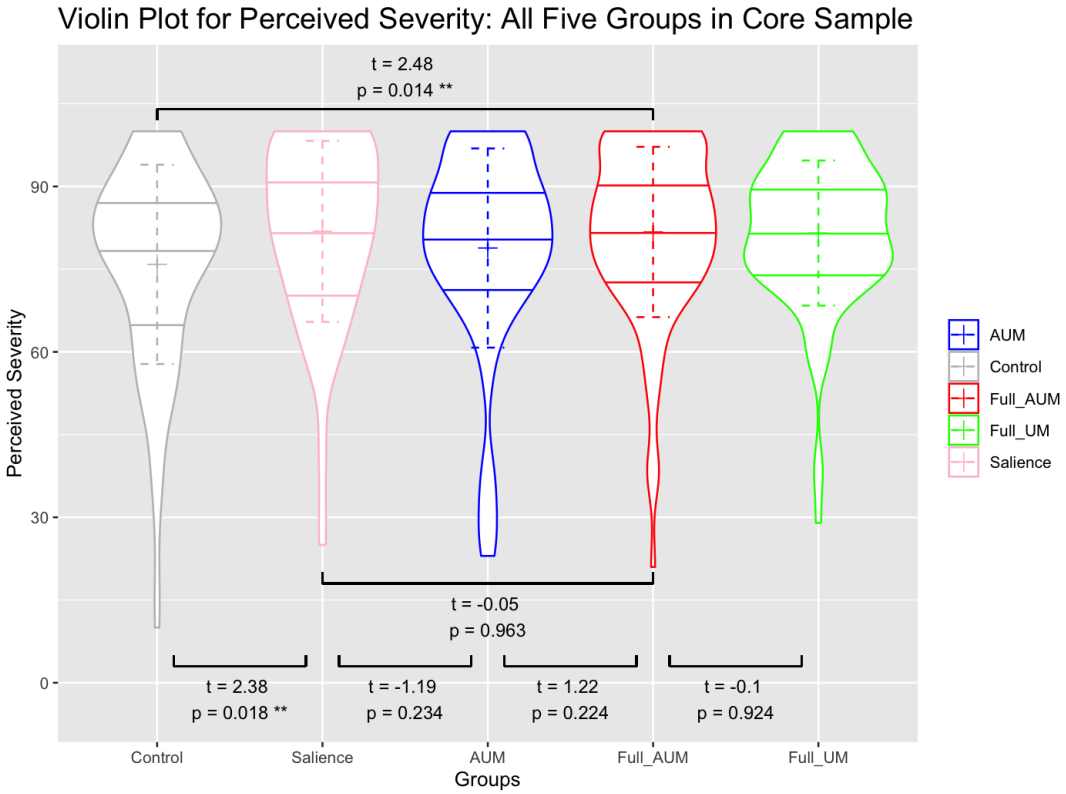
Figure 6B above reveals that the full AUM process yields the highest average Log-WTP. It is strongly higher than the two control groups that do not mention monetary loss. For instance, in terms of median, the WTP value in the AUM group is about fourfold of those in *Control* and *Salience* groups. It is also significantly higher ($p < 0.05$ in the core sample (N=490), approximately 0.10 in the full sample (N=636)) than the incomplete

AUM group. It is slightly higher (statistically insignificant) than the UM group, where the answer is directly provided to the user. This analysis further demonstrates the statistical significance of the AUM effect and its substantial economic significance.

Moreover, it indicates that the net impact solely from Active engagement is not substantial or that the positive impact of Active engagement may be offset by the authority effect of directly providing the answer and a more pronounced anchoring effect.

Pairwise comparisons between the groups mentioning amount or not indicate that numbers matter; they have set up a good benchmark for people to file in their WTP for prevention, though the existence of other biases may still block people from inputting a sufficiently large amount. This is clear evidence for the evaluability effect, which is more obvious when combining our finding in Experiment 2 that pure anchoring did not significantly matter. Finally, pairwise comparisons show that for WTP, a money reference is fundamental, while the direct salience effect is relatively limited.

Figure 7: Perceived Severity and Log-WTP for a Prevention in Five Groups



The violin plot indicates the density function at each value on the Y-axis. The solid horizontal lines are 25%, 50% (median), and 75% quantiles. The crossings in center positions are means. The dashed horizontal lines are +1 and -1 standard deviations.

However, when it comes to severity rating (Figure 7), the differences are lower across

groups. The only group with a significantly low severity rating is the non-salient control group. This shows that the salience effect plays a more crucial role when the dependent variable is subjective severity.

Combining the analysis on Log-WTP and severity in Experiments 2 and 3, we can reach the conclusion that both salience and evaluability play crucial roles in shaping risk perception, but their dominance varies based on the outcome measure. Salience seems more influential for subjective severity ratings, while evaluability has a more pronounced effect on WTP. Nevertheless, the evidence effects of Active engagement may be thin if we offer the strongest non-active intervention (directly giving the numerical answer).

To further investigate the mechanisms and check the robustness, we conduct heterogeneous analyses by interacting with our crucial pairwise comparisons such as income, economic resilience, self-reported numeracy, and correctness in AUM calculation. No strong and robust moderation is detected, and details will be seen in the forthcoming Online Appendix.

4.4 Experiment 4: AUM Enhances Sensitivity to Diseases' Time Horizon

Experiment 2-3 discussed how AUM helps resolve the opportunity cost neglect problems regarding prevention and enhanced risk perception. In the following three experiments, we will show that AUM not only increases the level of health risk perception but also the sensitivity to the numerical difference underlying health risks. Particularly, Experiments 4-5 focus on duration neglect, and Experiment 6 on probability insensitivity.

Duration neglect (time horizon insensitivity) is a commonly detected decision bias that the evaluation of unpleasant experiences responds insufficiently to time duration (Fredrickson and Kahneman, 1993; Holyoak and Morrison, 2005; Morewedge et al., 2009), and this insensitivity is particularly stronger for unfamiliar scenarios (Morewedge et al., 2009). However, the real impact and course of many health issues are closely related (or even directly proportional) to duration, considering factors such as medical expenses for certain treatments or wage losses due to sick leave. In the context where these losses play a significant role, it is necessary to consider the duration factor when assessing potential health consequences to protect their own interest. If people exhibit significant duration neglect when faced with these hypothetical scenarios, they may adopt inadequate or

excessive preventive measures, leading to welfare losses.

Experiment 4 and Experiment 5 employed a between-subject and a within-subject (between-subject analysis still available) experimental design, respectively, to explore the alleviation of duration neglect through AUM. We provided robust evidence to demonstrate that AUM is indeed effective in addressing duration neglect. Furthermore, we discussed why such improvements will likely lead to welfare improvements.

4.4.1 Experimental Procedures

We designed a 2×3 factorial design to separate out the effect of pure calculation (priming a calculative mindset) and the direct effects of AUM. In the control group, we investigated the scope sensitivity with a description mentioning only the physical consequences.

Experiment 4 was conducted on MTurk through CloudResearch, and the data was collected on Apr. 3, 2023. After informed consent, participants were randomly assigned to one of the six groups (Control-8, Control-16, Calculation-8, Calculation-16, AUM-8, and AUM-16). 959 participants took part in the experiment, and we acquired 846 effective responses after eliminating ineligible and potentially inattentive[11] or extreme responses[12].

Every participant saw a description of 8- or 16-week-long Angina Pectoris adapted from the WHO document. The four arms' participants in Control and Calculation groups saw the description with only symptom and duration information. However, the two AUM groups saw the same description plus a sentence showing information about financial losses. Then, the AUM group was asked to calculate the hypothetical financial loss from losing three months' personal income based on the income elicitation at the beginning of the experiment. The Calculation groups, designed to tease out numerical priming effects, did an irrelevant placeholder task which is the same as that in Experiment 2. The control group did no calculation. In this experiment, we allowed for information difference for two reasons: (1) we are studying the evaluability effect, not the salience effect, which is justified by Experiments 2 and 3; and (2) we intend to study this effect as if it is a real-world comparison between a description without mentioning money and an AUM scenario.

After the treatment, participants imagined the hypothetical scenario that the health outcome would occur to themselves and elicited their perceived severity and WTP for

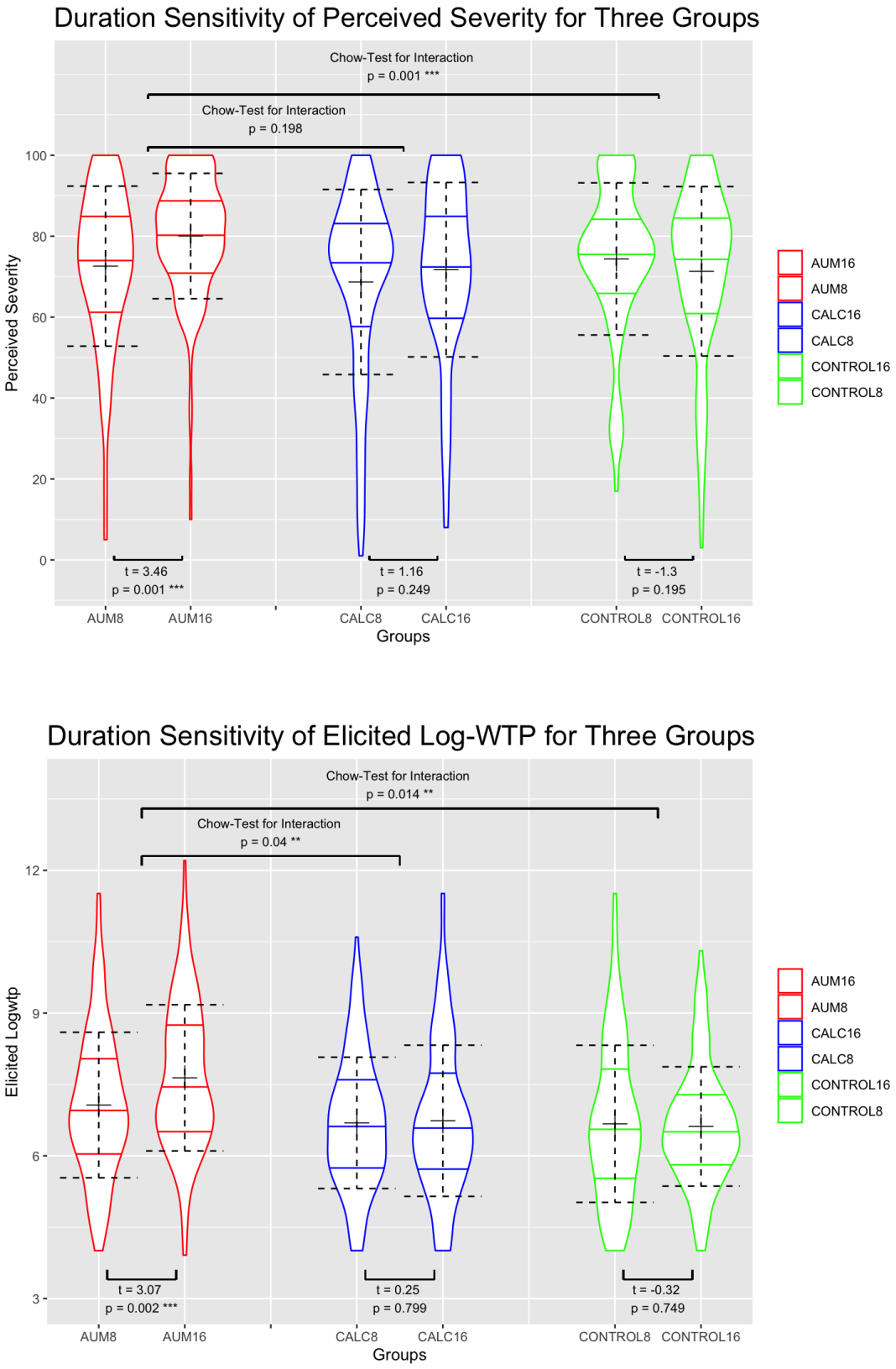
[11] Response time <75s
[12] WTP<$50 or >$500K

a guaranteed preventive measure following Experiment 2. Necessary demographics were collected.

4.4.2 Results

We begin by conducting a pairwise check on the time duration sensitivity across the three pairs of comparisons. Using T-tests, we found that the time horizon differences in the Control and Calculation groups did not result in any statistically significant differences in severity perception (mean (Control8) = 74.4, mean (Control16) = 71.3, t=-1.30, p=0.19; mean (Calculation8) = 68.7, mean (Calculation16) = 71.7, t=1.16, p=0.25) and the logarithm of WTP for prevention (Log-WTP) (mean (Control8) = 6.67, mean (Control16) = 6.62,t = -0.32, p = 0.75; mean (Calculation8) = 6.69, mean (Calculation16) = 6.73, t=0.255, p=0.80). This shows that without mapping with money, consumers may show insensitivity to the time duration of an unfamiliar disease in a hypothetical scenario. Moreover, there is evidence that this insensitivity cannot be mitigated by merely priming a calculative mindset. Hypotheses 1A and 1B are rejected. On the contrary, we find that the AUM treatment was effective. The t-tests showed significant horizon sensitivity in both severity perception (mean(AUM8)=72.9, mean(AUM16)=80.0, t=3.46, p<0.001) and LogWTP (mean(AUM8)=7.07, mean(AUM16) = 7.64, p=0.014).

Figure 8: Perceived Severity and Log-WTP for a Prevention in Three Groups



The base of the logarithm is e. The violin plot indicates the density function at each value on the Y-axis. The solid horizontal lines are 25%, 50% (median) and 75% quantiles. The crossings in center positions are means. The dashed horizontal lines are +1 and -1 standard deviations.

To ensure robustness, we evaluated the statistical significance of the interaction term using 2×2 comparisons for severity, both with and without demographic control variables. The results suggest that except for the perceived severity rating of the AUM-Calculation comparison, all regressions generate a positive interaction effect, indicating a relatively stable causal relationship between AUM and time duration sensitivity. This relationship is robust to adding various control variables. Results are shown in the following table:

Table 2: Interaction Analysis (Chow-Test) of AUM Effect on Duration Sensitivity

Table 2: Interaction Analysis (Chow-Test) of AUM's Effect on Duration Sensitivity

	AUM vs. Control				AUM vs. Calculation			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	LogWTP	Severity	LogWTP	Severity	LogWTP	Severity	LogWTP	Severity
AUM x Duration	0.626**	10.50***	0.653***	9.599***	0.526**	4.40	0.478*	3.07
	(2.46)	(3.29)	(2.64)	(3.03)	(2.06)	(1.29)	(1.91)	(0.90)
Duration = 16w	(0.06)	(3.05)	(0.11)	(2.64)	0.04	3.04	0.06	3.70
	(-0.31)	(-1.37)	(-0.64)	(-1.20)	(0.25)	(1.27)	(0.37)	(1.56)
AUM	0.395**	(1.79)	0.395**	(0.91)	0.376**	3.92	0.380**	5.079**
	(2.23)	(-0.80)	(2.29)	(-0.41)	(2.07)	(1.62)	(2.13)	(2.09)
Economic Resilience			0.341***	-2.990***			0.321***	-2.820**
			(3.94)	(-2.70)			(3.62)	(-2.33)
Health Status			(0.04)	(0.44)			0.03	-1.796*
			(-0.53)	(-0.48)			(0.39)	(-1.81)
Log-Income			0.118*	2.144**			0.12	0.59
			(1.69)	(2.41)			(1.62)	(0.57)
Age			0.0179***	0.187***			0.0138***	0.138**
			(3.70)	(3.03)			(2.76)	(2.02)
No Paid Leave			0.364**	1.96			0.345**	1.65
			(2.20)	(0.93)			(1.97)	(0.69)
Paid Leave>=3 Months			0.24	-5.913*			0.447*	(0.42)
			(0.98)	(-1.86)			(1.96)	(-0.14)
Constant	6.672***	74.39***	3.893***	51.49***	6.692***	68.69***	3.826***	68.43***
	(54.38)	(48.25)	(4.97)	(5.14)	(52.40)	(40.27)	(4.56)	(5.98)
Observations	560	560	560	560	558	558	558	558

t statistics in parentheses, * p<0.10, ** p<0.05, *** p<0.01

Experiment 4 showcases that people may exhibit strong duration neglect in regular, money-free judgments regarding preventive health. This insensitivity is arguably deviating from optimal decisions and will make decision-makers insensitive to information regarding time duration, which may be economically and physically important for consumers. Plenty of evidence shows that AUM helps mitigate this effect. However, the findings in Experiment 4 are purely between subjects. People were exposed to only one stimulus in the experiment, which corresponds to the separate evaluation case as mentioned in Hsee (1996). In the real world, there are possible cases in which people may compare their choice with other alternatives (joint evaluation), and in the next experiment, we investigate our intervention's effectiveness within subjects, which resembles joint evaluation scenarios.

4.5 Experiment 5: AUM Enhances Sensitivity to Diseases' Time Horizon - Further Evidence

Experiment 5 primarily acts as a within-subject extension of Experiment 4. A within-subject variation is helpful for us to under the effects of AUM under joint evaluation when it is possible to make comparisons among the alternatives. Simultaneously, it can further illustrate our sensitivity narrative on a broader numerical scale. We explore participants' willingness-to-pay for a guaranteed preventive measure against Angina Pectoris, examining varying time horizons of 5, 7, 10, and 15 weeks.

4.5.1 Experimental Procedures

We posit that exposing subjects to these four scenarios can enhance within-person sensitivity to the time horizon through the AUM procedure. From a theoretical standpoint, the within-person evaluation represents a joint evaluation (JE) scenario (Hsee, 1996; Hsee and Zhang, 2010), where individuals are anticipated to exhibit heightened sensitivity to time horizons compared to separate evaluations. We further hypothesize that the AUM process will amplify within-person sensitivity during joint evaluations. In other words, the AUM practice is expected to magnify the effects observed in joint evaluations. Theoretically, this expectation stems from the idea that monetary values provide participants with a tangible reference, aiding them in formulating willingness-to-pay reports and comparing relative severities across different JE scenarios. We depict the within-subject WTP response function to the time duration, both with and without the AUM practice.

Additionally, we analyze the between-subject pattern to both extend and validate the findings of Experiment 4. Since participants were initially unaware of subsequent questions with varying time horizons, their initial responses can serve as a between-subject measure for time horizon sensitivity under separate evaluations. Certain factors, such as the experiment's suggested duration of 4-5 minutes, might lead participants to suspect the existence of other groups. Thus, we primarily interpret the between-subject findings as both an extension and a validation of our results.

Experiment 5 was conducted on MTurk through CloudResearch, and the data was collected between May 23rd and June 2nd, 2023. 1200 Participants participated in this experiment, and we finally got 877 effective responses after eliminating duplicates, poten-

tially inattentive[13] subjects, and subjects who had reported at least one extreme value that may be significant outliers (too low or too high income[14], too extreme willingness-to-pay[15], too extreme rating[16]).

During the survey, we gathered data on participants' annual individual income, demographics, and basic economic conditions. As in Experiment 4, each participant was presented with a description of Angina Pectoris, adapted from a WHO document, detailing durations of 5, 7, 10, or 15 weeks. Apart from the duration, descriptions within each group remained consistent. Participants in the four Control arms saw the description with only symptom and duration information. In contrast, those in AUM arms received an additional sentence about potential financial losses. Then, the AUM group did the same calculation procedures as in Experiment 4, and the Control groups did no calculation. Finally, subjects reported their severity rating of these health conditions and elicited their WTP for a guaranteed preventive measure.

4.5.2 Analysis and Results

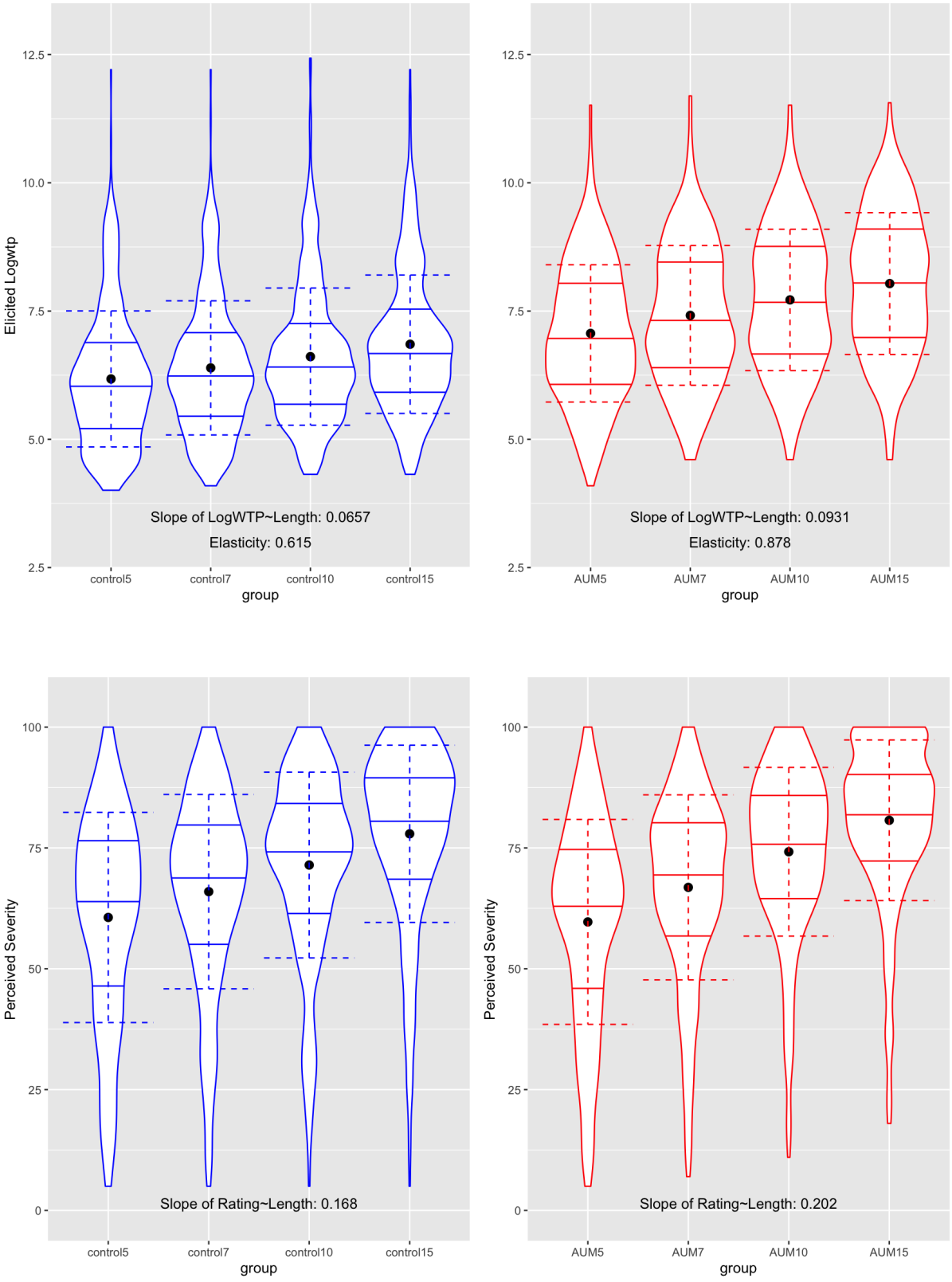
This experiment examines if AUM heightens participants' sensitivity to illness duration when presented with varying disease duration information. Our initial test determines whether there is a statistically significant difference in the "sensitivity coefficient" between the AUM and control groups. Mathematically, this is equivalent to the following Fixed Effect Regression:

$$LogWTP_{i,l} = \alpha_1(length_{i,l} \times AUM_i) + \alpha_2 length_{i,l} + \alpha_3 AUM_i + FE_i + \epsilon \qquad (1)$$

In the aforementioned regression, the interaction term arises from the product of time length and the AUM treatment. The regression coefficient α_1 reflects the main treatment effect.

[13] too low completion time, <120s
[14] <$1k or >$1M
[15] <$50 or >$500k
[16] 0 or 100 in severity rating

Figure 9: Perceived Severity and Log-WTP for a Prevention in Three Groups



The base of the logarithm is e. The violin plot indicates the density function at each value on the Y-axis. The solid horizontal lines are 25%, 50% (median) and 75% quantiles. The crossings in center positions are means. The dashed horizontal lines are +1 and -1 standard deviations.

The graphs indicate two key insights: (1) A joint evaluation setting, where participants can readily compare alternatives, typically enhances sensitivity to time duration, and (2) within this joint evaluation, AUM further amplifies this sensitivity. Following standard procedures, we define the *elasticity* of WTP with respect to the time horizon as:

$$e_{WTP,t} = \frac{\partial Log(WTP)}{\partial t} \tag{2}$$

This can be estimated by regressing $LogWTP$ against $Log(time)$, incorporating within-subject control variables. We reveal that this elasticity rises from 0.615 to 0.878 with the introduction of AUM, a statistically significant difference ($p<0.001$). When shifting to perceived severity as the dependent variable, the difference in sensitivity persists, albeit with diminished numerical magnitude and statistical significance ($p<0.05$).

Another crucial metric within subjects is response monotonicity. Assuming constant factors such as disease severity, a rational individual's willingness-to-pay for prevention should consistently rise with extended disease duration. Upon examining participants' response patterns we observed that in the absence of AUM, only 35.3% of reported WTP sequences exhibited a strict increase from 5 to 15 weeks. AUM elevated this rate to 46.8% ($t=3.48$, $p<0.001$). Adopting a more lenient, non-strict monotonicity criterion, the rates still saw a notable rise: 78.4% for those without AUM and 84.7% for those with AUM ($t=2.41$, $p=0.016$). This result compellingly indicates that the AUM intervention heightens participants' sensitivity to disease duration, yielding more refined responses.

As a concluding step, we conducted a between-subject robustness check to complement the findings of Experiment 4. Given that participants faced the four questions in a randomized sequence across distinct screens, it's plausible they were unaware of other groups during their initial elicitation task. Therefore, by considering only the participants' initial responses, we can approximate a between-subject analysis. The between-subject analysis revealed that the elasticity of willingness-to-pay concerning the time horizon was 0.23 (non-significant) in the control group. In contrast, this elasticity was 0.74 in the treatment group ($p<0.001$), with the Chow test confirming a significant difference ($p<0.05$). This is an effective robustness check of Experiment 4, underscoring AUM's overall efficacy in counteracting duration neglect.

Detailed robustness checks in Experiment 5 will be available soon the Online Appendix.

4.6 Experiment 6: AUM Enhances Probability Sensitivity

4.6.1 Experiment Procedures

A notable limitation of Experiments 2-5 is the assumption that preventive measures are guaranteed to be effective, which realistically only applies to specific cases like vaccines offering lifetime immunity, such as the measles vaccine. In many cases, preventive measures are not guaranteed to be effective; rather, they reduce the risk of disease onset probabilistically. For instance, the COVID-19 and influenza vaccines have efficacy rates that range from 60% to 95%. To investigate the role of AUM in probabilistic preventive decision-making, we designed Experiment 6. The goal of this experiment is to examine the impact of AUM on decision-makers' sensitivity to varying efficacy rates, particularly in terms of their willingness-to-pay for preventive measures.

In Experiment 6, we focused on Angina Pectoris, the same disease examined in Experiments 4 and 5, and maintained a constant potential duration of three months. Our numerical settings were based on prospect theory. Prospect theory posits that individuals are overly sensitive to changes in extreme probabilities (close to 0 or 1) while showing insufficient sensitivity to changes in intermediate probabilities (Wu and Gonzalez, 1996; Prelec, 1998). To avoid interference between decisions made at extreme and intermediate probabilities, we set the upper and lower limits for the disease occurrence probability at 80% and 20%, respectively. Theoretically, this implies that people's decisions within this range will always exhibit the bias of probability insensitivity. Our hypothesis is that AUM can mitigate this bias, enhancing sensitivity to probability changes.

We employed a 2×4 factorial design similar to that of Experiment 5. In this experiment, all participants were informed that the likelihood of contracting the disease was 80%, with a duration of three months. Participants were then randomly assigned to either one of the four experimental groups or one of the four control groups. To better isolate the effect of AUM, we implemented a strong control measure: both the AUM and control groups received identical information about a potential loss of three months' income. The only difference was that the AUM group was asked to calculate the exact financial loss equivalent to three months' income, whereas the control group was not.

Each participant was presented with four different preventive measures, each having a different efficacy rate. These measures were designed to reduce the risk of contracting

the disease by 20, 30, 40, and 60 percentage points, translating to reduced probabilities of 60 percentage points (pp), 50pp, 40pp, and 20pp, respectively. Participants were asked to state their WTP for each of these preventive plans. The presentation order of these measures was randomized for each participant. As in Experiment 5, in this experiment, we can conduct a within-subject analysis for the full sample with the Fixed Effect model. Additionally, we can perform a between-subject analysis using the first WTP value reported by each participant.

The analytical approach for this experiment differs somewhat from the previous ones. Neoclassical health economics typically assumes that individuals act based on expected utility and quasi-linear utility (Mas-Colell et al., 1995). This suggests that, when outcomes are constant and only probabilities change, a rational individual's willingness-to-pay should be proportional to the changes in probability values. Therefore, it is theoretically justified to use raw WTP values instead of logarithms in the main regression. The regression function for within-subject analysis is as follows:

$$WTP_{i,j} = \alpha_1(Reduction_{i,j} \times AUM_i) + \alpha_2 Reduction_{i,j} + \alpha_3 AUM_i + FE_i + \epsilon_{i,j} \quad (3)$$

in which i denotes the i-th participant, and j denotes the preventive measure seen by the participant. For the between-subject analysis, we exclude the subscript j from the regression, as it considers only the initial response. In this scenario, we incorporate subject-level controls denoted by $Controls_i$. The between-subject regression is as follows:

$$WTP_{i,j} = \alpha_1(Reduction_i \times AUM_i) + \alpha_2 Reduction_i + \alpha_3 AUM_i + Controls_i + \epsilon_i \quad (4)$$

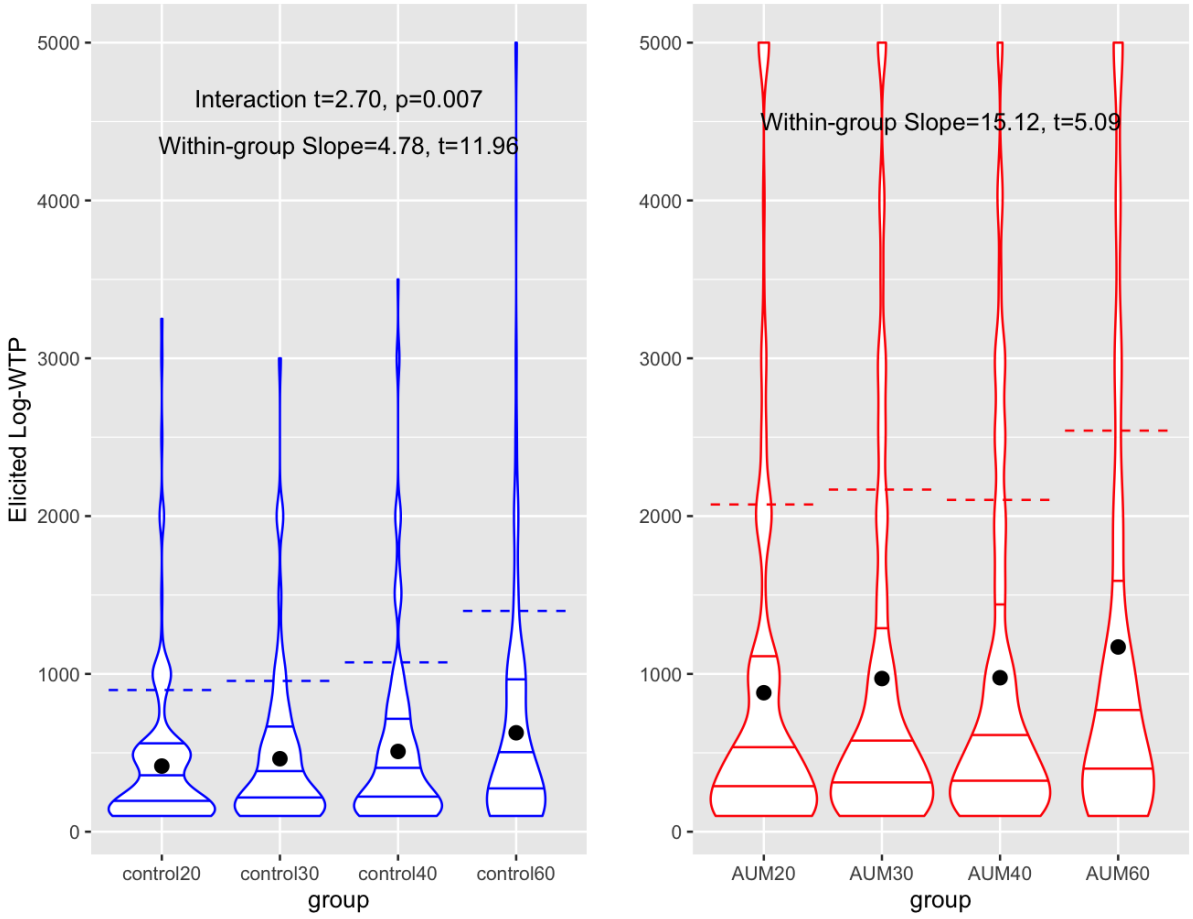
We commenced our data collection for this experiment on July 28, 2023, and concluded it on August 8, 2023. Our sample comprised 1,136 compliant individuals. After excluding plausibly inattentive samples[17], we were left with 957 participants. Since the baseline analysis of this study uses WTP rather than Log-WTP as the dependent variable, it was necessary to eliminate outliers in a more systematic way. Thus, all samples with first responses falling below the 10th percentile or exceeding the 95^{th} percentile were excluded, yielding a final sample of 850 individuals.

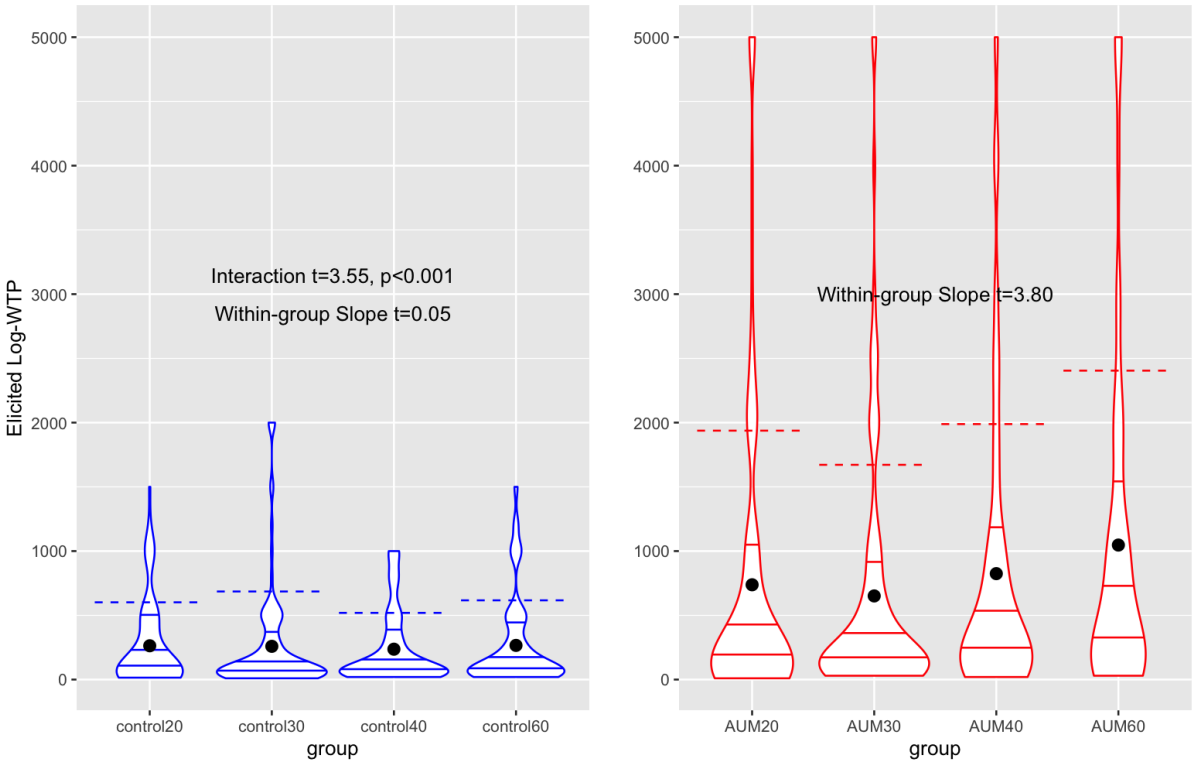
[17] completed in less than 90 seconds, or wtp <= \$1

4.6.2 Results

The ensuing graphs illustrate the results of this experiment, encompassing both within-subject and between-subject analyses.

Figure 10: Perceived Severity and Log-WTP for a Prevention in Three Groups





The violin plot indicates the density function at each value on the Y-axis. The solid horizontal lines are 25%, 50% (median) and 75% quantiles. The black solid circles in the center positions are means. The dashed horizontal lines are +1 standard deviations. However, if we deduct a standard deviation from the mean, it turns below 0 and, therefore not displayed in the graph.

The results showcase patterns that share some features with Experiment 5. In the separate evaluations, as reflected in the between-subject analysis, there's a notable lack of sensitivity to the preventive measure's effectiveness without AUM involvement. Simultaneously, many participants reported extremely low WTP values. Such insensitivity to probability cannot rationalized by *ANY utility function*, implying irrational behavior. Our findings highlight that AUM effectively mitigates this issue, significantly enhancing participants' sensitivity to probabilities.

As in Experiment 5, joint evaluation, as depicted by the within-subject analysis, enhances sensitivity to probabilities. There is also significant evidence that AUM groups lead to higher sensitivity than control groups in this condition, but the effect size seems smaller than in separate evaluation.

It is noteworthy that the WTP responses exhibit considerable skewness, suggesting a robustness check with logarithms. When expressing both WTP and probabilities in logarithmic forms, the interaction term remains significant in the between-subject analysis (t=2.55, p=0.011). However, the interaction term in the within-subject analysis is not

significant (t=0.11, p=0.91). Based on these findings, we posit that the between-subject results are notably robust, whereas the within-subject analysis yields a less pronounced treatment effect, lacking significance across all regression configurations..

Our observations indicate that while the between-subject treatment effects are notably robust and pronounced, the within-subject treatment effects appear more subdued. We theorize that this discrepancy might stem from the evaluability of probabilistic data and individuals' interpretation of probability within the framework of Expected Utility Theory. Once the participants engage in joint evaluation, the distinctions among probabilities like 20%, 30%, 40%, and 60% become glaringly clear. Within this context, irrespective of the actual outcome, the prominence of the numerical data in the outcome is significantly eclipsed by the inherent salience of the joint evaluation, making the treatment effect somewhat muted. On the other hand, in separate evaluation scenarios (as in between-subject conditions), the interpretability of probability data in Experiment 6 probably falls short of that of disease duration data, primarily due to the non-intuitive nature of prevention efficacy figures. Hence, without a comprehensive understanding of the outcome, participants lack a foundational reference for assessment, prompting them to offer responses that are not only significantly low but also lack sensitivity to probability variations. However, when individuals compute economic losses as part of the outcome, it assists them in estimating the expected magnitude of the loss, thereby markedly improving their evaluability. This, consequently, results in heightened sensitivity to probability variations.

5 General Discussion

5.1 Summary of Results

Our study methodically examines behavioral biases in preventive judgment, presenting AUM as an innovative intervention that encourages participants to actively quantify the economic implications of illnesses, acting as a remedy for these biases. In Experiments 2 and 3, we experimentally assessed the proposition that AUM amplifies risk perception. By emphasizing the salience and evaluability of the economic impacts of diseases, AUM counters biases like opportunity cost neglect, enhancing individuals' risk awareness and their willingness-to-pay (WTP) for preventive actions. Experiments 4, 5, and 6 provide

additional evidence for AUM's benefits from the mechanism of evaluability. Utilizing economic loss as a reference point and scale, AUM heightens the clarity of disease-related duration and probability data, countering biases such as duration neglect and probability insensitivity. This approach ensures individuals more effectively integrate this information into their health-related decisions. Our results demonstrate significant consistency and effect magnitude, providing a firm basis for exploring both its theoretical insights and practical applications.

5.2 Strengths

The novelty of this design lies in its ingenious use of monetary loss calculations as a mediator, achieving two objectives simultaneously. On one hand, it heightens the salience of economic losses in people's perception, addressing the issue of limited attention. On the other, it translates the otherwise abstract and emotional health losses into quantifiable monetary values, improving the evaluability problem. To our knowledge, this paper is the first to systematically categorize health decision biases into these two dimensions and propose interventions that effectively address both simultaneously.

Another notable strength of our study is its extensive scale and resilience to experimental setups. We gathered health evaluations from nearly 5,000 participants, examining their decision-making behaviors concerning various diseases. For lab-based research, this scale is substantial, allowing us to scrutinize the primary experimental effect with strong statistical power and stable effect size across diverse demographics, socioeconomic statuses (such as the availability of paid sick leave and whether living paycheck to paycheck), and individual decision-making characteristics (such as numeracy and risk preferences). This approach enables a thorough examination of AUM's relatively unanimous impact for different groups of individuals. Our findings, which remain consistent across different model specifications, bolster our confidence in AUM's potential real-world applicability.

Additionally, this paper's practical orientation provides actionable insights for future implementation. We can provide recommendations for two scenarios. The first scenario involves the most common web platform disease descriptions.Using COVID-19 and Long COVID as illustrations, health communicators could emphasize the potential economic impacts of diseases, such as long COVID, whose major symptom is fatigue that may lead to months of inability to work. In the absence of extended paid sick leave, this

could translate to significant wage loss. Health communicators could engage users with a prompt, suggesting, "You can input your monthly income figure to estimate the financial loss you could potentially avoid, on average, if you evade long COVID." TThe crux of this strategy is to swiftly make website visitors aware of the potential economic implications of the illness. This feature should be prominently positioned on the website, ensuring users can easily access and utilize it. (Evidently, the UM approach might not be as apt, given that websites can't pre-access users' income data.) The second context pertains to direct medical communication, such as interactions between insurance companies and clients, doctor-patient conversations, or caregivers guiding at-risk individuals. Given the substantial difficulties of casting field experiments with a novel nudge, this paper did not undertake in-depth empirical tests for this context. Nonetheless, we offer actionable insights grounded in our findings and theoretical framework. Generally, if AUM is conducted between two parties in close proximity, the process can be further personalized. The communicator can facilitate the other party's calculations while also prompting deeper reflection on the economic implications of illnesses, including potential medical expenses and wage losses, and their potential size compared to the cost of prevention. This thoughtful engagement not only promotes salience and evaluability, but also stimulates System II thinking, empowering the other party to make well-informed decisions aligned with their best interests.

Lastly, it is particularly important to note that our consistent use of online experiments doesn't strongly compromise external validity. In a digital world, many individuals source disease prevention information from reputable online platforms, such as the websites of the CDC and Mayo Clinic. Users often browse these sites fast, without deep reflection. Many prevention decisions, simultaneously, are made after this brief information acquisition. Therefore, internet-based 'System 1' decisions may play an important role in real life. The online experiments regarding the AUM in this paper are a good exemplification that preliminarily explores the potential of AUM as a concise online intervention.

5.3 Limitations and Future Perspectives

This study has several significant limitations. Given constraints related to paper length and research costs, it's challenging to exhaustively address all facets in one paper, prompt-

ing this section to outline future prospects for AUM interventions as well.

A primary limitation is the study's exclusive focus on the U.S. context and the concept of sick leave. In the U.S., most people have health insurance, but statutory paid sick leave is less common. Hence, for individuals with adequate insurance, the absence of sick leave might represent a significant economic setback, leading to our emphasis on sick leave as a primary economic concern. In practical scenarios, particularly in developing regions like China, challenges like exorbitant medical costs and healthcare accessibility might be more pronounced, potentially heightening the local populace's demand for AUM. Conversely, given the steep healthcare costs in such regions, financial setbacks might be viewed more as direct health-related losses, with decision-makers already recognizing their gravity. Such disparities strongly motivate us to expand our research methodically, delving into the contexts of developing nations.

An additional limitation pertains to the variability in income, financial stability, and sick leave statuses, along with the ensuing practical challenges. While our findings indicate that factors like income and economic resilience primarily act as covariates without significant interaction with our treatments, this could be attributed to an inadequate sample size preventing a comprehensive exploration of all potential interactions. Despite its good emulation of real-world scenarios, the current version of AUM is still largely conceptual and does not fully incorporate many details in preventive decision-making. For example, the detailed breakdown doesn't factor in aspects like paid sick leave status or anticipated medical expenses, given the constraints of online survey methodologies. This drives our ambition to refine AUM, tailoring it more personally and aligning it closer to real-world decision-making scenarios in the future. Specifically, for economically disadvantaged individuals or those living paycheck to paycheck, there's potential for enhancement. A structured breakdown of their risk profile could foster a deeper understanding and heightened preventive awareness.

A potential critique is that AUM, designed as a cognitive tool to amplify risk awareness, could have adverse welfare consequences in specific contexts for particular individuals. While welfare considerations in Experiments 2-5 indicate promising potential for welfare benefits and Experiment 6 implies a guaranteed increase (because the insensitivity in Experiment 6 is not rationalizable), the outcomes aren't uniformly positive across all participants and scenarios. In the future, more comprehensive and stringent welfare

evaluations will be crucial.

Additionally, AUM might intensify individuals' anxieties, potentially harming their psychological well-being. This critique, This line of criticism, frequently labeled as "emotional taxes", interventions.(O'Donoghue and Rabin, 2006; Just, 2017; Thunström et al., 2018), poses a recurrent and formidable challenge to the wider domain of behavioral nudges and information provision interventions. HYet, recent empirical data indicates that, when viewed from a holistic welfare standpoint, the outcomes of such nudges are generally favorable, with the advantages surpassing any emotional drawbacks (Thunström, 2019; Anderson et al., 2022). This particular trade-off still exhibits heterogeneity across different populations (Kelly and Sharot, 2021). Thus, prior to the formal adoption of AUM as a policy measure, it's essential to delve deeper into its psychological impact, particularly on economically vulnerable individuals who may already grapple with financial stress or a scarcity-driven mindset (Mullainathan and Shafir, 2013; Kaur et al., 2021), which might compromise emotional well-being or productivity.

6 Conclusion

Imperfect preventive decision-making significantly impacts individuals' health, and precise risk evaluation is crucial for improving these decisions. The various experiments in this paper consistently demonstrate the pervasive presence of opportunity cost neglect of financial losses, which may lead to insufficient risk perception, and numerical insensitivity, which may encompass both duration and probability neglect, when assessing potential health outcomes. These biases usually stem from a lack of salience and evaluability. To combat these biases, we introduced the AUM intervention that utilizes monetary values. While it accentuates the prominence of financial losses, this increased salience also boosts the evaluability of health-related numerical data. Our findings indicate that this approach enhances decision-making across different contexts. The method not only heightens risk awareness but also fosters a keen understanding of vital numerical details like disease duration and efficacy rates.

Our research notably contributes to the literature of both behavioral and health economics. We've integrated the evaluability theory into economic decision-making, merging it with the concept of salience. We've examined the prevalence and significance of this

phenomenon in health decisions, outlining its psychological underpinnings and suggesting intervention strategies. Additionally, we present an innovative and straightforward nudging method to address the widespread biases in real-world preventive decisions. This approach demonstrates robust intervention outcomes and a pronounced effect size. In the end, we elucidate the potential policy implications of AUM and considerations for its practical application.

In conclusion, we envision a forward-looking theoretical and empirical research trajectory for AUM and its foundational economics. Theoretically, there's a compelling case for exploring the modeling and application of the overarching evaluability theory in the economic landscape. As an example, in a separate study of ours, we meld the general evaluability theory with the sparsity model. Employing a framework grounded in the Lasso algorithm, we weave in aspects of familiarity, mode, and nature, culminating in a refined sparsity model anchored in evaluability. Additional models could delve deeper into the nexus between evaluability and numerical acuity, drawing parallels in economics with the interplay between hedonic analysis and notions of elasticity.

From an empirical and experimental perspective, upcoming research endeavors could sculpt more bespoke and individualized AUM methodologies to bolster results. Scholars could embark on practical field tests of AUM and orchestrate comprehensive welfare assessments. The versatility of AUM also lends itself to non-financial decision arenas, like nudging students to ponder over potential GPA dips due to class absences. Such an approach paves the way for a richer, more nuanced probe into strategies that refine preventive decision-making, uplift community health selections, and realize superior results.

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes*, 50(2):179–211.

Anderson, M. L., Hyun, M., and Lee, J. (2022). Bounds, benefits, and bad air: Welfare impacts of pollution alerts. Technical report, National Bureau of Economic Research.

Baicker, K., Congdon, W. J., and Mullainathan, S. (2012). Health insurance coverage and take-up: Lessons from behavioral economics. *The Milbank Quarterly*, 90(1):107–134.

Baillon, A., Bleichrodt, H., Emirmahmutoglu, A., Jaspersen, J., and Peter, R. (2022). When risk perception gets in the way: Probability weighting and underprevention. *Operations Research*, 70(3):1371–1392.

Banerjee, A. V. and Duflo, E. (2011). *Poor economics: A radical rethinking of the way to fight global poverty*. Public Affairs.

Blackwell, C., Goya-Tocchetto, D., and Sturman, Z. (2017). Nudges in the restroom: How hand-washing can be impacted by environmental cues. *Available at SSRN 3007866*.

Bleichrodt, H. and Quiggin, J. (1999). Life-cycle preferences over consumption and health: when is cost-effectiveness analysis equivalent to cost–benefit analysis? *Journal of health economics*, 18(6):681–708.

Blumenthal-Barby, J. S. and Burroughs, H. (2012). Seeking better health care outcomes: the ethics of using the "nudge". *The American Journal of Bioethics*, 12(2):1–10.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2012). Salience theory of choice under risk. *The Quarterly journal of economics*, 127(3):1243–1285.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2013). Salience and consumer choice. *Journal of Political Economy*, 121(5):803–843.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2020). Memory, attention, and choice. *The Quarterly journal of economics*, 135(3):1399–1442.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2022). Salience. *Annual Review of Economics*, 14:521–544.

Brnstrm, R. and Brandberg, Y. (2010). Health risk perception, optimistic bias, and personal satisfaction. *American journal of health behavior*, 34(2):197–205.

Brooks, M. E. and Pui, S. Y. (2010). Are individual differences in numeracy unique from general mental ability? a closer look at a common measure of numeracy. *Individual Differences Research*.

Brown, Z. Y. and Jeon, J. (2021). Product proliferation under rational inattention: Application to health insurance. In *AEA Papers and Proceedings*, volume 111, pages 554–559. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Bucher, T., Collins, C., Rollo, M. E., McCaffrey, T. A., De Vlieger, N., Van der Bend, D., Truby, H., and Perez-Cueto, F. J. (2016). Nudging consumers towards healthier choices: a systematic review of positional influences on food choice. *British Journal of Nutrition*, 115(12):2252–2263.

Cohen, D. R., Henderson, J. B., et al. (1988). *Health, prevention and economics.* Oxford University Press.

Conlisk, J. (1996). Why bounded rationality? *Journal of economic literature*, 34(2):669–700.

Cui, Z., Liu, L., and Shen, Y. (2023). Technical report.

Cutler, D. M. (2022). The costs of long covid. In *JAMA Health Forum*, volume 3, pages e221809–e221809. American Medical Association.

Dai, H., Saccardo, S., Han, M. A., Roh, L., Raja, N., Vangala, S., Modi, H., Pandya, S., Sloyan, M., and Croymans, D. M. (2021). Behavioural nudges increase covid-19 vaccinations. *Nature*, 597(7876):404–409.

De las Cuevas, C., Motuca, M., Baptista, T., and de Leon, J. (2018). Skepticism and pharmacophobia toward medication may negatively impact adherence to psychiatric medications: a comparison among outpatient samples recruited in spain, argentina, and venezuela. *Patient preference and adherence*, pages 301–310.

Dercon, S. and Hoddinott, J. (2004). Health, shocks and poverty persistence. *Insurance against poverty*, pages 123–136.

Dror, A. A., Eisenbach, N., Taiber, S., Morozov, N. G., Mizrachi, M., Zigron, A., Srouji, S., and Sela, E. (2020). Vaccine hesitancy: the next challenge in the fight against covid-19. *European journal of epidemiology*, 35:775–779.

Dubé, E., Laberge, C., Guay, M., Bramadat, P., Roy, R., and Bettinger, J. A. (2013). Vaccine hesitancy: an overview. *Human vaccines & immunotherapeutics*, 9(8):1763–1773.

Epley, N. and Gilovich, T. (2006). The anchoring-and-adjustment heuristic: Why the adjustments are insufficient. *Psychological science*, 17(4):311–318.

Fagerlin, A., Zikmund-Fisher, B. J., Ubel, P. A., Jankovic, A., Derry, H. A., and Smith, D. M. (2007). Measuring numeracy without a math test: development of the subjective numeracy scale. *Medical Decision Making*, 27(5):672–680.

Frederick, S., Novemsky, N., Wang, J., Dhar, R., and Nowlis, S. (2009). Opportunity cost neglect. *Journal of Consumer Research*, 36(4):553–561.

Fredrickson, B. L. and Kahneman, D. (1993). Duration neglect in retrospective evaluations of affective episodes. *Journal of personality and social psychology*, 65(1):45.

Gabaix, X. (2014). A sparsity-based model of bounded rationality. *The Quarterly Journal of Economics*, 129(4):1661–1710.

Gertler, P. and Gruber, J. (2002). Insuring consumption against illness. *American economic review*, 92(1):51–70.

Gielen, A. C. and Sleet, D. (2003). Application of behavior-change theories and methods to injury prevention. *Epidemiologic reviews*, 25(1):65–76.

Hoenink, J. C., Mackenbach, J. D., Waterlander, W., Lakerveld, J., Van Der Laan, N., and Beulens, J. W. (2020). The effects of nudging and pricing on healthy food purchasing behavior in a virtual supermarket setting: a randomized experiment. *International Journal of Behavioral Nutrition and Physical Activity*, 17(1):1–12.

Hoffman, D. D. and Singh, M. (1997). Salience of visual parts. *Cognition*, 63(1):29–78.

Holyoak, K. J. and Morrison, R. G. (2005). *The Cambridge handbook of thinking and reasoning*. Cambridge University Press.

Hsee, C. K. (1996). The evaluability hypothesis: An explanation for preference reversals between joint and separate evaluations of alternatives. *Organizational behavior and human decision processes*, 67(3):247–257.

Hsee, C. K. and Kunreuther, H. C. (2000). The affection effect in insurance decisions. *Journal of Risk and Uncertainty*, 20:141–159.

Hsee, C. K., Loewenstein, G. F., Blount, S., and Bazerman, M. H. (1999). Preference reversals between joint and separate evaluations of options: A review and theoretical analysis. *Psychological bulletin*, 125(5):576.

Hsee, C. K., Yang, Y., Li, N., and Shen, L. (2009). Wealth, warmth, and well-being: Whether happiness is relative or absolute depends on whether it is about money, acquisition, or consumption. *Journal of Marketing Research*, 46(3):396–409.

Hsee, C. K. and Zhang, J. (2010). General evaluability theory. *Perspectives on Psychological Science*, 5(4):343–355.

Itti, L. (2007). Visual salience. *Scholarpedia*, 2(9):3327.

Janz, N. K. and Becker, M. H. (1984). The health belief model: A decade later. *Health education quarterly*, 11(1):1–47.

Johannesson, M. and Jönsson, B. (1991). Economic evaluation in health care: is there a role for cost-benefit analysis? *Health policy*, 17(1):1–23.

Johnson, E. J., Häubl, G., and Keinan, A. (2007). Aspects of endowment: a query theory of value construction. *Journal of experimental psychology: Learning, memory, and cognition*, 33(3):461.

Johnson, E. J., Hershey, J., Meszaros, J., and Kunreuther, H. (1993). Framing, probability distortions, and insurance decisions. *Journal of risk and uncertainty*, 7:35–51.

Just, D. R. (2017). The behavioral welfare paradox: practical, ethical and welfare implications of nudging. *Agricultural and Resource Economics Review*, 46(1):1–20.

Kahneman, D. and Tversky, A. (1984). Choices, values, and frames. *American psychologist*, 39(4):341.

Kan, K. (2007). Cigarette smoking and self-control. *Journal of health economics*, 26(1):61–81.

Kaufmann, C., Müller, T., Hefti, A., and Boes, S. (2018). Does personalized information improve health plan choices when individuals are distracted? *Journal of economic behavior & organization*, 149:197–214.

Kaur, S., Mullainathan, S., Oh, S., and Schilbach, F. (2021). Do financial concerns make workers less productive? Technical report, National Bureau of Economic Research.

Kelly, C. A. and Sharot, T. (2021). Individual differences in information-seeking. *Nature communications*, 12(1):7062.

Kenkel, D. S. (1994). The demand for preventive medical care. *Applied economics*, 26(4):313–325.

Kimball, A., Torrone, E., Miele, K., Bachmann, L., Thorpe, P., Weinstock, H., and Bowen, V. (2020). Missed opportunities for prevention of congenital syphilis—united states, 2018. *Morbidity and Mortality Weekly Report*, 69(22):661.

Krishna, A. and Thompson, T. L. (2021). Misinformation about health: a review of health communication and misinformation scholarship. *American behavioral scientist*, 65(2):316–332.

Kroese, F. M., Marchiori, D. R., and De Ridder, D. T. (2016). Nudging healthy food choices: a field experiment at the train station. *Journal of Public Health*, 38(2):e133–e137.

Kruger, J. and Evans, M. (2004). If you don't want to be late, enumerate: Unpacking reduces the planning fallacy. *Journal of Experimental Social Psychology*, 40(5):586–598.

Kunreuther, H., Pauly, M., et al. (2006). Insurance decision-making and market behavior. *Foundations and Trends® in Microeconomics*, 1(2):63–127.

Linnemayr, S. (2015). Hiv prevention through the lens of behavioral economics (draft). *Journal of acquired immune deficiency syndromes (1999)*, 68(4):e61.

Loewenstein, G., Friedman, J. Y., McGill, B., Ahmad, S., Linck, S., Sinkula, S., Beshears, J., Choi, J. J., Kolstad, J., Laibson, D., et al. (2013). Consumers' misunderstanding of health insurance. *Journal of Health Economics*, 32(5):850–862.

MacDonald, N. E. et al. (2015). Vaccine hesitancy: Definition, scope and determinants. *Vaccine*, 33(34):4161–4164.

Mas-Colell, A., Whinston, M. D., Green, J. R., et al. (1995). *Microeconomic theory*, volume 1. Oxford university press New York.

McGraw, A. P., Shafir, E., and Todorov, A. (2010). Valuing money and things: Why a 20itemcanbeworthmoreandlessthan 20. *Management Science*, 56(5):816–830.

Meltzer, D. (2001). Addressing uncertainty in medical cost–effectiveness analysis: implications of expected utility maximization for methods to perform sensitivity analysis and the use of cost–effectiveness analysis to set priorities for medical research. *Journal of health economics*, 20(1):109–129.

Michaelidou, N. and Dibb, S. (2008). Consumer involvement: a new perspective. *The Marketing Review*, 8(1):83–99.

Milkman, K. L., Patel, M. S., Gandhi, L., Graci, H. N., Gromet, D. M., Ho, H., Kay, J. S., Lee, T. W., Akinola, M., Beshears, J., et al. (2021). A megastudy of text-based nudges encouraging patients to get vaccinated at an upcoming doctor's appointment. *Proceedings of the National Academy of Sciences*, 118(20):e2101165118.

Morewedge, C. K., Kassam, K. S., Hsee, C. K., and Caruso, E. M. (2009). Duration sensitivity depends on stimulus familiarity. *Journal of Experimental Psychology: General*, 138(2):177.

Mullainathan, S. and Shafir, E. (2013). *Scarcity: Why having too little means so much*. Macmillan.

Muncy, J. A. and Hunt, S. D. (1984). Consumer involvement: definitional issues and research directions. *ACR North American Advances*.

Nielsen, L. and Phillips, J. W. (2008). Health economic choices in old age: Interdisciplinary perspectives on economic decisions and the aging mind. In *Neuroeconomics*, pages 227–270. Emerald Group Publishing Limited.

Norman, P., Boer, H., Seydel, E. R., and Mullan, B. (2015). Protection motivation theory. *Predicting and changing health behaviour: Research and practice with social cognition models*, 3:70–106.

O'Donoghue, T. and Rabin, M. (2006). Optimal sin taxes. *Journal of Public Economics*, 90(10-11):1825–1849.

Olsen, L., Saunders, R. S., Yong, P. L., et al. (2010). The healthcare imperative: lowering costs and improving outcomes: workshop series summary.

Pachur, T., Hertwig, R., and Wolkewitz, R. (2014). The affect gap in risky choice: affect-rich outcomes attenuate attention to probability information. *Decision*, 1(1):64.

Prelec, D. (1998). The probability weighting function. *Econometrica*, pages 497–527.

Reñosa, M. D. C., Landicho, J., Wachinger, J., Dalglish, S. L., Bärnighausen, K., Bärnighausen, T., and McMahon, S. A. (2021). Nudging toward vaccination: a systematic review. *BMJ global health*, 6(9):e006237.

Rief, W. (2021). Fear of adverse effects and covid-19 vaccine hesitancy: recommendations of the treatment expectation expert group. In *JAMA Health Forum*, volume 2, pages e210804–e210804. American Medical Association.

Rogers, R. W. and Prentice-Dunn, S. (1997). Protection motivation theory.

Rottenstreich, Y. and Hsee, C. K. (2001). Money, kisses, and electric shocks: On the affective psychology of risk. *Psychological science*, 12(3):185–190.

Savitsky, K., Van Boven, L., Epley, N., and Wight, W. M. (2005). The unpacking effect in allocations of responsibility for group tasks. *Journal of Experimental Social Psychology*, 41(5):447–457.

Shiraly, R., Khoshdel, N., Jeihooni, A. K., and McLaws, M.-L. (2022). Nudging physical distancing behaviors during the pandemic: a field experiment on passengers in the subway stations of shiraz, iran. *BMC Public Health*, 22(1):1–7.

Sims, C. A. (2003). Implications of rational inattention. *Journal of monetary Economics*, 50(3):665–690.

Sims, T., Tsai, J. L., Jiang, D., Wang, Y., Fung, H. H., and Zhang, X. (2015). Wanting to maximize the positive and minimize the negative: Implications for mixed affective experience in american and chinese contexts. *Journal of Personality and Social Psychology*, 109(2):292.

Soofi, M., Najafi, F., and Karami-Matin, B. (2020). Using insights from behavioral economics to mitigate the spread of covid-19. *Applied health economics and health policy*, 18:345–350.

Southwell, B. G., Niederdeppe, J., Cappella, J. N., Gaysynsky, A., Kelley, D. E., Oh, A., Peterson, E. B., and Chou, W.-Y. S. (2019). Misinformation as a misunderstood challenge to public health. *American journal of preventive medicine*, 57(2):282–285.

Spiller, S. A. (2011). Opportunity cost consideration. *Journal of Consumer Research*, 38(4):595–610.

Sunstein, C. R. (2014). *Why nudge?: The politics of libertarian paternalism*. Yale University Press.

Sunstein, C. R. (2017a). Misconceptions about nudges. *Available at SSRN 3033101*.

Sunstein, C. R. (2017b). Nudges that fail. *Behavioural public policy*, 1(1):4–25.

Sunstein, C. R. (2018). On preferring a to b, while also preferring b to a. *Rationality and Society*, 30(3):305–331.

Suter, R. S., Pachur, T., and Hertwig, R. (2016). How affect shapes risky choice: Distorted probability weighting versus probability neglect. *Journal of Behavioral Decision Making*, 29(4):437–449.

Sydnor, J. (2010). (over) insuring modest risks. *American Economic Journal: Applied Economics*, 2(4):177–199.

Thaler, R. H. and Benartzi, S. (2004). Save more tomorrow$^{\text{TM}}$: Using behavioral economics to increase employee saving. *Journal of political Economy*, 112(S1):S164–S187.

Thaler, R. H. and Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.

Thunström, L. (2019). Welfare effects of nudges: The emotional tax of calorie menu labeling. *Judgment and Decision making*, 14(1):11–25.

Thunström, L., Gilbert, B., and Ritten, C. J. (2018). Nudges that hurt those already hurting–distributional and unintended effects of salience nudges. *Journal of Economic Behavior & Organization*, 153:267–282.

Trippas, D., Handley, S. J., and Verde, M. F. (2014). Fluency and belief bias in deductive reasoning: new indices for old effects. *Frontiers in Psychology*, 5:631.

Tversky, A. and Koehler, D. J. (1994). Support theory: A nonextensional representation of subjective probability. *Psychological review*, 101(4):547.

Vallgårda, S. (2012). Nudge—a new and better way to improve health? *Health policy*, 104(2):200–203.

Van Boven, L. and Epley, N. (2003). The unpacking effect in evaluative judgments: When the whole is less than the sum of its parts. *Journal of experimental social psychology*, 39(3):263–269.

Wang, Y. and Sloan, F. A. (2018). Present bias and health. *Journal of risk and uncertainty*, 57(2):177–198.

Wason, P. C. (1968). Reasoning about a rule. *Quarterly journal of experimental psychology*, 20(3):273–281.

Waters, E. A., Weinstein, N. D., Colditz, G. A., and Emmons, K. M. (2007a). Aversion to side effects in preventive medical treatment decisions. *British journal of health psychology*, 12(3):383–401.

Waters, E. A., Weinstein, N. D., Colditz, G. A., and Emmons, K. M. (2007b). Reducing aversion to side effects in preventive medical treatment decisions. *Journal of Experimental Psychology: Applied*, 13(1):11.

Weijers, R. J. and de Koning, B. B. (2021). Nudging to increase hand hygiene during the covid-19 pandemic: A field experiment. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 53(3):353.

Wettstein, C., Mugglin, C., Egger, M., Blaser, N., SALAZAR, L., Estill, J., Bender, N., Davies, M.-A., Wandeler, G., and Keiser, O. (2012). Missed opportunities to prevent mother-to-child-transmission in sub-saharan africa: systematic review and meta-analysis. *AIDS (London, England)*, 26(18):2361.

Wilson, A. L., Buckley, E., Buckley, J. D., and Bogomolova, S. (2016). Nudging healthier food and beverage choices through salience and priming. evidence from a systematic review. *Food Quality and Preference*, 51:47–64.

Wolff, K., Larsen, S., and Øgaard, T. (2019). How to define and measure risk perceptions. *Annals of Tourism Research*, 79:102759.

Wright, B. J., Garcia-Alexander, G., Weller, M. A., and Baicker, K. (2017). Low-cost behavioral nudges increase medicaid take-up among eligible residents of oregon. *Health Affairs*, 36(5):838–845.

Wu, G. and Gonzalez, R. (1996). Curvature of the probability weighting function. *Management science*, 42(12):1676–1690.

Zikmund-Fisher, B. J. (2019). Helping people know whether measurements have good or bad implications: increasing the evaluability of health and science data communications. *Policy Insights from the Behavioral and Brain Sciences*, 6(1):29–37.

Zikmund-Fisher, B. J., Fagerlin, A., and Ubel, P. A. (2004). "is 28% good or bad?" evaluability and preference reversals in health care decisions. *Medical Decision Making*, 24(2):142–148.

Zikmund-Fisher, B. J., Fagerlin, A., and Ubel, P. A. (2010). Risky feelings: why a 6% risk of cancer does not always feel like 6%. *Patient education and counseling*, 81:S87–S93.

Zikmund-Fisher, B. J., Smith, D. M., Ubel, P. A., and Fagerlin, A. (2007). Validation of the subjective numeracy scale: effects of low numeracy on comprehension of risk communications and utility elicitations. *Medical Decision Making*, 27(5):663–671.

Online Appendix. Placeholder

We arrange the appendices sequentially by the order they were first referenced:
(1) Appendix A showcases the important frames of the graphical interfaces for all six experiments.
(2) Appendix B includes the additional results, robustness checks, and additional results interpretation for all six experiments.
(3) Appendix C demonstrates a suggestive online interface design for using AUM in real-world websites.

A Appendix A

A.1 Appendix A1: Instructions for Experiment 1

Figure A.1: Disease Description

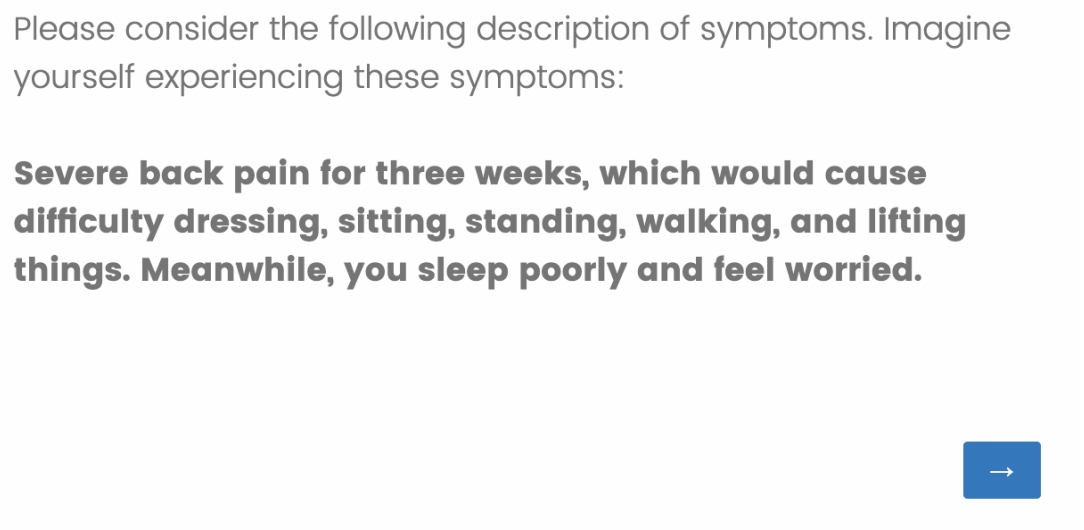


Figure A.2: Initial Open-ended Report

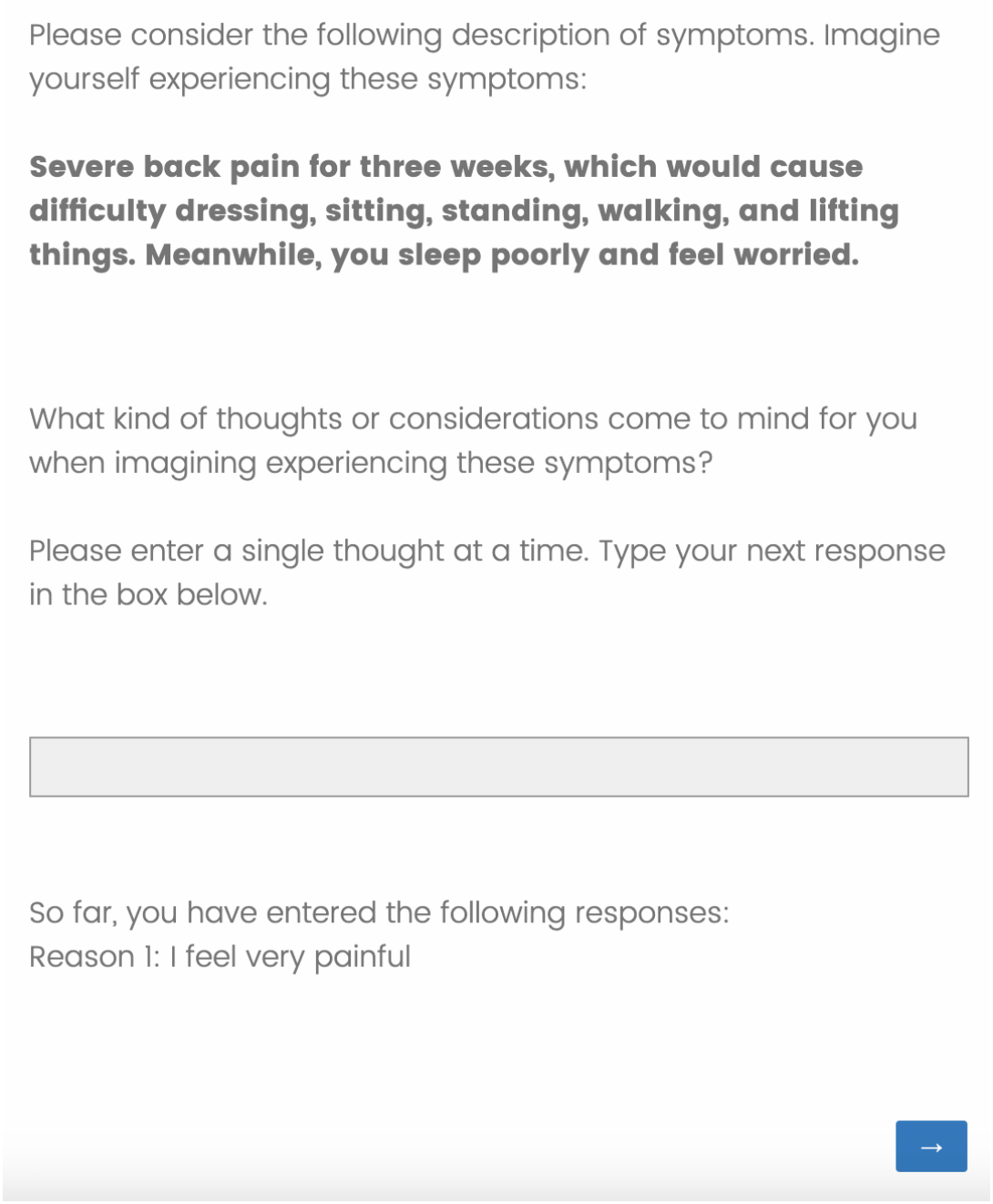


Figure A.3: Open-ended Report after 3 Entries

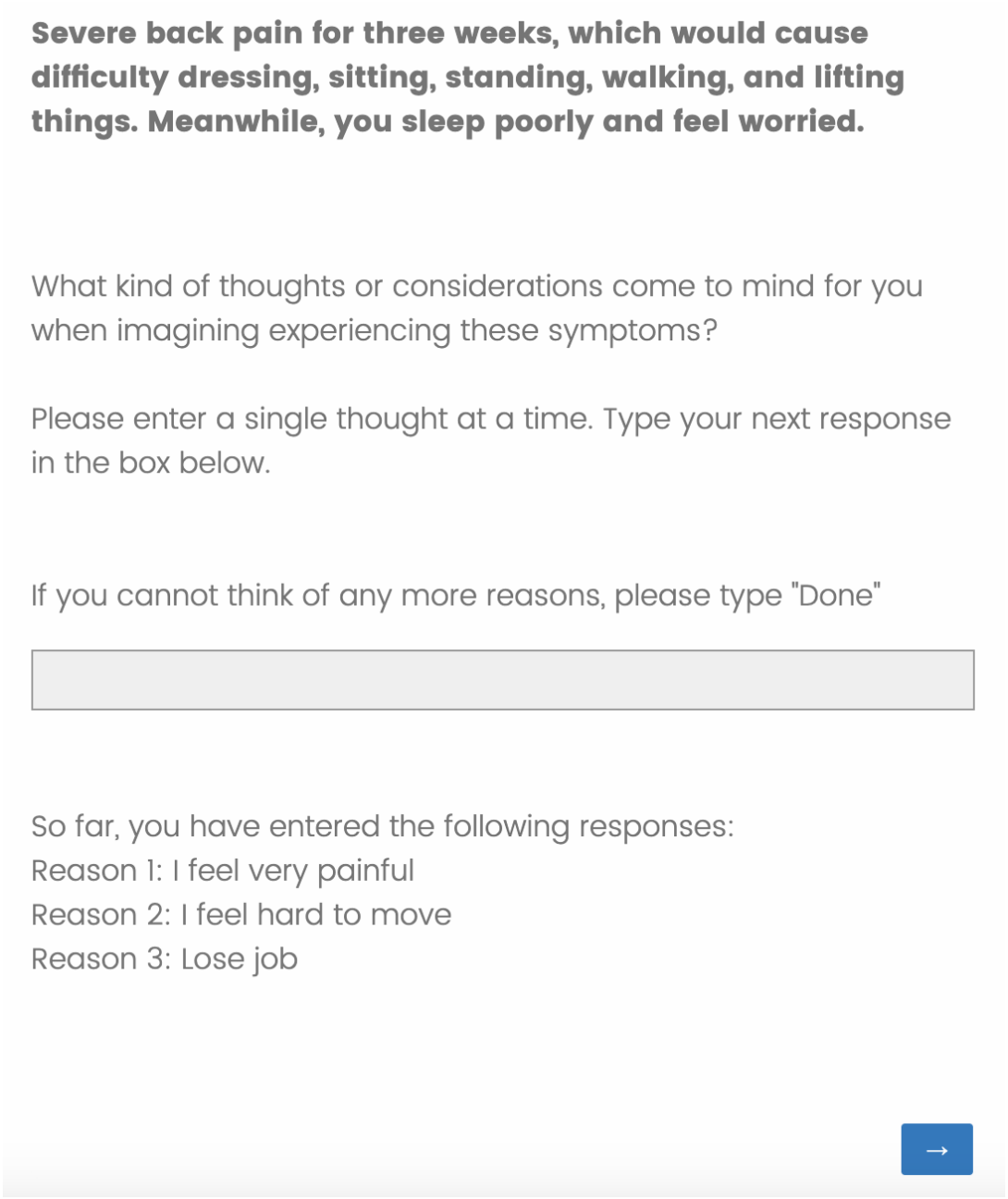


Figure A.4: Self-categorization on Open-ended Descriptions

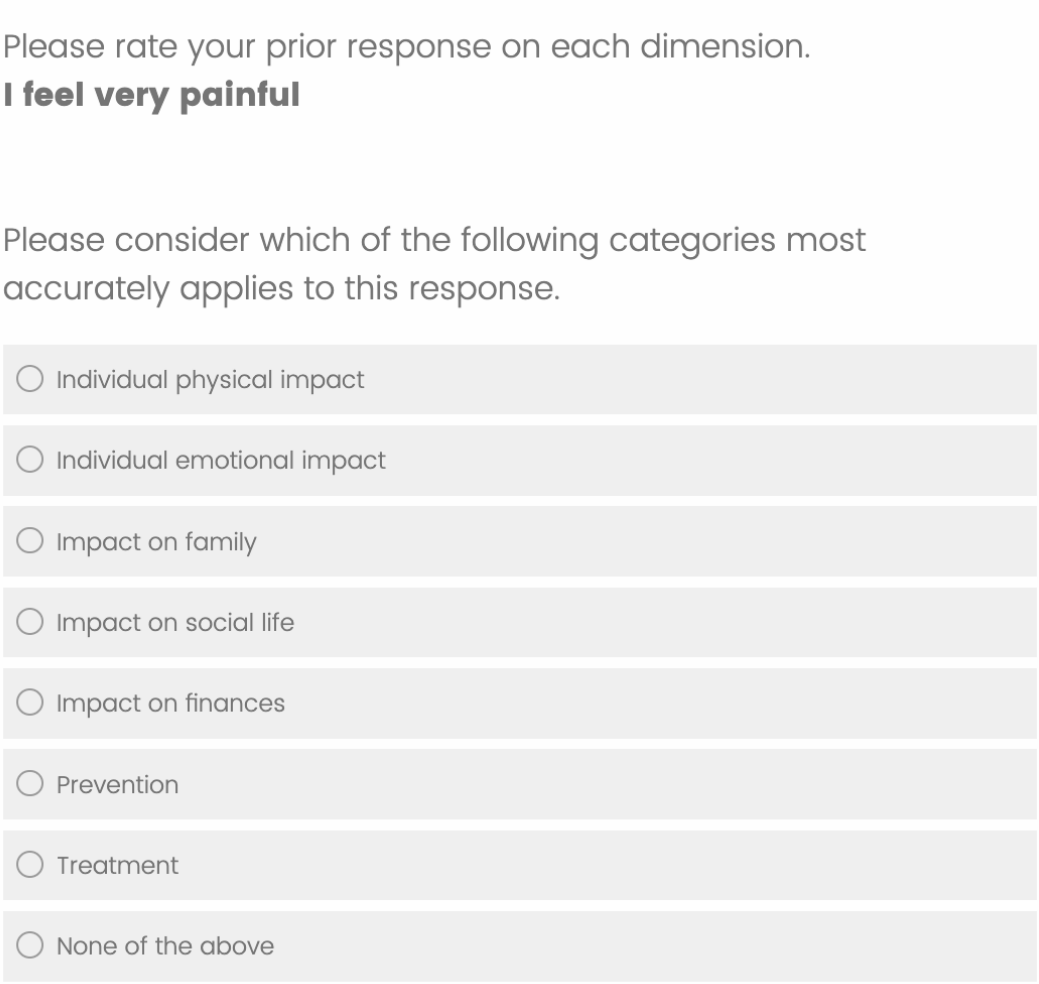
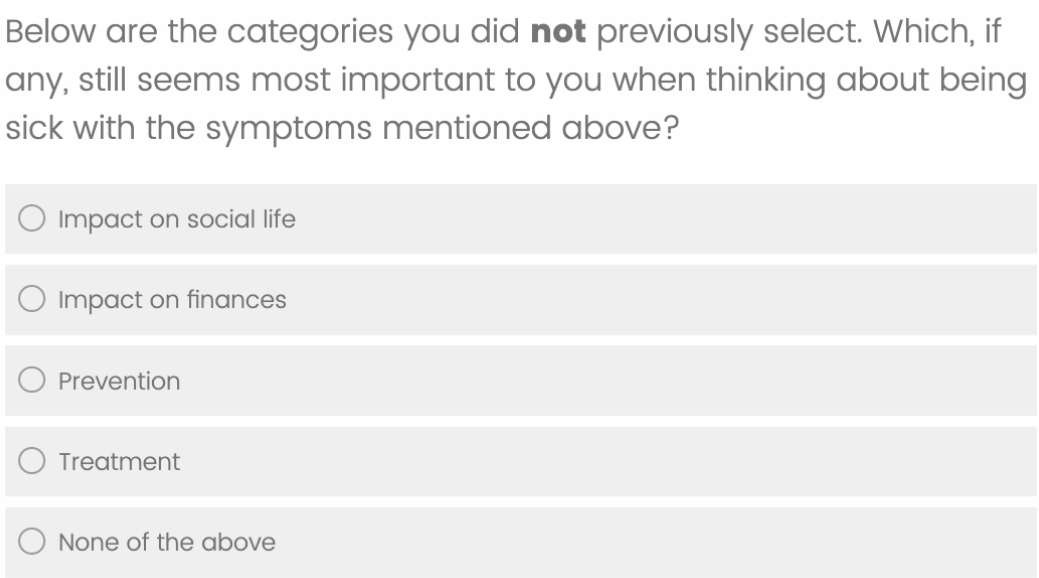


Figure A.5: Last Call Question on Non-selected Categories



A.2 Appendix A2: Instructions for Experiment 2

Figure A.6: General Instructions for Self-Reports (Similar for Experiments 2-5

The following questions will present you with a set of symptoms associated with a health condition. For all questions, the instructions are the same. After reading the health description, you will make the following two ratings:

Overall life impact: Read each health description and rate the severity of its life impact between **0** and **100.**

- **0** indicates that this health condition has the **smallest, least severe** impact on quality of life.
- **100** indicates that this health condition has the **largest, most severe** impact on quality of life.

Money you would pay to avoid this health outcome: Supposing this health outcome would happen to **YOU**, how much money would you pay to avoid it altogether? For example, imagine you are paying for medical prevention that would guarantee you do not experience this negative health outcome without causing undesirable side effects or changing your lifestyle. Please consider the **MAXIMUM** dollar amount you are willing to pay to avoid this health outcome. For instance, if you would pay $500 to avoid getting a cold, you should type in "500" in the box.

Figure A.7: Description and Task for the Control Group

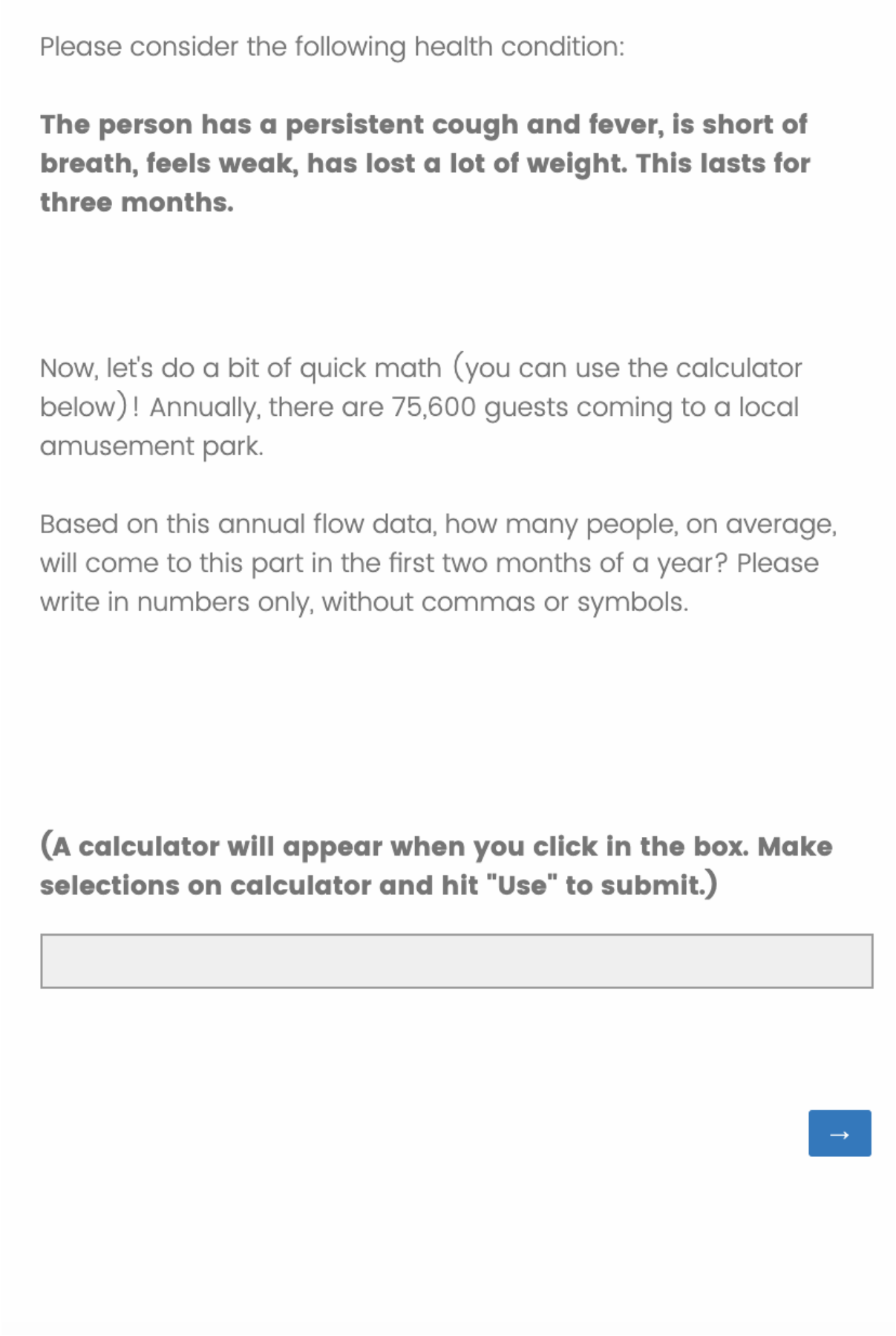


Figure A.8: Description and Task for the INFO Group

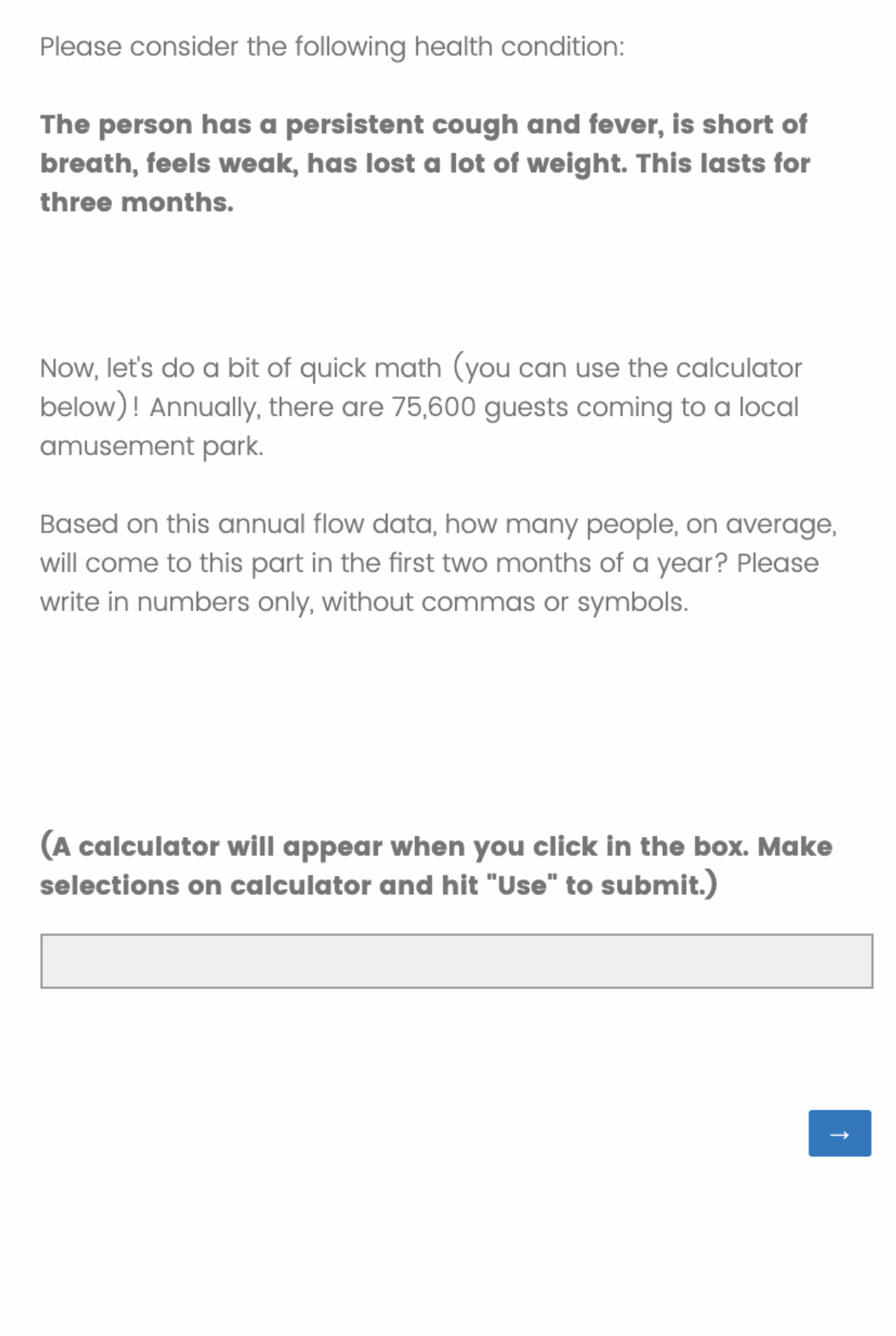


Figure A.9: Description and Task for the AUM Group

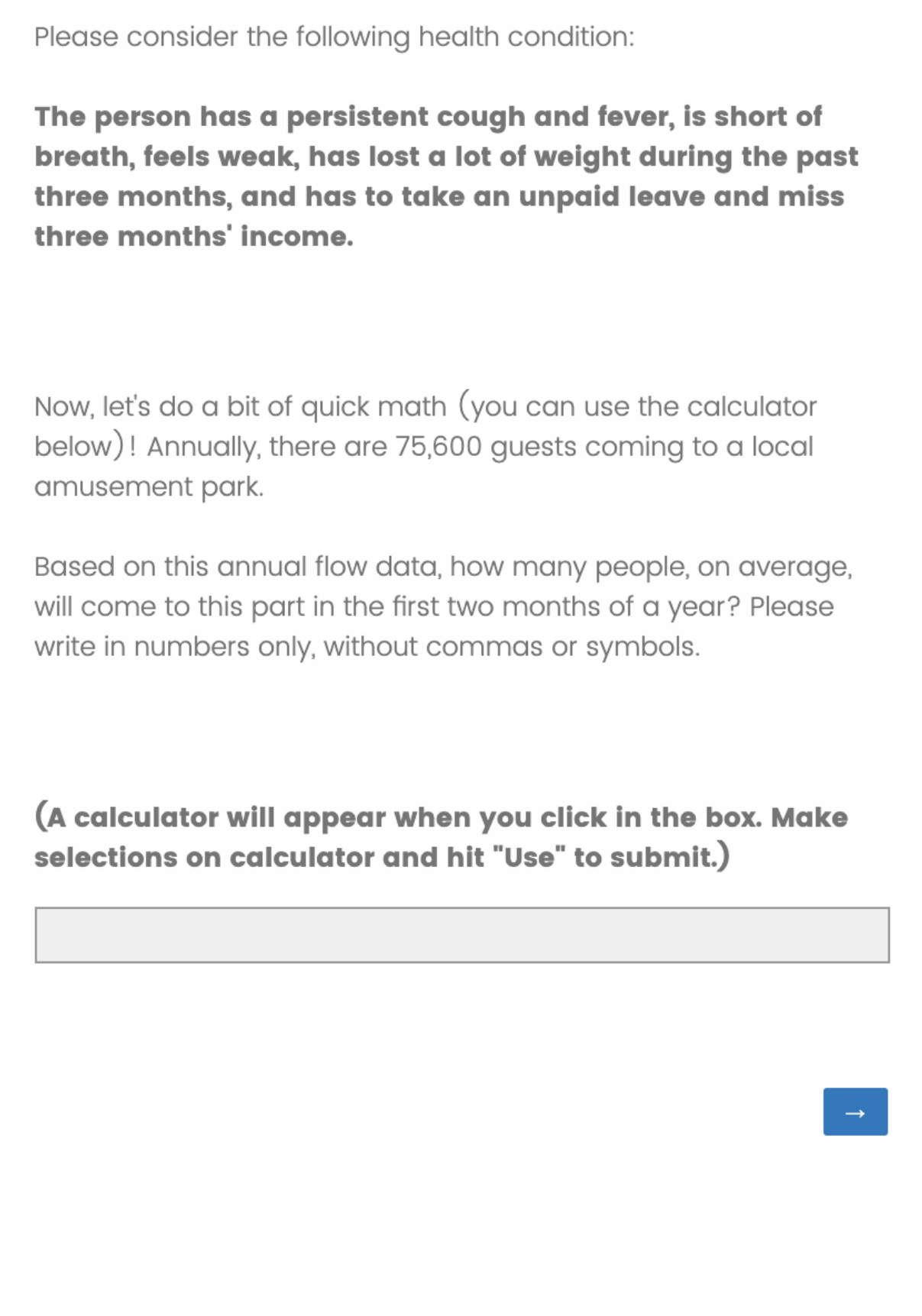
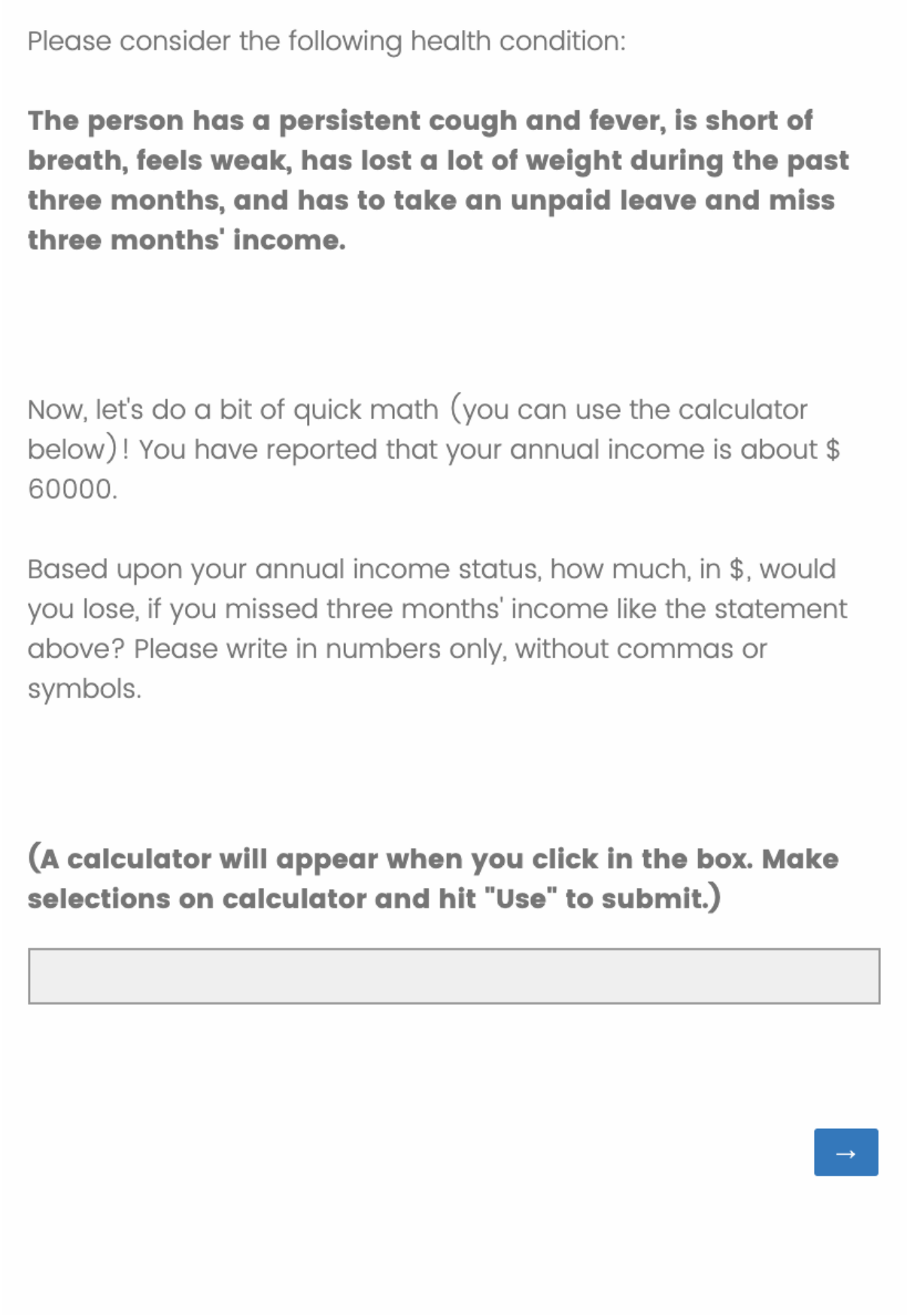


Figure A.10: Perceived Severity and WTP Elicitation



A.3 Appendix A3: Instructions for Experiment 3

A.4 Appendix A4: Instructions for Experiment 4

A.5 Appendix A5: Instructions for Experiment 5

A.6 Appendix A6: Instructions for Experiment 6

B Appendix B

C Appendix C

Active Unpacking of the Financial Consequences of Health Outcomes Improves Preventive Decisions

Zhihan Cui[†] Lu Liu[‡]

Current Version: September 2023

Abstract

Traditional health economic models assume that decision-makers (DMs) incorporate all dimensions of information regarding potential health outcomes when making preventive decisions. However, behavioral sciences suggest that individuals might deviate from this assumption in two key aspects. The first is opportunity cost neglect due to limited attention, whereby DMs may overlook some indirect or non-salient dimensions of loss from illness. The second is evaluability issues, whereby DMs may fail to comprehend the actual relevance of numerical health information to their circumstances. To increase people's attention to financial losses and make health outcomes more evaluable, we designed an intervention entitled "Active Unpacking with Money" (AUM), which directs DMs to *actively calculate the monetary losses from experiencing a negative health condition over a specified time period.* Through a series of six consecutive online experiments, we demonstrate that: (1) AUM amplifies people's perceived severity of health risks and their willingness-to-pay (WTP) for a guaranteed preventive measure; (2) AUM heightens people's sensitivity to the length of a disease; and (3) AUM bolsters sensitivity to probabilistic information. We clarify how AUM assists in mitigating both salience and evaluability challenges within these contexts, and why AUM is likely welfare-improving. Finally, we discuss its practical implications in health communication.

JEL Codes: C91, D14, I10, I12

*We are extraordinarily grateful for the guidance, support and encouragement of my advisors Geoffrey Heal, Jeffrey Shrader, Daniel Benjamin and Kathleen Nganouge without whom we can by no means complete the research. This project also gets important help and suggestions from the conversations with Ilana Brody, Eugene Caruso, Hengchen Dai, Mark Dean, David Dolifka, Evan Friedman, Craig Fox, Hal Hirshleifer, Xilin Li, RC Lim, Adriana Lleras-Muney, Jackson Lu, Yuhang Pan, Xuezheng Qin, Rui Sun, Yiwen Shen, Fanqi Shi, Junze Sun, Xu Tan, Megan Weber, Jueyu Wu, Mu Zhang, and Castiel Chen Zhuang.

[†]UCLA Anderson School of Management and SIPA, Columbia University; zhihan.cui@anderson.ucla.edu

[‡]New York University Shanghai; ll4748@nyu.edu

Online Appendix

We arrange the appendices sequentially by the order they were first referenced:

(1) Appendix A showcases the important frames of the graphical interfaces for all six experiments.

(2) Appendix B includes the additional results, robustness checks, and additional results interpretation for all six experiments.

(3) Appendix C demonstrates a suggestive online interface design for using AUM in real-world websites.

A Appendix A

A.1 Appendix A1: Instructions for Experiment 1

Figure A.1.1: Disease Description

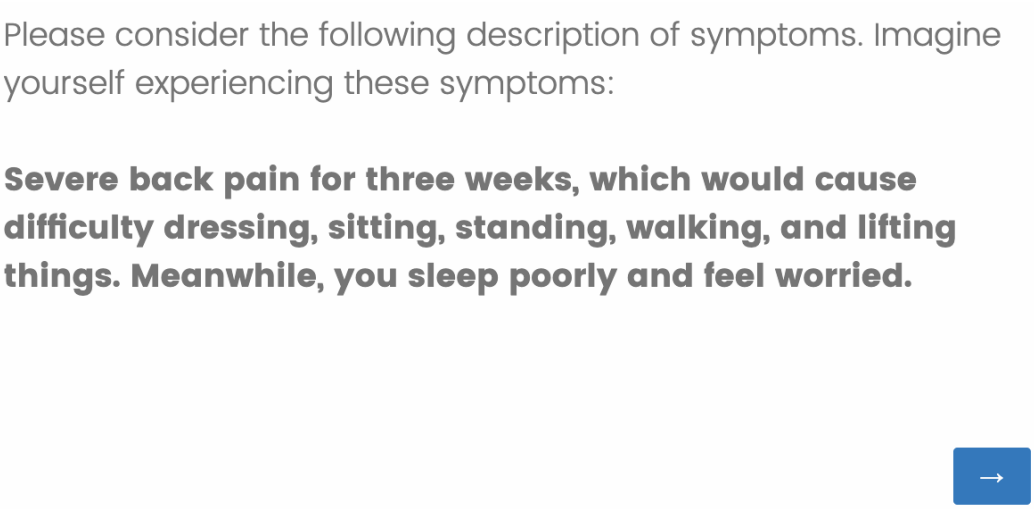


Figure A.1.2: Initial Open-ended Report

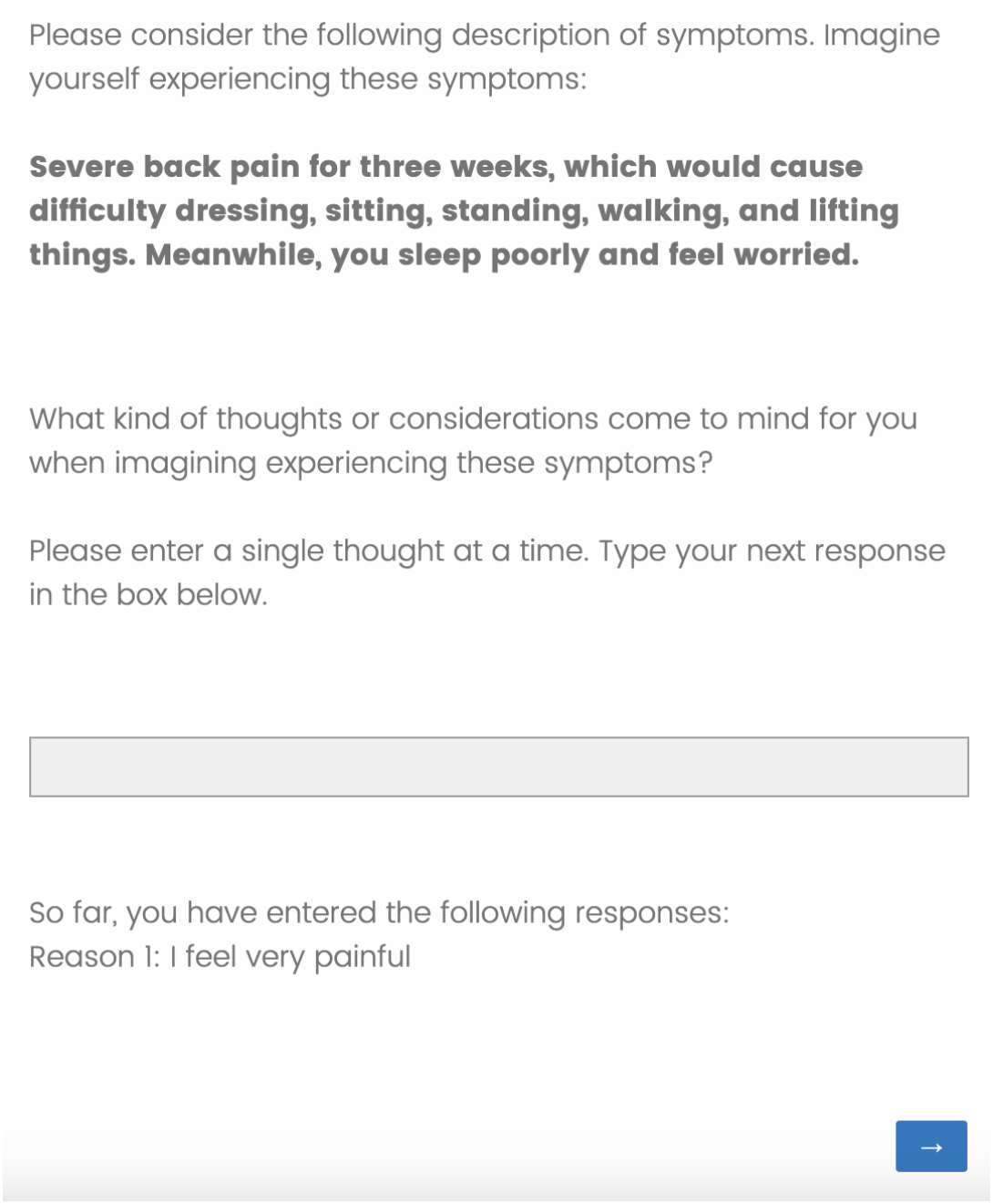


Figure A.1.3: Open-ended Report after 3 Entries

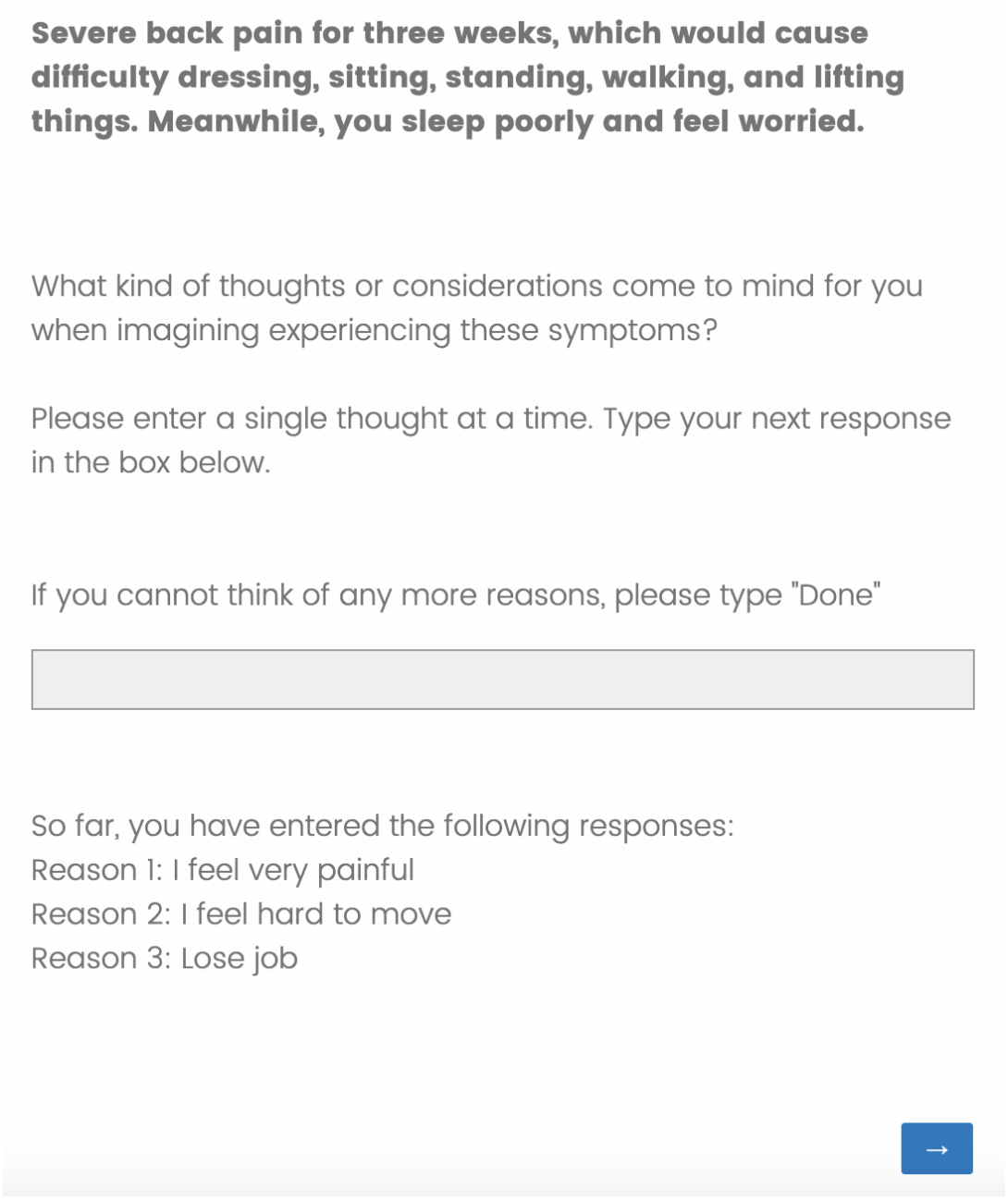


Figure A.1.4: Self-categorization on Open-ended Descriptions

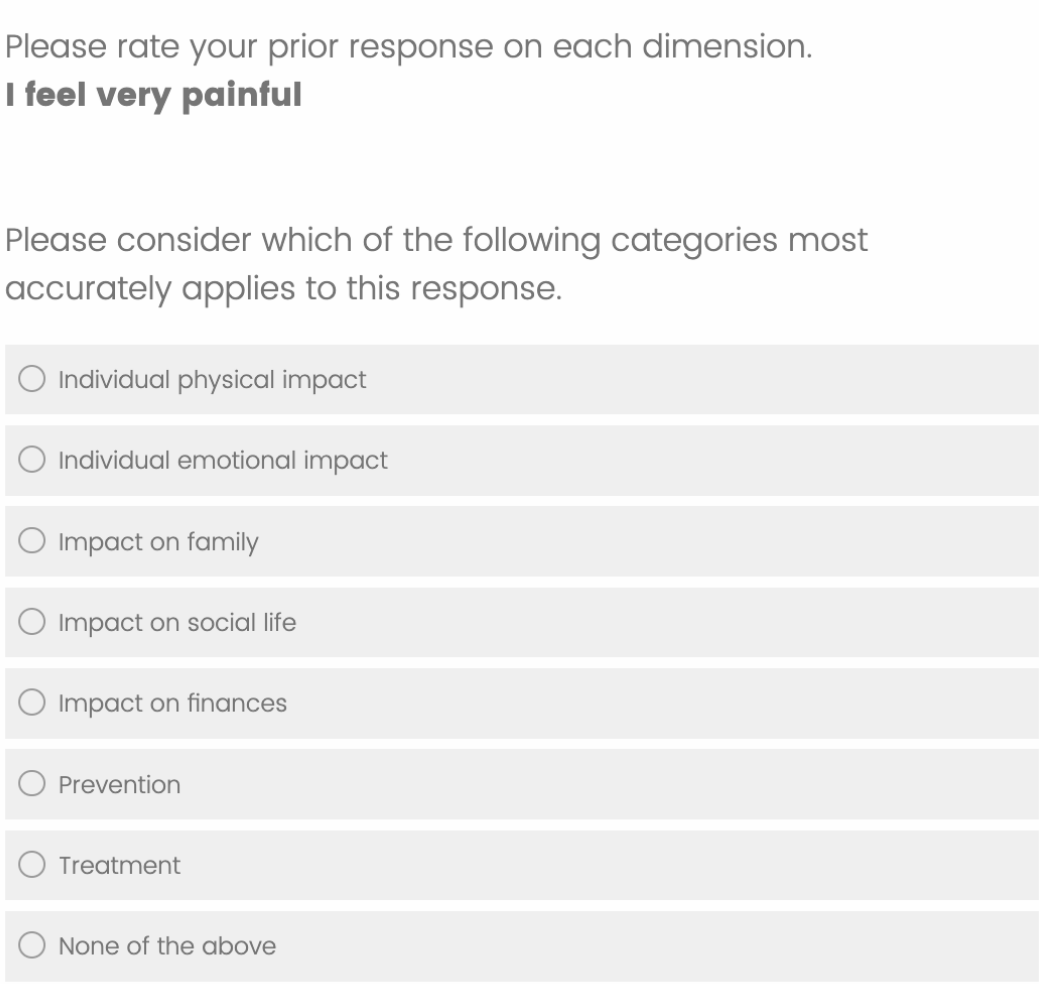


Figure A.1.5: Last Call Question on Non-selected Categories

A.2 Appendix A2: Instructions for Experiment 2

Figure A.2.1: General Instructions for Self-Reports (Similar for Experiments 2-5

The following questions will present you with a set of symptoms associated with a health condition. For all questions, the instructions are the same. After reading the health description, you will make the following two ratings:

Overall life impact: Read each health description and rate the severity of its life impact between **0** and **100.**

- **0** indicates that this health condition has the **smallest, least severe** impact on quality of life.
- **100** indicates that this health condition has the **largest, most severe** impact on quality of life.

Money you would pay to avoid this health outcome: Supposing this health outcome would happen to **YOU**, how much money would you pay to avoid it altogether? For example, imagine you are paying for medical prevention that would guarantee you do not experience this negative health outcome without causing undesirable side effects or changing your lifestyle. Please consider the **MAXIMUM** dollar amount you are willing to pay to avoid this health outcome. For instance, if you would pay $500 to avoid getting a cold, you should type in "500" in the box.

Figure A.2.2: Description and Task for the Control Group

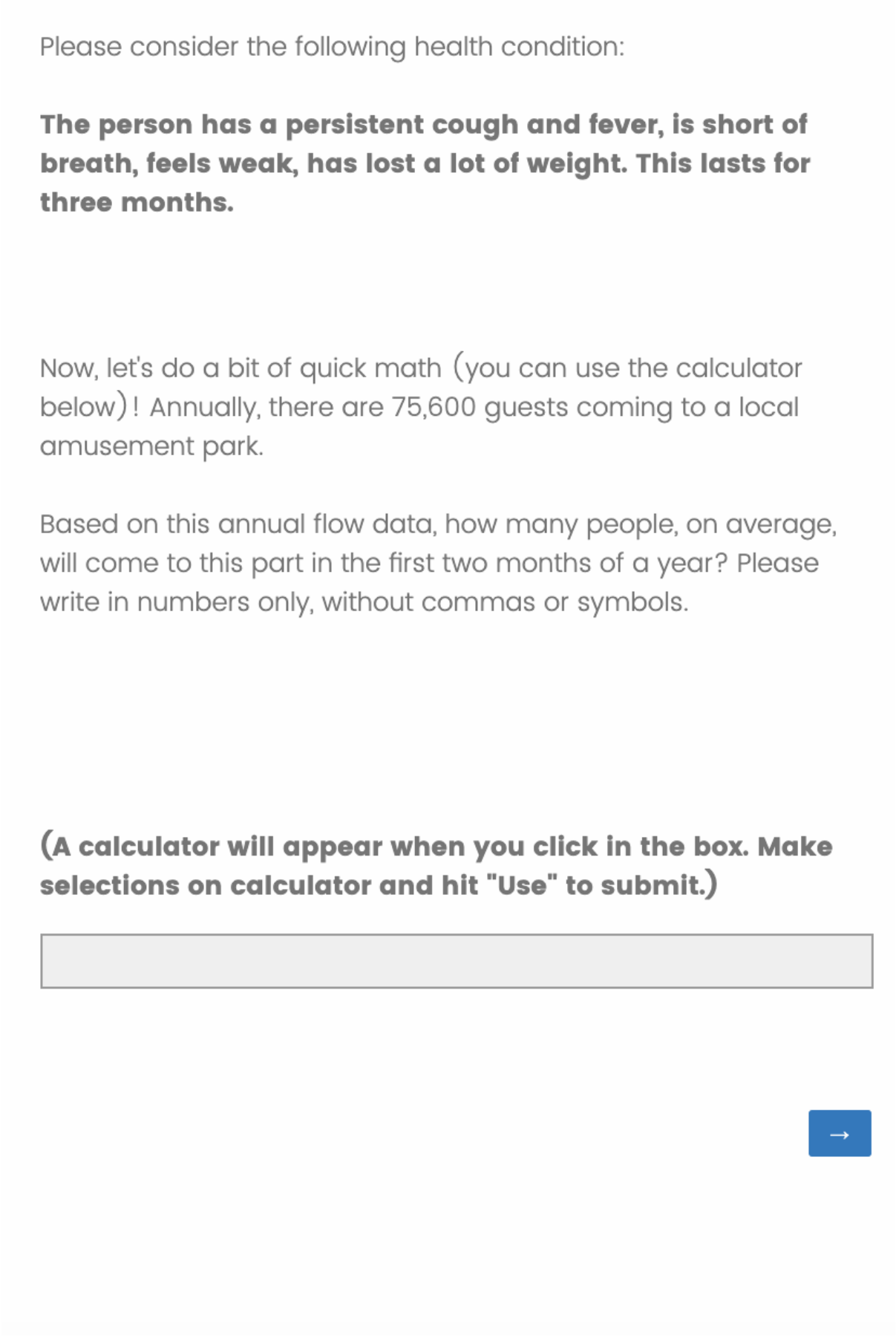


Figure A.2.3: Description and Task for the INFO Group

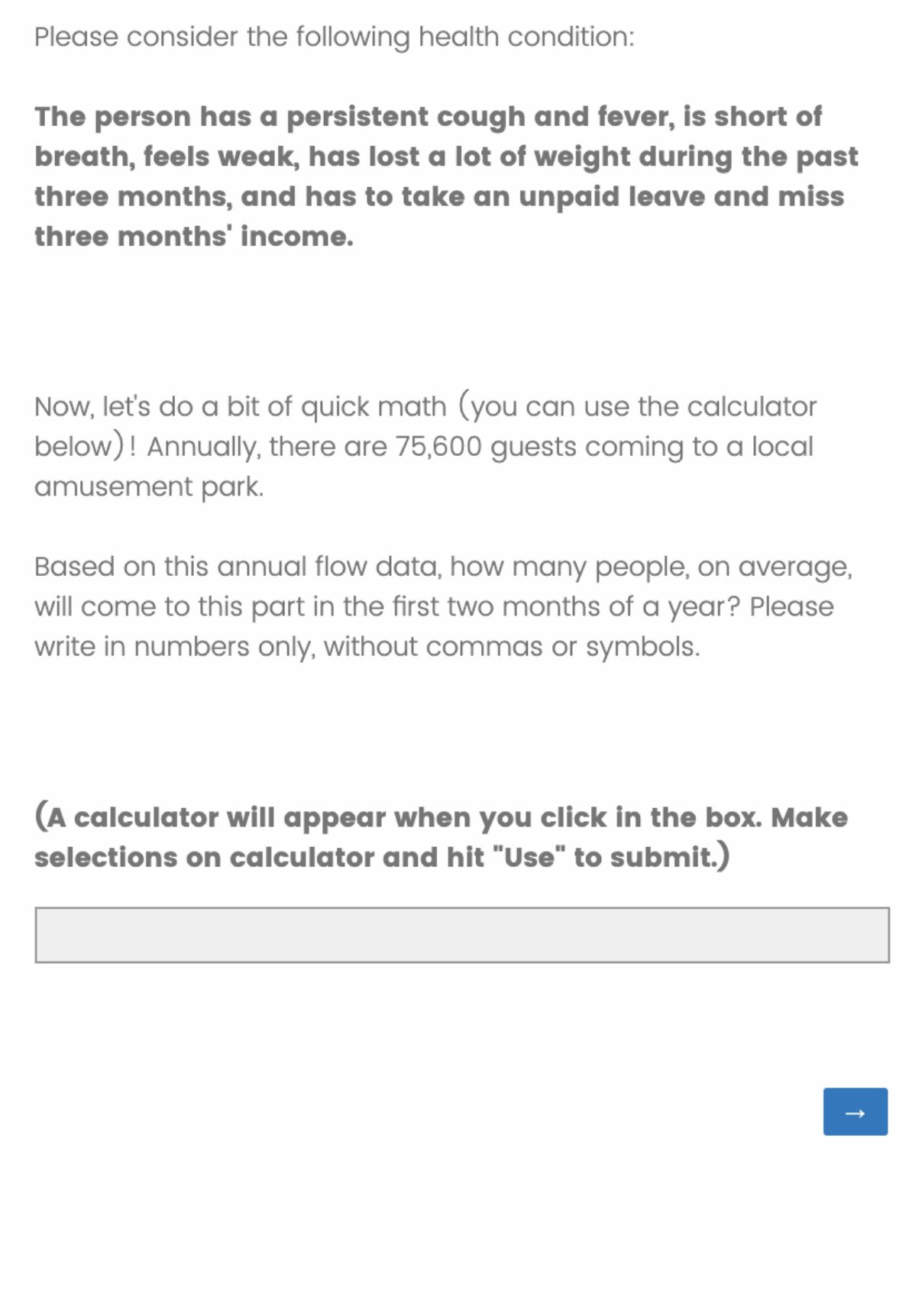


Figure A.2.4: Description and Task for the AUM Group

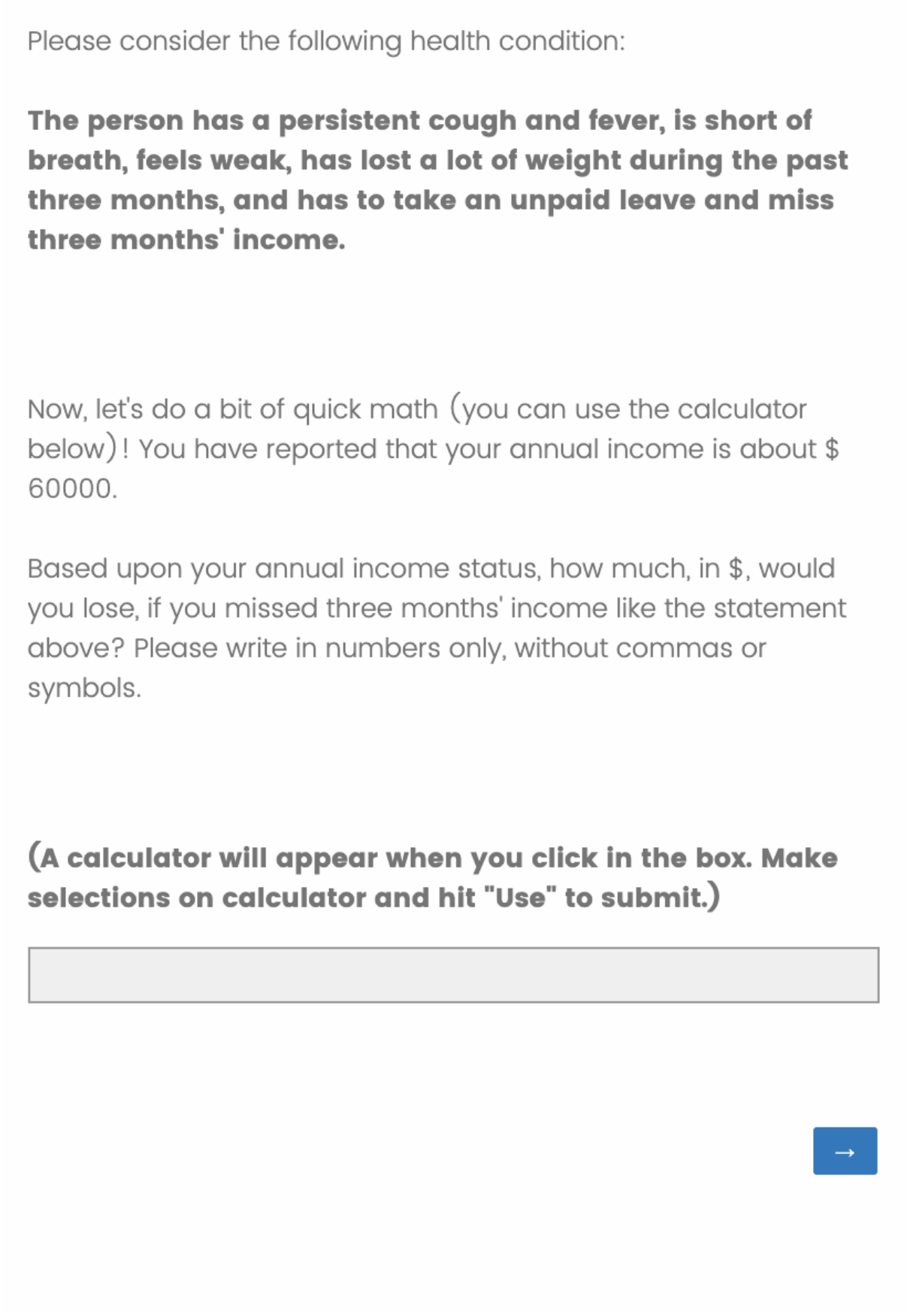
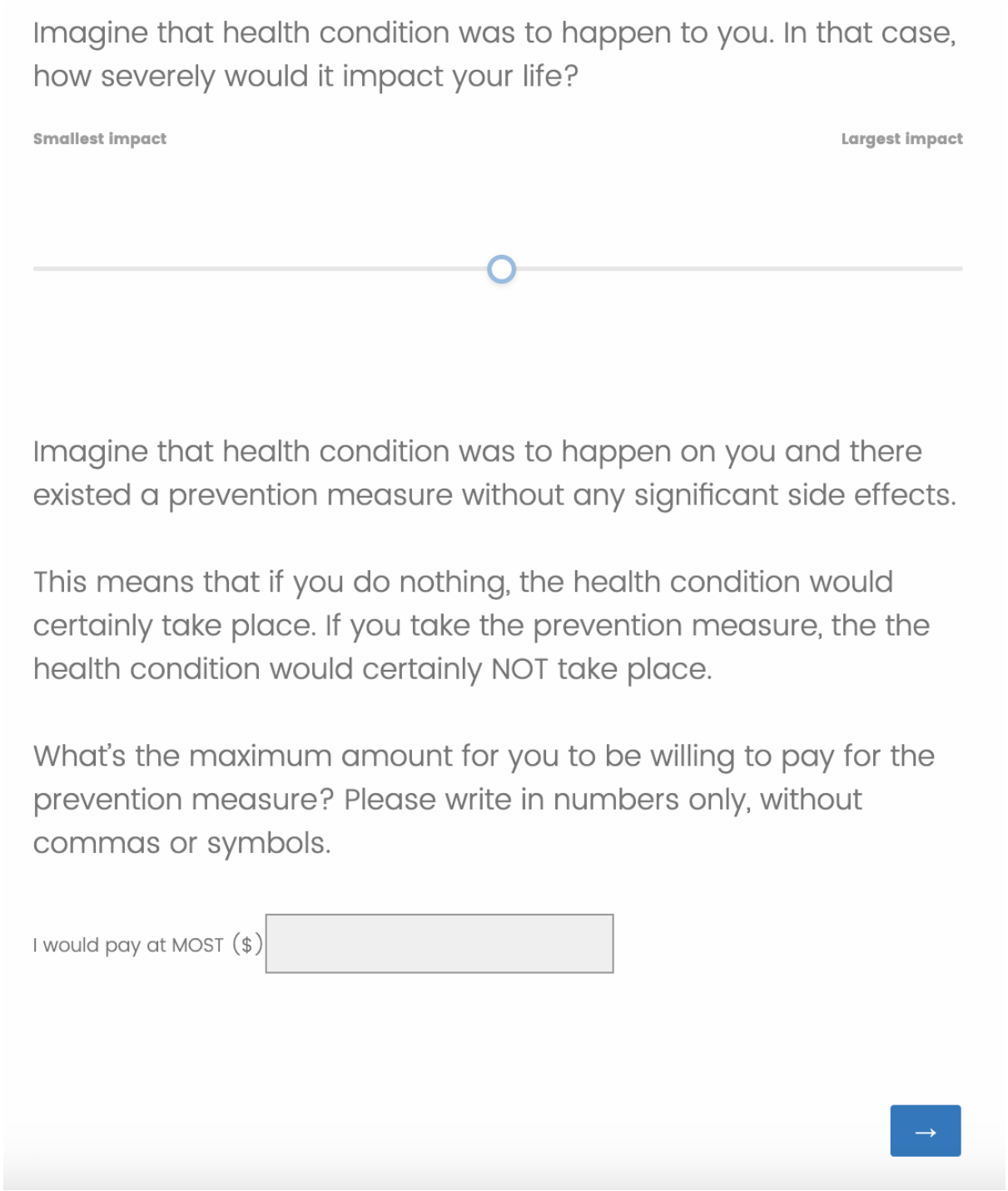


Figure A.2.5: Perceived Severity and WTP Elicitation (same for Exp. 2 and 3)



A.3 Appendix A3: Instructions for Experiment 3

Figure A.3.1: Disease Description for Experiment 3: Control Group

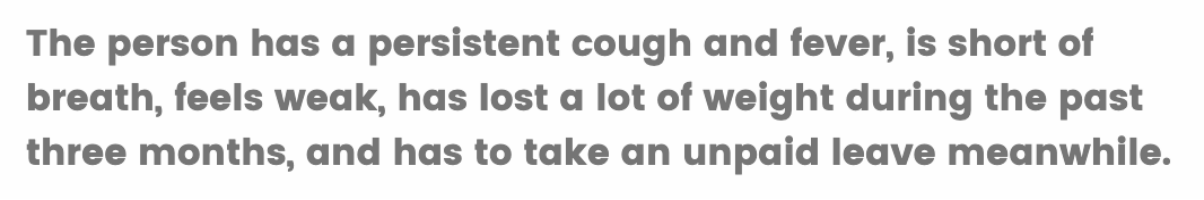


Figure A.3.2: Disease Description for Experiment 3: Salience and AUM Groups

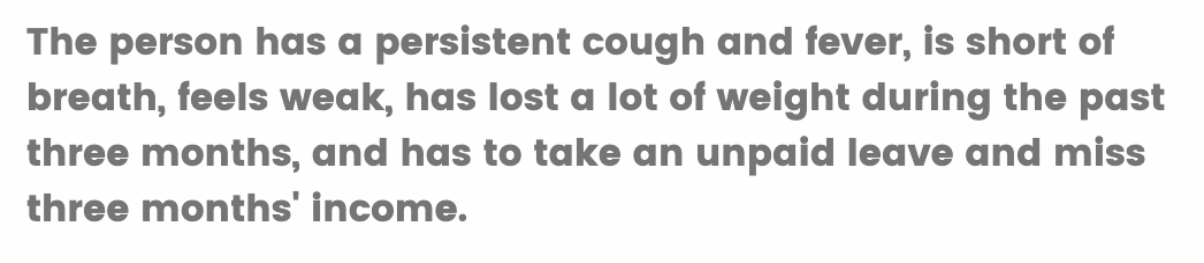


Figure A.3.3: Disease Description for Experiment 3: UM Group

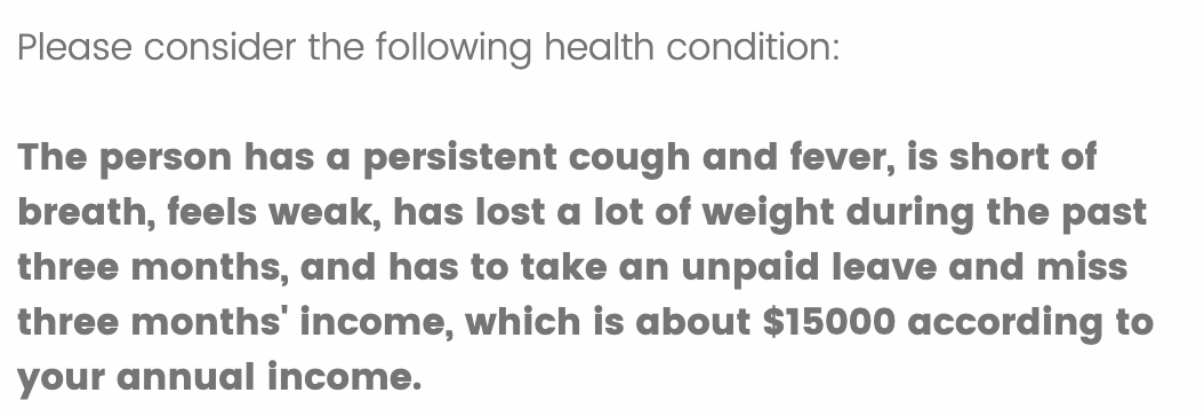


Figure A.3.4: Experiment 3: AUM Calculation

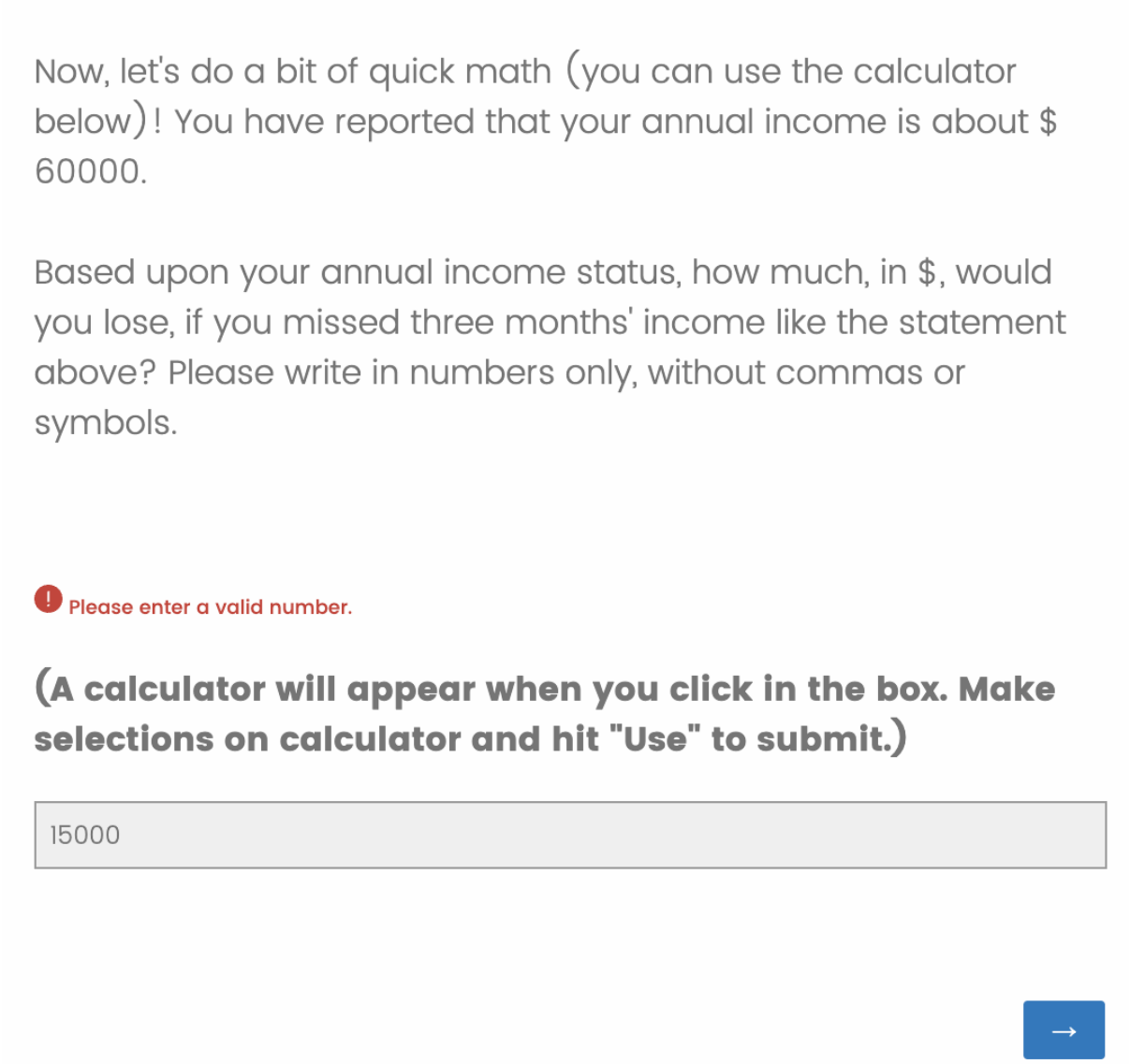


Figure A.3.5: Full AUM - Interface for Right Answer

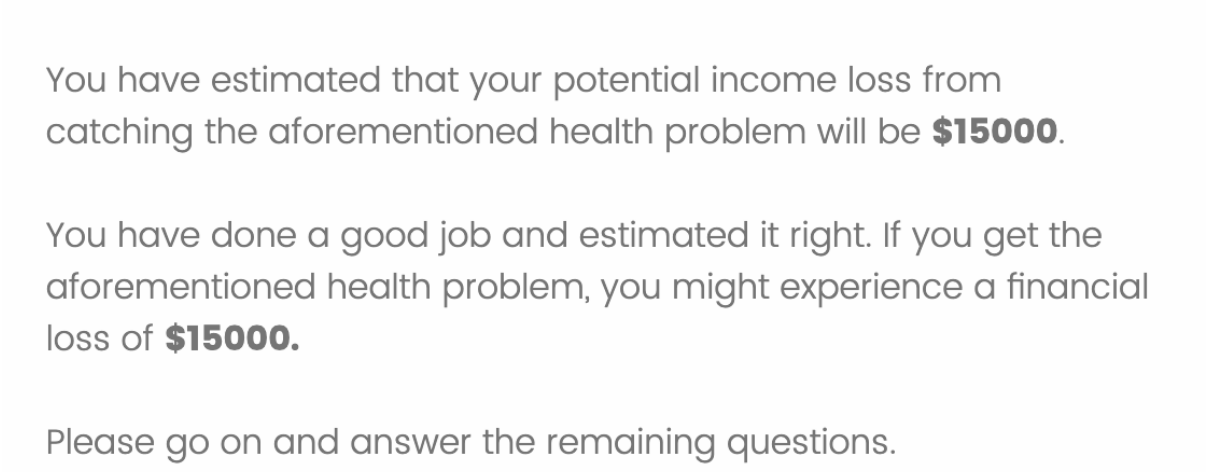
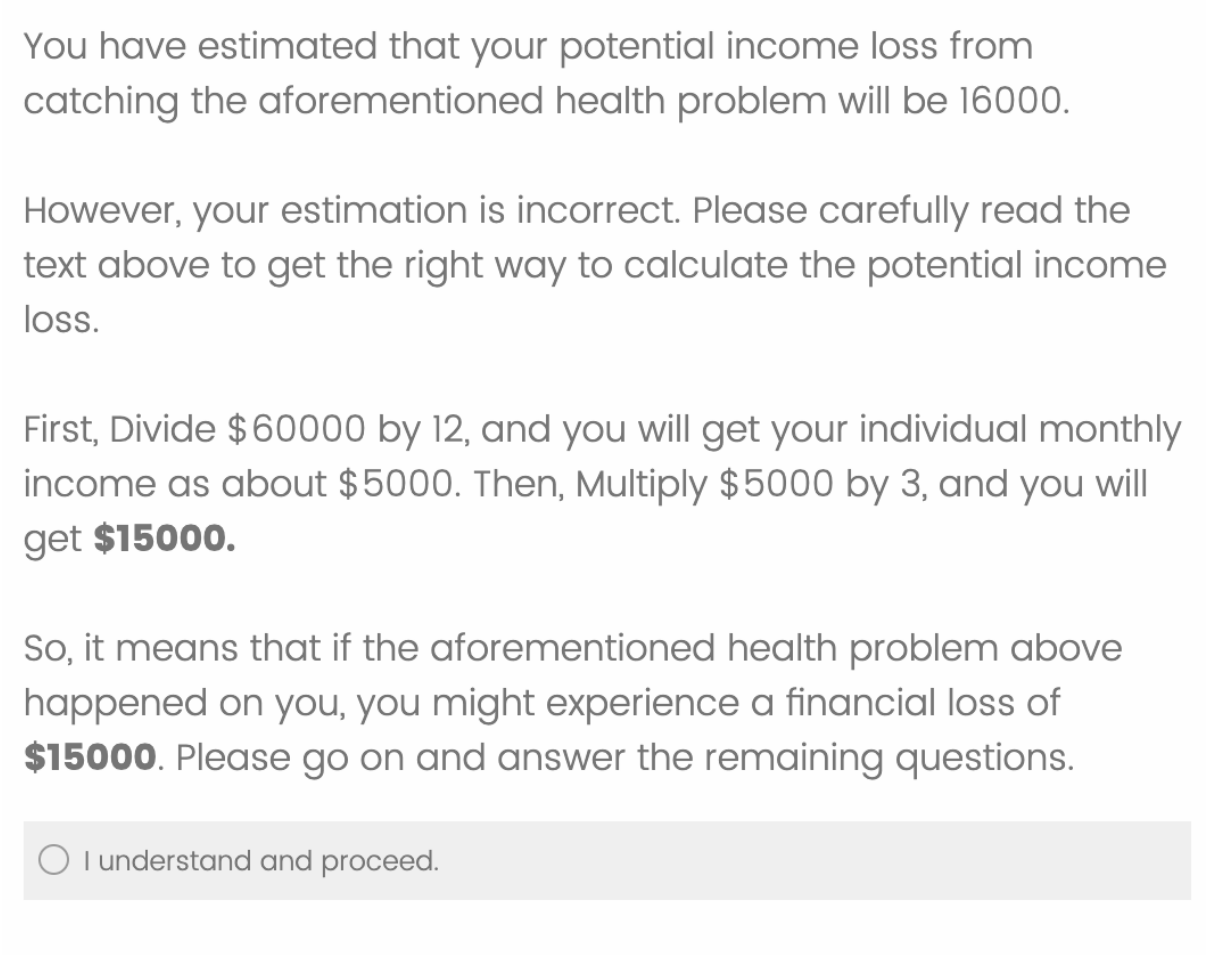


Figure A.3.6: Full AUM - Interface for Wrong Answer



A.4 Appendix A4: Instructions for Experiment 4

The demonstration uses the time duration of 16 weeks. The only difference in the 8-week group is in the description language.

Figure A.4.1: Disease Description for Experiment 4: Control Group

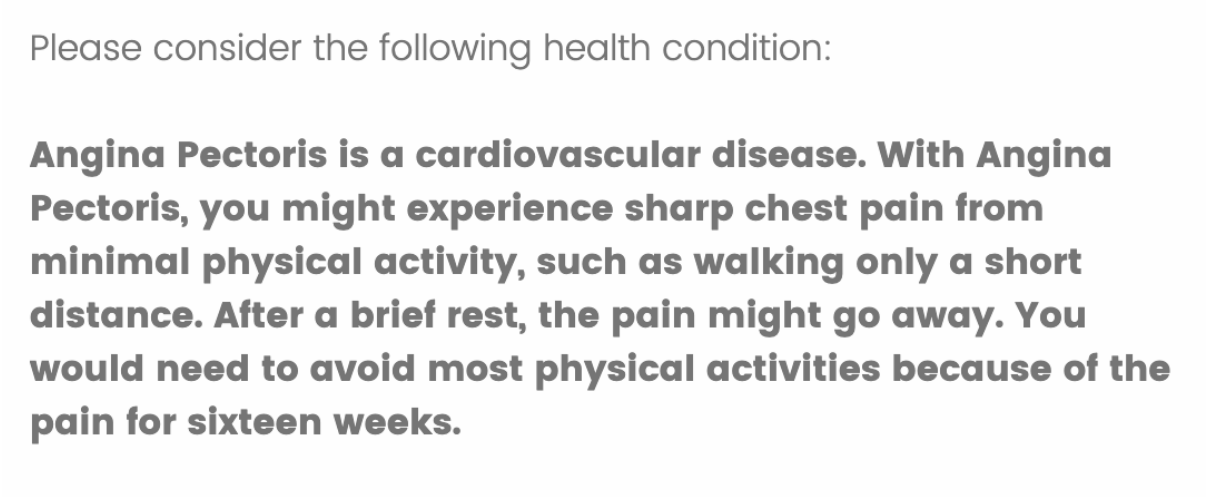


Figure A.4.2: Disease Description for Experiment 4: Calc Group

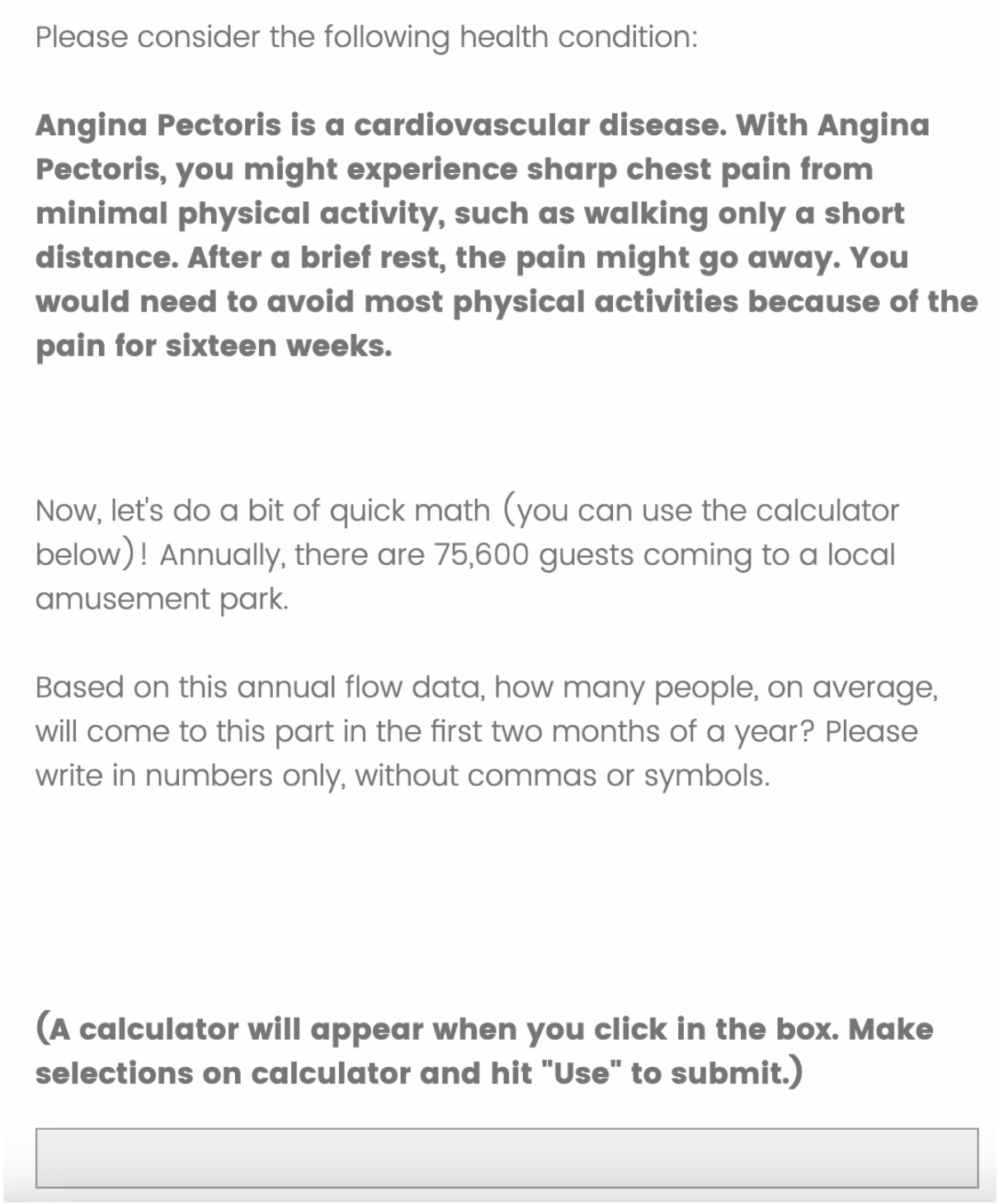


Figure A.4.3: Disease Description for Experiment 4: AUM Group

Please consider the following health condition:

Angina Pectoris is a cardiovascular disease. With Angina Pectoris, you might experience sharp chest pain from minimal physical activity, such as walking only a short distance. After a brief rest, the pain might go away. You would need to avoid most physical activities because of the pain, making you miss sixteen weeks' work and income.

Now, let's do a bit of quick math (you can use the calculator below)! You have reported that your annual income is about $ 60000.

Based upon your annual income status, how much, in $, would you lose, if you missed sixteen weeks' income like the statement above? Please write in numbers only, without commas or symbols. (A year contains 52 weeks.)

(A calculator will appear when you click in the box. Make selections on calculator and hit "Use" to submit.)

A.5 Appendix A5: Instructions for Experiment 5

The demonstration uses the time duration of 15 weeks. The only difference in the 8-week group is in the description language.

Figure A.5.1: Disease Description for Experiment 5

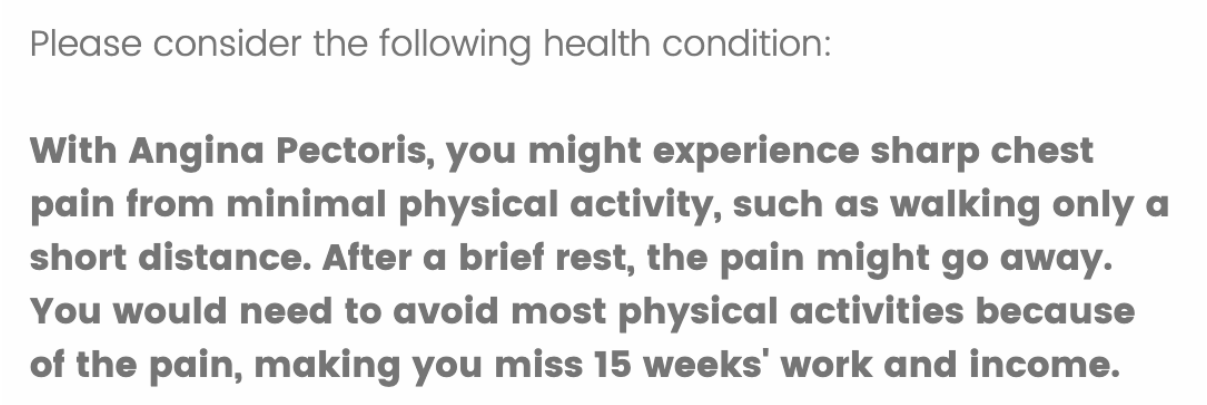


Figure A.5.2: AUM Procedure for Experiment 5

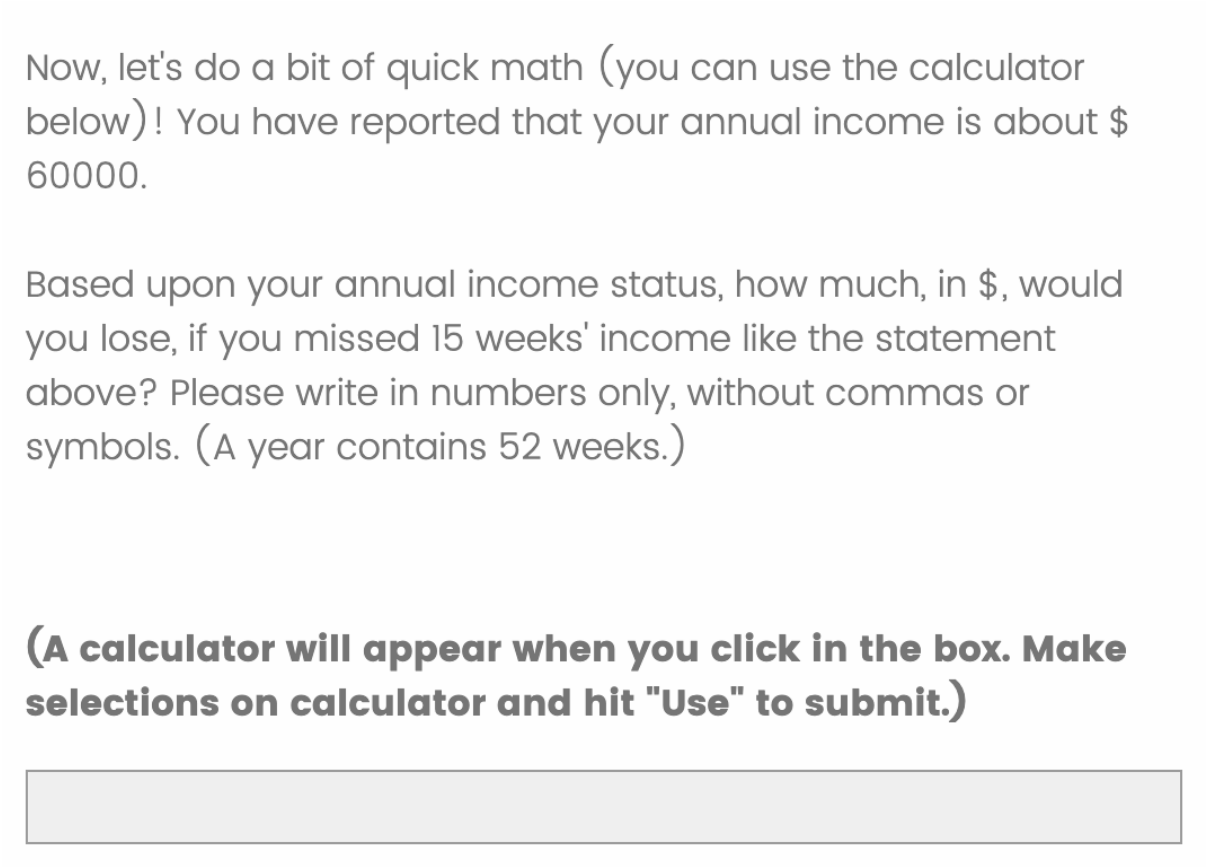
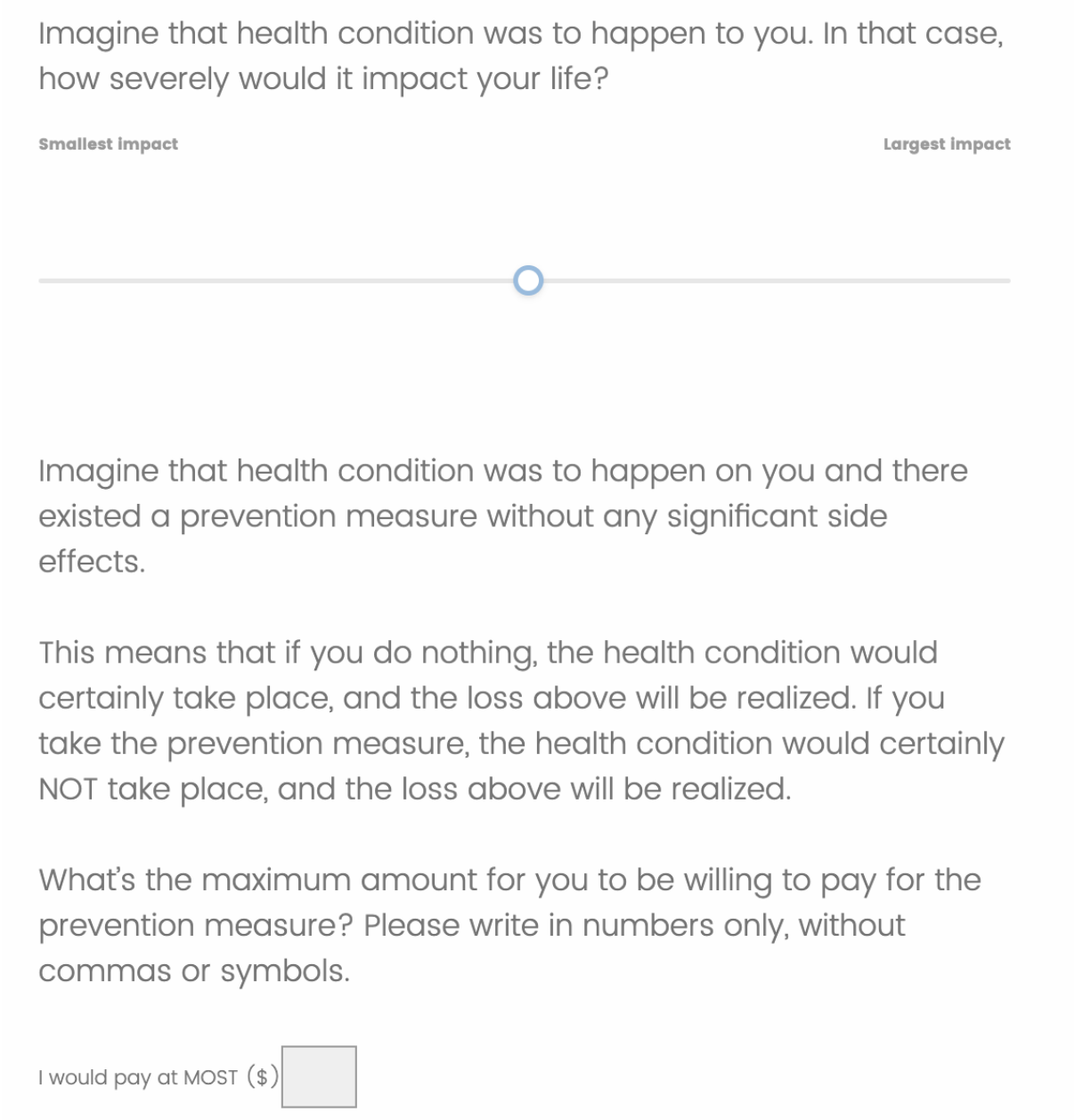


Figure A.5.3: Elicitation for Experiment 5



A.6 Appendix A6: Instructions for Experiment 6

Figure A.6.1: Disease Description for Experiment 6

With Angina Pectoris, you might experience sharp chest pain from minimal physical activity, such as walking only a short distance. After a brief rest, the pain might go away. You would need to avoid most physical activities because of the pain, making you miss 3 months' work and income.

Figure A.6.2: Description of Initial Probability for Experiment 6

Imagine that if you do nothing, the health condition above would take place with an **80%** chance, which is about **8 out of 10.**

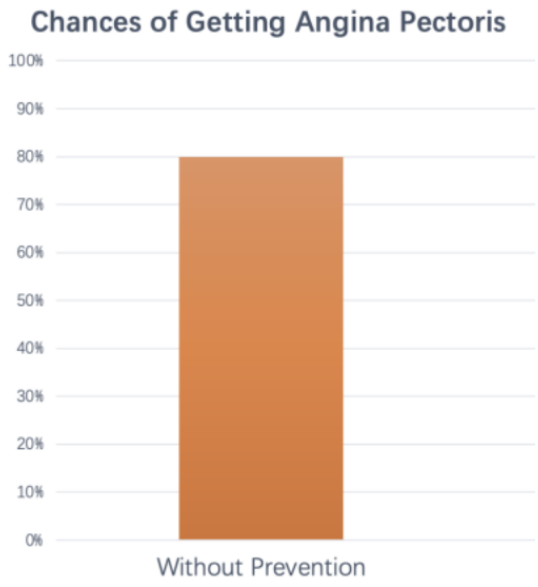


Figure A.6.3: Description of Initial Probability for Experiment 6

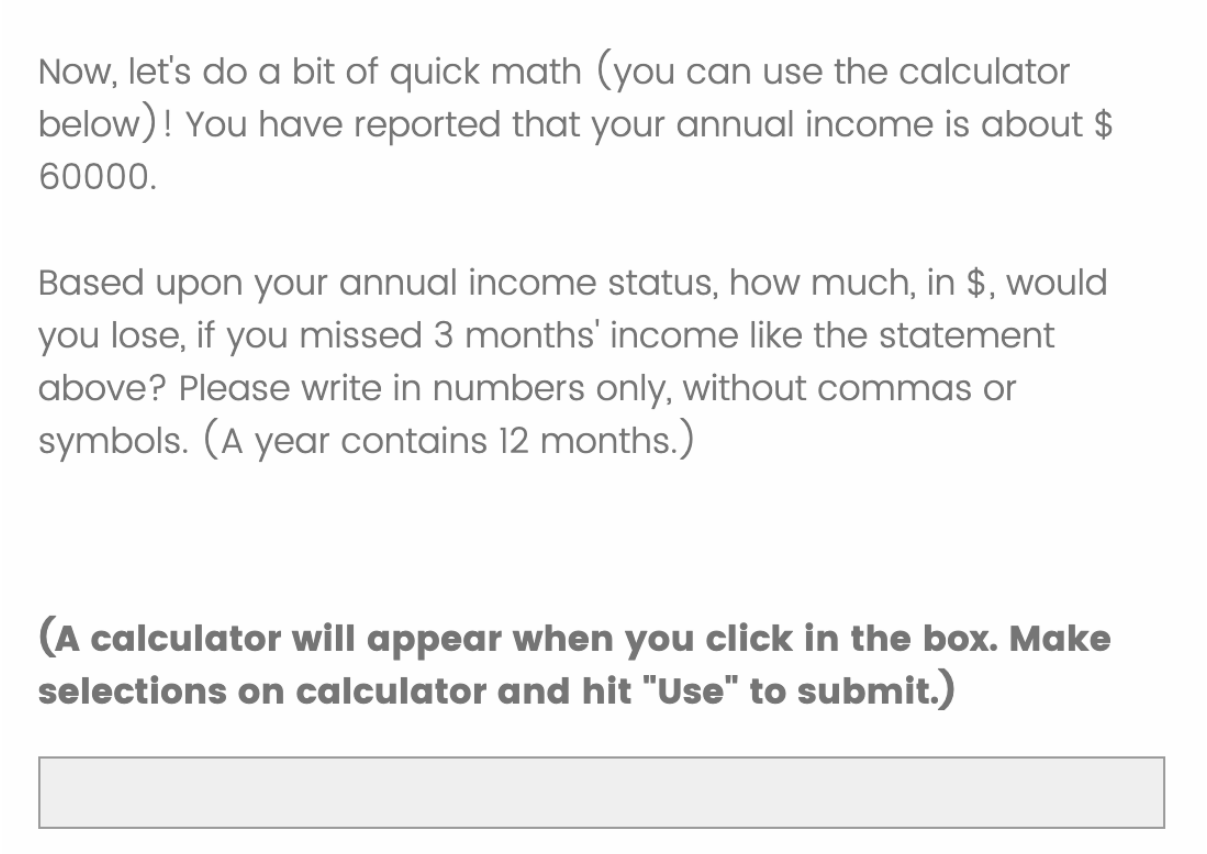
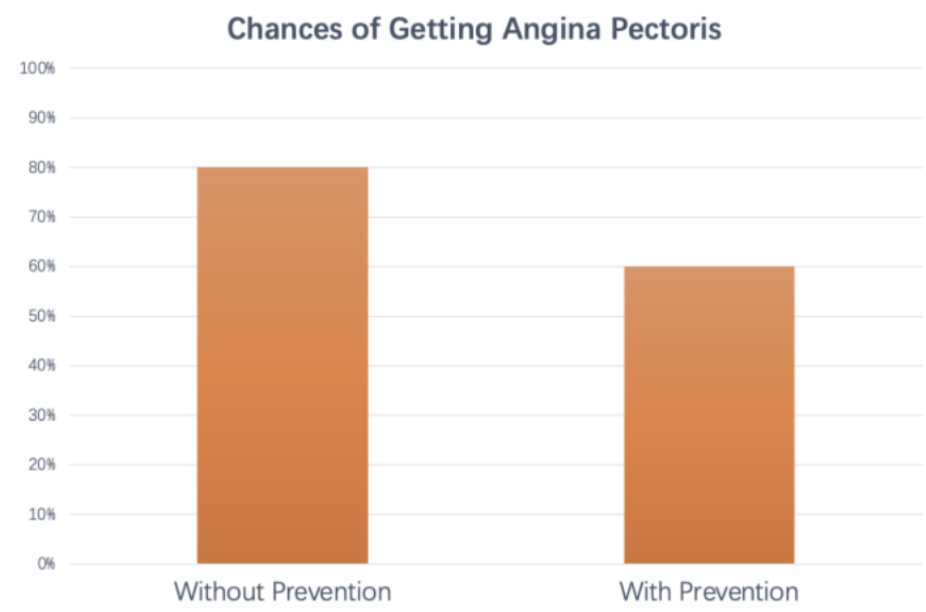


Figure A.6.4: Willingness-to-pay Elicitation for Experiment 6

Now, suppose a new side-effect-free prevention measure against Angina Pectoris becomes available. If you take the prevention measure, the chance that you get this illness would be reduced by **20** percentage points. This means that it would occur with only a **60%** chance, about **6 out of 10.**

A more intuitive illustration is here:



What's the maximum amount, out of your own pocket (on payments not covered by insurance, Medicare, Medicaid, etc.) for you to be willing to pay for the prevention measure? Please write in numbers only, without commas or symbols.

I would pay at MOST ($) [ ]

The

B Appendix B: Additional Results and Discussion

To be finished

C Appendix C: Further Discussion for Practical Applications

To be finished