University of Exeter
Business School

Decision Analysis under the Threat of Disaster

Di Mu

Submitted by Di Mu to the University of Exeter as a thesis for the degree of Doctor of Philosophy in Economics, March 2022.

I certify that all material in this thesis which is not my own work has been identified and that any material that has previously been submitted and approved for the award of a degree by this or any other University has been acknowledged.

Signed: ....

Abstract

This thesis studies the behavior of the public under the threat of an extreme event with the use of laboratory and on-line experiments. Chapter 2 and 3 are about decision making in the possibility of a natural climate disaster. Chapter 4 is about decisions in an infectious disease pandemic. Chapter 2 relates to the impact of different warning systems on public risk preferences during sudden natural disasters. It shows that while increasing the information content of warnings is usually beneficial and increases trust in the warning system, it must be done with caution since better decisions (judged by higher profits) are not always made with an increase in information. Chapter 3 is about changing the mode and content of information distribution influences the reaction time of the public. We have conducted research on this topic and found that adjusting the emergency warning system can reduce unnecessary waiting by the public, accelerate their response time, and thus reduce unnecessary casualties and property losses. Chapter 4 is about the influence of an early warning system and information intervention on public behavior in major public health events (such as COVID-19). This result shows that the public act selfishly during the pandemic but can be influenced by suggestions. This finding explains why many countries have been mired in the effects of the pandemic, with massive job losses and labor shortages.

Acknowledgements

First of all I would like to thank my first supervisor Todd R. Kaplan; he is the best person I have ever known and he always teaches me patiently and has a heart like a diamond. I would also like to thank my second supervisor, Cecilia Chen, for all the valuable guidance and insightful comments she has provided in my PhD.

I would also like to thank all the faculty members of the Business School at the University of Exeter for all the help and support I have received. I would like to thank the University of Exeter Business School and the Israeli Ministry of Science and Technology for financial support that enabled me to run my experiments.

Finally, I want to thank my family; I could not have finished my dissertation without your support.

Contents

List of Figures

List of Tables

1. Introduction

In this thesis we analyze people's behavior under the threat of a crisis. We mainly analyze two types of crisis: high-impact weather events and highly dangerous infectious disease. High-impact weather events significantly affect our lives and have done so throughout history. In 2005, based on the report of NHC (2008), Hurricane Katrina in the United States caused 75 billion dollars' worth of damage and over 1200 deaths. In the final report of the IPET Draft (Committee on New Orleans Regional Hurricane Protection Projects, 2009), the pre-event risk assessment of Katrina ASCE (2007) did not have a clear presentation of the risks. This may have made decision making more difficult. In 2021, mass flooding occurred in Germany and Belgium, causing 180 and 38 deaths, respectively (BBC, 2021). The recriminations continued into the autumn in Belgium with debates about dam management and even an investigation into the possibility of involuntary manslaughter charges over authorities' alleged failures to react to alerts issued by the EU's flood warning system. This is why we want to investigate how to improve the effectiveness of warning reactions by improving the way of communicating the warnings.

We also analyze the decisions made when there is a highly dangerous infectious disease. Globally, as of 22 December 2021, there have been 274,628,461 confirmed cases of COVID-19, including 5,358,978 deaths, reported to WHO (2020). It is important to analyze the decision-making process of citizens when society is under such a threat.

In this thesis,the main aim of my thesis is understanding of how to communicate with the public under the threat of diaster. In Chapters 2 and 3, we use university students in a laboratory experiment. The use of students as participants in this type of laboratory experiment is widely accepted as standard (Harrison and List, 2004). As the variances in age, education and income are small, internal validity can be assured (any difference in results between treatments is caused by the difference between the treatments themselves rather than any difference between the subjects allocated to each treatment). Non-students would have a larger variance of background knowledge and experience which would affect the internal validity. To assure external validity (i.e., that the results apply to decisions elsewhere), an important element is experimental realism. We do this by using a real payoff in cash with the decisions of the participants affecting their payoffs. Furthermore, a warning of severe weather is also something that participants may face in their real life and the experiment is framed as such.

In the first two chapters we use students as participants. There is a debate as to whether using students is comparable to a random sample of the general population. External validity is defined as "the inferences about the extent to which a causal relationship holds over variations in persons, settings, treatments, and outcomes" (Stanton and Weiss, 2002). Exadaktylos, Espín, and Branas-Garza (2013) compare the behavior of students to non-students in ultimatum, dictator, and trust games. They find that university students are a suitable subject pool for experiments in such games. Since most people in society, including students, potentially face the problem of decision making when there is an extreme weather event this means that university students are in and will be in a similar situation to the general public. Therefore, using students in an experimental study for weather warning will not reduce experimental realism or affect external validity (Mu, Kaplan, and Dankers, 2018). In Chapter 4, we conduct the experiment online. We want to find out how best to communicate information during a pandemic, so we collect data from the United States, the country with the most confirmed COVID -19 cases. Participants come from different age groups and professional, political, and religious backgrounds, so the result is general to the entire public.

In Chapter 2, I design experiment and run the experiment with students. then collect the data and do the data analysis. the we study decisions under different weather warning systems that vary in format and/or information conveyed using a laboratory experiment. Participants have to decide between a safe but costly option (spending to protect from a storm) and a risky option (of not spending for protection). We ran three treatments based upon the severe weather warning system for the UK that the Met Office has been using since 2011: a risk matrix to communicate the impact and likelihood of an event. In Treatment 1, participants received a colored table with a check in the box of the matrix that showed the likelihood and impact level of the warning. In Treatment 2, participants had the colored table and the color of the warning communicated but without a check in the exact box. In Treatment 3, participants only had the color of the warning communicated without seeing the associated table. Overall, our work shows that while increasing the information content of warnings is usually beneficial and increases trust in the warning system, it must be done with caution since better decisions (judged by higher profits) are not always made with an increase in information.

In Chapter 3, I come up with the research question and design the experiment and run the experiment with students. then collect the data and do the data analysis. we use laboratory experiments to study decisions with a weather warning system that varies in the accuracy of early warnings. The warning system is based on the one used by the Met Office since 2011: a risk matrix where the rows and columns are likelihood levels and impact levels, respectively. Participants can choose between a safe but costly option or a risky but free option. They can make their decisions based on an early warning that may be wrong or wait for a correct warning that comes later; however, waiting makes the safe option become more costly. Overall, our paper shows that issuing the early warning more likely to be right usually causes participants to be less likely to wait, and more surprisingly, contrary to theory, the early warning they have before has a significant effect on the decision they make after waiting and receiving the correct warning.

In Chapter 4, I decide to change the research questions to public decisions in a pandemic. This problem has many similarities to communicating weather risks (e.g., storms) to the public. Then I come up with the research question myself and learn how to program the experiment on oTree and run it online. in the context of continuous major public health events (like an epidemic), we study the impact of an early warning system and information intervention on public behavior. The innovative point is how the early warning system and different types of information sharing can be adapted to influence the public's decision between their own interests and the interests of the public in major public health events. When an epidemic is severe, some people tend to continue working because they are under pressure to preserve their lives, leading to a more rapid spread of the epidemic. Once the epidemic is brought under control, a significant number of people still work passively, slowing economic recovery. If the government does not intervene and direct people, they will behave selfishly, which is detrimental to the overall interests of society. It can even be said that people's social psychological changes have a far greater impact on the economy and society than the virus itself.

2. Decision Making with Risk-Based Weather Warnings

2.1 Introduction

In this chapter study decisions under different weather warning systems that vary in format and/or information conveyed using a laboratory experiment. Participants have to decide between a safe but costly option (spending to protect from a storm) and a risky option (of not spending for protection). We ran three treatments based upon the severe weather warning system for the UK that the Met Office has been using since 2011 - a risk matrix to communicate the impact and likelihood of an event. In Treatment 1, participants received a colored table with a check in the box of the matrix that showed the likelihood and impact level of the warning. In Treatment 2, participants had the colored table and the color of the warning communicated but without a check in the exact box. In Treatment 3, participants only had the color of the warning communicated without seeing the associated table. Overall, our work shows that while increasing the information with content of warnings is usually beneficial and increases the trust in the warning system, it must be done with caution since better decisions (judged by higher profits) are not always made with an increase of information.

High-impact weather events significantly affect our lives and have done so throughout history. In 2005, Hurricane Katrina in the US caused 75 billion dollars' worth of damage and over 1200 deaths (NHC, 2008). In the final report of the IPET Draft (Committee on New Orleans Regional Hurricane Protection Projects, 2009), the pre-event risk assessment of Katrina (ASCE, 2007) did not have a clear presentation of the risks. This may have made decision making more difficult.

Early Warning Systems (EWS), combined with effective communication and emergency preparedness at national and local levels, have huge potential to limit the human and economic losses from natural disasters such as Hurricane Katrina. Around the world, EWS have been developed for a wide range of hazards, most notably for extreme weather, floods and tsunamis, but also for other hazards like snow avalanches, wildfires, landslides, earthquakes and volcanic eruptions, and more complex hazards such as drought, food security and desertification (UNISDR, 2006). For hydrometeorological hazards, in particular, scientific and technological advances in recent decades have resulted in a marked improvement in the quality, timeliness, and lead time of hazard warnings. The adoption of ensemble prediction methods has resulted in a shift toward quantitative probabilistic forecasts, providing a range of possible outcomes and indicating the probability or chance of a particular weather or flood event happening.

Sivle and Kolstø (2016) has proven that both the likelihood of an event and its potential impact play an important role in weather forecast systems, but also provide additional challenges in communication and decision making ((Demeritt et al., 2013),(Dale et al., 2014),). Deciding whether to evacuate a large number of people or to issue a warning only to the most vulnerable needs to be risk-based, meaning the decision should be based not just on the likelihood of an event happening but also on its consequences. A warning system can be enhanced by providing both these types of information (Casteel, 2016).

However, in many EWS currently in operation the focus is on forecasting the hazard, and impacts are not commonly assessed. Already in 2006 the UNISDR noted a lack of knowledge of risks and vulnerabilities, and the limited engagement of relevant social sciences, as an area of weakness for EWS (Basher, 2006). This also includes the integration of risk information into the hazard warning itself. A better understanding of the risks and potential impact could help improve decision making, which in turn could mitigate the damage. The current EWS is accessible to the public but rarely accessed by them. Therefore, finding relationship between information in a warning system and public understanding of that warning system is a valuable avenue to investigate. Then, the Met office can choose which information system should be introduced to the public. It could also result in warnings that are more relevant to users and easier to communicate to the public at risk.

Within the UK, the National Severe Weather Warning Service (NSWWS) adopted a risk-based approach in 2011. Before then, warnings were issued based on the chance of ‘widespread disruption’, which was loosely defined and required a subjective judgement on the expected disruption depending on the vulnerability within a county (Neal et al., 2014). The current NSWWS system improves on this by using a risk matrix, combining the likelihood of a particular severe weather event on the vertical axis and its potential impact on the horizontal axis. Meteorologists use their expert knowledge of the area forecast to be affected during the weather event to assess the level of impact expected and therefore where to issue the ‘check’ on the horizontal axis. For the vertical axis, outputs from various models are analyzed to assess the probability of the event occurring. These assessments are combined to determine the color of the warning that is issued.

The specific level of risk – the check in the matrix – is predominantly used in communications with emergency managers, but is also available to the public. The headline message to the public, however, is based around the overall warning color.

This chapter’s main objective is to evaluate this approach to early warnings using students making decisions in an experimental laboratory.

Understanding how users make decisions with uncertain information is of paramount importance, and past research has focused on understanding and improving the decision-making process with weather forecasts and warnings. Our experimental methodology has been deployed previously. Joslyn et al. (2007) showed when participants need to decide whether to issue a warning without the full probability information, the participants tended to post too many warnings in the low probability situations and too few warnings in the high probability situations. This problem was reduced when the full probability information was provided. Roulston and Kaplan (2009) reported that by giving participants graphical information (uncertainty information of the forecast) they were able to make better decisions than participants who were not provided with uncertainty information. Marimo et al. (2015) found that participants make decisions faster when uncertainty information is provided in a graph compared to a table. Abraham et al. (2015) also showed that uncertainty information could affect participants' decisions. Previous research has also shown that providing more information will encourage people to make safer decisions (Wickens, Gempler, and Morphew, 2000). These results demonstrate that information can not only affect participants' decisions but also their risk preferences. There is also a desire among users to receive more information in the hope that it will lead them to make better decisions. For instance, when uncertainty information is not part of the initial forecast, 85% of the participants choose to have additional information to help them make decisions (Nadav-Greenberg and Joslyn, 2009). However, Jacoby, Speller, and Berning (1974) found that when people had more information, actually, they were making worse decisions even though they were more likely to feel satisfied and less confused. Sivle and Kolstø (2016) found that when people need to make a quick decision, the complexity of the warning must be reduced. In extreme weather events, people need to make decisions at short notice, so the right amount of information will help the public as well as emergency managers and responders to understand the warning and make better decisions.

Participants' decisions are not only related to the information conveyed. Winett and Kagel (1984) found that when messages had the same information, the presentation format of the information affected participants' decisions.

A key element in most warning systems, including the NSWWS, is the warning color. The color itself can significantly affect participants' behavior; Ryan (1990) showed the warning color red was associated with the highest level of hazard followed by orange, yellow and green. In Braun and Silver (1995)'s study a red warning color resulted in a higher compliance rate than green.

Based on the literature, more information is beneficial in most cases, but more information may also cause confusion. As the weather warning is based on an enormous amount of professional information, the information in the warning should be neither too simple nor too complicated. Therefore, two main interrelated questions that we wish to answer in the present paper: (1) To what extent does the amount of information that should be included in the warning affect decisions? (2) In which situations should the weather warning contain more details and in which situations should we keep the weather warning simple? By analyzing these two questions, we can establish the link between progress in public decision making and the Met Office's improvement of the weather warning system

We try to answer these two questions with a laboratory-based decision-making experiment. The findings of this paper have potential for application in communication of risk in other fields, as risk matrices are used not just in weather warnings but more widely in project management and corporate planning, including by organizations such as the US Federal Highway Administration and the US Federal Aviation Association.

We also try to see if there are behavior biases in decision making. Tversky and Kahneman (1973) indicated that the decisions participants make are based on the results they have in earlier periods, and people may develop a bias. For instance, a person who has suffered serious damage from a previous storm will have a biased perception of the likelihood and impact level. We consider whether or not there is a 'Cry Wolf Effect.' LeClerc and Joslyn (2015) show that there is an effect in that subjects tend to trust the warning system more and there is a significant improvement in decision quality as the false alarm level decreases.

2.2 Experimental design/method

In total, 341 University of Exeter undergraduate students from various disciplines were recruited to participate in the experimental sessions. There were 145 male and 196 female students; 149 year one students, 105 year two students and 87 year three students. We used the recruitment software ORSEE ((Greiner, 2015)) and all sessions used software programmed in z-Tree ((Fischbacher, 2007)). The sessions were computer based in sizes of 30 and took place in the Finance and Economics Experimental Laboratory (FEELE). Each participant's payoff only depended upon his/ her decisions and not those of other participants. The experiment consisted of 60 rounds of questions plus three test questions at the start and a questionnaire afterwards.

The use of students as participants in this type of laboratory experiment is widely accepted as standard (Harrison and List, 2004). As the variances in age, education and income are small, internal validity can be assured (any differences in results among treatments are caused by the difference in the treatments themselves rather than difference between the subjects allocated to each treatment). Non-students would have a larger variance of background knowledge and experience which would affect the internal validity. To assure external validity (i.e., that the results apply to decisions elsewhere), an important element is experimental realism. We do this by using a real payoff in cash with the decisions of the participants affecting their payoffs. Furthermore, a warning of severe weather is also something that the participants may face in their real life and the experiment is framed as such. Thus, using students in an experimental study will not reduce experimental realism or affect external validity ((Druckman and Kam, 2011)).

In each round of the experiment, participants are asked a question framed as deciding whether or not to move a product into a warehouse after receiving a warning of a coming storm. Moving the product involves a fixed cost but no risk from the storm. Not moving it has no fixed cost but risks damage from the storm. Participants a redivided into three treatments according to the information given to them.

Treatment 1: Colored table with a check in the box that shows both what the likelihood and potential damage is.

Treatment 2: Colored table without a check in the exact box, but with the color of the warning communicated.

Treatment 3: Color of the warning communicated without the associated table.

The warning in Treatment 1 shows impact level and likelihood level information. The warning in Treatment 2 shows only information about the color of the warning, but participants are visually shown how this color can vary according to likelihood and impact. Participants in Treatment 3 receive information about the color of the warning only.

The reasoning behind the three treatments is as follows:

T1: Impact/likelihood information (although no precise probabilities): this is similar to the information that is available from the current Met Office warning system.

T3: Color only: this is the top-level information that is mostly communicated to the public; most "traditional" warning systems (also for other areas beyond weather) use a simple color level approach.

T2: Participants are visually reminded that the warning is a combination of impact and likelihood, but only the color is communicated. This is an intermediate step that allows us to better explain some of the differences between T1 and T3.

For instance, when we compare T1 and T2, we can find out whether including the impact/likelihood information leads to better decisions. When we compare T2 with T3 we learn whether awareness of uncertainty in impact/likelihood leads to better decisions. When we compare T1 with T3 we can test how information on impact and likelihood affect the decisions of the participants. This risk-based warning system that is tested in Treatment 3 is based upon the warning system for severe weather in the UK. This system targets different audiences; for example, the main warning message to the public is based on the color only. The impact/likelihood information (i.e., the position to the matrix) is also available to the public (on the Met Office website) but is used more in the communication to emergency managers and the responder community.

The information given to the participants was kept to a minimum to allow them to answer a large number of questions within a reasonable time. Clearly, an experimental setting like ours is always going to be different from decision making in the real world; however, in the real world the amount of information received by the public through TV, radio and newspapers may well also be limited to headlines only. The Met Office provides more detailed information on weather warnings, for example on its website; however, this information is only viewed by a portion of the public, although admittedly professional users may well obtain more detailed information. Sivle (2016) indicated that the general public may not have enough background knowledge to understand the more detailed information in weather warnings. In our study, we used university students from diverse backgrounds as experiment participants, so arguably our results should be representative of the general public rather than professional decision makers.

The format of the warning is the same for every participant in the same treatment, but the impact level, likelihood and whether or not the storm actually happens are randomly generated by the computer. The probability of receiving a green, yellow, amber or red warning is set at 10%, 30%, 30% and 30%, respectively. The reason that the green warning only has a 10% chance of being chosen is because the green warning is the default status and we are more interested in how people behave while receiving an actual warning, that is, yellow, amber, or red. Once a color is selected, there is an equal probability of each box within the matrix of that color being selected. Participants were not told the probability of each warning or the likelihood of each box being selected and were not informed about the precise likelihood of each storm level occurring corresponding to the columns in the risk matrix (which were 20%, 40%, 60%, and 80%), but were told the cost of the damage associated with each storm level. Hence, they needed to form their own judgment of the warning system to make decisions with different payoff functions.

At the beginning of each period (each period has one question to answer), participants received 2000 tokens and then had to choose whether or not to move the product to a warehouse after they received a warning of a storm. The payoff function for the participants at each period is below in (2.1). The decision that participants made and their payoff in the previous period do not affect the initial tokens after the period.

Table 2.1: Damage of each impact level in each payoff function and the cost of moving.

Type/level	level 1	level 2	level 3	level 4	Moving cost
Low damage	300	600	900	1200	500
High damage	600	1000	1400	1800	500
Likelihood sensitive	925	975	1025	1075	500
Impact sensitive	300	400	1600	1700	500

Table 2.2: Expected profit per period and percentages of conditional HEPDs in Treatments 1, 2 and 3.

	Expected profit per period	Standard deviation	Percentage of conditional HEPDs	Standard deviation
Treatment 1	1541	147	82.9%	37.6
Treatment 2	1515	159	78.8%	40.8
Treatment 3	1520	160	80.1%	39.9

$$\text{payoff} = \begin{cases} \text{2000(initial endowment) - moving cost} & \text{if chose to move} \\ \text{2000 (initial endowment) - damage of the storm} & \text{if stayed and storm occurred} \\ \text{2000 (initial endowment)} & \text{if stayed and storm did not occur} \end{cases} \tag{2.1}$$

There were four different payoff functions: low damage, high damage, likelihood sensitive and impact sensitive, determining the amount of damage associated with each level of storm. The damages of every type and level are shown in Table 2.1. When making decisions, in the impact-sensitive payoff function, the difference in damage between impact level 2 and level 3 is larger than in any other payoff function, so participants will be more sensitive to the impact level.

In reality, people will be heterogeneously affected by weather events, and the level of impact may be different for different users. In real life, the cost of a decision will almost never be as clearly defined as in our experiment, and the impact depends on the level of exposure to the hazard as well as the vulnerability of the people, area or assets being affected. The vulnerability in turn may depend on a range of factors, including the capacity to cope with the effects of the hazard. Participants can also be affected by other factors as well, for example the maximum value of the damage; the real payoff function can be different for each person and much more complicated. However, the choices made are a trade-off between being able to test the differences between payoff structures and not introducing too much variability between payoffs or too many different payoff structures. In our experiment, exposure is implicitly assumed to be there, while the varying levels of impact in the different payoff functions reflect different sensitivities of the people to the weather, which may result in different responses.

Each participant faces 15 periods of each type of payoff, and the order is random to control for any learning effect.

Table 2.3: Expected profit per period, percentages of conditional HEPDs and percentages of HEPDs in different warning levels.

	Expected profit per period	Percentages of conditional HEPDs	Percentages of HEPDs
Green warning	1719	79.6%	79.3%
Yellow warning	1566	62.8%	61.7%
Amber warning	1474	85.5%	80.0%
Red warning	1468	94.8%	94.8%

Participants need to read the instructions first. If they answer a test question wrong, participants are asked to stop and wait for explanation before they move on to the main part of the experiment. Participants are given the warning (Fig. 2.1) and the payoff information (Fig. 2.2) at the same time. After choosing whether to move or not move the product they are responsible for, they see the results in a screen similar to Fig. 2.3: the initial endowment, the level of the storm and whether or not the storm occurred, their decision at this period, and the profit after each question.

At the end of the experiment participants are paid £5 plus £1 for every 2000 tokens they made in five randomly chosen periods. The average participant profit per period was 1525 tokens. This was slightly below what they were expected to make (1527 tokens) given their decisions. This contrasts with the 1561 tokens they would make if they chose the highest expected payoff decision every period in Treatment 1.

The cash reward is standard in the field of experimental economics. The benchmark is the incentive version with researchers concerned that a lack of monetary incentives (or insufficient incentives) will affect results.

2.3 Hypotheses

Hypothesis 2.1 *Offering information with likelihood and impact level is beneficial as they have more information about the weather warning.*

Based on the previous research, Joslyn et al. (2007) found that when probability information is provided, participants more often make economically rational decisions. there for participants should make the decision with higher expected payoff by providing impact information and likelihood information. We plan to test this hypothesis by comparing participants' expected payoff between treatment with probability and impact information and treatment with warning color information only.

Hypothesis 2.2 *If a warning system has payoff information and warning color at the same time, the warning color and payoff information both affect participants' decisions.*

For participants in treatment 1, they will have the warning color and payoff information at the same time, even participants can make decisions based on the payoff information. Silic et al. (2017) found the warning color carry information and can significantly affect the decision-making process. So, we believe the warning color and risk information both affect participants' decision. We can treat the warning color and payoff information as independent variables and test how the decision to move was affected by these variables in treatment 1

Hypothesis 2.3 *More information will make people follow the warning more, but contradictory information will hurt decision making.*

Since the warning in treatment 1 still contains payoff information and the warning color, there is a possibility that participants do not want to move because of the payoff information, but the warning color is amber. Fong, Hammond, and Hitchman (2009) found that the combination of the picture and text was more persuasive than the text. When treating with probability, impact level, and warning color. Participants should be more likely to follow the warning if all the information is present, but Mitroff (1988) shows that more information in crisis management can lead to confusion if the information is contradictory. We will compare participants' behavior when the warning color and payoff provide different suggestions, e.g., participants should decide to move based on the payoff information, but the warning color is yellow, which is a signal for not moving.

2.4 Results

We define the highest expected payoff decision (HEPD) as the choice with higher expected monetary payoff given full information about the potential impact and likelihood including the impact level, likelihood level and likelihood of each storm level occurring corresponding to the columns in the risk matrix. The reason that we want to measure this variable is that a higher fraction of HEPD can represent a better warning system (judged by higher profits).

We define the conditional highest expected payoff decision (conditional HEPD) as the choice with higher expected monetary payoff given information received by the participant and likelihood of each storm level occurring corresponding to the columns in the risk matrix. This also assumes the ability of the participants to learn the parameters of the experiment. For Treatments 1 and 2, they must learn that each likelihood level corresponds to a likelihood probability indicated for each box, 20–80%. For Treatment 2, they must learn that each box with a color is selected with an equal chance. For Treatment 3, they must learn the overall likelihood and impact that a warning indicates. Hence, making the conditional HEPDs could be challenging. This variable can be used to measure how well participants understand/follow the warning.

We are also interested in how participants behave when they do not made decisions that were the HEPD. To do so, we define risk direction as follows. If participants choose the conditional HEPD, then the risk direction is assigned 0. If the conditional HEPD is 'move' and the participants choose 'not move', thenthe risk direction is assigned 1. If the conditional HEPD is 'not move' and participants choose 'move', then the risk direction is assigned -1. Risk direction represents a participant's risk-aversion level. High risk direction means that participants favor a more risky option (not move) over the conditional HEPD, while low risk direction means participants more often choose a risk-free option (move) over the conditional HEPD.

We also tested reaction time but found that there was no significant difference in reaction time between the different treatment groups, and there was no significant relationship between reaction time and the proportion of HEPD.This result contrasts with the work of Kahneman and Tversky (2013) and Rubinstein (2007), who found that in many cases better decisions are made when there is more cognitive reasoning which takes longer response time – "thinking slow". We also test the risk preference of the participants but did not observe any significant result.

In this section, we analyze the responses of the participants. In order to reduce noise in the analysis, we use the expected profit based upon a participant's decision rather than actual payments received. This eliminates the noise induced from whether or not the storm occurred.

Result 2.1 *The participants that received information on both the likelihood and impact levels not only had higher expected profits but made decisions with higher expected payoffs given their information than participants receiving only the warning color.*

To test Hypothesis 2.1 we analyze the difference between different treatments, we compare the expected profit per period. One-way ANOVA shows the difference between treatments is significant ($F(2, 20, 457) = 61.04, p < 0.001$) in expected profit per period. Participants in Treatment 1 had significantly (significant is defined at the 5% level) higher expected payoffs than participants in Treatment 2 ($t = -10.52, p < 0.001$) and 3($t = -7.49, 0.001$). However, there is no significant difference between Treatment 2 and Treatment 3 ($t = 1.45, p = 0.146$).

A similar relationship holds on the fraction of conditional HEPDs. Participants in Treatment 1 are more likely to make the HEPDs than participants in Treatment 2 ($t = -6.54, p < 0.001$) and Treatment 3 ($t = -3.76, p < 0.001$). The difference between Treatments 2 and 3 is not significant ($t = 1.80, p = 0.072$). When we look at the fraction of HEPDs, the differences between Treatment 1 and the other two treatments (Treatment 2 ($t = -11.28, p < 0.001$), Treatment 3 ($t = -8.25, p < 0.001$)) are significant. The difference between treatments 2 and 3 is not significant ($t = 1.34, p = 0.181$).

Result 2.2 *Higher warning levels increased the percentages of both highest and conditional HEPDs, but the expected profit still decreased.*

It is important to find out how participants behaved under different situations. When presented with a yellow warning participants' expected profit per period was significantly higher (t=38.04, p < 0.001) with amber warnings, while the fraction of conditional HEPD was significantly lower ($t = –33.69, p < 0.001$). Participants had the highest expected profit when the warning color was green (1719) and the lowest expected profit (1467) when the color was red. To a large extent, this was due to the fact that the expected damage when not moving is higher at higher warning levels.

Participants are more likely to make conditional HEPDs when the warning color is red (94.8%) than with other colors. Participants made the lowest fraction of conditional HEPDs (62.8%) when the warning color was yellow. Moreover, when looking at the HEPDs, participants were more likely to make HEPDs with red warnings (94.8%) than other colors. The lowest fraction of conditional HEPDs (61.7%) was obtained when the warning color was yellow.

These results show that participants behaved differently based upon the level (color) of the warning they received. With green warnings, participants made more profit than in any other situation, not surprising given that the expected damage is lowest. With warnings ranging from yellow to red, the fraction of conditional HEPD increased from 62.8% to 94.8%. This shows that if expected damage increases, participants will behave more according to the HEPDs. On the other hand, participants made less profit as the expected damage increased.

2.4.1 Differences between treatments for yellow warnings

For yellow warnings, one-way ANOVA shows the differences among all three treatments are significant with respect to the fraction of conditional HEPDs ($F(2, 6221) = 46.54, p < 0.001$) as well as the fraction of HEPDs ($F(2, 20, 116) = 57.43, p < 0.001$). Participants in Treatment 1 made more conditional HEPDs than participants in Treatment 2 ($t = 9.61, p < 0.001$) and Treatment 3 ($t = 4.76, p < 0.001$), and participants in Treatment 3 made significantly more conditional HEPDs than participants in Treatment 2 ($t = 3.42, p < 0.001$). A similar pattern was observed with the HEPDs.

2.4.2 Differences between treatments for amber warnings

For amber warnings, one-way ANOVA shows the differences between treatments are significant with respect to the fraction of conditional HEPDs ($F(2, 6026) = 10.15, p < 0.001$) as well as the fraction of HEPDs ($F(2, 6026) = 11.64, p < 0.001$). These results also show that differences in participants' behavior between treatments decreased (F value decreased) as the warning color went up from yellow to amber. Participants in Treatment 1 chose conditional HEPDs significantly less often than participants in Treatment 2 ($t = –3.83, p < 0.001$) and Treatment 3 ($t = -3.77, p < 0.001$). Participants in Treatments 2 and 3 had no significant differences in the fraction of conditional HEPDs. However, as participants in Treatment 1 had full information about likelihood and impact level, they made significantly more HEPDs than participants in Treatment 2 ($t = 4.71, p < 0.001$) and Treatment 3 ($t = 2.95, p = 0.003$) while there was no significant difference between Treatments 2 and 3. The HEPDs are defined under the assumption that if participants all have full information about likelihood and impact level, conditional HEPDs are the highest payoff decision based on the information they have. Participants in Treatment 1 had higher fraction of highest payoff decisions and lower fraction of conditional highest payoff decisions, meaning overall they more often chose HEPDs as they had more information but worse understanding of the warning than the other two groups.

2.4.3 Differences between treatments for red warnings

This result supports the Hypothesis 2.3 when the participants are issued with a red warning, the HEPD is always move (risk-free decision), so the fractions of conditional HEPDs and HEPDs are always the same. One-way ANOVA showed that the differences between treatments are significant for the fraction of HEPDs and conditional HEPDs ($F(2, 6000) = 14.99, p < 0.001$). Participants in Treatment 1 made significantly more conditional HEPDs than participants in Treatment 2 ($t = 4.98, p < 0.001$) and Treatment 3 ($t = 4.18, p < 0.001$). Participants in Treatments 2 and 3 do not show a significant difference in the fraction of HEPDs.

These results show that participants in each treatment respond differently to a given warning. Participants in Treatment 1 had the highest fraction of HEPDs across all four warning colors; however, they had the lowest fraction of conditional HEPDs for amber warnings. There is another interesting point: when issued with a red warning, participants in Treatments 1 and 2 had exactly the same information (there is only one box in the warning matrix which is red). However, the data show that participants in Treatment 1 made significantly more HEPDs than participants in Treatment 2, while there were no differences between participants in Treatments 2 and 3. In other words, people follow the warning more when more information is provided about the impact level and likelihood level.

Result 2.3 *Participants' behavior varied depending on their payoff functions. They made more rational decisions when the magnitude of the impact was more important (Impact-sensitive payoffs questions), or the expected damage was higher (High-damage payoffs questions).*

Participants behaved differently under different types of payoff functions. One-way ANOVA shows the differences between types of payoffs are significant in expected payoffs per period ($F(3, 20, 456) = 119.02, p < 0.001$) and in the fraction of conditional HEPDs ($F(3, 20, 456) = 105.26, p < 0.001$). With a low-damage payoff function (the damage is lower for each impact level compared with a high-damage payoff), the participants' expected payoff per period is significantly higher ($t = 16.91, p < 0.001$)

than with a high-damage payoff, while the fraction of conditional HEPDs is significantly lower ($t = -7.54, p < 0.001$).

Comparing all four types of payoff functions, participants were more likely to make conditional HEPDs with impact-sensitive payoffs (85.3%), and had the lowest fraction of conditional HEPDs (73.4%) when they were answering likelihood-sensitive payoff questions.

We see that payoff type has a strong effect on both the fraction of conditional HEPDs and earnings per period (Table 2.4). By increasing the impact level (damage) of the storm with the moving cost unchanged, the expected profit per period decreases; however, with higher damage participants are more likely to choose the conditional HEPD. The results also indicate that participants more often behave risk neutral with an impact-sensitive payoff compared to a likelihood-sensitive payoff, potentially proving participants are more sensitive to the impact level rather than the overall likelihood. Participants chose to move significantly more often with the impact-sensitive payoff than with the other payoff functions (Table 2.5 and Table 2.6).

Participants in different treatments groups behaved differently in each individual question type. To find out how different treatments affect participants under different payoff functions, we looked at participants' expected payoff based on their decisions, expected payoff if they only made the highest expected decisions based on the information they had, and the fraction and percentage of HEPDs.

Table 2.4: Expected profit per period, percentages of conditional HEPDs and percentages of HEPDs for different payoff functions.

	Expected profit per period	Percentages of conditional HEPDs	Percentages of HEPDs
low damage	1552	79.1%	76.5%
high damage	1500	84.9%	82.1%
likelihood sensitive	1514	73.4%	72.4%
impact sensitive	1540	85.3%	83.6%

Table 2.5: Expected profit per period based on participants' decisions, HEPD strategy based on the information they have in each treatment for each payoff function, percentage of conditional HEPDs and percentage of HEPDs in each treatment for each payoff function.

	Participants' decision	HEPD strategy	Percentages of conditional HEPD	Percentages of HPED
low damage				
Treatment 1	1566	1593	78.9%	78.9%
Treatment 2	1539	1582	78.4%	73.7%
Treatment 3	1551	1586	80.6%	77.2%
High damage				
Treatment 1	1512	1545	84.8%	84.8%
Treatment 2	1494	1520	85.6%	80.6%
Treatment 3	1491	1522	84.1%	80.0%
Likelihood sensitive				
Treatment 1	1538	1566	82.0%	82.0%
Treatment 2	1500	1525	69.3%	67.7%
Treatment 3	1495	1533	65.2%	63.6%
Impact sensitive				
Treatment 1	1549	1587	85.7%	85.7%
Treatment 2	1528	1580	81.8%	80.3%
Treatment 3	1542	1578	90.6%	85.7%

Table 2.6: Average risk direction in each treatment.

	Average risk direction	Standard deviation
Average	-0.0516	0.4364
Treatment 1	-0.0553	0.4101
Treatment 2	-0.0540	0.4574
Treatment 3	-0.0406	0.4441

2.4.4 Low-damage payoffs

For low-damage payoffs, one-way ANOVA showed the differences between treatments are significant ($F(2, 5111) = 16.61, p < 0.001$) in expected profit per period. Participants in Treatment 1 had significantly higher expected profit per period compared to participants in Treatment 2 ($t = 2.76, p = 0.006$) and Treatment 3 ($t = 5.75, p < 0.001$). However, there is no significant difference between Treatments 1 and 3 in HEPDs. Similar to the overall result, there is no significant difference between Treatments 2 and 3 in profit per period. However, unlike the overall result there is no significant difference between treatments in the fraction of conditional

HEPDs; one-way ANOVA shows the differences between treatments are not significant ($F(2, 5111) = 1.08, p = 0.3408$). When we look at the fraction of HEPDs, one-way ANOVA shows that in this case the difference between treatments is significant ($F(2, 5111) = 7.59, p < 0.001$). Similar to the expectation profit, participants in Treatment 1 significantly more often chose the HEPDs compared to participants in Treatment 2 ($t = 3.86, p < 0.001$), but there is no significant difference between Treatments 1 and 3.

2.4.5 High-damage payoffs

When participants were presented with a high-damage payoff function, differences between treatments were significant in expected profit per period, as shown by one-way ANOVA ($F(2, 5111) = 9.77, p < 0.0001$). Participants in Treatment 1 had higher expected profit than participants in Treatment 2 ($t = 3.71, p < 0.001$) and Treatment 3 ($t = 3.74, p < 0.001$). Even though participants in Treatment 1 had higher payoff per period, the fraction of conditional HEPDs is not higher than for the other groups. When we look at the fraction of HEPDs, one-way ANOVA showed the differences between treatments are significant ($F(2, 5111) = 8.33, p < 0.001$). Participants in Treatment 1 chose the HEPDs significantly more often than participants in Treatment 2 ($t = 3.48, p < 0.001$) and Treatment 3 ($t = 3.40, p < 0.001$).

2.4.6 Likelihood-sensitive payoffs

With likelihood-sensitive payoffs, participants reacted differently compared to participants with other payoffs. One-way ANOVA shows the differences between treatments are significant ($F(2, 5111) = 67.98, p < 0.001$) in the fraction of conditional HEPDs. To analyze the effect of treatments, we used a regression model to estimate the fraction of conditional HEPDs. These results show the differences c between different treatments (group 2 coefficient = 12.7%, $t = 9.19, p < 0.001$, group 3 coefficient = 16.9%, $t = 10.39, p < 0.001$) are significantly higher than in the general case (across all payoff functions) (group 2 coefficient = 4.1%, $t = 6.54, p < 0.001$, group 3 coefficient = 2.76, $t = 3.76, p < 0.001$). Looking at the fraction of HEPDs, one-way ANOVA shows the differences between treatments are significant ($F(2, 20, 457) = 74.55, p < 0.001$). The difference between Treatment 1 and Treatments 2 and 3 increased compared with conditional HEPDs (group 2, $t = 9.79, p < 0.001$, group 3, $t = –10.75, p < 0.001$) and the difference between Treatments 2 and 3 does not significantly change.

2.4.7 Impact-sensitive payoffs

With impact-sensitive questions, one-way ANOVA shows the differences between treatments are significant ($F(2, 20, 457) = 22.43, p < 0.001$) in the fraction of conditional HEPDs. However, unlike all the other question types and the general case participants in Treatment 3 made more profit and had a higher fraction of conditional HEPDs than the other treatments. Participants in Treatment 3 significantly more often chose HEPDs compared to participants in Treatment 1 ($t = 3.50, p < 0.001$), and participants in Treatment 1 significantly more often chose HEPDs than participants in Treatment 2 ($t = -3.70, p < 0.001$). One-way ANOVA shows these differences in the fraction of HEPDs between treatments are significant ($F(2, 5111) = 13.29, p < 0.001$).There was no significant difference between participants in Treatments 1 and 3 ($t = -0.02, p < 0.984$), but the difference between participants in Treatments 1 and 2 is still significant ($t = 4.68, p < 0.001$). One-way ANOVA shows the differences in the fraction of expected profit between treatments are also significant between treatments ($F(2, 20, 457) = 5.58, p < 0.0038$).

Based on these results, participants in Treatment 1 had at least one of the highest fractions of HEPDs across all four payoff functions. When participants were presented with likelihood-sensitive questions participants in Treatment 1 did not show significant change compare with other payoff functions; however, participants in Treatments 2 and 3 had significantly lower fractions of HEPDs and conditional HEPDs. Participants in Treatment 3 had the highest fraction in conditional HEPDs and one of the highest fractions of HEPDs with impact-sensitive payoff decisions.

Result 2.4 *Participants' behavior can be negatively affected when the optimal decision given their information and the warning color are not consistent. (For instance, when optimally they should not move, yet the warning color is amber.)*

Different parts of the information in a warning system could convey different messages to the public. In Treatment 1, the warning contains the warning color and the likelihood level of the impact. Therefore, we need to test how the public behaves when these two parts of the warning system give different suggestions to participants.

To test Hypothesis 2.2 it is important to know how participants' behavior is biased. On average, participants more often chose the risk-free decision than the risky decision (risk direction = -0.052, standard deviation = 0.437). In other words, they had a slight preference for making a decision to move the product over the conditional HEPD. This also holds for each individual treatment. Risk direction did not vary between treatments: one-way ANOVA showed the differences in risk direction between treatments is not significant ($F(2, 20, 457) = 1.83, p = 0.16$).

To analyze, we assume that amber and red warnings give a signal of 'move' (risk-free), and green and yellow warnings give a signal of 'not move' (risky). We classify the fraction of amber and red warnings when the HEPD is 'not move' as the rate of false alarms, as the warning shows a signal of 'move' when the HEPD is 'not move'; and we classify the fraction of green and yellow warnings as missed events, as the warning shows a signal of 'not move' when the HEPD is 'move'. We use these two factors to evaluate warning systems in different treatments. We define the error level as the fraction of false alarms and missed events of the total warnings issued.

By comparing low-damage and high-damage payoff functions, we examine whether the fraction of error level relates not only to the probability of the event but the impact level as well.

Firstly we want to evaluate the warning matrix, as we defined green and yellow as signs of safe, and amber and red warnings as signs of risky. Hence, the optimal warning system should present green and yellow when the HEPD is 'not move', and amber and red warnings when the HEPD is 'move'.

We calculate the probability of a false alarm (the probability of an amber or red warning if the HEPD is 'not move' over the probability of the HEPD being 'not move') and missed event (the probability of a green or yellow warning if the HEPD is 'move' over the probability of the HEPD being 'move').

In the previous discussion, we saw that when participants answered with high-damage payoffs they chose the HEPD significantly more often than when they were answering with low-damage payoffs ($t = 7.72, p < 0.001$). The error level has a significant effect on conditional HEPDs, with one-way ANOVA showing the differences between treatments are significantly negative ($t = –14.37, p < 0.001, F(1, 20, 458) = 206.64, p < 0.001$), and a significant effect on the risk direction, with one-way ANOVA showing the differences between treatments are significantly negative ($t = -14.37, p < 0.001, F(1, 20, 458) = 288.88, p < 0.001$).

Table 2.7: For HEPDs of move/not move, percentages of green, yellow / amber, red warnings and their corresponding percentages of conditional HEPDs and risk directions.

	Probability of false alarm	Probability of missed event	Percentage of conditional HEPDs	Risk direction
Low-damage	15.8%	0%	79.1%	-0.0760
High-damage	0%	23.1%	84.9%	0.0606
Likelihood-sensitive	23.3%	22.6%	73.4%	-0.1673
Impact-sensitive	0%	9.1%	85.3%	-0.0236

Table 2.7 shows that participants significantly less often chose conditional HEPDs ($t = –15.41, p < 0.001, F(1, 20, 458) = 234.36, p < 0.001$) and significantly more often made risk-free decisions while they did not choose conditional HEPDs ($t = –22.16, p < 0.001, F(1, 20, 458) = 491.19, p < 0.001$) with the likelihood-sensitive payoff function.

With impact-sensitive payoffs, as participants were sensitive to the impact level in these questions, the error level is high on the likelihood level but low on the impact level. For example, when the impact level is 3 and the likelihood level is 40%, the warning color is yellow but the HEPD is to move due to the high damage associated with this impact level (see Table 2.1). One-way ANOVA shows participants significantly more often chose conditional HEPDs ($t = 9.65, p < 0.001, F(1, 20, 458) = 93.06, p < 0.001$) and significantly less often chose to make the risk-free decisions while they were not the conditional HEPDs ($t = 5.30, p < 0.001, F(1, 20, 458) = 28.05, p < 0.001$) with the likelihood-sensitive payoff function.

We use ANOVA test to analyze the effect of false alarms and missed events.

In the model we assume two new variables: false alarm and missed event, if there is a false alarm, the variable false alarm is 1, otherwise it is 0. If there is a missed event, the variable missed event is 1 otherwise it is 0. One-way ANOVA shows the fraction of false alarms and missed events have a significant effect on the fraction of conditional HEPDs ($F(1, 20, 458) = 206.64, p < 0.001$). Both the fractions of false alarms ($t = –17.53, p < 0.001$) and missed events ($t = –2.28, p = 0.023$) have a significantly negative effect on the fraction of conditional HEPDs; in other words, the higher the error level, the fewer conditional HEPDs participants will make. One-way ANOVA shows the fractions of false alarms and missed events have a significant effect on the participants' risk direction ($F(1, 20, 458) = 288.88, p < 0.001$). The fraction of false alarms ($t = -25.33, p < 0.001$) has a significant negative effect on the risk direction level; in other words, the higher the false alarm level, the more risk-free decisions will be made by the participants. The fraction of missed events ($t = 1.92, p = 0.055$) has a positive effect on the risk direction but is not significant.

If we look deeper into each treatment, we can find out that both false alarms and missed events have an effect on the risk direction in each treatment. Similar to the overall case, false alarms have a negative effect on the risk direction in each treatment. In Treatment 1, the missed events had a positive effect on the risk direction ($t = 8.56, p < 0.001$). This means the higher the level of missed events, the more likely participants are to choose risky decisions. Missed events have a significant negative effect in both Treatments 2 ($t = –2.45, p = 0.005$) and 3 ($t = 3.16, p < 0.001$).

Our results show how false alarms and missed events affect participants' behavior in the fraction of conditional HEPDs and risk direction. The higher the error level, the lower the fraction of conditional HEPDs. This may be caused by the level of trust that participants have in this warning system.

Trainor et al. (2015) showed that participants felt there were fewer false alarms than the actual number of false alarms, and in our experiment the false alarm rate was too low (maximum 23.3%) which dampened any cry wolf effect and we find no significant evidence for one.

Result 2.5 *When participants received both a warning color and risk information (impact and likelihood level), they followed both the warning color and risk information contained in the warning, but followed the color more.*

To analyze how warning color and information affect the participants' behavior at the same time, we built a probit regression model to show the relations of dependent variables, whether or not participants chose the risk-free decision (not move, $y = -1$) or the risky decision (move, $y = 1$), and independent variables, warning color and information. As only participants in Treatment 1 had the full information and the warning color, we only looked at the data of these participants.

We define two new variables is (information suggestion) and cs (color suggestion). When the conditional HEPD is not move, it is equal to 0, and when the conditional HEPD is move, it is equal 1. If the warning color is green or yellow, then $cs = 0$, and if the warning color is amber or red, then $cs = 1$. The probability predictor of participants' behavior is shown in Table 9.

$$Pr(y = 1, (notmove)) = \iota(\beta_0 + \beta_1 cs + \beta_2 is)$$

The regression results show that both variables are significant and influence the decision but the coefficient for cs is significantly higher than that of is.

2.5 Discussion and conclusion

In this chapter, we used experimental methodology to compare participants' behavior under different warning formats that varied based upon the information sent and the presentation format. We discovered that giving the impact level and likelihood level helps participants make better decisions (as indicated by the fractions of conditional HEPDs and HEPDs). However, when participants cared more about impact than likelihood because their damages rose steeply at higher impact levels, those given only the warning color (Treatment 3) more often chose the option that had higher expected payoff conditional on their information. These results suggest that an effective warning system should not just have one presentation format, but should vary the format based upon the needs of those receiving it.

Broken down by warning colors, participants more often chose the conditional HEPDs when they had amber or red warnings than when they had green or yellow warnings. Participants that were given the likelihood and impact levels (Treatment 1) had a higher fraction of conditional HEPDs and higher profit in all warnings except amber, and a higher fraction of HEPDs for all warnings. In other words, participants more often chose the HEPDs when the warning was more serious and the potential damage was high.

It is interesting that when the warning color was red, participants in Treatments 1 and 2 had exactly the same information (as there is only one red box in the matrix); however, the participants in Treatment 1 had a significantly higher fraction of HEPDs. Thus, more information caused more participants to follow the warning. The reason could be that, while they had the same information, participants in Treatment 2 were less likely to follow the warning because they had less information about the meaning of the warning. Thus, more information could increase trust in the warning system.

Table 2.8: Summary of how information suggestion and color suggestion affect participants' decisions.

y	Coefficient	Std err	z	p $\geq z$	95% confident	interval
is	0.8037	0.0273	29.39	0.000	0.7502	0.8573
cs	1.0127	0.0276	36.68	0.000	0.9586	1.0668
cons	-0.3988	0.0156	-25.49	0.000	-0.4295	-0.3681

Participants made different decisions under different types of payoff function (low damage, high damage, likelihood sensitive and impact sensitive). With high-damage payoffs, participants more often chose conditional HEPDs than when they had low-damage payoffs (independent of treatment). This is more evidence that participants more often chose the HEPDs when the potential damage was high.

With likelihood-sensitive payoffs where there was little difference in damage between the impact levels, participants were more sensitive to the likelihood level than the impact level. Under such payoffs, participants in Treatment 1 more often chose HEPDs and conditional HEPDs than participants in the other two treatments, while participants in Treatment 2 more often chose HEPDs and conditional HEPDs than participants in Treatment 3. The reason for this is that the design of the warnings is not the most suitable for those with likelihood-sensitive payoffs. Ideally, for such payoffs, the warning colors in the matrix should be stripes with the red in the top row and the green in the bottom row.

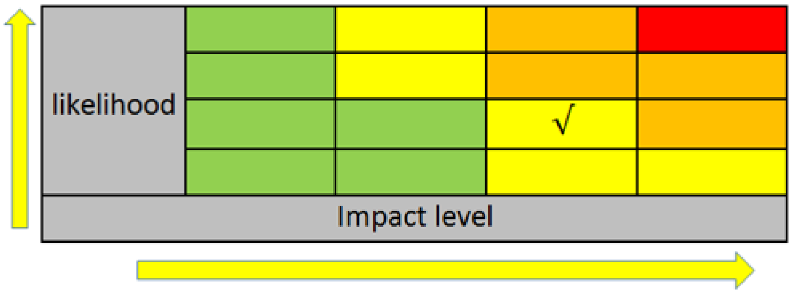
With impact-sensitive payoffs where damage increases steeply at higher impact levels, participants in Treatment 3 had the highest fraction of conditional HEPDs and HEPDs of all three treatments. This result shows that when participants cared more about impact than likelihood, just giving the warning color helped them to make more HEPDs. This also shows that giving people more information will not always help people make better decisions. In this case, the color was almost sufficient to make the HEPD. The one difference from the ideal warning is that a level 3 storm with a likelihood of 40% should be an amber warning rather than a yellow warning within the parameters of the current experiment.

On average, participants more often choose the safe option when they are not choosing the conditional HEPDs. We find that the error level significantly affects participants' risk directions and fraction of conditional HEPDs. The results show that the higher the fraction of false alarms (amber and red warnings when the HEPD is ('not move') and missed events (green and yellow warnings when the HEPD is 'move'), the more participants chose the safe option over the highest expected payoff one.

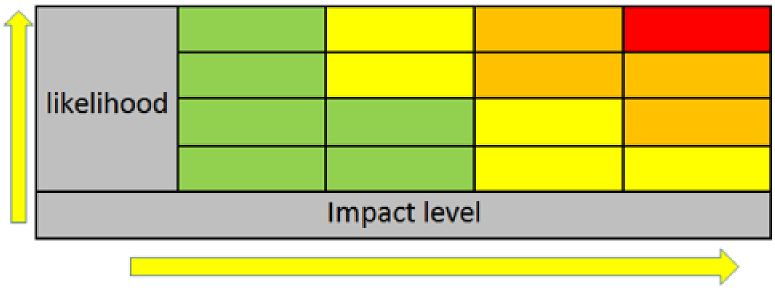
For participants in Treatment 1, false alarms had a negative effect on the risk direction (as with the overall effect on all three treatment groups) and missed events had a positive effect on a participants' risk direction (in contrast to the overall effect on all three treatment groups). This could be caused by participants using both the warning color and the information contained, so it is easier for participants in Treatment 1 to find out when it is a false alarm or missed event. Our probit regression model shows that when participants had both information (impact level and likelihood level) and a warning color, participants made use of both but followed the warning color more than the information.

Overall, Met office meteorologists can benefit from the results of this chapter. Our work shows that introducing the information content of warnings at the level of probability and impact is beneficial to the public. A greater amount of information can be beneficial, but more information also means that the likelihood of conflicting information increases. Each individual has a different understanding of the information when we provide an alert with multiple types of information, but each type of information can provide a conflicting message to the individual. For instance, public may want to choose to move if they only consider the payoff information, but the warning color is yellow which is a signal of not move. Therefore, we suggest that we provide a warning with more information when the situation is complicated and there is heterogeneity among the users of the information. In addition, to keep the warning simple when the situation is urgent and the actions that they need to take are homogeneous.

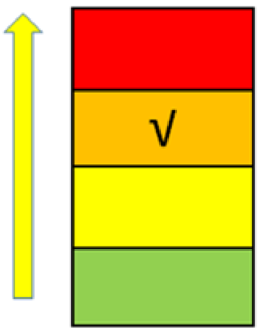
There are many directions of future research. We can continue to run similar experiments with different parameters for the payoffs in order to better pinpoint shifts of behavior. Moreover, in our experiments, the weather warning perfectly matched the actual probability and level of the storm. While we did this for simplicity in our initial experiment, actual storms may occur at a higher or lower level rather than at a specific level or not at all. In addition, there could be ambiguity about the likelihood of the storm given the chaotic nature of the weather. Hence, giving warnings that have these additional realistic errors could be useful in determining how people behave and adjust to mistakes.



(Treatment 1)

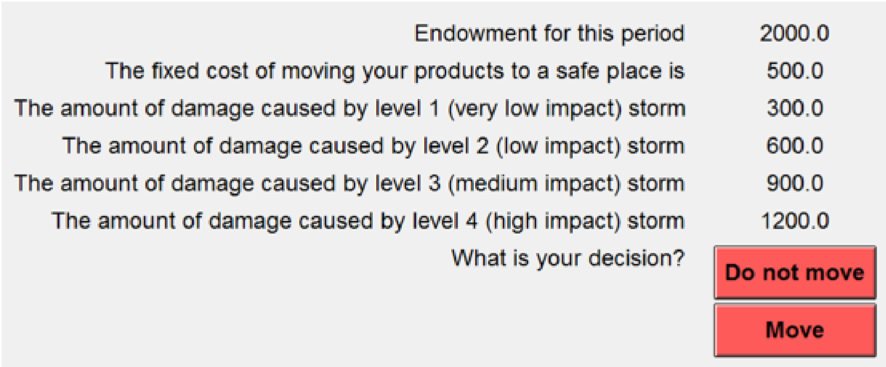


(Treatment 2)

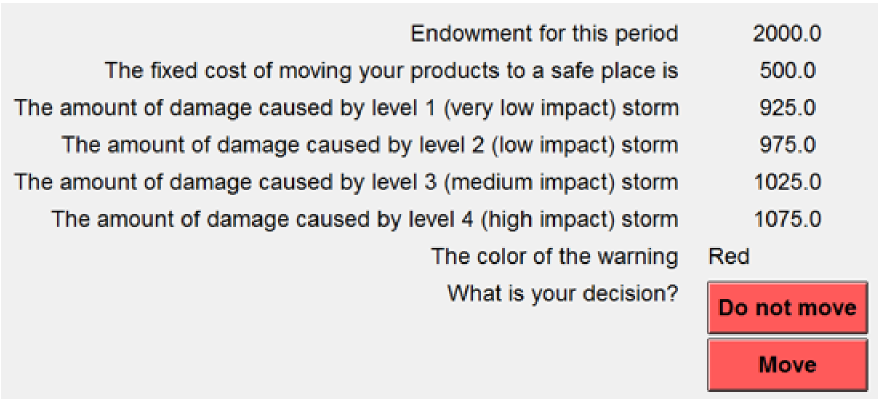


(Treatment 3)

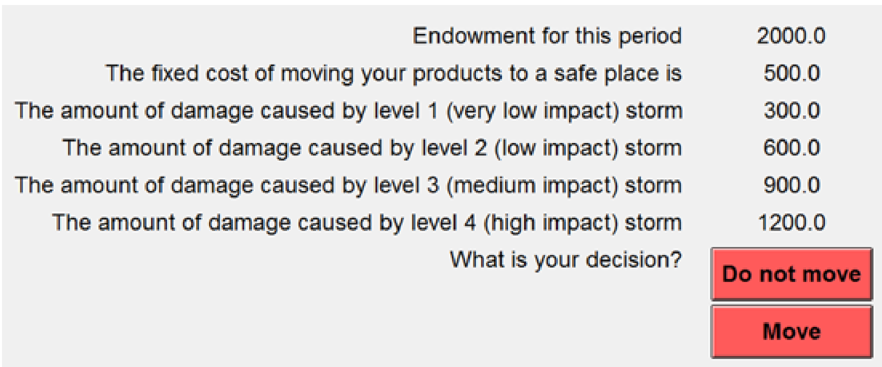
Figure 2.1: The form of the warnings presented in Treatments 1, 2 and 3. In the matrix, likelihoods were 20%, 40%, 60%, and 80% (this information is not available to participants) and the impact level corresponded from 1 to 4.



(Treatment 1)



(Treatment 2)



(Treatment 3)

Figure 2.2: Payoff information shown to participants in Treatments 1, 2 and 3. (The payoff function and probability of each payoff type is consistent through all treatments.)

Endowment for this period	2000.0
The level of the storm	level 3 (medium impacct)
The storm	Did not occur
Your decision was	Not move
Your profit from this round will be	2000(endowment)
Your payoff for this period	2000.0

Figure 2.3: End of period results presented to every participant. (If a storm did not occur it may be that the storm occurred elsewhere and thus did not affect the decision maker.)

3. Decisions with Weather Warnings when Waiting is an Option

3.1 Introduction

There are many research papers that focus on improving the decision-making process during an extreme weather event (e.g., Roulston et al., 2006, Joslyn and Nichols, 2009, Joslyn and Savelli, 2010, Joslyn and LeClerc, 2012, Mu, Kaplan, and Dankers, 2018, and Grounds and Joslyn, 2018). In this paper, we have focused on how to reduce harm by finding the optimal decision time

There are many advantages to acting at an earlier time. The cost of action is lower when the public acts earlier, such as the cost of evacuating to a safe location. But there are also advantages to waiting: The reliability of the warning increases over time, and people can make decisions based on better information. When the likelihood and magnitude of the effects of a natural disaster are uncertain, the weather bureau must decide when to issue a weather warning and with what degree of reliability.

There is also research on decision making time and the cost of waiting. One of the basic theories of waiting states that people gather new information until the cost of new information is higher than the expected value of the information, (Wald, 1945). However, people are not always rational when given the opportunity to wait and gather new information. Descamps, Massoni, and Page (2022) designs a game where participants must guess the correct combination of balls, during the experiment participants can wait for more clues or guess based on the information they already have. They find that people gather too much information when they acutely need less information, and gather too little information when they need more information. In weather forecasting, new information arrives naturally, so waiting for more information always seems to be the better strategy, but we must keep in mind that the cost of actions can increase significantly in a very short time. This increase means that the potential cost of waiting for more information is not free during extreme weather events. For this reason, in this chapter, I want to focus on how individuals balance between acting earlier, but potentially wasting time and money, and acting late, but potentially causing more damage.

In this chapter, we use laboratory experiments to study decisions with a two-time stage weather warning system that varies in the reliability of early warnings. The warning system is based on one used by the Met office since 2011: a risk matrix where the rows and columns represent probability levels and impact levels, respectively. Participants can choose between a safe but costly option and a risky but free option. They can make their decisions based on an unreliable early warning or wait for a more reliable warning, however, where waiting makes the safe option more expensive. Overall, our work shows that increasing the reliability of the early warning generally leads participants to wait less and, more surprisingly, that contrary to theory, the reliability of the early warning has a significant impact on the decision they make after waiting and receiving the more reliable.

Disasters caused by natural hazards have caused significant damage throughout history. Many can be predicted in advance to some extent. If accurately warned far enough in advance, people can ameliorate much of the negative impact. In 2021, mass flooding occurred in Germany and Belgium, causing 180 and 38 deaths, respectively. The recriminations continued into the autumn in Belgium with debates about dam management and even an investigation into the possibility of involuntary manslaughter charges over authorities' alleged failures to react to alerts issued by the EU's flood warning system. This is why we want to investigate how to improve the effectiveness of warning reactions (Tafel et al., 2021).

There are many uncertainties in a natural disaster, and the potential impact of the disaster can significantly increase in a very short time span. The flood investigation report of Braithwaite (Council (2015)) focuses on the village of Braithwaite, situated around Coledale Beck. The section of Coledale Beck upstream of the Coledale High Bridge is heavily vegetated with large trees, growing within proximity of the river channel. During the flood in December 2015 many of these trees were washed out into the watercourses, causing multiple blockages and resulting in a wider impact of flooding on the village. Based on the technology we have now, it is impossible to predict the exact timing when trees will be washed out and what will happen when large trees hit a bridge. Therefore, we must accept the idea that (1) we cannot have perfectly accurate warnings, and (2) the impact and likelihood of a disaster might quickly change over a short duration of time.

A lot of work at the Met Office is aimed at estimating the uncertainty in an evolving weather situation. Nonlinear prediction (NLP) ((Porporato and Ridolfi, 1997), (Islam and Sivakumar, 2002)), nonlinear time series approaches including hidden Markov models (HMMs) (Ayewah, 2003), and artificial neural networks (ANNs) ((Hopfield, 1988),(Yegnanarayana, 2009)) have been applied to forecasts. In these models the reliability of the predictions can be increased with more data collected from the past, and the reliability decreases if we want to predict something further away from the present.

We can now issue warnings before the event; for example, the unusual path and intensification of Hurricane Sandy in October 2012 was predicted 8 days ahead, the 2010 Russian heatwave and the 2013 US cold spell were forecast 1—2 weeks ahead, and tropical sea surface temperature variability following the El Ninõ/Southern Oscillation phenomenon can be predicted 3—4 months in advance (Bauer, Thorpe, and Brunet, 2015).

The current time limit for small-scale events is between hours and days, for accurate and reliable prediction of high-impact weather events about 1—2 weeks, for prediction of large-scale weather patterns and regime transitions about a month, and for global circulation anomalies about a season (Hoskins, 2013). In most situations the reliability increases with time.

There are many things that everyone can do to reduce potential damage. Cox, Losee, and Webster (2022) has proves that the weather damage can be reduce by appropriate preparation. For instance, when a flood is coming, one can stockpile emergency building materials to prevent leaking, assemble supplies in case the electricity is cut off, and gather water and food that does not need to be cooked or kept in a refrigerator. A better solution is to stop excessive sediment from entering a stream by maintaining roads, streambanks and surfaces, and by maintaining vegetation buffers that soak up runoff during storm events. Educating one's family and preparing one's home can help reduce water damage (Kreibich et al., 2005).

Disaster risk reduction efforts traditionally focus on long-term preventative measures or post-disaster response. Outside of these, there are many short-term actions, such as evacuation, that can be implemented in the period of time between a warning and a potential disaster to reduce the risk of impacts. However, this precious window of opportunity is regularly overlooked in the case of climate and weather forecasts (Perez et al., 2015) as these actions can be costly and time-consuming, which means following an early warning with low reliability can be a waste of money and time. On the other hand, most of the actions, such as preparing supplies, are always cheaper if people choose to follow them earlier, and maintaining and educating always work better if people do them earlier.

To improve the effectiveness of early warning systems, we need to better understand to what extent people are willing to take action based on early warnings that may come with higher uncertainty, or whether they prefer to wait for more certainty at the expense of higher costs and/or potential damage, and what factors influence this preference.

3.2 Model and theory

3.2.1 Model

We analyze how participants behave before a natural disaster through an individual decision model with an option of waiting. We assume a scenario in which a flood is coming, and to simplify the scenario we assume there will only be two outcomes: the flood happens, or does not happen. Every participant is assigned the role of a manager of a warehouse. As the manager they need to make decisions between moving their goods to a safe place or doing nothing. We also want to analyze how participants behave when they have a chance to wait for information more likely to be correct. We build a scenario with two stages. In stage 1, participants will have a warning with low reliability. Participants can choose to move at stage 1 or do nothing and wait for better information, then make a decision based upon the better information at stage 2. The participants' decision problem is an individual choice problem, that is, each participant's payoff is not affected by the decisions of other participants.

We use a colored 4-by-4 risk-likelihood matrix to present the warning (see Figure 3.2). There are four different impact levels on the x-axis and four different likelihood levels on the y-axis; the impact level shows the damage if the flood happens, and the likelihood level shows the probability that the flood happens. There are 16 boxes in the matrix in total and these boxes are colored with four different colors: green, yellow, amber and red. Each box is colored with only one color, so the warning color is the possible range of the box. There are 6 green boxes, 5 yellow boxes, 4 amber boxes and 1 red box. As we want to investigate the significance of waiting for information more likely to be right, we assume at stage 1 participants will have the right warning color. The warning color gives participants some information that they can use to make a decision, but the likelihood and impact may not be correctly shown. All the information in stage 2, the warning color, likelihood and impact level, are all accurately presented. Therefore, we introduce a new variable accuracy level (A) to describe the probability that the likelihood and impact level are correctly presented. A higher accuracy level means the warning in stage 1 is more likely to be the final warning. A lower accuracy level means the warning in stage 1 is less likely to be the final warning.

The information in stage 2 is the most accurate but there is a potential cost for waiting. At the beginning of each round, each participant is given I endowment. They need to decide whether to move with cost C_1 or wait with no cost at stage 1; they cannot change their decision if they choose to move at stage 1, but if they choose to wait, they can choose to do nothing with no cost or move with cost C_2 at stage 2. However, if the flood occurs then participants who choose to do nothing will suffer damage (D); the damage depends on the impact level and the probability of occurrence depends on the likelihood level. The moving cost at stage 2 is always higher than the moving cost at stage 1, as the cost of action always increases if you want to do it in a very short period; the difference between the moving cost in stage 1 and stage 2 can also be thought of as a cost of waiting. The payoff for each decision is shown in Figure 3.1.

The payoffs for the participants in each round were determined as follows based upon the decision and outcome of the storm:

1. Endowment minus moving cost at stage 1 [if decided to move at stage 1]

2. Endowment minus moving cost at stage 2 [if decided NOT to move at stage 1 and move at stage 2]

3. Endowment [decided not to move at both stages and the storm did not occur]

4. Endowment minus damage from the storm [decided not to move at both stages and the storm occurred]

3.2.2 Equilibrium Strategies

In the experiment, participants need to make decisions at different stages; here we use backward induction to analyze their optimal decision-making process. We assume the participants are risk neutral and always choose the decision with higher expected payoff. We try to find the **higher expected payoff decision (HEPD)**

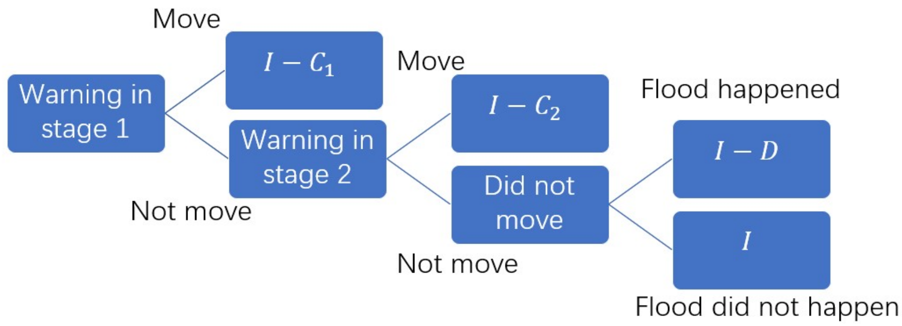


Figure 3.1: The payoffs in a decision tree.

based upon the information that participants have at the time of making the decision.

To find the HEPD in stage 1, we need to start with analyzing the decision-making process when participants choose to wait at stage 1 and have the correct warning at stage 2, and then compare the expected payoff of choosing to wait at stage 1 and the moving cost at stage 1.

As participants have the correct impact level and likelihood level, there are two possible scenarios: the expected damage based on the warning at stage 2 is higher than the moving cost at stage 2 ($D(W_2) > C_2$) or less and equal to the moving cost at stage 2 ($D(W_2) \leq C_2$). When ($D(W_2) > C_2$) participants should choose to move at stage 2, otherwise participants should choose not to move. Based on the calculation above we can define the HEPD at stage 2.

Although the warning at stage 1 might be different to the warning at stage 2, the color of the warning will remain the same. We thus need to consider the expected damage for every possible warning in the region. When the decision with higher payoff is not to move at stage 2, the expression $E[D(W_2) \mid D(W_2) < C_2\&color]$ represents the expected damage over all the possible warnings in one color region when the expected damage is lower than C_2 based on the warning at stage 2. The expression $P(D(W_2) < C_2 \mid color)$ represents the probability over all the possible warnings in one color region when the expected damage is lower than C_2 based on the warning at stage 2. For simplicity let us denote it as $P_{nm}(color)$ where nm is short for **not move**.

If the stage 1 warning is not correct, the probability for each box is identical.

The expected damage in this color region when the warning is not correct is the probability that the expected damage of the warning at stage 2 is lower than the move cost at stage 2 times the expected damage plus the probability that the expected damage of warning at stage 2 is higher than the move cost times the move cost at stage 2. Formally, this is:

$$S_2(color) = E[D(W_2) \mid D(W_2) < C_2 \& color] * P_{nm}(color) + C_2 * (1 - P_{nm}(color)) \tag{3.1}$$

We denote the expression as $S_2(color)$, which stands for the spending at stage 2 which is equal to the cost of not moving at stage 1 given the color. We also note that

$$S_2(color) = E[min[D(W_2), C_2] \mid color] \tag{3.2}$$

where the expectation is over possible stage 2 warnings given the color. We denote color (W) as the color of the warning.

Therefore, the expectation of the spending at stage 2 given the warning at stage 1, $S_2(W_1)$, is based upon the warning at stage 1 and is equal to the probability that warning 1 is accurate denoted by A (accuracy level) times the damage plus the expected damage when it is not correct $(S_2(color(W_1))$ times the probability $(1 - A)$:

$$S_2(W_1) = A \cdot min\{D(W_1), C_2\} + (1 - A) \cdot S_2(color(W_1)). \tag{3.3}$$

Hence, participants will compare the expected cost of choosing not to move in stage 1 based on the warning at stage 1 $(S_2(W_1))$ with the moving cost at stage 1 C_1; if $C_1 < S_2(W_1)$ then they should choose to move at stage 1, and if $C_1 \geq S_2(W_1)$ then they should choose not to move at stage 1. Therefore, we can define the HEPD at stage 1 based on the warning at stage 1.

We also find that when participants choose not to move at stage 1 and make the decision again at stage 2, they only need to compare the expected cost of choosing not to move at stage 2 $(D(W_2))$ and the moving cost at stage 2 (C_2), both of which are independent of any additional information contained in the stage 1 warning. Hence, when participants are making decisions at stage 2, they do not need to consider any information in stage 1 and they know this when they are making decisions at stage 1.

3.2.3 Experiment design

In this experiment, participants were asked over several rounds to make decisions when receiving flood warnings. In each round, they were asked whether to move products into a warehouse. Each round had two stages with a flood warning in each stage. Both warnings were about the same flood with matrices showing the impact levels in columns and likelihood levels in rows. All the boxes in the matrices were colored green, yellow, amber or red, indicating the expectation of the damage. There was a check mark in the matrix with the predicted likelihood level and impact level. It was a 4-by-4 matrix with four impact levels, 1, 2, 3 and 4, increasing from left to right. There were four likelihood levels, 20%, 40%, 60% and 80%, with the likelihood level increasing from the bottom to the top. Participants were not told the exact likelihood level during the experiment.

The warning color in each round is always correct, but the check mark may not be in the correct box of that color. The probability of the check mark being in the right box depends upon the accuracy level. If the accuracy level is 1 (max), the check mark is 100% certain to be in the right box. If the accuracy level is 0 (min), the check mark carries no additional information over the color of the warning in that the probability for each box of that color is the same. The participants were told the accuracy level of each warning, out of low, medium, and high, and the moving cost involved with each impact level, but were not told what the likelihood level corresponded to.

Participants were told before deciding at stage 1 that the moving cost at stage 2 (750) is always higher than the moving cost at stage 1 (either 400, 550, or 700); the color of the warning is always correct at both stages, and the warning at stage 2 is always 100% correct in the warning color and the check mark. If they chose to move their product at stage 1, they could not change their decision in stage 2. Each impact level indicates a different level of damage: level 1 is 600, level 2 is 800, level 3 is

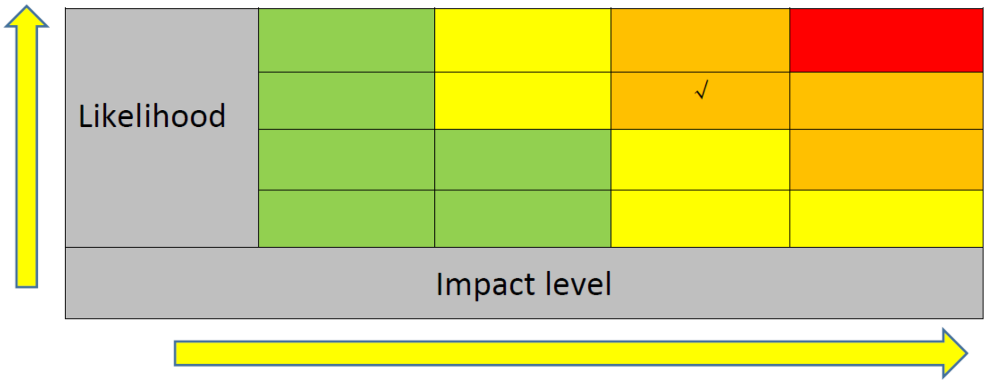


Figure 3.2: The warning matrix.

1000 and level 4 is 1200. Each likelihood level indicates a different level of likelihood: level 1 is 20%, level 2 is 40%, level 3 is 60% and level 4 is 80%. Participants did not have the exact amount of damage and likelihood they needed to learn what the real likelihood and impact level were.

In Treatment 1, the accuracy level information was presented as low accuracy level (25%), medium accuracy level (50%) and high accuracy level (75%). Participants needed to learn what the real accuracy level was. As there was only one red box, participants knew the exact likelihood and impact level. For participants we assumed they had a warning with 100% accuracy when they had the red warning. In Treatment 2, participants knew the warning color at stage 1 and the exact likelihood and impact level at stage 2. In other words, the accuracy level was 0%. The difference is that the participants in Treatment 2 did not have any inaccurate information. In other words, the participants in Treatment 1 had an inaccurate check mark at stage 1 and a correct check mark at stage 2, but the participants in Treatment 2 had an accurate warning color at stage 1 and an accurate check mark at stage 2. The information that participants 2 received was always reliable, which means they did not have the idea of reliability.

Participants were told the information below:

In stage 1, participants were told the endowment for this period, the moving cost in stage 1, the moving cost in stage 2, the damage for each impact level, the accuracy level (only for participants in Treatment 1) and information that the flood may happen only after stage 2. In stage 2, participants were told the moving cost in stage 2 and the damage for each impact level. After they finished the decision-making process in stages 1 and 2, they were given the payoff information, the initial endowment, the level of the storm, participants' decisions, whether the flood happened, how the payment was calculated, and the payoff in this period.

3.3 Experimental procedure

The participants in this study were 161 undergraduate students from the University of Exeter consisting of 53 year one students, 60 year two students, and 48 year three students. There were 100 female participants and 61 male participants. The potential payment range was £5 to £10, while the actual payments varied from £7.80 to £9.65 with the average payment £8.75.

It took participants on average 5 minutes to read the instructions. Afterwards, there were test questions. Participants that incorrectly answered any questions needed to wait for an explanation before starting. There were 90 questions in Treatment 1 consisting of 30 questions for each accuracy level of 25%, 50% and 75%. In the experiment they were presented as low accuracy level (25%), medium accuracy level (50%) and high accuracy level (75%). Participants needed to learn what the real accuracy level is. In Treatment 2, participants only knew the warning color at stage 1 and found out the accuracy, likelihood and impact level at stage 2. In other words, the accuracy level was 0%. (Participants were not informed of the accuracy level.) The probability for all three stage 1 moving costs was the same (participants did not know the probability).

As for green and red warnings, the decisions are more straightforward, so we had a lower percentage of those questions. Overall, the percentages for green, yellow, amber and red warnings were 10%, 40%, 40% and 10%, respectively (participants were not informed of these percentages).

The order of each type of question was randomly generated by the computer and different for each participant. There were 90 questions in Treatment 2 as well. For all questions in Treatment 2, the accuracy level was 0% and there was no check mark, just the warning color. Participants could choose again after seeing the correct warning in stage 2 if they chose not to move at stage 1, or by moving they could just see the actual warning in stage 2 and jump to the result part. In the result stage, there was information about whether the flood happened, the impact level, what the damage was and the tokens they had for this question.

When participants were making decisions in stage 2, as the warning correctly presented the likelihood level and impact level, they did not need to make their decisions based on any information they had in stage 1; in other words, the information in stage 1 could not affect the decision in the next stage.

3.4 Hypothesis

In this section, we propose several hypotheses.

Hypothesis 3.1 *The warning color itself conveys a message, so participants may choose to move when the warning color is yellow or red, even if the decision to move does not yield the highest payoff.*

Generally people are risk averse and the warning color is determined by the impact level and the likelihood level. There has been research shows that the warning color especially the red warning significantly affect people's risk preference (Leonard, 1999). Hence, the warning color itself provides no additional information. However, the color does give some indication of a possible action. Green and yellow warnings tend to indicate a sign of safety, while amber and red warnings indicate a sign of danger. Therefore, we can expected that the warning color may cause biased behavior; namely, participants choose to move more often than they are supposed to when the warning color is amber or red, and less often than they are supposed to when the warning color is green or yellow. We plan to compare the proportion of participants who choose the decision with the highest gain between different warning colors.

Hypothesis 3.2 *Warning color should also affect participants' decision time. A higher warning color may induce participants to move earlier than necessary.*

It is more general that the decision making time will be effected by the warning color. There has been research shows color of the report can affect participants' decision making time (Benbasat, Dexter, and Todd, 1986). As the red warning is a general signal of danger for human being (Leonard, 1999). We can expect participants to choose to move more often in stage 1 when they have a higher warning color level warning. In other words, the warning color can encourage participants to make a decision quicker.

As a higher warning color level means higher expected cost, this means choosing to move is more likely to be the HEPD when the warning color level is higher. Hence, to test if there is a bias we need to compare the percentage of moving in stage 1 over all the participants who choose to move eventually. It should be higher when the warning color is higher.

Hypothesis 3.3 *A lower reliability of a warning may encourage participants to move later, even if the decision to move earlier has a higher expected payoff.*

The reliability of the warning tells people how likely it is that the warning they have is correctly presented in the earlier stage. There has been research shows when reliability of the information is low, people tend to choose to wait(Banbury et al., 1998). We expect that the higher the reliability of the warning they have, the more likely they are to make their decision at stage 1 as they already have the information they need to decide. We plan to test the proportion of those who choose to move in Phase 1 compared to participants who choose to move, given warnings with different levels of reliability.

Hypothesis 3.4 *The effect of the accuracy level depends on whether the expected damage is more likely to increase at stage 2 or decrease at stage 2 if the check mark changes position within the color.*

As the warning color is always correctly presented in stage 1, the check mark they have at stage 1 may be in the lower-damage or the higher-damage part of the region. If the check mark they have is in the lower-damage part, the expected damage at the next stage decreases with the accuracy level; a low accuracy level means participants are more likely to have a higher expected damage warning at stage 2, and a high accuracy warning means the warning is more likely to remain the same at stage 2. Based on the calculation in Section 3.2.2 we can expect that when the warning is in the lower-expected-damage part, participants more often choose to stay if the reliability increases as the expected damage is not likely to increase. For the same reason, participants should choose to move more often if the check mark is in the higher-expected-damage region and the accuracy level is high. We will test participants' decision with different levels of warning accuracy in Stage 1.

Hypothesis 3.5 *The decision in Stage 2 should be independent of the warning in Stage 1, but there may be a survivorship bias.*

As participants at stage 2 always have the warning correctly presented, the accuracy level should not have any effect on the participants' behavior at stage 2. However, there is also a potential for a survivorship bias. There are three types of participants: always choose to move, always choose not to move and choosing between moving and not moving. A low-accuracy warning at stage 1 will encourage participants choosing to wait (Banbury et al., 1998). This means that the proportion of those always choosing not to move increases at stage 2. Participants should less often choose to move at stage 2 if they have a higher accuracy level warning. We will test participants decision with different accuracy level warning in stage 1.

3.5 Results

3.5.1 Warning color and moving cost

The warning color is the key piece of information whose influence on behavior we want to investigate. There are four warning colors: green, yellow, amber and red. We assign a warning color level from 1 to 4 as $Green < Yellow < Amber < Red$ where 1, assigned to Green, is the lowest warning color level and 4, assigned to Red, is the highest warning color level.

Firstly, we want to investigate how the warning color affects participants' decision making in stage 1, stage 2 and the decision-making process overall. We do so by looking at the frequency with which participants choose the decision with higher expected payoff (HEPD).

We use a fixed-effects logistic model to examine the effect of color on decision making:

$$Decision = c_1 colorlevel + U_i + \varepsilon$$

$$HEPD = b_1 colorlevel + U_i + \varepsilon$$

The results of fixed-effects logistic regression for the decision in stage 1, HEPD in stage 1, decision in stage 2 and HEPD in stage 2 are summarized in table 3.1 below. Decision is 0 if participants choose to wait at this stage, and 1 if participants choose to move at this stage. HEPD is 1 if a participant chooses the decision with higher expected payoff based on the information they have at this point, and 0 if they choose the decision with lower expected payoff. For decision in stage 2 and HEPD in stage 2 we only use the data when participants choose to wait at stage 1. After we analyze the behavior at stages 1 and 2, we need to analyze how the participants behave overall. We define three new variables; decision overall is equal to 1 if participants choose to move eventually, including moving at stage 1 and moving at stage 2, and it is 0 if they ultimately choose to stay.

Table 3.1: Fixed-effects logistic regression where color is the independent variable.

	Decision of moving[1]			Choosing HEPD[2]	
	in stage 1	in stage 2	eventually	in stage 1	in stage 2
color level	1.78*** (0.0329)	1.44*** (0.0581)	1.94*** (0.0352)	-0.010 (0.0352)	-1.21*** (0.055)
Number of observations	14651	14651	6588	6853	14651
Number of subjects	161	161	149	154	161

Note: Standard errors in parentheses. *** indicate $p < 0.01$ $**$ indicate $0.05 > p > 0.01$.
1: Value of decision is 1 if participants choose to move, and 0 if they choose to stay.
2: Value is 1 if participants' decision is same with HEPD, and 0 if it is not the same.

We are also interested in how the moving cost affects participants' behavior in stage 1, stage 2 and decision overall, as the moving cost in stage 2 is the same for every participant and they know the moving cost for both stages in stage 2, so we only run the regression for the moving cost in stage 1. We use a fixed-effects logistic model to examine the effect of moving cost in stage 1 on decision making; the results are summarized in Table 3.2 below, where the dependent variables are the same as in Table 3.1:

Table 3.2: Fixed-effects logistic regression where moving cost in stage 1 is the independent variable.

	Decision of moving[1]			Choosing HEPD[2]	
	in stage 1	in stage 2	eventually	in stage 1	in stage 2
moving cost in stage 1	-0.24***	0.109***	-0.175***	-0.084***	-0.108***
unit(100)	(0.0142)	(0.0264)	(0.0142)	(0.01553)	(0.0269)
Number of observations	14651	14651	6588	6853	14651
Number of subjects	161	161	149	154	161

Note: Standard errors in parentheses. *** indicate $p < 0.01$ $**$ indicate $0.05 > p > 0.01$.
1: Value of decision is 1 if participants choose to move, and 0 if they choose to stay.
2: Value is 1 if participants' decision is same with HEPD, and 0 if it is not the same.

It can be seen in the regressions from Table 3.1 that the coefficient of color is positive for decision in stage 1, decision in stage 2 and decision overall, which means that participants choose to move more often if they have a higher color level warning. It can also be seen from Figure 3.3 (decision is 0 if participants choose to stay, 1 if they choose to move; HEPD is 1 if participants choose the decision with higher expected payoff, 0 otherwise); the proportion of people that choose to move (blue bar) significantly increases with the color warning level.

Result 3.1 *In stage 1, participants more often choose the HEPD when the warning color is green or red than when the color is yellow or amber.*

While it can be seen from Table 3.1 that there is no significant linear relationship between HEPD in stage 1 and the warning color, the proportion of participants that choose the HEPD does vary in a U shape with the warning color level as seen in Figure 3.3. This can explain why the coefficient of the regression did not show a significant monotonic relationship between the warning color and frequency of choosing HEPD in stage 1.

Result 3.2 *In stage 2, participants less often choose the HEPD when the warning color is higher.*

In Table 3.1, the coefficient of color level for HEPD is significantly negative, which means participants less often choose the HEPD when the warning color level

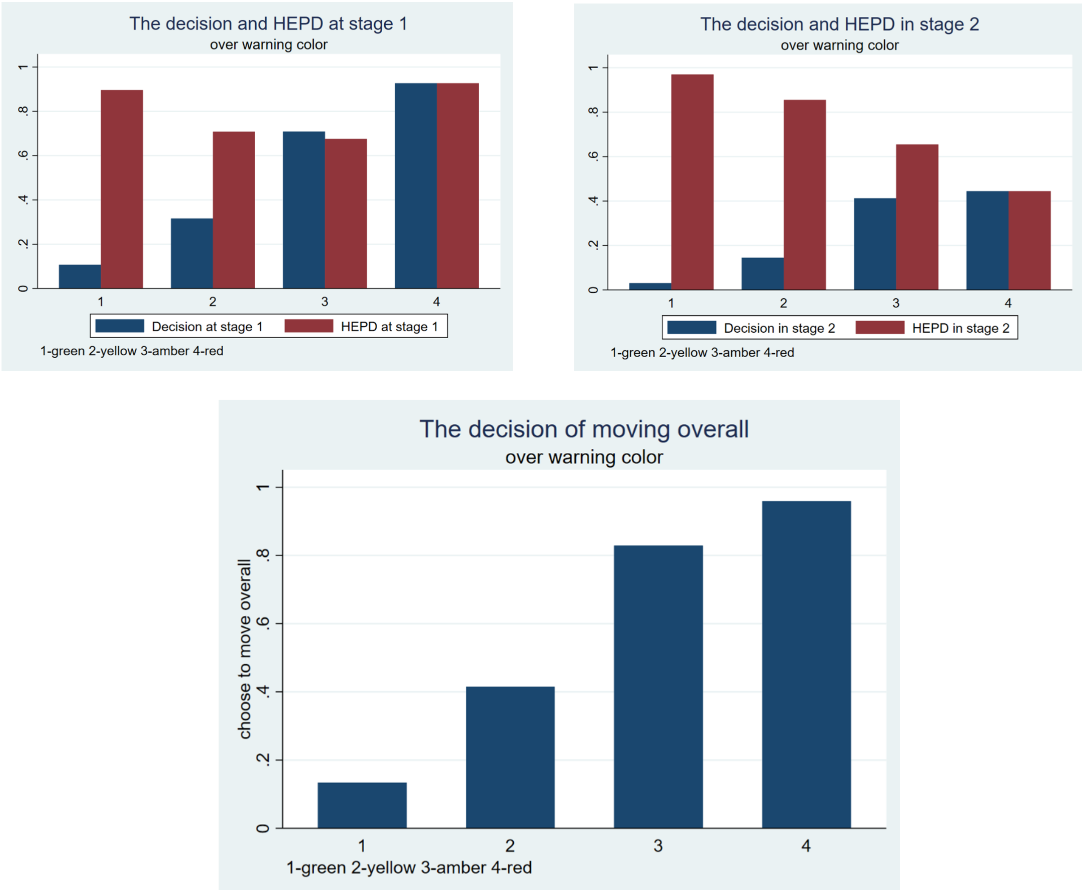


Figure 3.3: Decision and HEPD at stage 1 and at stage 2

is high. Figure 3.3 also clearly shows that the proportion of participants choosing HEPD decreases continuously with the warning color level. For the red warning, unlike the behavior at stage 1, the proportion of those choosing the HEPD does not increase when the warning color is red.

The HEPD should more often be moving when the warning color is high. As participants can choose to move in stage 1, most of the risk-averse and risk-conscious participants should already choose to move in stage 1. That is why participants choosing to stay and make a decision again in stage 2 are more risk-loving than average. This explains why participants less often choose the HEPD in stage 2 as the warning color increases. It is also interesting that over 40% of participants who had a red warning at stage 1 changed their mind, as there is only one red box which means participants will not have any extra information for waiting. Participants know the likelihood level and impact level at stage 1, which is understandable for some of the participants who are risk-loving, but it is pointless to change their decision at stage 2 as the cost of moving is always higher at stage 2 and they know this at stage 1. The only explanation is that some people just want to wait when they have the chance to wait.

Result 3.3 *For participants who choose to move eventually, participants more often choose to move in stage 1 if the warning color level is high.*

This result is consistent with Hypothesis 3.2. Figure 3.3 shows the participants who eventually choose to move. In Figure 3.3, it is 1 if participants choose to move in stage 1 and 0 if they choose to move in stage 2; the participants' choice of stage is not included. Participants more often choose to move in stage 1 when the warning color is high. This result can prove that a higher warning color level can increase the response time of the disaster.

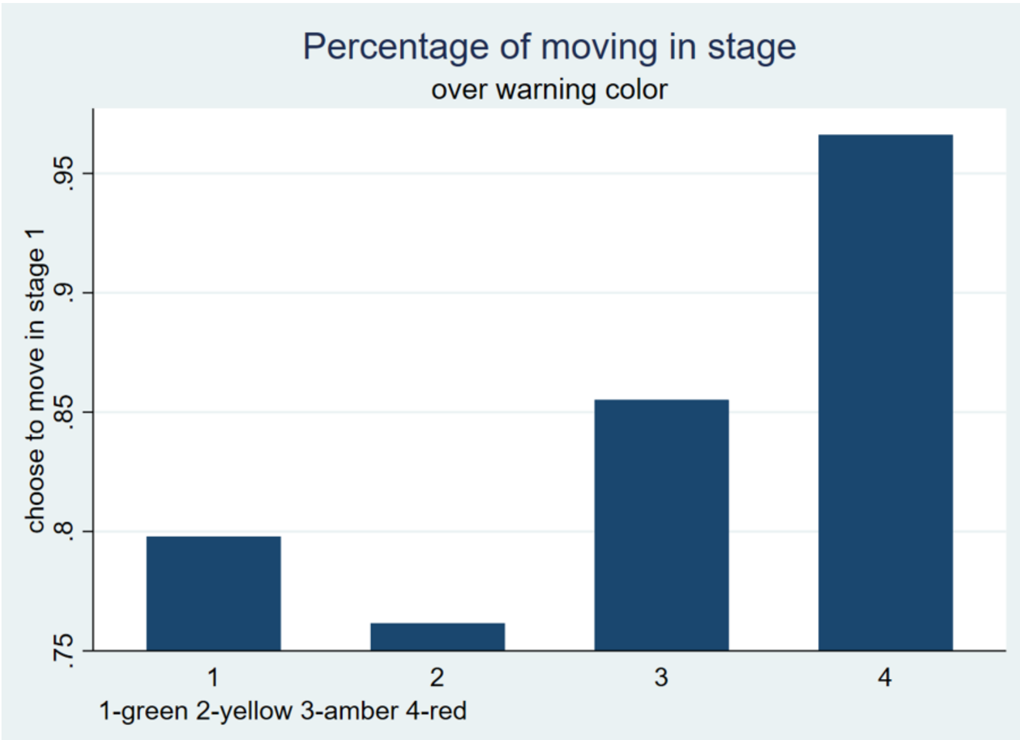


Figure 3.4: Percentage of participants choosing to move in stage 1 out of all the participants who choose to move eventually.

Result 3.4 *In stage 1, when choosing to stay is the HEPD, participants less often choose the HEPD the higher the warning color level. When choosing to move is the HEPD, participants more often choose the HEPD the higher the warning color level.*

This result is consistent with Hypothesis 3.1. Figure 3.4 shows that the warning color has a significant effect on the probability of participants choosing the HEPD. In other words, when the warning color is low, participants more often choose to stay even if the HEPD is to move, and when the warning color increases participants more often choose to move even if the HEPD is to stay. This explains why the proportion of people choosing the HEPD decreases with warning color and increases when the warning color is red.

Participants more often to choose to move which is the HEPD and reduces the probability of deciding again which increases the potential for making mistakes.

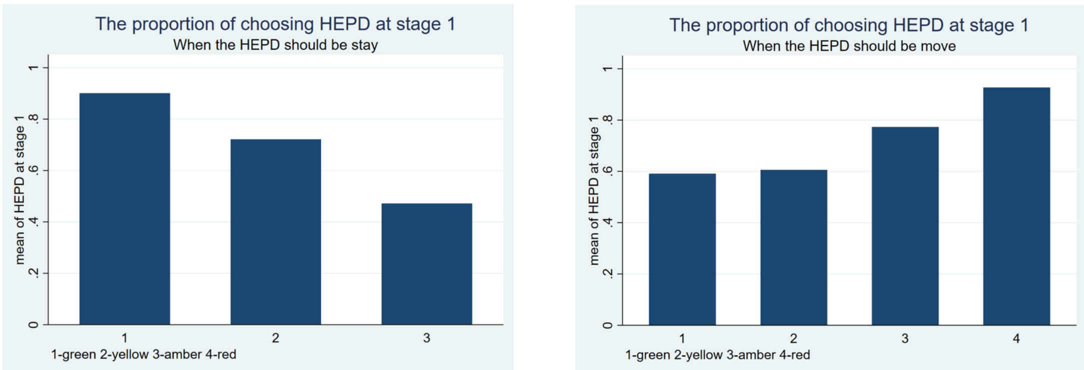


Figure 3.5: Proportions of participants choosing the HEPD when the HEPD decision is to stay and to move.

Result 3.5 *A higher moving cost in stage 1 makes people more likely to wait in stage 1 and increases the probability of moving in stage 2 (given that they haven't moved in stage 1).*

It can be seen from Table 3.2 that the moving cost in stage 1 significantly decreases the probability of choosing to move in stage 1, which means that participants more often choose to wait in stage 1 if the moving cost in stage 1 is high. As a high moving cost in stage 1 makes more participants choose to wait in stage 1, it significantly increases the proportion of participants who choose to move in stage 2. It can also be seen that a higher moving cost has a negative effect on the HEPD in both stage 1 and stage 2, which means a higher moving cost always decreases the chance of choosing the lower-expected-payoff decision in each stage.

Overall, participants less often choose to move eventually, which means if the moving cost in the earlier stage increases, it will decrease the probability of taking action for all the participants.

3.5.2 Accuracy level with different current situation and future expectation

We want to test how information with different accuracy levels affects participants' behavior. In this experiment, the warning color was always correct, but the exact box checked from all those boxes with that color was not necessarily correct. For instance, a yellow warning can be any one of five boxes. The accuracy level is defined as the probability of the impact level and likelihood level being 'correct' in stage 1. If the warning is not correct, there is an equal chance of it being any box in the color region. Note that this includes the initial box as well. The accuracy levels are 75%, 50%, 25% (Treatment 1) and 0% (Treatment 2). Since there is only a single box with a red warning, a participant knows with certainty which box will be chosen even though the red warning technically cannot vary in the color region. We also care about the expected payoff based on decision, which is the expected payoff based on the decision participants make; the expected payoff overall is the initial endowment minus the moving cost at each stage if they choose to move at stage 1 or stage 2, or if they choose to stay eventually it is the initial endowment minus the likelihood times the damage for the corresponding impact level at stage 2. Then we compare it with the expected payoff based on the HEPD and that based on non-HEPD; these are the expected payoffs if participants always choose the HEPD and if participants always choose the opposite decision of the HEPD.

As the accuracy levels present how likely participants are to have the correct warning, we also need to consider the situation when they do not have the correct warning in stage 1. As participants always have the right warning color (and they know this), they will have expectation of the situation if the box with the color is not correct. There are three possible future situations for the warning in stage 1.

Worse at present with better future (Situation = 1): if the expected damage of the warning participants receive in stage 1 is lower than the average expected damage in this color region, then the situation is more likely to be better in stage 2. In this situation, a higher accuracy level means higher expected damage in the future.

Better at present with worse future ($Situation = -1$): if the expected damage of the warning participants receive in stage 1 is higher than the average expected damage in this color region, then the situation is more likely to be worse in stage 2. In this situation, a higher accuracy level means lower expected damage in the future.

Same future ($Situation = 0$): if the expected damage of the warning in stage 1 is same as the average expected damage in the color region, then there is no change independent of accuracy levels. We also note this is the case when the warning is red since there is only one red box. This is also the case in Treatment 2. There is no accuracy level in this situation.

To verify the relationship, we run fixed-effects logistic regression for the role of accuracy level with decision and HEPD first:

$$Decision = c_1 accuracylevel + U_i + \varepsilon$$

$$HEPD = b_1 accuracylevel + U_i + \varepsilon$$

Table 3.3: Fixed-effects logistic regression of the role of accuracy level

	Decision of moving[1]			Choosing HEPD[2]	
	in stage 1	in stage 2	eventually	in stage 1	in stage 2
Accuracy level	1.227***	-1.337***	0.621***	0.0922	1.167***
	(0.101)	(0.201)	(0.101)	(0.109)	(0.204)
Number of observations	9996	9996	4744	4829	9996
Number of subjects	122	122	113	115	122

Note: Standard errors in parentheses. *** indicates $p < 0.01$, ** indicates $0.05 > p > 0.01$.
1: Value of decision is 1 if participants choose to move, and 0 if they choose to stay.
2: Value is 1 if participants' decision is same with HEPD, and 0 if it is not the same.

Then we run fixed-effects logistic regression for the role of accuracy level and situation with decision and HEPD:

$$Decision = c_1 accuracylevel + c_2 situation + c_3 accuracylevel * situation + U_i + \varepsilon$$

$$HEPD = b_1 accuracylevel + b_2 situation + b_3 accuracylevel * situation + U_i + \varepsilon$$

As there is no accuracy level when the warning is red and in Treatment 2, we only investigate the relation when the situations are worse future and better future (no red warning and participants in Treatment 2). Table 3.2 and Table 3.3 summarize the results; the definitions of the dependent variables are the same as those in Table 3.1.

Table 3.4: Fixed-effects logistic regression of the role of accuracy and situation which is the location in the color region for better and worse future situations.

	Decision of moving[1]			Choosing HEPD[2]	
	in stage 1	in stage 2	eventually	in stage 1	in stage 2
Accuracy level	1.258***	-1.261***	0.644***	0.0383	1.097***
	(0.104)	(0.201)	(0.103)	(0.111)	(0.205)
Situation	0.157***	-0.138	0.056	-0.274***	0.119
	(0.056)	(0.097)	(0.111)	(0.060)	(0.099)
Accuracy level	0.482***	0.553***	0.622***	0.354***	-0.543***
*Situation	(0.104)	(0.202)	(0.104)	(0.111)	(0.205)
Number of observations	9996	9996	4744	4829	9996
Number of subjects	122	122	113	115	122

Note: Standard errors in parentheses. *** indicates $p < 0.01$, ** indicates $0.05 > p > 0.01$.
1: Value of decision is 1 if participants choose to move, and 0 if they choose to stay.
2: Value is 1 if participants' decision is same with HEPD, and 0 if it is not the same.

Result 3.6 *In stage 1 participants choose to move more when the accuracy level is high, regardless of the HEPD and situation.*

This result is consistent with Hypothesis 3.3. It can be seen in Table 3.4 that the coefficient of accuracy level is significantly positive and the accuracy level $\times$ situation is also significantly positive, and the absolute value of the coefficient of the accuracy level $\times$ situation is smaller than the absolute value of the coefficient of the accuracy level, which means participants choose to move more when the accuracy level is high.

In Figure 3.6 decision is 0 if participants choose to stay, and 1 if participants choose to move. It can be seen from Figure 3.6 that the percentage of participants choosing to move significantly increases with the accuracy level; at the same time, the accuracy level does not have a significant effect on the proportion of moving being the HEPD.

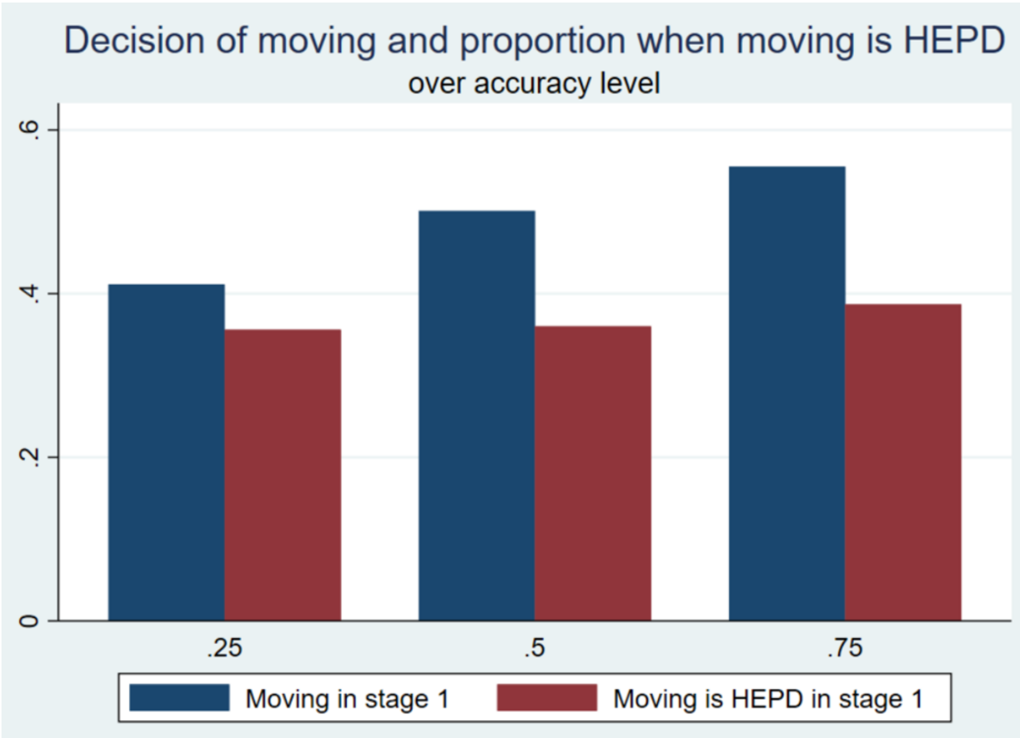


Figure 3.6: Decision in stage 1 (blue bar) and proportion of moving being the HEPD (red bar) over accuracy.

To investigate how participants' decisions were affected by the accuracy level when the HEPD was moving and staying, we split their decisions into two different situations: worse at present with better future situation ($Situation = -1$) and better at present with worse future situation ($Situation = -1$). We can see from Figure 3.7 that in each situation participants more often choose to move in both situations when the HEPD is moving and staying.

Result 3.7 *In stage 1, if high accuracy level means higher expected damage in stage 2 the accuracy level has opposite effects on the proportion of choosing the HEPD compared with the situation when high accuracy level leads to lower expected damage.*

This result can partially explain Hypothesis 3.4. From Table 3.4 the relationship between accuracy level and HEPD in stage 1 is not significant. The coefficient of

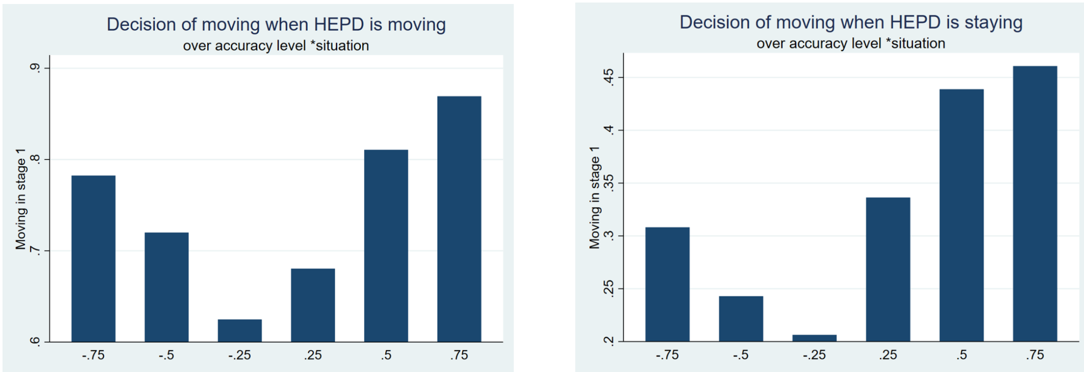


Figure 3.7: Decision of moving over accuracy × situation when HEPD is moving and when HEPD is staying.

accuracy level × situation is significantly positive. This result means when they have a situation of better at present with worse future, participants less often choose the HEPD when the accuracy level is high; on the other hand, participants more often choose the HEPD with a high accuracy warning level when they have a worse at present with better future situation. This result can also be seen from Figure 3.8. In Figure 3.8, the x-axis is accuracy × situation; it is positive if the situation is worse at present with better future, and negative if the situation is better at present with worse future.

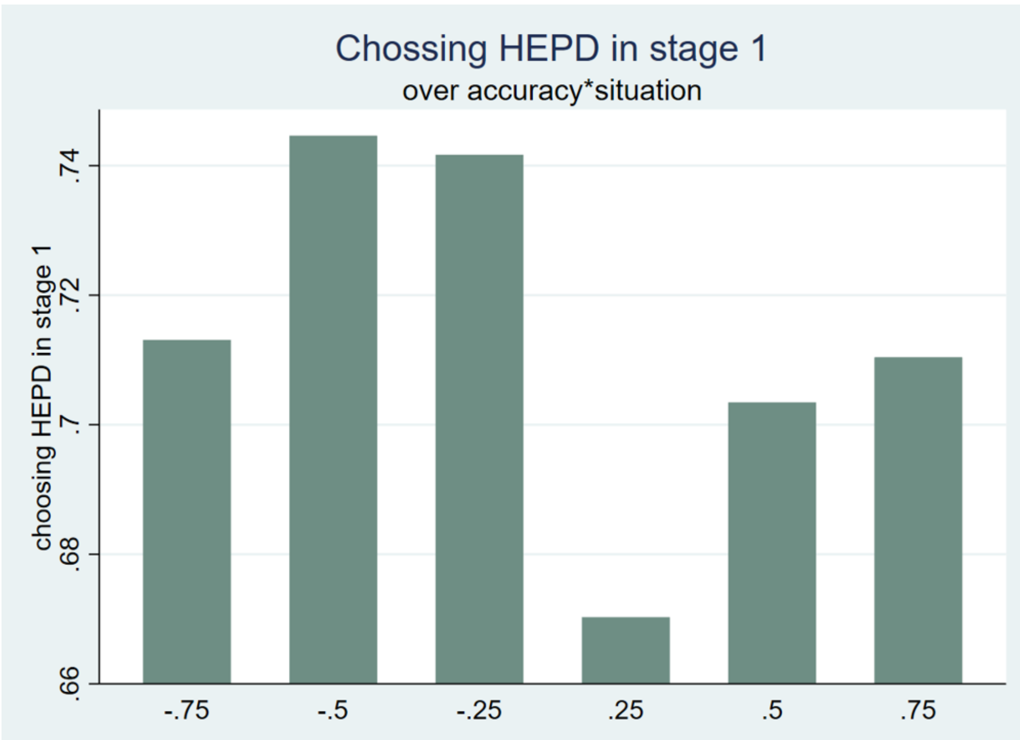


Figure 3.8: Proportion of participants choosing the HEPD in stage 1, based on the accuracy level and the situation.

Result 3.8 *If the accuracy level is high, people move less than they should in stage 2.*

In Table 3.4, the coefficient of accuracy level is significantly negative and accuracy level $\times$ situation is significantly positive, and the absolute value of the coefficient of the accuracy level $\times$ situation is significantly smaller than the absolute value of the coefficient of the accuracy level. This means participants choose to stay more when the accuracy level in stage 1 is high, and the frequency of participants choosing to move is more sensitive with the accuracy level when they are more likely to have a worse situation. In Figure 3.9, decision is 0 if participants choose to stay, and 1 if participants choose to move. It can be seen from Figure 3.9 that the percentage of participants choosing to move significantly decreases with the accuracy level; at the same time, the proportion of moving being the HEPD decreases with the accuracy level but less sensitively compared with the decision of moving.

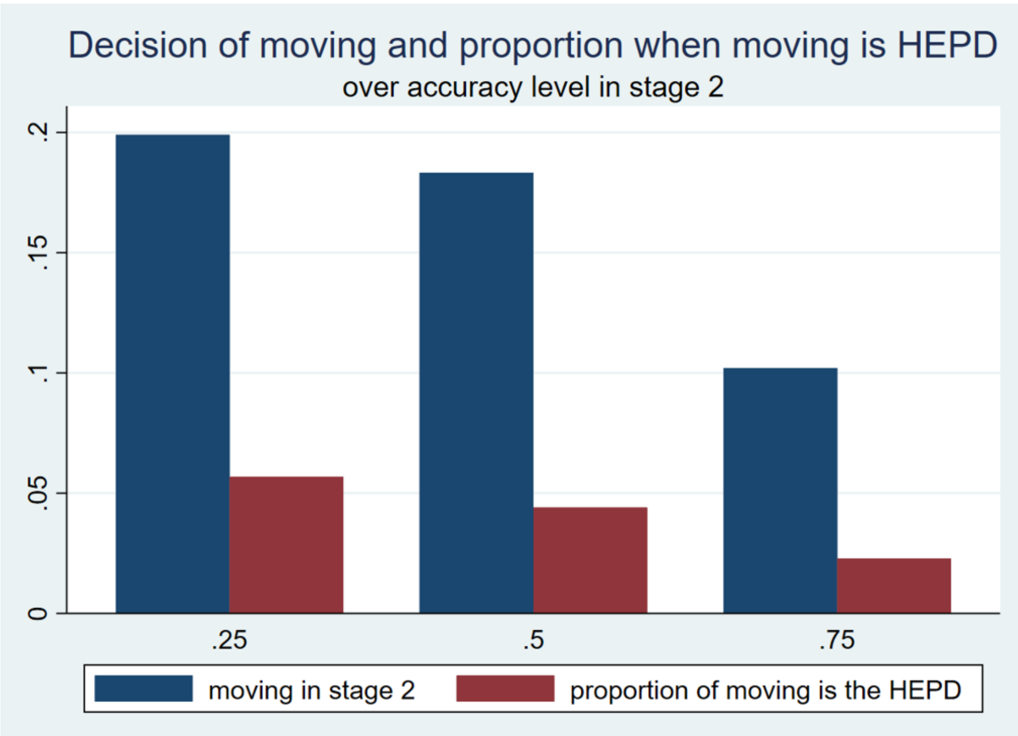


Figure 3.9: Decision in stage 2 (blue bar) and proportion of moving being the HEPD (red bar) over accuracy.

Result 3.9 *In stage 2 participants choose the HEPD more often when the accuracy level is high, regardless of the situation.*

This result is consistent with Hypothesis 3.5. In Table 3.4, the coefficient of the accuracy level is significantly positive and accuracy level × situation is significantly negative, and the absolute value of the coefficient of the accuracy level × situation is significantly smaller than the absolute value of the coefficient of the accuracy level. This proves that participants choose to move HEPD when the accuracy level is high, and the frequency of choosing the HEPD is more sensitive with the accuracy level when they are more likely to have a worse future situation. This result can also be proven from Figure 3.10. In Figure 3.10, the x-axis is accuracy × situation; it is positive if the situation is worse at present with better future, and negative if the situation is better at present with worse future.

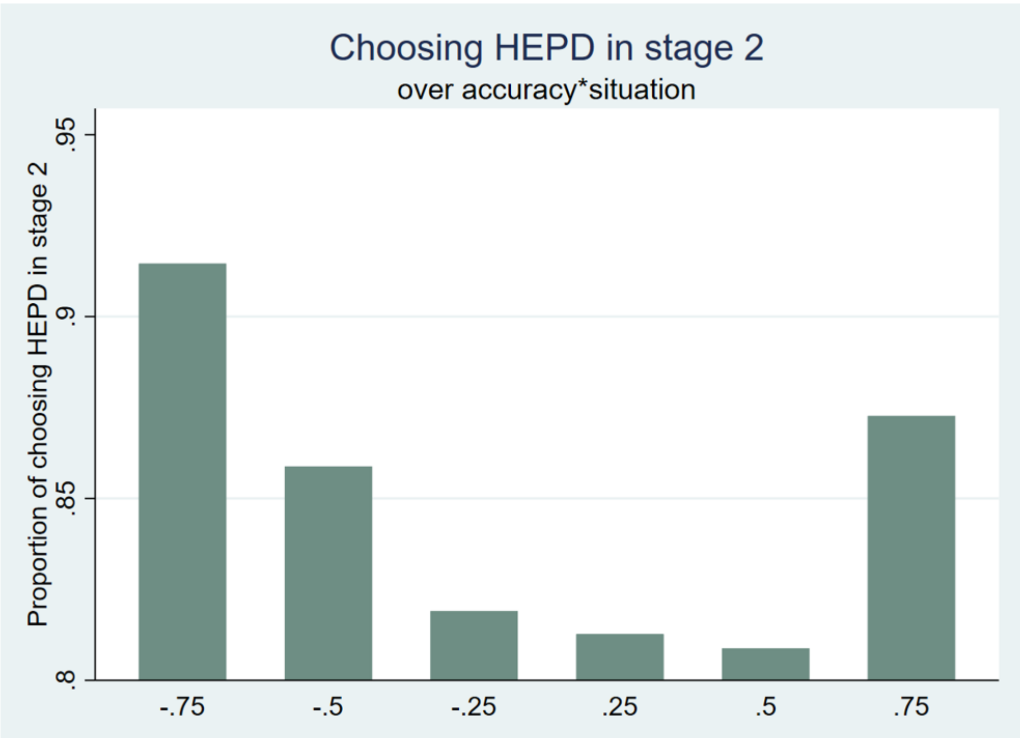


Figure 3.10: Proportion of participants choosing the HEPD in stage 2.

Result 3.10 *Overall, participants significantly more often choose to move with high accuracy level.*

In Table 3.4, the coefficient of accuracy level is significantly positive for accuracy level and the coefficient of accuracy level × situation is also significantly positive, and there is no significant difference between the absolute value of the coefficient of the accuracy level × situation and the coefficient of the accuracy level. This means participants choose to move more when the accuracy level is high.

Result 3.11 *Of the participants who choose to move eventually, higher accuracy is positively related to participants choosing to move earlier (in stage 1).*

This result is consistent with Hypothesis 3.4. In Figure 3.11, it is 1 if participants choose to move in stage 1 and it is 0 if they choose to move in stage 2; the participants' choice of stage is not included. It can be seen from Figure 3.11, for the participants who eventually choose to move, that participants more often choose to move in stage 1 when the accuracy level is higher. This proves that issuing a warning with high accuracy level can increase the response time. As the moving cost in stage 1 is always lower than the moving cost in stage 2, issuing a warning with high accuracy level can reduce the cost of reaction.

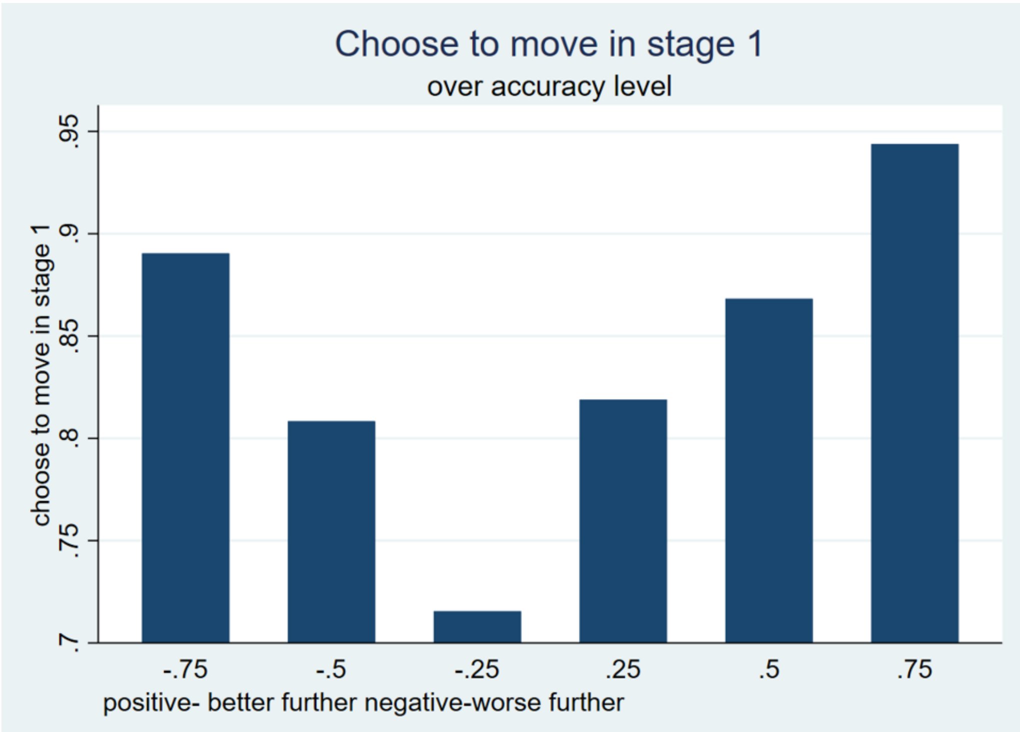


Figure 3.11: Percentage of participants choosing to move in stage 1.

Result 3.12 *Providing an accuracy level for warnings increases the expected payoff of the participants.*

It can be seen from the left part of Figure 3.12 that the expected payoff increases with the accuracy level overall. We also consider the fact that participants expected payoff to increase because the situation improved due to the increase in accuracy level. Therefore, we plot the expected payoff if participants always choose the HEPD and if they always choose the non-HEPD. We can see from Figure 3.12 that the expected payoff of the HEPD has same level of increase with the accuracy level compared with the expected payoff of a real decision and the expected payoff did not change. This means the expected payoff based on the real decision is closer to the optimal expected payoff when the accuracy level increases.

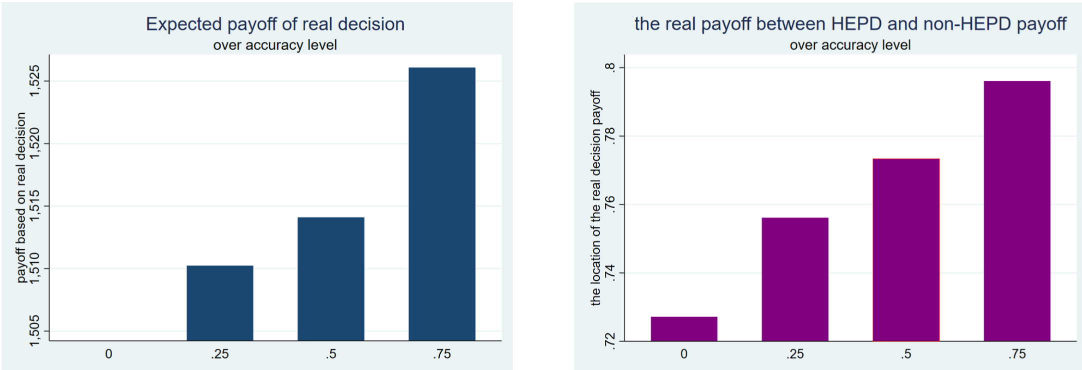


Figure 3.12: Percentage of optimal expected payoff gain achieved by real decisions, that is, the gap between the HEPD payoff and non-HEPD payoff.

There are two treatments in this experiment: a treatment with accuracy level and a treatment that only has the warning color level. The reason that we have two treatments is because we want to test whether having the accuracy level in the warning can improve the communication efficiency. We summarize the expected payoff overall and in each warning color in Figure 3.12.

3.6 Discussion and conclusion

In this paper we focus on investigating how participants behave when there is a chance to wait for better information. The main objective of our study is to determine the impact on decision making of providing information about the reliability of a warning. We frame the decision-making task in terms of a possible flood and participants can move their goods early or wait; afterward, if they wait, they can again decide whether or not to move their goods. Our analysis focuses on the participants' decisions and whether they are choosing the decision with higher expected payoff, and to what extent mistakes affect the participants' expected payoffs.

We find that when there is a higher warning color present (red or amber), participants more often choose to act even if it may not be the best action; on the other hand, when there is a lower warning color present, these participants more often choose to wait for more information even if they should act now.

There is also some clear-cut irrational behavior. When the warning color is red, for the participants who choose to stay at the earlier stage, over 40% of the participants change their decision to move at the later stage. This is interesting as there is only one red box in the warning matrix, so participants will not gain any extra information by waiting and the moving cost at the later stage is always higher. It could be optimal for some participants to choose to stay if they are risk-loving, but there is no point waiting at the earlier stage and then choosing to move with a higher cost at the later stage. The explanation may be that some people just want to wait when they have the chance and then make a final decision when they must choose.

There are interesting relationships between behavior at the earlier stage and later stage. Providing a more reliable warning can encourage people to take action. It is also interesting to see that the reliability of the warning at the earlier stage can significantly affect the decision at the later stage. People more often choose the decision with higher expected payoff at the later stage if they have a high reliability warning at the earlier stage. These results prove that issuing a warning with high reliability can help people to make better decisions.

The expectation of the future is also one of the important factors that we want to test. If people are expecting the future to be better in a disaster, it always means that the current situation is much worse. We find that if the participants have a higher damage warning in the earlier stage and are expecting a lower damage warning in the later stage, they choose to move more when the accuracy level is high. However, if they have a lower damage warning in the earlier stage and are expecting a higher damage warning in the later stage, the accuracy level does not have an effect on participants' decisions. Therefore, when we issue the warning, we can encourage participants to take action by issuing the warning with the upper bound of the expectation damage and a high-accuracy-level warning can enhance the effect.

The results in this chapter are useful for the design of warning systems. We have shown that providing reliability information can lead to some biased behavior, but overall participants benefit from reliability information. By comparing the expected gain of participants who received a warning with reliability information to participants who received a warning without reliability information, we find that participants always have a higher expected gain when they have information about the reliability of the warning, even when reliability is low. This shows that including reliability information in the warning system can benefit society. Therefore, we should include information about reliability in the warning system, especially for events where the cost of action increases significantly within a short period of time.

4. The Conflict between Personal Interests and Group Interests during an Epidemic

4.1 Introduction

COVID-19 rapidly developed into a global pandemic. There were over 1 million confirmed cases in the first three months, 10 million confirmed cases within six months, and over 100 million confirmed cases within a year. Over 1 percent of those confirmed to have COVID died (up to the delta wave)(Organization, 2022). Countries have enforced lockdowns, closed restaurants and shops, moved education online, and recommended various degrees of social distancing.

To contain the spread of an infectious disease, individuals can take costly action. Taking such action not only reduces their own risk of being infected, but also that of others, which is not fully internalized. Therefore, the equilibrium level of actions taken by individuals will be less than optimal with respect to society (Toxvaerd, 2019).

When the potential risks of a disease are uncertain, a government may have better information than the public. It is then a matter of policy how the government should communicate this risk to the public. It is also necessary to determine how much information should be made public, since it is best for society if individuals act as if the situation is worse for them than it actually is. If the individual has an aversion to ambiguity, they might act as if the situation is worse than it actually is.

In this chapter, we use both theoretical and experimental techniques (using

Internet-based interactive surveys) to determine the optimal government policy in releasing information to the public in the face of an epidemic and/or pandemic such as the current COVID -19. Using a stylized model, we aim to examine the conditions under which a rational individual is willing to control his or her exposure to infection by engaging in costly preventive behavior, such as staying home. We focus on the role of the government as the main source of information about the epidemic and examine how the type and accuracy of information conveyed by the government can strongly influence individual behavior and overall social welfare. To scientifically validate our model and its predictions, we use randomized control experiments in which participants make incentive decisions and interact online in real time.

We find that if the government does not intervene and direct people, they will behave selfishly, which is detrimental to the overall interests of society. It can even be said that people's social psychological changes have a far greater impact on the economy and society than the virus itself.

Staying at home has been proven to be one of the most efficient methods to stop the spread of COVID (and infectious diseases in general) (HMS, 2021). First, quarantine can stop contagious people from infecting others, which will significantly reduce the number of COVID cases. Secondly, for people who have not been infected, staying at home will reduce their probability of being infected. Finally, staying at home and resting could prevent those infected with no or very mild symptoms from getting worse.

China has contained the outbreak by quarantining and increasing medical resources. However, for most other countries, after a lockdown is declared, the number of infections may decrease during the lockdown but increase rapidly after the reopening of the economy, leading to another lockdown in the future. The French government, for example, announced its third lockdown on March 22, 2021, which lasted a month.

There are two different types of lockdowns: active lockdowns and passive lockdowns. An active lockdown is when everyone stays at home for two weeks, and during the lockdown, everyone is tested in order to find out who is infected and either quarantine them in full isolation or hospitalize them. After a long enough lockdown, everyone will either be cured, uninfected or remain in hospital. This outcome would then eliminate any new post-lockdown infections. In order to prevent new infections being spread by those arriving from outside the country, arrivals should be placed in quarantine. To date, many countries have tried this strategy with varying success. Only China has been fully successful at the time of writing this paper. Depending upon policy enforce-ability and contagiousness of a virus, by a following zero-covid policy, society can use medical resources more efficiently,(Su et al., 2022). Nonetheless, this means that while it is possible for this strategy to work, it is currently politically impossible for most countries in the world.

A passive lockdown is designed to prevent the health care system from becoming overburdened. During a passive lockdown, hospitals will treat only severe cases while allowing mild cases to recover in peace. The government will reopen the market once the health care system is functioning again. After the lockdown, those infected and cured will be vaccinated so that when the markets reopen, there will be fewer infections than before the lockdown, allowing hospitals to treat more patients with severe complications. Finally, the number of new infections will remain low because most people were infected and have recovered and developed herd immunity. This herd immunity is based on two assumptions: that most people will be immune after infection and that they will recover. The likelihood of severe cases is very low, so hospitals are able to treat patients who develop severe symptoms. However, a rapid mutation rate may reduce the success of such an approach. This could prevent herd immunity by reducing the effectiveness of vaccines and increasing the number of reinfections.

Staying at home is the most effective way to control the number of infections during an epidemic (albeit, with a cost for those that do so). For this reason, most countries recommended that the public stay at home as much as possible during the early stages of the epidemic. but people oppose it for several reasons. We also need to consider the economic consequences of encouraging people to stay at home; according to theoretical studies, if people stay at home for more than 3 months, it will cause economic recession (Bairoliya and Imrohoroglu, 2020). This is one reason why in most locked-down countries there are demonstrations. The first concern is income. According to Kreamer, Stock, and Rogelberg (2021), only 48% of Americans work remotely. For many going out to work, staying at home is extremely costly in terms of lost income. If they go out and work, they would also be more likely to be super spreaders. For instance, a waitress might not be able to afford missed income for a week and might continue to work when feeling unwell. She would be in contact with many customers, potentially spreading the virus. A study by the Institute for Fiscal Studies (IFS) suggests that under-25s are 2.5 times more likely to have no income than all employees. The second worry is that staying at home means people cannot access services that would make them happy, such as hairdressers, pubs, and restaurants. Staying at home means that the lack of communication with others also increases the likelihood of mental illness (Pfefferbaum and North, 2020). A final problem is that lockdowns can only reduce infection through herd immunity, which means that eventual infection seems inevitable. Therefore, if you cannot get the number of infections down to zero, then lockdowns can only delay infections, at great cost.

We also need to focus on another group of people who think they would be better off staying at home, but society needs them to work; for example, health workers during a pandemic. During the pandemic, their work has become more dangerous; they have to work with a very high risk of infection and they are also under high levels of emotional stress (Spoorthy, Pratapa, and Mahant, 2020).

This does not just apply to medical work during a pandemic; we also need to think about it after the pandemic is over. Late in the pandemic, the likelihood of infection is low, but infection is still possible. Society needs people to work and spend to aid the recovery, but people are still worried that they could be infected if they choose to work or go out to get services.

The supply and demand model states that when demand remains constant, a reduction in supply leads to an increase in prices. During the lockdown, most people stay at home, which means fewer people work in society, but people still have needs, which means the less people choose to work in society, the higher the cost of living in society. The EUI report said the 2020 Global Cost of Living index was affected by this year's supply chain difficulties (Roggeveen and Sethuraman, 2020).

COVID is just one of countless infectious diseases occurring in our society, and there may be more serious ones than COVID. We need to investigate how people manage their own interests and society's interests during a pandemic. Then we can improve our health warning systems to reduce the impact of the next epidemic.

Toxvaerd (2019) state that rational and forward-looking individuals should be able to analyze the situations and decide to take costly preventive behavior by themselves. But there are two following questions we want to ask, firstly are individuals always rational and forward-looking, secondly are the behaviors best to individuals are also the best behaviors to the society?

There are still some debate about what is the best behavior to the society during epidemic. By the end of April, the number of new cases had started to decrease and remained at a low level afterwards (Lee et al., 2020). On the other hand, the USA had the highest number of confirmed cases by 26 March 2020. The USA issued the highest level 4 warning on 20 March, but the number of cases keep increasing until January 2021. however, Gibson (2022) shows in during the epidemic the lockdown in New Zealand is ineffective as they suffer larger economic lost compare with the benefit of the life they saved.

Another example is Sweden, which is ranked as having one of the best healthcare systems in the world (Tandon et al., 2000). During the COVID outbreak in 2020, the Swedish government refused to issue any warnings or even acknowledge that wearing a mask would reduce the chance of infection. So far, 13,000 people have died from COVID-19 in Sweden. In Norway, which has half the population of Sweden, only 700 people died possibly thanks to government warnings and lockdown policies (Andersen et al., 2020).

During the epidemic, there were many cases of irrational behavior, especial when authorities gave problematic advice. Ajzenman, Cavalcanti, and Da Mata (2020) showed that when the leader of the country publicly challenged scientific recommendations during Covid, it induce the followers to engage in risky behavior. Bursztyn et al. (2020) shows that TV shows that downplayed the threat of Covid significantly increased the number of deaths from Covid. This shows that communication can have an effect on behavior.

4.2 Model and theory

4.2.1 Model

In our experiment, we try to design a game that simulates the decision problem that people have to solve during the epidemic. In our experiment, participants must think about how to coordinate with other participants. More specifically, if everyone decides to work, the probability of getting infected is too high and the expected payoff for working is negative. Therefore, our new experiment is a combination of a public good game and a coordination game. The public good game is designed to test how people trade off between their own interests and the interests of the group, and in a coordination game, participants must make decisions based on the behavior of others rather than making their own decision.

The public- good game is an appropriate tool to study participants' behavior during the epidemic, Clark (2022) shows that there is a significant positive relationship between participants' behavior in a public-lab experiment and their attitude toward Covid 19.

Coordination game is also used to simulate disease control measures between different governments, Maggi et al. (2014) uses coordination game to simulate the trade-off problems of medical resources and social costs and proves that both governments can benefit if they can negotiate healthcare costs. In our work, we aim to test whether individuals can benefit from being able to better coordinate with a government proposal.

To analyze individual behavior in a society, we devised a group game to model how an infectious disease spreads in a society. We assume there are N participants in each independent group. We simplify the decision so that individuals must make a choice between a safe option and risky but more potentially more profitable option. As such, we assume it is simultaneously the choice between working and staying at home during the spread of an infectious disease. Working is risky but may be more profitable and staying at home is perfectly safe with a fixed payoff.

In our model people need to consider the consequences of staying at home for society. To simplify the situation, we assume everyone is perfectly healthy and not infected at the beginning. The decision-making process is shown in Figure 4.1.

If an individual chooses to work, they will receive a wage (W). As there is also an infectious disease, those choosing to work may get infected with probability (PI). If they get infected, there is a cost of getting infected (C) that will be taken from their payoff. For participants that choose to stay, they will not receive any wage and the PI is 0. Independent of choice, participants get an initial endowment (I) and must pay the living cost (LC). In this experiment, everyone has the same health conditions, so PI and C are the same for everyone.

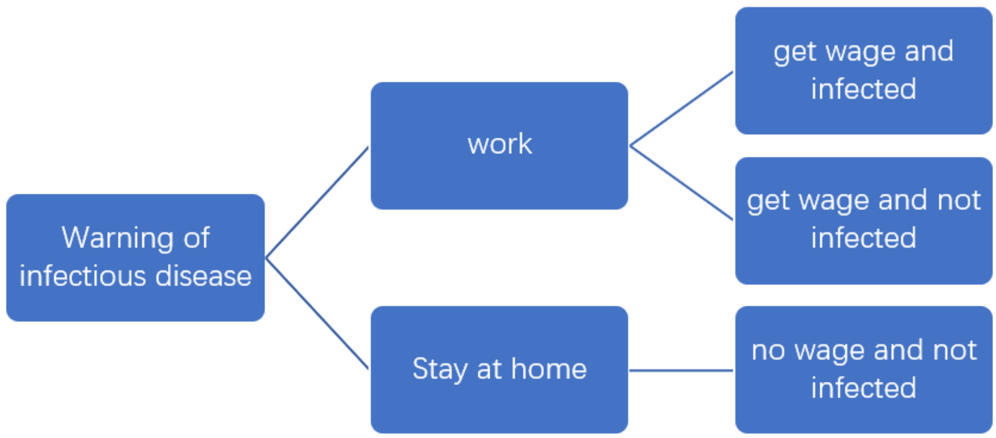


Figure 4.1: Decision-making process

In this experiment, W, C, and I are constants in each question, and PI and LC depend on the **number of people that choose to work in the group** (n). We assume that the fewer people choosing to work in a society, the higher the cost of living. Therefore, the living cost decreases or remains the same when more people choose to work, $LC_{(n+1)} \leq LC_n$. The difference between $LC_{(n+1)}$ and LC_n is constant in each individual question. To make it easier to calculate, the gap between the living cost is always the same ($LC_{(n+1)} - LC_n$ is constant). More people going to work leads to a higher probability of coming into contact with others for every person who chooses to go to work. Therefore, we assume that PI increases with n, $PI_{(n+1)} > PI_n$.

As their payoff and PI can be affected by the number of others choosing to work in the group, when participants make decisions, they also need to guess how many other participants are choosing to work in this situation. This is the reason why we ask them their **belief** of the number of others who choose to work (b) in every question. To encourage them to make their best guess, participants are rewarded if their guess is the same as the actual number of other people that choose to work.

Based on the assumption above, we calculate the expected personal payoff (Y_p) and expected social payoff (Y_s) based on their decision d. If there are m other participants (out of 9) who choose to work, the expected personal payoffs $E(Y_p)$ of choosing to work ($d = w$) and to stay at home ($d = h$) are:

$$E[Y_p(d)] = \begin{cases} W + I - LC_{(m+1)} - PI_{(m+1)} \cdot C & \text{if } d = w \\ I - LC_m & \text{if } d = s. \end{cases} \tag{4.1}$$

The expected social payoffs $E(Y_s)$ of choosing to work and to stay at home are:

$$E[Y_s(d)] = \begin{cases} (m+1) \cdot (W + I - LC_{(m+1)} - PI_{(m+1)} \cdot C) + (9 - m) \cdot (I - LC_{(m+1)}) & \text{if } d = w \\ m \cdot (W + I - LC_m - PI_m \cdot C) + (10 - m) \cdot (I - LC_m) & \text{if } d = s \end{cases} \tag{4.2}$$

4.2.2 Equilibrium

We start by analyzing the equilibrium in different situations and different utility of the participants. First, we describe the pure-strategy Nash equilibrium behavior and then we analyze the mixed-strategy Nash equilibrium. We do so for both when the utility of the participants is their own payoff and when the utility is the social payoff of all.

We first analyze the participants only care about their personal interest. To analyze whether participants should choose to work, firstly, we calculate the pure-strategy Nash equilibrium in different situations. The pure-strategy Nash equilibrium is a set of strategies (decisions) where no player has an incentive to deviate (can have a higher payoff by changing their decisions).

We can characterize the pure-strategy Nash equilibria into the class of equilibria where n^* people work and $10 - n^*$ choose not to work. Let m be the number of **other** participants that choose to work. In a pure-strategy equilibrium, if a participant chooses to work, then there should be $m = n^* - 1$ others that choose to work. For these participants, working should have a weakly higher payoff than not working. In equilibrium, if a participant chooses not to work, then there should be $m = n^*$ others that choose to work. For these participants, not working should have a weakly higher payoff than working.

Proposition 4.1 *There exists a pure-strategy Nash equilibrium for both objective functions.*

Proof: Let us characterize the pure-strategy Nash equilibrium by how many participants choose to work (rather than which participants work). For a group of 10, for participants who only care about their **personal interest**, participants will choose to work when n others choose to work if

$$W + I - LC_{(n+1)} - PI_{(n+1)} \cdot C \geq I - LC_n.$$

When $n = 0$ participants only choose to work if

$$W + I - LC_{(1)} - PI_{(1)} \cdot C \geq I - LC_{(0)}.$$

Otherwise, the pure-strategy Nash equilibrium will be always choose to stay if

$$W + I - LC_{(1)} - PI_{(1)} \cdot C < I - LC_{(0)}.$$

When everyone else chooses to work, participants will choose to stay if

$$W + I - LC_{(10)} - PI_{(10)} \cdot C \leq I - LC_{(9)}.$$

Otherwise, the pure-strategy Nash equilibrium will be always choose to work if

$$W + I - LC_{(10)} - PI_{(10)} \cdot C > I - LC_{(9)}.$$

We now need to explore the possibility that there are no equilibria where either (i) everyone chooses to work ($n = 10$) or (ii) everyone chooses to stay ($n = 0$). This implies that

$$W + I - LC_{(1)} - PI_{(1)} \cdot C \geq I - LC_{(0)}.$$

and

$$W + I - LC_{(10)} - PI_{(10)} \cdot C \leq I - LC_{(9)}.$$

First remember that $LC_{(n+1)} - LC_n$ is constant and that probability of getting infected is increasing in n.

For an equilibrium with n' working where $0 < n' < 10$, we must have

$$\begin{gathered} W + I - LC_{(n')} - PI_{(n')} \cdot C > I - LC_{n'-1} \textit{ and} \\ W + I - LC_{(n'+1)} - PI_{(n'+1)} \cdot C < I - LC_{n'}. \end{gathered} \tag{4.3}$$

Rearranging yields

$$\begin{gathered} W + I - PI_{(n')} \cdot C > I + LC_{(n')} - LC_{n'-1} \textit{ and} \\ W + I - PI_{(n'+1)} \cdot C < I + LC_{(n'+1)} - LC_{n'}. \end{gathered} \tag{4.4}$$

For $n' = 1$, the top equation of (equation 4.4) is true. For $n' = 9$, the bottom equation of (equation 4.4) is true. Note that (A) if the top equation is not true for n', the bottom equation must be true for $n' - 1$. Likewise, (B) if the bottom equation is not true for n', the top equation must be true for $n' + 1$. Since the LHS of both equations are decreasing in n' and the RHS is constant, we must have a unique $n' \in \{1, \ldots, 9\}$ such that both are true. One can see this by starting at $n' = 1$ and increasing n' by one each time. We know that for $n' = 1$, the top equation is true. If the bottom is true, then we have a unique value of n'. If not we can keep increasing n'. We know that at some point, we must have the bottom equation become true as well since it is for $n' = 9$. Note that we cannot go from the top being true and the bottom being false to the top being false and the bottom being true by point (B) from above.

We now prove existence of a pure-strategy Nash equilibrium when participants care about their **social interests** rather than their **personal interests** in a similar manner to the above. For a group of 10, for participants who only care about their **social interest**, participants will choose to work when n others choose to work if

$$(n+1)\cdot W+10\cdot I-10\cdot LC_{(n+1)}-(n+1)\cdot PI_{(n+1)}\cdot C \geq n\cdot W+10\cdot I-10\cdot LC_{(n)}-n\cdot PI_{(n)}\cdot C.$$

Note that the social payoff includes wages for all workers in addition to the probability of getting infected and the respective cost for *all* workers as well as the living costs and endowments of all participants. Thus, choosing to work will not only effect the person who made the decision, but will decreasing the living cost and increase the probability of getting infected for everyone else. For social expected payoff, on one hand choosing to work will increase the expected payoff for everyone, as it will decrease the living cost for everyone. On the other hand, it will decrease expected payoff for the other participants choosing to work, as it will increase their probability of getting infected. So, choose to work will have a positive effect for the participants choose to stay, and a uncertain effect for the participants choose to work.

When $n = 0$ participants only choose to work if

$$W + 10 \cdot I - 10 \cdot LC_{(1)} - PI_{(1)} \cdot C \geq 10 \cdot I - 10 \cdot LC_{(0)}.$$

Otherwise, the pure-strategy Nash equilibrium will be always choose to stay if

$$W + 10 \cdot I - 10 \cdot LC_{(1)} - PI_{(1)} \cdot C < 10 \cdot I - 10 \cdot LC_{(0)}.$$

When everyone else chooses to work, participants will choose to stay if

$$10 \cdot W + 10 \cdot I - 10 \cdot LC_{10} - 10 \cdot PI_{(10)} \cdot C \leq 9 \cdot W + 10 \cdot I - 10 \cdot LC_{(9)} - 9 \cdot PI_{(9)} \cdot C.$$

Otherwise, the pure-strategy Nash equilibrium will be always choose to work if

$$10 \cdot W + 10 \cdot I - 10 \cdot LC_{10} - 10 \cdot PI_{(10)} \cdot C > 9 \cdot W + 10 \cdot I - 10 \cdot LC_{(9)} - 9 \cdot PI_{(9)} \cdot C.$$

We now need to explore the possibility that there are no equilibria where either (i) everyone chooses to work ($n = 10$) or (ii) everyone chooses to stay ($n = 0$). This implies that

$$W + 10 \cdot I - 10 \cdot LC_{(1} - PI_{(1)} \cdot C \geq 10 \cdot I - 10 \cdot LC_{(0)}.$$

and

$$10 \cdot W + 10 \cdot I - 10 \cdot LC_{10} - 10 \cdot PI_{(10)} \cdot C \leq 9 \cdot W + 10 \cdot I - 10 \cdot LC_{(9)} - 9 \cdot PI_{(9)} \cdot C.$$

For an equilibrium with n' working where $0 < n' < 10$, we must have

$$\begin{gathered} n \cdot W + 10 \cdot I - 10 \cdot LC_n - n \cdot PI_n \cdot C \geq (n-1) \cdot W + 10 \cdot I - 10 \cdot LC_{(n-1)} - (n-1) \cdot PI_{(n-1)} \cdot C \\ and \\ (n+1) \cdot W + 10 \cdot I - 10_{(n+1)} - (n+1) \cdot PI_{(n+1)} \cdot C \leq n \cdot W + 10 \cdot I - 10 \cdot LC_{(n)} - n \cdot PI_{(n)} \cdot C \end{gathered} \tag{4.5}$$

Rearranging yields

$$\begin{gathered} W + 10 \cdot (LC_{n-1)} - LC_n) > n \cdot PL_n \cdot C - (n-1) \cdot PL_{n-1} \cdot C \\ and \\ W + 10 \cdot (LC_{n)} - LC_{n+1}) < (n+1) \cdot PL_{n+1} \cdot C - (n) \cdot PL_n \cdot C \end{gathered} \tag{4.6}$$

For $n' = 1$, the top equation of (4.6) is true. For $n' = 9$, the bottom equation of (4.6) is true. Note that (A) if the top equation is not true for n', the bottom equation must be true for $n' - 1$. Likewise, (B) if the bottom equation is not true for n', the top equation must be true for $n' + 1$. Since the RHS of both equations are increasing in n' and the LHS is constant, we must have a unique $n' \in \{1, \ldots, 9\}$ such that both are true. One can see this by starting at $n' = 1$ and increasing n' by one each time. We know that for $n' = 1$, the top equation is true. If the bottom is true, then we have a unique value of n'. If not we can keep increasing n'. We know that at some point, we must have the bottom equation become true as well since it is for $n' = 9$. Note that we cannot go from the top being true and the bottom being false to the top being false and the bottom being true by point (B) from above.

□

Therefore, there are pure-strategy Nash equilibria for participants that care about personal expected payoff and social expected payoff.

Proposition 4.2 *There also exists a mixed-strategy equilibrium under a certain situation.*

Proof:

For participants who want to maximize their personal expected payoff, let the probability of choosing to work be $k^{'}$; then the expected personal payoff is:

$$\mathrm{E}(\mathrm{Y}_p) = \begin{cases} \sum_{i=0}^{9}(B(i,9,k') \cdot (W + I - LC_{(i+1)} - PI_{(i+1)} \cdot C))) & \text{if } d = w \\ \sum_{i=0}^{9}(B(i,9,k') \cdot (I - LC_i)) & \text{if } d = s \end{cases}$$

(4.7)

In a mixed-strategy equilibrium, by the indifference principle, the payoff from choosing to work should equal the payoff from choosing to stay at home. Hence, the expected personal payoffs of both decisions should be equal to each other.

When $k = 0$, $B(i, 9, k) = 1$ when $i = 0$ and equals 0 for all other i. Thus, the expected payoff of participants that choose to work when $k = 0$ is

$$(W + I - LC_1 - PI_1 \cdot C)$$

and the expected payoff of participants that choose not to work when $k = 0$ is

$$(I - LC_0)$$

Participants will choose to work if

$$(W + I - LC_1 - PI_1 \cdot C) > (I - LC_0)$$

simplified to

$$(W + LC_0 - LC_1 - PI_1 \cdot C) > 0$$

When $k = 1$, $(B(i, 9, k) = 1$ when $i = 9$ and it is 0 for all other i, so the expected payoff of choosing to work is

$$(W + I - LC_{10} - PI_{10} \cdot C)$$

and the expected payoff of choosing not to work is

$$(I - LC_9)$$

Participants will choose not to work if

$$(W + I - LC_{10} - PI_{10} \cdot C) < (I - LC_9)$$

simplified to

$$(W + LC_9 - LC_{10} - PI_{10} \cdot C) < 0$$

As the living cost difference is constant, when

$$PI_1 \cdot C < W + LC_i - LC_{i+1} < PI_{10} \cdot C$$

the expected payoff of choosing to work is higher when $k = 0$ and that of choosing not to work when $k = 1$. As these are continuous functions in k, there exists a k where they are equal.

For participants who want to maximize their social expected payoff, let the probability of choosing to work be k^*; then the expected personal payoff is:

$$E(Y_s) = \begin{cases} \sum_{i=0}^{9}(B(i,9,k^*) \cdot (i+1) \cdot (W + I - LC_{(i+1)} - PI_{(i+1)} \cdot C) + (9-i) \cdot (I - LC_{(m+1)})) \\ \qquad \text{if } d = w \\ \sum_{i=0}^{9}(B(i,9,k^*) \cdot i \cdot (W + I - LC_i - PI_i \cdot C) + (10-i) \cdot (I - LC_i)) \quad \text{if } d = s \end{cases}$$

(4.8)

When $k = 0$, $(B(i,9,k) = 1$ when $i = 0$ and it is 0 for all other i, so the expected social payoff of choosing to work is

$$(W + I - LC_1 - PI_1 \cdot C) + 9 \cdot (I - LC_1)$$

and the expected social payoff of choosing not to work is

$$10 \cdot (I - LC_0)$$

Participants will choose to work if

$$(W + I - LC_1 - PI_1 \cdot C) + 9 \cdot (I - LC_1) > 10 \cdot (I - LC_0)$$

simplified to

$$(W + 10 \cdot LC_0 - 10LC_1 - PI_1 \cdot C) > 0$$

When $k = 1$, $(B(i,9,k) = 1$ when $i = 9$ and it is 0 for all other i, so the expected payoff of choosing to work is

$$10 \cdot (W + I - LC_{10} - PI_{10} \cdot C)$$

and the expected social payoff of choosing not to work is

$$(W + I - LC_9 - PI_9 \cdot C) \cdot 9 + (I - LC_9)$$

Participants will choose not to work if

$$(W + I - LC_9 - PI_9 \cdot C) \cdot 9 + (I - LC_9) > 10 \cdot (W + I - LC_{10} - PI_{10} \cdot C)$$

simplified to

$$(W + 10 \cdot LC_{10} - 10 \cdot LC_9 - 10 \cdot PI_{10} \cdot C + 9 \cdot PI_9 \cdot C) < 0$$

As the living cost difference is constant, when

$$(W + 10 \cdot LC_{i+1} - 10 \cdot LC_i - 10 \cdot PI_{10} \cdot C + 9 \cdot PI_9 \cdot C) < 0$$

the expected payoff of choosing to work is higher when $k = 0$ and that of choosing not to work when $k = 1$. As these are continuous functions in k, there exists a k where they are equal.

□

Proposition 4.3 *(i) If a participant would choose to work when believing m other participants are working, then the participant would choose to work when believing $m' < m$ other participants are working.*
(ii) Likewise, if a participant would choose not to work when believing m other participants are working, then the participant would choose not to work when believing $m' > m$ other participants are working.

Proof:

We assume a participant chooses to work and believes m others will choose to work in this situation, which means

$$W + I - LC_{(m+1)} - PI_{(m+1)} \cdot C > I - LC_m.$$

Simplifying yields:

$$W - LC_{(m+1)} + LC_m - PI_{(m+1)} > 0.$$

As the gap between $LC_{(m+1)}+LC_{(m)}$ is independent of m (with one exception), and $PI_{(m+1)}$ increases with m, which means $-PI_{(m+1)}$ increases when m decreases,

$$W - LC_{(m'+1)} + LC_{m'} - PI_{(m'+1)} > 0$$

They will still choose to work for any $m' < m$.

We assume a participant believes m others will choose not to work in this situation, which means

$$W + I - LC_{(m+1)} - PI_{(m+1)} \cdot C < I - LC_m$$

$$W - LC_{(m+1)} + LC_m - PI_{(m+1)} < 0.$$

As the gap between $LC_{(m+1)}+LC_{(m)}$ is the same, and $PI_{(m+1)}$ increases with m, which means $-PI_{(m+1)}$ decreases when m increases,

$$W - LC_{(m'+1)} + LC_{m'} - PI_{(m'+1)} < 0$$

They will still choose to work for any $m' < m$.

$\square$

4.3 Experiment design

In each group there are 10 participants. Participants make their decisions simultaneously between working and staying at home when there is an infectious disease spreading in society. To decrease the difficulty of their decision and prevent any bias that may be caused by the names of the variables, we do not show participants how their payoff is calculated. They can only see the payoff and the probability of getting infected in each situation with the corresponding number of people choosing to work (excluding themselves). An example of information given to a participant before a decision is shown in Figure 4.2.

Work Payoffs Table										
Number of people working (excluding yourself)	0	1	2	3	4	5	6	7	8	9
Payoff if infected (in points)	-82	-79	-76	-73	-70	-67	-64	-61	-58	-55
Payoff if **NOT** infected (in points)	43	46	49	52	55	58	61	64	67	70
Probability of infection	18%	20%	22%	24%	26%	28%	30%	32%	34%	36%
Stay at Home Payoffs Table										
Number of people working (excluding yourself)	0	1	2	3	4	5	6	7	8	9
Payoff (in points)	10	13	16	19	22	25	28	31	34	37

Figure 4.2: Example of information given to participants before they make their decision whether or not to work.

4.3.1 Choosing parameters

To test the participants' behavior in different situations, we set up 13 questions with different pure-strategy and mixed-strategy Nash equilibria for expected personal payoff and expected social payoff. There is one question in which the personal expected payoff and social expected payoff are both maximized when only 5 out of 10 participants choose to work. In 8 out of 13 questions, maximizing the social payoff requires more participants to work than would in a pure-strategy Nash equilibrium where the participants maximize their own personal payoff. In 4 out of 13 questions, maximizing the social payoff requires fewer participants than when they are maximizing their personal payoff. To control for any possible order effect, all the questions are randomly ordered. The details are shown in Table 4.1 below. PSP is the pure-strategy Nash equilibrium for expected personal payoff, which means for participants who only care about their personal interest they should choose to work if they believe number of others choose to work is less than the PSP. PSS is the pure-strategy Nash equilibrium for expected social payoff, which means for participants who only care about their social interest they should choose to work if they believe number of others choose to work is less than the PSS. MSP is the mixed-strategy Nash

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13
group	1	2	2	2	2	3	3	3	3	4	4	4	4
PSP	5	5	4	2	1	10	8	5	2	0	0	0	0
PSS	5	10	8	5	2	5	4	2	1	10	8	5	2
MSP	0.49	0.58	0.43	0.26	0.08	1	0.74	0.44	0.2	0	0	0	0
MSS	0.52	0.84	0.82	0.54	0.22	0.58	0.35	0.22	0.1	0.84	0.81	0.52	0.22

Table 4.1: The PSP, PSS, MSP and MSS for each question.

equilibrium for expected personal payoff, which is the probability of choosing to work if they only want to maximize their personal payoff, and MSS is the mixed-strategy Nash equilibrium for expected social payoff, which is the probability of choosing to work if they only want to maximize their social payoff. The mixed-strategy condition will be satisfied To test participants' behavior under conflicts between personal and social interest, we set up four types of questions in this experiment. In group 1, the number of people who need to work to maximize the social and personal interest is the same. In group 2, the number of people who need to work to maximize the social interest is about twice that of the number of people who need to work to maximize the personal interest. In group 3, the number of people who need to work to maximize the social interest is about half the number of people who need to work to maximize the personal interest.

4.3.2 The treatments

To investigate how participants react to a suggestion of behavior in different situations, we introduce two types of suggestions: color warning and word suggestion. These are two types of messages that we normally receive in our real lives. In the treatments with suggestions, participants always have the suggestion alone with the payoff table, and all participants answer the same questions in random order.

To test participants' behavior with different suggestions, we design the following five treatments: Con Treatment (control treatment, no suggestion only payoff table), WP Treatment (word suggestion based on personal mixed Nash equilibrium with the payoff table), WS Treatment (word suggestion based on social mixed Nash equilibrium with the payoff table), CP Treatment (color warning based on personal mixed Nash equilibrium with the payoff table), and CS Treatment (color warning based on social mixed Nash equilibrium with the payoff table).

Word suggestion

Word suggestions are private suggestions sent to individuals. In the treatments with word suggestion, for each question, the participants are told the proportion of participants who have a suggestion of working. Then they have a suggestion of working or staying at home; they will also be told the suggestion is for their own good or for that of the group. The probability of having a suggestion of work is based on the mixed-strategy Nash equilibrium in the question. Thus, different participants may have different suggestions in the same questions.

Color warning

Color warnings are public warnings sent to all participants, which is commonly known. In the treatments with color warnings, participants have a color warning with green, yellow, amber or red. In the instructions, there is a table showing participants the meaning of each color (Figure 4.3). When we think the mixed-strategy Nash equilibrium is between 0% and 25% the color is red, it is amber between 25% and 50%, yellow between 50% and 75%, and red between 75% and 100%. In the questions of the main experiment, participants have a highlighted color warning to show the warning color, and they can always check the meaning of each warning color during the experiment.

	Alert level			
Warning color	**Green**	**Yellow**	**Amber**	**Red**
Advice to public	Choosing to work is recommended in all the situations	Choosing to work is recommended in most of situations	Staying at home is recommended in most of situations	Staying at home is recommeded in all the situations

Figure 4.3: Explanation of each color warning in the introduction

4.4 Experimental procedure

The control treatment (50 subjects), color warning based on MSS (70 subjects), color warning based on MSP (70 subjects), word suggestion based on MSS (72 subjects), and word suggestion based on MSP (70 subjects) treatments were conducted in late 2020 and only for subjects who currently live in the US. Subjects were recruited and selected through Prolific. The sessions were programmed and conducted in Otree (Chen, Schonger, and Wickens, 2016). In total there were 332 participants: 152 were male, 174 were female, 3 selected another gender option, and 3 did not want to disclose their gender.

In each treatment participants needed to answer six test questions before the main experiment, and they could not proceed unless they got the right answer for all the test questions. In the main experiment they need to answer 13 questions in a random order. Four of the 13 questions were randomly selected for bonus payment. After subjects finished the main experiment, we elicited their risk preferences, tested their understanding of statistics and gave them a demographic questionnaire.

After they finished all the questions, participants were grouped by their finish time. Each participant who finished all the questions received a £1.67 showing-up fee, and their bonus depended on the selected results of the main experiment: whether they guessed the right number of others who chose to work in the group, the risk preference question and questions about their understanding of statistics; they got $1 for every 100 points they got in the experiment. Average earnings in treatments Con, WP, WS, CP and CS were $3.95, $4, $4.1, $3.9 and $4.2, respectively.

4.5 Hypotheses

Hypothesis 4.1 *The social interest and personal interest should both have positive relationship with participants utility, and participants' utility should be more sensitive to their personal interest.*

In the control treatment, we expect participants to care about the personal and social expected payoff, Barasch et al. (2014) shows that people are selfish and selfless at the same time even with out any external reward. So, we expect that when a decision can increase the expected personal and social payoff at the same time people always want to choose that decision. In other words, when there is a gap between mixed-strategy Nash equilibrium for personal payoff and social payoff, we expect the frequency of those choosing to work to be between these two mixed-strategy Nash equilibria.

We also expect the frequency of choosing to work to be closer to the MSP than MSS, as the participants care more about their own payoff than the social payoff, so there is a conflict between personal interest and social interest; they will more often choose the decision that increases their personal payoff.We will test the relationship between participants decision with MSP and MSS in different situations.

Hypothesis 4.2 *Having a suggestion of work should encourage people to choose to work more, and a word suggestion for each individual should have more of an effect than a color warning.*

Ajzenman, Cavalcanti, and Da Mata (2020) proves that during epidemic, suggestion of authorities can significantly affect people's behavior. We expect participants to choose to work more often if they have a suggestion of working; the frequency of choosing to work should significantly increase when participants have a suggestion of working (green and yellow warnings in color warning treatments) and significantly decrease when they have a suggestion of staying at home (amber and red warnings in color warning treatments).

We also expect suggestions in words should have more effect on participants' decisions as they deliver more accurate information than a warning color. In word suggestion treatments, participants know the proportion of those having a suggestion of work and the word suggestion is a clear message of what participants should choose. In color warning treatments, the meaning of each warning color may be different for each participant, and the range of mixed-strategy Nash equilibrium is not revealed; they must learn it during the experiment.We will test the how participants decision was affected in treatment with different suggestions.

Hypothesis 4.3 *Having a universal suggestion of work makes people believe others will choose to work more. If there are individual suggestions for work, beliefs about the total number of people who choose to work should not depend upon one's individual suggestion.*

Based on the calculation we have in Proposition 4.3 We expect when participants have a universal suggestion of work it will increase participants' belief about the number of others choosing to work, as once they have received a universal suggestion of work, everyone else will also have received the same suggestion, which will increase the probability of choosing to work for everyone. If the suggestion is specific for each participant and the probability of having a suggestion of work is the same and has been told to everyone, whether or not they receive a suggestion of work should not affect the belief of the total number of people choosing to work in the group.We will test the difference of participants decision in treatment with universal suggestion and individual suggestions.

Hypothesis 4.4 *Participants in treatments with a suggestion based on social interest should have higher payoffs, as they more likely to cooperate with each other.*

Based on the calculation we have in Proposition 4.2, their expected payoff should be maximized if they follow the suggestion based on mixed strategy based on social interest. We expect the expected personal and social payoff to increase when participants have suggestions based on a mixed strategy of expected social payoff. As a decision increasing expected social payoff will increase the expected payoff for everyone in the group, more participants will follow the suggestion to increase the expected payoff.

Suggestions based on a mixed strategy for expected personal payoff should have a positive effect on the expected personal payoff. Although the suggestion encourages participants to choose the decision with higher expected personal payoff, there is also a negative effect because if more participants choose the decision to increase expected personal payoff but lower social payoff then the expected personal payoff will decrease. Therefore, we expect the suggestion based on a personal mixed strategy to have less effect than the suggestion based on a social mixed strategy.We will compare the expected payoff based on participants' decision in treatment with different suggestion and the control treatment.

4.5.1 Results

We start to analyze the participants' behavior in the control group. To make it easy to understand, we categorize the mixed-strategy equilibrium for social interest and personal interest MSS/MSP into three different levels. The MSS/MSP level is 1 (low) when MSS/MSP is between 0% and 33%, 2 (medium) when MSS/MSP is between 33% and 66%, and 3 (high) when MSS/MSP is between 66% and 100%. We also define the IES as 1 if based on the participants' belief that choosing to work can increase the expected social payoff, and 0 otherwise; IEP is 1 if based on the participants' belief that choosing to work can increase the expected personal payoff, and 0 otherwise. The work is the percentage of people choosing to work, and belief is the proportion of people they believe will choose to work, which is the number of others they believe will choose to work out of the group size.

Result 4.1 *In the control treatment, participants' decisions were not affected by whether the society needed more or fewer people to work, but they chose to work more if the selfish Nash equilibrium had more choosing to work.*

This result is partially consistent with Hypothesis 4.1; the difference is that the social interest does not affect their decisions; in other words, participants are more selfish than we expected. It is clear to see that in Fig. 4.4 in the control treatment the mean of work always increases when the MSP level increases in every MSS level. In Fig. 4.4 the MSS/MSP level is 1 (low) when MSS/MSP is between 0% and 33%, 2 (medium) when MSS/MSP is between 33% and 66%, and 3 (high) when MSS/MSP is between 66% and 100%. The work is the percentage of people that choose to work, and belief is the proportion of people that they believe choose to work, which is the number of others they believe choose to work out of the group size. However, the mean of the decision does not have any significant relationship with the MSS level in each MSP level. This evidence proves that the frequency of choosing to work is always sensitive to the MSP level and does not have a significant relationship with the MSS level.

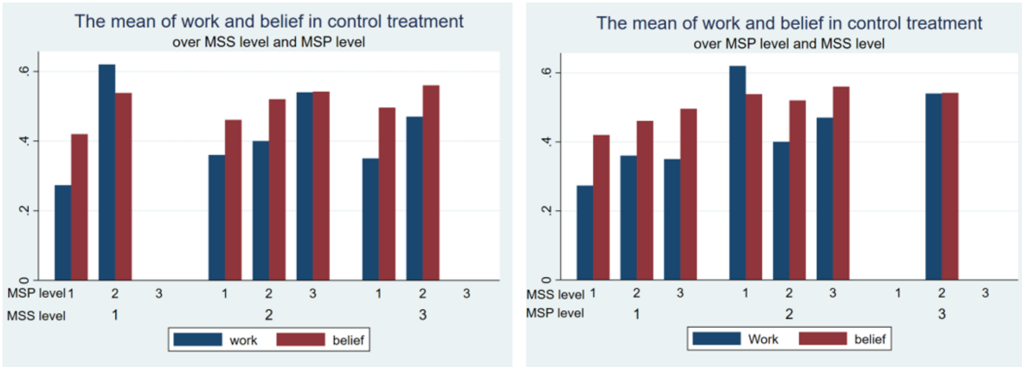


Figure 4.4: Bar plot of the mean of work and belief over different MSS and MSP levels in the control treatment.

Result 4.2 *Participants only care about their personal expected payoff even when there is no bias of belief of the situation.*

This result is also partially consistent with Hypothesis 4.1. We test how participants make decisions during the experiment. We determine the information that they have in the experiment in three parts: the IES and IEP based on the payoff information they have, and the belief of how many others choose to work for the same question. For participants in treatments with suggestions, their decision will be affected by the message in the suggestion.

To analyze the relation between frequency of work and payoff information, the frequency of choosing to work in the control should be calculated as

$$f = c_1 IES + c_2 IEP + U_i + \varepsilon$$

We define new variables: increase expected personal payoff (IEP) and increase expected social payoff (IES). IEP is 1 if the belief is less than PSP, which means based on their belief that choosing to work can increase their expected personal payoff, otherwise IEP is 0. IES is 1 if the belief is less than PSS, which means based on their belief that choosing to work can increase their expected social payoff, otherwise IEP is 0. U_i represents the effect caused by different participants.

We run a linear fixed-effects logit model for the frequency of choosing to work in each treatment on IES and IEP in each treatment; the results are summarized in table 4.2.

Table 4.2: Fixed-effects logit model of the role of payoff information in decisions

Treatment	Con	WP	WS	CP	CS
IEP	0.621***	1.068***	-0.432**	1.219***	-0.641***
	(0.228)	(0.194)	(0.175)	(0.188)	(0.166)
IES	-0.195	-0.759***	0.652***	-0.155	0.386**
	(0.202)	(0.170)	(0.163)	(0.163)	(0.158)
Number of observations	533	845	806	819	845
Number of subjects	41	65	62	63	65

Note: Standard errors in parentheses. *** indicates $p < 0.01$, ** indicates $0.05 > p > 0.01$.There is no suggestion in the control treatment.

We observe that the coefficient on IEP is significantly positive in Con treatment and the coefficient on IES is not significant. The result shows that when participants do not have any suggestions they significantly more often choose to work if based on their belief that choosing to work can increase their personal payoff, but whether choosing to work can increase the social payoff has no significant effect on participants' decisions. This result proves our first hypothesis, which is that participants only care about their personal interest and do not care about the social interest.

Result 4.3 *Having a specific suggestion significantly affects participants' decisions and beliefs, and the decision to work is more sensitive than the belief is to the suggestion.*

This result is consistent with Hypothesis 4.2, but not consistent with Hypothesis 4.3. It is interesting to see that the participants' beliefs are also affected by a specific suggestion. As the participants in treatments with color warnings have a uniform warning under the same situation, in this part we only look at the treatments with specific suggestions, WS (word suggestion based on social interest) and WP (word suggestion based on personal interest). In the WS and WP treatments, participants have the overall proportion of participants having a suggestion of work, and a specific suggestion just for themselves. The probability of having a suggestion of working depends on the overall proportion.

Figure 4.5 clearly shows that participants more often choose to work when they have a suggestion to work (1) than when they have a suggestion to stay (0) in both WP and WS treatments. This holds for all MSP/MSS levels. This shows that having a specific suggestion based on personal interest and a suggestion based on group interest both significantly affect the participant's decision.

It is also clear to see that the specific suggestion affects participants' belief; participants believe fewer other people chose to work when they have a suggestion to stay. This shows that even though they know the suggestion is a specific suggestion just for them and they know they have the same overall suggestion rate, the suggestion still affects participants' belief. This means even though someone knows only a proportion of people were asked to work, the specific suggestion will make them believe more people will choose to work.

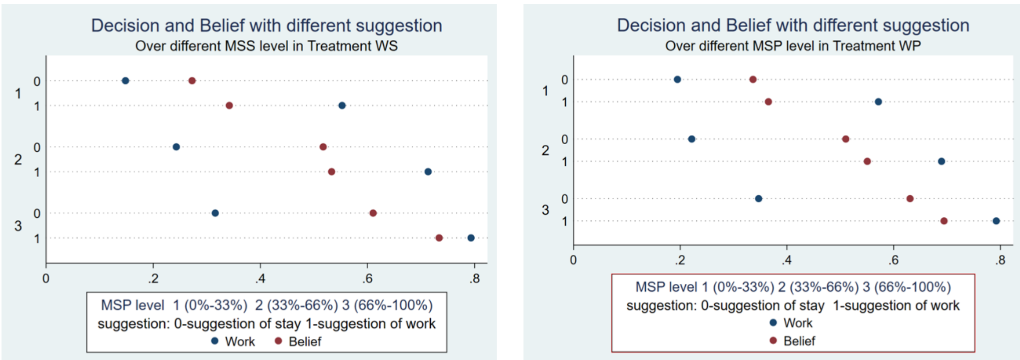


Figure 4.5: Dot chart for frequency of working and belief with suggestion of work, suggestion of staying in treatment with word suggestion over MSS and MSP.

Result 4.4 *In the treatments with suggestions based on social interest, participants care significantly more about the social interest.*

From the results of Table 4.3, in the treatments with suggestions based on social interest (WS, PS) the coefficient of IES is significantly positive and the coefficient of IEP is significantly negative. These results show that when participants have a suggestion based on social interest they choose to work more when it can increase the social interest, and choose to work less even if it can increase their personal interest.

Compared with the control treatment, in the treatment with word suggestion based on personal interest, participants choose to work more if work can increase the personal interest and less if it can increase the social interest. This result shows that word suggestion based on personal interest makes people increasingly selfish about their own interest. For the treatment with color suggestion based on personal interest, participants choose to work more often if it can increase the personal interest, but whether it will increase the social payoff does not have a significant effect.

Result 4.5 *Participants are more likely to choose to work the more they believe others will work.*

We run a linear fixed-effects logit model for the frequency of choosing to work in each treatment on belief in each treatment; the results are summarized in table 4.

Table 4.3: Fixed-effects logit model of the role of belief in decisions.

Treatment	Con	WP	WS	CP	CS
belief	0.190***	0.426***	0.515**	0.455***	0.441***
	(0.047)	(0.038)	(0.046)	(0.043)	(0.037)
Number of observations	533	845	806	819	845
Number of subjects	41	65	62	63	65

Note: Standard errors in parentheses. *** indicates $p < 0.01$, ** indicates $0.05 > p > 0.01$. There is no suggestion in the control treatment.

It is clear to see from Table 4.3 that the coefficient of belief is positive in all treatments, which means that participants who have high belief always choose to work more than participants with low belief.

It can also be observed that the coefficient of belief in the control treatment is significantly lower than the coefficient in all treatments with suggestions. This result proves that in the treatment with suggestions participants choose to work more often if they have the same level of belief.

Result 4.6 *For the treatments with color warnings, the warning color can significantly affect participants' decisions and beliefs.*

We categorize the percentage of people who choose to work and the belief into different MSP/MSS levels and different warning colors; it is clear to see that the percentage of participants choosing to work significantly decreases when the warning color changes from green to red. This also holds for the belief. These results can also be shown in Figure 4.6. The MSS/MSP level is 1 (low) when MSS/MSP is between 0% and 33%, 2 (medium) when MSS/MSP is between 33% and 66%, and 3 (high) when MSS/MSP is between 66% and 100%. The work is the percentage of people who choose to work, and belief is the proportion of people that they believe choose to work, which is the number of others they believe choose to work out of the group size. The color is 1 – green, 2 – yellow, 3 – amber, 4 – red.

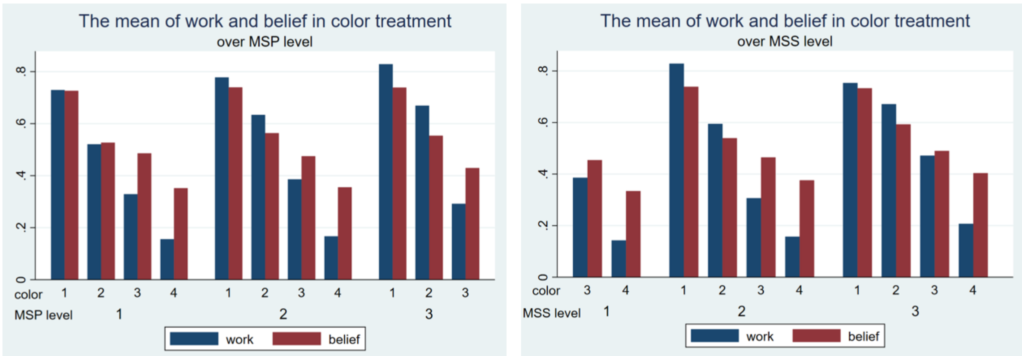


Figure 4.6: Bar chart of the mean of work and belief over different MSS and MSP levels in treatments with color warning (CP, CS).

Result 4.7 *In the control group participants believe that others care about the social interest.*

As the belief is also based on the payoff information, to analyze the relation between belief and the payoff information, in the control treatment, the belief in control should be calculated as

$$belief = c_1 MSP + c_2 MSS + U_i + c_3$$

It is interesting that in the regression both MSP and MSS have a significant positive effect on the belief. This means participants believe there will be more participants choosing to work if either the society needs more people to work or more people choosing to work can increase the payoff for themselves. This proves that they think others care about their personal interest and care about the social interest.

Table 4.4: Fixed-effects logit model of payoff information and suggestion for belief.

Treatment	Con	WP	WS	CP	CS
MSS	2.364***	-0.383	9.361***	3.222***	10.18***
	0. 846	(0.498)	(2.66)	(0.791)	(2.039)
MSP	-0.486	11.450***	3.410***	6.341***	0.300
	(0.202)	(0.170)	(0.163)	(0.163)	(0.158)
Number of observations	208	559	806	806	780
Number of subjects	16	43	62	62	60

Note: Standard errors in parentheses. *** indicates $p < 0.01$, ** indicates $0.05 > p > 0.01$.There is no suggestion in the control treatment.

Result 4.8 *In the treatment with the suggestion participants believe others will follow the suggestion.*

Table 4.4 shows that in the WP and CP treatments, the coefficients of MSP are significantly positive and significantly greater than the coefficient in the Con treatment. This result shows that when participants have suggestions based on their personal interest, they believe more people will choose to work to increase their personal payoff, and others will care about their personal interest more compared with the Con treatment. In the WS and CS treatments, the coefficients of MSS are significantly positive and significantly larger than the coefficients in the Con treatment, which proves that when participants have suggestions based on the interest of the society, they believe more people will choose to work if the society needs more people to choose to work, and more others will care about their social interest than if there is no suggestion.

It is interesting that only in the CP treatment do participants think others care about both the personal interest and social interest, but in all other treatments the belief is only significantly relevant to either their personal interest or the social interest.

Result 4.9 *(i) Both word suggestions and color warnings based on the social interest significantly increase payoffs. (ii) Word suggestions based on personal payoffs significantly decrease payoffs.*

This result is also partially consistent with Hypothesis 4.4. It is clear to see in Figure 4.7 that the expected personal/social payoff in Treatments 3 (WS) and 5 (CS) are significantly higher than the expected personal payoff in the control treatment, which proves that when giving participants suggestions based on the MSS, there is no significant difference between the expected personal and social payoff. Word suggestions based on MSP significantly decrease the expected personal/social payoff. There is no significant difference between the control treatment and the CP treatment, and a color warning based on MSS has no significant effect on the expected personal/social payoff.

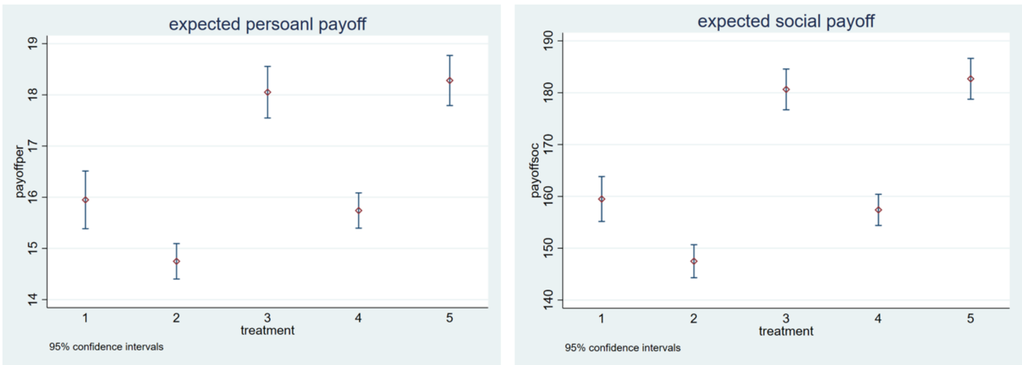


Figure 4.7: Expected personal payoff and expected social payoff in the control treatment (Treatment 1), WP treatment (Treatment 2), WS treatment (Treatment 3), CP (Treatment 4), and CS (Treatment 5).

4.6 Discussion

Overall we find that in the control treatments without any suggestion, participants only care about their personal interest; they do not care about the social interest when they make their decisions. However, it is interesting that they believe others will choose choose to work more/less if the society needs it. In the treatments with suggestions, we find the personal interest and social interest can be significantly increased by providing a suggestion based on social interest.

In the control treatment participants choose to work significantly more/less if more/less choosing to work can increase their personal expected payoff, but the social interest does not affect their decisions. As the expected payoff is highly related to the number of others choosing to work in the same group, we investigate whether their decisions are due to the wrong belief of the number of others choosing to work or whether they are just selfish. We find they choose to work significantly more when choosing to work can increase their expected payoff based on their belief, but whether choosing to work will increase the group expected payoff does not affect their decision, which means they are selfish when there is no suggestion.

There is also some irrational behavior which contradicts the theoretical model. Our theoretical model proves that participants should choose to work less if they believe more people will choose to work in the same group with them, as people care about themselves and people care about the group interest of both. We then use experiments to examine participants' behavior during the pandemic scenario and find that the number of others who choose to work has a significant positive relationship with the probability of participants choosing to work. This positive relation is true in all treatments, which means participants are more likely to choose to work if they think more people will choose to work in the same group. The theoretical prediction and experimental data are contradicted.

Herd mentality is defined as the tendency of the people in a group to think and behave in ways that conform with others in the group rather than as individuals. This definition is very widely used to explain economic bubbles; for instance, when a market is overheating and the price is much higher than the real value, but everyone is still buying it. Even the investors realize the market price is much higher than the real price, but the herd mentality makes them want to invest in this market as everyone else is investing in this market. This is very similar to our experiment. In our experiment, when more participants choose to work the living cost decreases, so even though the probability of getting infected increases and the personal and social expected payoff decrease, participants still want to work more when others also choose to work.

We find evidence that when there are no suggestions, participants only want to increase their personal interest and do not care about the group interest. However, when they are guessing how many others will choose to work in the same scenario, participants believe that others will care about both the personal interest and the social interest. This means participants are selfish themselves but believe others care about the group and their own interest at the same time.

In the initial phase of the covid epidemic, not all governments made recommendations based upon what was in the best interest for the people according to science. (Recommendations were based upon political considerations instead). Our result shows that the government's recommendations have an effect and thus should be based on the interests of society, and that this will significantly increase the welfare of each individual and society.

This result will also help the economic recovery, and if a similar infectious disease occurs again, the government should issue warnings based on the group interest. In addition, the government should promote the measures that are beneficial to society and persuade people to take other measures as well. For example, most countries in the world focus on economic recovery. Therefore, it is important to communicate with the public in the right way, without causing excessive panic among the population.

In this work, we also try to find out the difference between individual suggestions and uniform suggestions, because during the epidemic, the many governments/schools/companies divided the employees/students into several shifts and tried to keep the society functioning with the least number of people working at the same time. However, in our experiment, both the individual suggestions and the uniform suggestions had a significant effect on participants' behavior, and we found no significant difference. The reason could be that in our experiment, the payoff function is the same for everyone, so the coordination is not so important in our experiment, we can give different payoff functions to different participants in a future study.

5. Summary and Conclusion

In this dissertation, we analyze decision making when there is an emergency event (or a strong possibility of one) and information about the likelihood or strength is communicated.

Firstly, we study the behavior of individuals under the threat of a disaster and how to influence their behavior through information release. In Chapter 3, we study how different ways of communication affect the timing of individuals' decisions. In Chapter 4 we study the influence of early warning systems and information intervention on public behavior in the context of major public health events (such as epidemics). The main research question is how people weigh personal interests and social interests in the event of major public events. On the basis of this, we should carry out further research. We find that when a crisis comes, people tend to maximize their own interests and/or make non-optimal decisions. However, we can influence their decisions through the combination of color and text warnings to improve the overall interests of society.

The main research directions in the future are as follows. In our previous studies, subjects all made decisions simultaneously, and subjects could not observe each other's decisions. However, in real life, whether faced with natural disasters or epidemics, people can make decisions after observing others' decisions. Therefore, in future studies, we should study how people's behavior can be affected by the behavior of others by asking subjects to make decisions at different times. This can build upon the work of Chakravarty, Fonseca, and Kaplan (2014), who found that a banking panic can spread if the participants can observe the behavior of others.

When we study the conflict between the interests of individuals and the interests of society in the context of an infectious disease pandemic, our subjects are all people living in the US. However, the US now has the highest number of confirmed cases and the highest number of deaths of any country. Hence, we in the UK have expanded our research further, applying the same experiments to countries that have fared better in the pandemic, such as China, Japan and Singapore. In this way, we can compare whether different ways of communication can improve the decision-making behavior of the masses in these regions with different cultural backgrounds.

Bibliography

Abraham, S., R. Bartlett, M. Standage, A. Black, A. Charlton-Perez, and R. McCloy (2015). "Do location-specific forecasts pose a new challenge for communicating uncertainty?" In: *Meteorological Applications* 22.3, pp. 554–562 (Cited on page 8).

Ajzenman, N., T. Cavalcanti, and D. Da Mata (2020). "More than words: Leaders' speech and risky behavior during a pandemic". In: *Available at SSRN 3582908* (Cited on pages 69 and 86).

Andersen, A. L., E. T. Hansen, N. Johannesen, and A. Sheridan (2020). "Pandemic, shutdown and consumer spending: Lessons from Scandinavian policy responses to COVID-19". In: *arXiv preprint arXiv:2005.04630* (Cited on page 69).

ASCE (2007). "The New Orleans hurricane protection system: what went wrong and why". In: American Society of Civil Engineers (Cited on pages 1 and 5).

Ayewah, N (2003). "Prediction of spatial temporal events using a hidden Markov model". In: *Department of Computer Science and Engineering, Southern Methodist University, Dallas, TX* (Cited on page 37).

Bairoliya, N. and A. Imrohoroglu (2020). "Macroeconomic consequences of stay-at-home policies during the COVID-19 pandemic". In: *COVID Economics* 13, pp. 71–90 (Cited on page 68).

Banbury, S., S. Selcon, M. Endsley, T. Gorton, and K. Tatlock (1998). "Being certain about uncertainty: How the representation of system reliability affects pilot decision making". In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 42. 1. SAGE Publications Sage CA: Los Angeles, CA, pp. 36–39 (Cited on pages 47 and 48).

Barasch, A., E. E. Levine, J. Z. Berman, and D. A. Small (2014). “Selfish or selfless? On the signal value of emotion in altruistic behavior.” In: *Journal of personality and social psychology* 107.3, p. 393 (Cited on page 85).

Basher, R. (2006). “Global early warning systems for natural hazards: systematic and people-centred”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 364.1845, pp. 2167–2182 (Cited on page 6).

Bauer, P., A. Thorpe, and G. Brunet (2015). “The quiet revolution of numerical weather prediction”. In: *Nature* 525.7567, pp. 47–55 (Cited on page 37).

BBC (2021). *Europe’s floods: Lessons from German tragedy.* `https://www.bbc.co.uk/news/world-europe-58992093/`. Accessed April 5, 2021 (Cited on page 1).

Benbasat, I., A. S. Dexter, and P. Todd (1986). “An experimental program investigating color-enhanced and graphical information presentation: an integration of the findings”. In: *Communications of the ACM* 29.11, pp. 1094–1105 (Cited on page 47).

Braun, C. C. and N. C. Silver (1995). “Interaction of signal word and colour on warning labels: differences in perceived hazard and behavioural compliance”. In: *Ergonomics* 38.11, pp. 2207–2220 (Cited on page 8).

Bursztyn, L., A. Rao, C. P. Roth, and D. H. Yanagizawa-Drott (2020). *Misinformation during a pandemic.* Tech. rep. National Bureau of Economic Research (Cited on page 70).

Casteel, M. A. (2016). “Communicating increased risk: An empirical investigation of the National Weather Service’s impact-based warnings”. In: *Weather, Climate, and Society* 8.3, pp. 219–232 (Cited on page 6).

Chakravarty, S., M. A. Fonseca, and T. R. Kaplan (2014). “An experiment on the causes of bank run contagions”. In: *European Economic Review* 72, pp. 39–51 (Cited on page 98).

Chen, D. L., M. Schonger, and C. Wickens (2016). “oTree—An open-source platform for laboratory, online, and field experiments”. In: *Journal of Behavioral and Experimental Finance* 9, pp. 88–97 (Cited on page 85).

Clark, T. B. (2022). “The Relationship Between Adherence to COVID-19 Regulations and The Public Goods Game”. In: (Cited on page 70).

Council, C. C. (2015). “Braithwaite Flood Investigation Report”. In: *Flood Event 5–6th December 2015* (Cited on page 37).

Cox, D. J., J. E. Losee, and G. D. Webster (2022). “Discounting under severe weather threat”. In: *Weather, climate, and society* 14.1, pp. 65–79 (Cited on page 38).

Dale, M., J. Wicks, K. Mylne, F. Pappenberger, S. Laeger, and S. Taylor (2014). “Probabilistic flood forecasting and decision-making: an innovative risk-based approach”. In: *Natural Hazards* 70.1, pp. 159–172 (Cited on page 6).

Demeritt, D., S. Nobert, H. L. Cloke, and F. Pappenberger (2013). “The European Flood Alert System and the communication, perception, and use of ensemble predictions for operational flood risk management”. In: *Hydrological Processes* 27.1, pp. 147–157 (Cited on page 6).

Descamps, A., S. Massoni, and L. Page (2022). “Learning to hesitate”. In: *Experimental Economics* 25.1, pp. 359–383 (Cited on page 35).

Druckman, J. N. and C. D. Kam (2011). “Students as experimental participants”. In: *Cambridge Handbook of Experimental Political Science* 1, pp. 41–57 (Cited on page 10).

Exadaktylos, F., A. M. Espín, and P. Branas-Garza (2013). “Experimental subjects are not different”. In: *Scientific reports* 3.1, pp. 1–6 (Cited on page 2).

Fischbacher, U. (2007). “z-Tree: Zurich toolbox for ready-made economic experiments”. In: *Experimental Economics* 10.2, pp. 171–178 (Cited on page 9).

Fong, G. T., D. Hammond, and S. C. Hitchman (2009). “The impact of pictures on the effectiveness of tobacco warnings”. In: *Bulletin of the World Health Organization* 87, pp. 640–643 (Cited on page 16).

Gibson, J. (2022). “Government mandated lockdowns do not reduce COVID-19 deaths: implications for evaluating the stringent New Zealand response”. In: *New Zealand Economic Papers* 56.1, pp. 17–28 (Cited on page 69).

Greiner, B. (2015). “Subject pool recruitment procedures: organizing experiments with ORSEE”. In: *Journal of the Economic Science Association* 1.1, pp. 114–125 (Cited on page 9).

Grounds, M. A. and S. L. Joslyn (2018). “Communicating weather forecast uncertainty: Do individual differences matter?” In: *Journal of experimental psychology: applied* 24.1, p. 18 (Cited on page 35).

Harrison, G. W. and J. A. List (2004). “Field experiments”. In: *Journal of Economic Literature* 42.4, pp. 1009–1055 (Cited on pages 1 and 10).

HMS, H. M. S. (2021). *As coronavirus continues to spread, many questions and answers.* `https://www.health.harvard.edu/diseases-and-conditions/coronavirus-resource-center/`. Accessed December 22, 2021 (Cited on page 66).

Hopfield, J. J. (1988). “Artificial neural networks”. In: *IEEE Circuits and Devices Magazine* 4.5, pp. 3–10 (Cited on page 37).

Hoskins, B. (2013). “The potential for skill across the range of the seamless weather-climate prediction problem: a stimulus for our science”. In: *Quarterly Journal of the Royal Meteorological Society* 139.672, pp. 573–584 (Cited on page 37).

Islam, M. and B Sivakumar (2002). “Characterization and prediction of runoff dynamics: a nonlinear dynamical view”. In: *Advances in Water Resources* 25.2, pp. 179–190 (Cited on page 37).

Jacoby, J., D. E. Speller, and C. K. Berning (1974). “Brand choice behavior as a function of information load: Replication and extension”. In: *Journal of Consumer Research* 1.1, pp. 33–42 (Cited on page 8).

Joslyn, S., K. Pak, D. Jones, J. Pyles, and E. Hunt (2007). “The effect of probabilistic information on threshold forecasts”. In: *Weather and Forecasting* 22.4, pp. 804–812 (Cited on pages 7 and 15).

Joslyn, S. and S. Savelli (2010). “Communicating forecast uncertainty: Public perception of weather forecast uncertainty”. In: *Meteorological Applications* 17.2, pp. 180–195 (Cited on page 35).

Joslyn, S. L. and J. E. LeClerc (2012). “Uncertainty forecasts improve weather-related decisions and attenuate the effects of forecast error.” In: *Journal of experimental psychology: applied* 18.1, p. 126 (Cited on page 35).

Joslyn, S. L. and R. M. Nichols (2009). “Probability or frequency? Expressing forecast uncertainty in public weather forecasts”. In: *Meteorological Applications: A journal of forecasting, practical applications, training techniques and modelling* 16.3, pp. 309–314 (Cited on page 35).

Kahneman, D. and A. Tversky (2013). “Prospect theory: An analysis of decision under risk”. In: *Handbook of the Fundamentals of Financial Decision Making: Part I*. World Scientific, pp. 99–127 (Cited on page 17).

Kreamer, L., G. Stock, and S. Rogelberg (2021). *Optimizing virtual team meetings: Attendee and leader perspectives* (Cited on page 68).

Kreibich, H., A. H. Thieken, T. Petrow, M. Müller, and B. Merz (2005). “Flood loss reduction of private households due to building precautionary measures—lessons learned from the Elbe flood in August 2002”. In: *Natural Hazards and Earth System Sciences* 5.1, pp. 117–126 (Cited on page 38).

LeClerc, J. and S. Joslyn (2015). “The cry wolf effect and weather-related decision making”. In: *Risk Analysis* 35.3, pp. 385–395 (Cited on page 9).

Lee, Z. D., D. L. C. Yeu, B. T. Ang, W. H. Ng, and W. T. Seow (2020). “COVID-19 and its impact on neurosurgery: our early experience in Singapore”. In: *Journal of Neurosurgery* 133.1, pp. 24–25 (Cited on page 69).

Leonard, S. D. (1999). “Does color of warnings affect risk perception?” In: *International Journal of Industrial Ergonomics* 23.5-6, pp. 499–504 (Cited on pages 46 and 47).

Maggi, L., F. De Pellegrini, A. Reiffers, P. J.-J. Herings, and E. Altman (2014). “Coordination of epidemic control policies: A game theoretic perspective”. In: *2014 7th International Conference on NETwork Games, COntrol and OPtimization (NetGCoop)*. IEEE, pp. 180–187 (Cited on page 70).

Marimo, P., T. R. Kaplan, K. Mylne, and M. Sharpe (2015). "Communication of uncertainty in temperature forecasts". In: *Weather and Forecasting* 30.1, pp. 5–22 (Cited on page 8).

Mitroff, I. I. (1988). "Crisis management: Cutting through the confusion". In: *MIT Sloan Management Review* 29.2, p. 15 (Cited on page 16).

Mu, D., T. R. Kaplan, and R. Dankers (2018). "Decision making with risk-based weather warnings". In: *International journal of disaster risk reduction* 30, pp. 59–73 (Cited on pages 2 and 35).

Nadav-Greenberg, L. and S. L. Joslyn (2009). "Uncertainty forecasts improve decision making among nonexperts". In: *Journal of Cognitive Engineering and Decision Making* 3.3, pp. 209–227 (Cited on page 8).

Neal, R. A., P. Boyle, N. Grahame, K. Mylne, and M. Sharpe (2014). "Ensemble based first guess support towards a risk-based severe weather warning service". In: *Meteorological Applications* 21.3, pp. 563–577 (Cited on page 7).

NHC, C. (2008). *Hurricanes in History*. `https://www.nhc.noaa.gov/outreach/history/#katrina/`. Accessed April 4, 2021 (Cited on pages 1 and 5).

Organization, W. H. (2022). *WHO Coronavirus (COVID-19) Dashboard*. Tech. rep. World health organization (Cited on page 65).

Perez, E Coughlan de, B. van den Hurk, M. Van Aalst, B Jongman, T Klose, and P Suarez (2015). "Forecast-based financing: an approach for catalyzing humanitarian action based on extreme weather and climate forecasts". In: *Natural Hazards and Earth System Sciences* 15.4, pp. 895–904 (Cited on page 38).

Pfefferbaum, B. and C. S. North (2020). "Mental health and the COVID-19 pandemic". In: *New England Journal of Medicine* 383.6, pp. 510–512 (Cited on page 68).

Porporato, A. and L. Ridolfi (1997). "Nonlinear analysis of river flow time sequences". In: *Water Resources Research* 33.6, pp. 1353–1367 (Cited on page 37).

Roggeveen, A. L. and R. Sethuraman (2020). "How the COVID-19 pandemic may change the world of retailing". In: *Journal of Retailing* 96.2, p. 169 (Cited on page 69).

Roulston, M. S., G. E. Bolton, A. N. Kleit, and A. L. Sears-Collins (2006). "A laboratory study of the benefits of including uncertainty information in weather forecasts". In: *Weather and Forecasting* 21.1, pp. 116–122 (Cited on page 35).

Roulston, M. S. and T. R. Kaplan (2009). "A laboratory-based study of understanding of uncertainty in 5-day site-specific temperature forecasts". In: *Meteorological Applications: A journal of forecasting, practical applications, training techniques and modelling* 16.2, pp. 237–244 (Cited on page 7).

Rubinstein, A. (2007). "Instinctive and cognitive reasoning: A study of response times". In: *The Economic Journal* 117.523, pp. 1243–1259 (Cited on page 17).

Ryan, J. P. (1990). "Design of warning labels and instructions." In: (Cited on page 8).

Silic, M., D. Cyr, A. Back, and A. Holzer (2017). "Effects of Color Appeal, Perceived Risk and Culture on User's Decision in Presence of Warning Banner Message". In: *Silic, M., Cyr, D., Back, A., & Holzer, A.(2017, January). Effects of Color Appeal, Perceived Risk and Culture on User's Decision in Presence of Warning Banner Message. In Proceedings of the 50th Hawaii International Conference on System Sciences* (Cited on page 15).

Sivle, A. D. (2016). "Oh no, it's raining! A study of how information in online weather reports is interpreted, integrated, and used in everyday decision-making by laypeople". In: (Cited on page 12).

Sivle, A. D. and S. D. Kolstø (2016). "Use of online weather information in everyday decision-making by laypeople and implications for communication of weather information". In: *Meteorological Applications* 23.4, pp. 650–662 (Cited on pages 6 and 8).

Spoorthy, M. S., S. K. Pratapa, and S. Mahant (2020). "Mental health problems faced by healthcare workers due to the COVID-19 pandemic—A review". In: *Asian Journal of Psychiatry* 51, p. 102119 (Cited on page 68).

Stanton, J. M. and E. M. Weiss (2002). "Online panels for social science research: An introduction to the StudyResponse project". In: *Syracuse, NY: Syracuse University, School of Information Studies, Tech. Rep* 13001 (Cited on page 2).

Su, Z., A. Cheshmehzangi, D. McDonnell, J. Ahmad, S. Šegalo, Y.-T. Xiang, and C. P. da Veiga (2022). “The Advantages of the Zero-COVID-19 Strategy”. In: *International Journal of Environmental Research and Public Health* 19.14, p. 8767 (Cited on page 67).

Tafel, L., E. Ott, M. Brillinger, C. Schulze, and B. Schröter (2021). “Attitudes of administrative decision-makers towards nature-based solutions for flood risk management in Germany”. In: *Sustainability Science*, pp. 1–15 (Cited on page 36).

Tandon, A., C. J. Murray, J. A. Lauer, and D. B. Evans (2000). “Measuring overall health system performance for 191 countries”. In: *Geneva: World Health Organization* (Cited on page 69).

Toxvaerd, F. (2019). “Rational disinhibition and externalities in prevention”. In: *International Economic Review* 60.4, pp. 1737–1755 (Cited on pages 65 and 69).

Trainor, J. E., D. Nagele, B. Philips, and B. Scott (2015). “Tornadoes, social science, and the false alarm effect”. In: *Weather, Climate, and Society* 7.4, pp. 333–352 (Cited on page 27).

Tversky, A. and D. Kahneman (1973). “Availability: A heuristic for judging frequency and probability”. In: *Cognitive Psychology* 5.2, pp. 207–232 (Cited on page 9).

UNISDR (2006). *Design of warning labels and instructions, UNISDR. Global Survey of Early Warning Systems: An Assessment of Capacities, Gaps and Opportunities Towards Building a Comprehensive Global Early Warning System for All Natural Hazards. United Nations Inter-Agency Secretariat of the International Strategy for Disaster Reduction.* `https://www.undrr.org/publication/global-survey-early-warning-systems/`. Accessed December 22, 2021 (Cited on page 6).

Wald, A. (1945). “Sequential method of sampling for deciding between two courses of action”. In: *Journal of the American Statistical Association* 40.231, pp. 277–306 (Cited on page 35).

WHO (2020). *WHO Coronavirus (COVID-19) Dashboard.* `https://covid19.who.int/`. Accessed December 22, 2021 (Cited on page 1).

Wickens, C. D., K. Gempler, and M. E. Morphew (2000). “Workload and reliability of predictor displays in aircraft traffic avoidance”. In: *Transportation Human Factors* 2.2, pp. 99–126 (Cited on page 8).

Winett, R. A. and J. H. Kagel (1984). “Effects of information presentation format on resource use in field studies”. In: *Journal of Consumer Research* 11.2, pp. 655–667 (Cited on page 8).

Yegnanarayana, B. (2009). *Artificial Neural Networks*. PHI Learning Pvt. Ltd. (Cited on page 37).