8. SUPPLEMENTAL MATERIALS

This paper encompasses several methods, each with their respective data sets. In this appendix of supplementary material, we aim to make our processes for data collection and analysis fully transparent. Study 1 is qualitative in nature and employs think-aloud protocols. Studies 2 and 3 are behavioral experiments, one of which is a preregistered replication. Below we present the tasks assigned to study participants in greater detail, as well as robustness checks for the clustering employed in Study 1. Further descriptions of the tasks and data are available at the Open Science Framework repository for this project at: https://osf.io/eh5m2/

In this repository, we included the data-collection and data-analysis protocols used in all the main studies of this project, as well as all preregistered data collections. For Study 1, we included a guide for collecting think-aloud protocols, and two fully coded verbal protocols. These protocols are the verbatim transcripts of Person A and B we present in Study 1. We further explain, step by step, how we coded the data, created sequences from the coded transcripts, and stored these sequences in transition matrices. We then provide the 49 transition matrices from Study 2, the code we use to create each table, and the robustness check of this study, including the robustness check available in these supplementary materials.

In the Open Science Framework repository, we also include the surveys used to collect the data for Studies 2 and 3. These surveys include the tasks employed in the study and all the extra questions we asked the participants. For the purpose of this project, we report the main variables used in hypothesis testing. However, we collected much more detailed data. Therefore, in the repository, we include this data, and the scripts used to transform it from time-traced mouse clicks to the aggregate variables used in hypothesis testing. This data is important to allow future researchers to trace the data analysis we did to the one we stated in the preregistration of Study 3—but, more generally, to allow future researchers to use the data more freely and not be bound to our variable definitions.

Below, we present the three tasks employed in this study. Furthermore, we explain when and how we collected data and how the data collection was designed and handled.

8.1 Tasks used in this paper

The three problems used in this paper have been published and validated. We present each below:

8.1.1 "Karabayos" problem (Laureiro-Martínez and Brusoni 2018)

In Study 1, we rely on a task that was used as an ill-structured problem by Laureiro-Martínez and Brusoni (2018). This task asks participants to imagine themselves as "the leader of the Karabayos," an Amazonian tribe. The Karabayos tribe faces a set of threats and challenges, and its leader's task is to imagine what to do to save the tribe. The full text and layout of the task as given to the participants are shown in Figure A.1.

Insert Figure A.1 about here.

8.1.2 "Winter survival" problem (Johnson and Johnson 1982, p. 111)

Participants in Study 2 were presented first with the "winter survival" problem of Johnson and Johnson (1982, p. 111). This task requires participants to imagine themselves at the site of a plane crash that they have just survived. It is now midday, the temperature is freezing, and they find themselves in a forest with a group of fellow survivors who will follow their commands. They have a list of 12 items they can use to survive the night, and they must rank them based on their importance to their survival. The participants are told there is a town 30 kilometers away; this information complicates the problem, as participants have to determine whether they prefer to stay near the crash site and await rescue or walk to the town in the freezing temperatures.

To finish the problem, participants need to move the 12 items into the order they deem appropriate. Figure A.2 presents the graphical user interface that participants used during the task, which includes the full text of the task. Figure A.2 shows the task before any moves have been made. In order to finish the task, the participant needs to drag and drop all items from the left column to the right column. After all the items have been placed, a button asking the participant to "confirm final ranking" appears on the screen. This button is shown in Figure A.3.

Insert Figure A.2 about here.

Survival experts gave Johnson and Johnson (1982) an optimal ranking of the items. They determined that, first of all, staying near the crash site was the best decision. To survive until help arrives, it is of utmost importance to prioritize heating and food supplies. Table A.1 shows the 12 items in the order they are shown to the participants and the correct positions in which they should be ranked, according to the experts.

Insert Table A.1 about here.

We adapted the units to the International System of Units and changed the names of the items slightly after pilot sessions showed us that some items were hard for participants to understand (e.g., "can of shortening"). We kept the original formulation from Johnson and Johnson (1982) and added short explanations in parentheses. Additionally, on the tiles that participants needed to drag and drop, we added a small image of the item to make it easier to recognize. We took the same approach in the case of the "NASA survival" problem presented in the next section.

8.1.3 "NASA survival" problem (Hall and Watson 1970)

After performing the "winter survival" problem, participants were shown the manipulations. The manipulations are explained in section 7.3. After the manipulation, participants solved the "NASA survival" task (Hall and Watson 1970). This task, like the "winter survival" problem, requires participants to imagine themselves at a crash site. In this case, however, they have crashed on the surface of the moon along with the crew of their lunar module. There is only one possible course of action: they must try and reach the meeting point. Participants are given 15 items that can help them reach the meeting point. The task is to rank these 15 items in terms of their importance in allowing the crew to reach the meeting point. As in the "winter survival" problem, Hall and Watson (1970) obtained an expert ranking and used this ranking to compare the participants' responses to an objective measure. Table A.2 shows the items in the order shown to the participants and each item should be ranked in the correct position.

Insert Table A.2 about here.

Figure A.3 presents the implementation-disposition condition's graphic interface. It shows the interface after all items are placed in the right-hand column. It also shows the "confirm final ranking" button that, once clicked, takes participants to the next stage in the experiment. The sentence in red text changed in every experimental condition, as the next section explains in more detail.

Insert Figure A.3 about here.

In Table A.3, we provide a point-by-point comparison of the "winter survival" and "NASA survival" problems.

Insert Table A.3 about here.

8.2. Coding of think-aloud protocols in Study 1

8.2.1 Content analysis

The think-aloud protocols were transcribed verbatim by research assistants following accepted standards for the process (Poland 1995; McLellan et al. 2003). We analyzed participants' verbalization using content-analysis techniques (Neuendorf 2002; Krippendorff 2012). Since we departed from verbalizations, we avoided punctuating the transcriptions and let the raters decide where each phase started and finished. This procedure is the same one used by Simon and Ericsson, who called the sets of words "chunks of thought". Since each rater defined the chunks of thought and added the label to each, it is likely to have a lower inter-rater reliability because selecting the exact words under each phase adds additional sources of differences. On the other hand, by not punctuating the text artificially this process stays closer to the idea of coding thinking processes.

The protocols were coded according to the seven phases of the combined problem-solving model presented in the theory section. Table 2 presents a more detailed view of how the coding was operationalized, including the seven problem-solving phases, a short description of the construct, the type of processes involved, and a quote representative of each specific code. The initial phase is frame stating (FS), in which the participant analyzes the problem by repeating or paraphrasing the data mentioned in the problem description (a text provided to each participant). Frame assuming (FA) occurs when the participant develops their own hypotheses and assumptions about the problem at hand

and begins taking them for granted, even when these were not mentioned in the problem description. Direction setting (DS) defines general paths one intends to follow without stating a specific proposal or generating alternative proposals for what to implement later on. Evaluation (EV) occurs when the participant judges the merits of a proposed path and considers the solution without evaluating its specific details. The decision (DE) phase is when the participant manifests a clear choice regarding what they intend to do. In implementation (IM), the participant designs the sequence of actions to carry out their proposals. The seventh stage is implementation evaluation (IE), where the participant evaluates the feasibility of their implementation. We coded any unintelligible sounds as "babble."

We should highlight that to achieve a more objective interpretation of the think-aloud protocols, the researchers were involved in refining and piloting the code but not in the actual coding process. We calculated the average percentage of agreement, which was 92.9%. The average agreement is helpful in the case of simple codes, but when the data are complex, prior studies recommend using Cohen's κ. For this metric, we obtained a moderate value of 0.51 (Cohen 1960). Both values are satisfactory for the type of text we studied (Neuendorf 2002; Lombard et al. 2002). However, further data cleaning is necessary to converge to a protocol where all processes are coded into one problem-solving phase. We engage in this process in the next section.

8.2.2 Code convergence

Each rater provided a fully coded transcript for each participant's protocol. Although we achieved moderate reliability, a perfect match for every word in every protocol was implausible. However, a prerequisite for sequence analysis is that each passage (representing chunks of thought) must be assigned a single code. Therefore, we followed a second content-analysis process where we compiled the coded transcripts of each rater and followed a simple process of code convergence. By code convergence, we mean taking the cases in which multiple codes were given by the raters for a single passage and converting them to a single code.

Our code-convergence process consisted of two steps. First, in cases of partial agreement (i.e., two raters select the same code, and one disagrees), we saved the value chosen by the majority. Second, in cases of complete disagreement among the raters (i.e., all three assign different codes), two authors conferred and selected the appropriate code for the passage in question from the three codes proposed by the raters. The output of these steps was a fully coded transcript in which every passage was coded into a single problem-solving phase. The result was a sequence of phases for each participant that represented their entire problem-solving process. At this stage, we removed the babble codes (which accounted for 2.8% of the protocols in total) from the sequence since they do not represent the problem-solving process.

8.2.3 Sequence analysis

Next, we shifted our attention from the content of the phases to the transitions between them. To do this, although the duration of phases can vary widely, we assigned them all the same unitary length for this analysis (Pentland 2003; Salvato 2009). Figure 3 illustrates the problem-solving sequences of two participants (Persons A and B). These two sequences provide granular data of the typical problem-solving strategies that emerge later in the study (Greve 2018). Each shows the problem-solving phases as color-coded rectangles in Figure 1. The two problem-solving sequences differ considerably, although both employ all seven problem-solving phases. Person A focuses more on frame stating and assuming and only spends time on implementation and implementation evaluation toward the end. Person B, in contrast, performs frame assuming and frame stating on far fewer occasions and focuses on implementation and implementation evaluation earlier and more often.

8.2.4 Transition matrices

We create a transition matrix from each sequence that reduces the variance between problem-solving processes and provides a more comparable data structure between participants (Gibbs et al. 1971, Pentland 2003). Following Lipshitz and Bar-Ilan (1996), each cell in the matrix represents a transition between phases. The starting phase of the transition is given by the cell's row, while the destination phase is denoted by its column. Given the seven phases in this study, there are 42 transitions between phases. The 42 values are entered in off-diagonal cells of the transition matrix and represent all the transitions made by each participant during the problem-solving process. We normalized these values to obtain transition numbers comparable among participants; for each protocol, the sum of all transitions (i.e., off-diagonal cells) sum to 1.

In addition to the transition between phases, we included the percentage of time spent in each of the seven problem-solving phases in the analysis. In sequence analysis, it is common to have transitions within the same phase. However, the thoughts recorded in the think-aloud protocols do not have clear transitions for coding within-phase transitions. As a proxy for within-phase transitions, we take the percentage of time spent on each of the seven phases. Therefore, for each participant, we have 49 individual values that describe their problem solving.

8.2.5 Clustering

Although we created 49 variables to characterize a problem-solving process, we performed cluster analysis on these data to reduce its dimensionality to a single categorical variable. We use clustering algorithms to identify common patterns of attention our participants used when solving a wicked problem—the strategies they followed (Ocasio and Joseph 2018, p. 289). We employed a clustering method called *partitioning around medoids* (Kaufman and Rosseeuw 1990; Hennig 2015). This method first selects k, the best number of clusters for the data set. It then groups the rest of the participants around the k most representative participants, called "medoids." Compared to others, the benefit of this method is that its clustering output is deterministic. After completing the clustering, we were left with a categorical variable that assigned each think-aloud protocol to a single cluster. In our case, the cluster is a dichotomous variable.

Note that this dichotomous variable is the key outcome of the processes used to structure and reduce the dimensions in the data. We use this variable to understand of how wicked problems are solved, how problem-solving strategies differ, their commonalities and differences.

8.2.6 Robustness checks

As a robustness check for the clustering procedure in Study 1, we reran the *pamk* clustering analysis in a specific way (Hennig 2015). We removed one participant from the sample and clustered the remaining 47. We repeated this 48 times, removing each participant once and clustering the remaining participants together. The motivation for this is that when participants are removed, the medoids can change, and the partitioning-around-medoids clustering method can potentially determine that a different number of clusters are needed, that new medoids are found, or that participants should be clustered together in very different ways. By removing one participant at a time, we could observe how robust the results of the full-sample clustering were.

We removed each participant from the clustering once. This left us with 48 categorical clustering variables. After doing this for each participant, we realized two things. First, in 46 of the 48 cases, the partitioning-around-medoids method selected two groups as the best number for the analysis. Thus, the results we show in Study 1 were not a fluke but the common number of clusters for our dataset. Of the two cases where more than two clusters were selected, we found that the protocols were separated into three groups in one case. Interestingly, only one person (originally classified as *solution-focused*) was placed alone in the third group. In the other case, the protocols were clustered into four groups. We found that the *problem-focused* group remained unchanged, but the *solution-focused* group was separated into three clusters. Of the 20 participants of the *solution-focused* group, 12 were selected into one group, seven into another, and the participant who was assigned their own one-person cluster before was selected into their own cluster once again. From this, we obtain our first finding—namely, that two clusters are most often the best way to separate the think-aloud protocols. Additionally, in the rare case when more clusters are required, the separation into *problem-focused* is stable, the only thing that changes is how the *solution-focused* cluster is defined—mainly due to a single protocol.

Our second finding goes down one level and studies the misclassifications. We find misclassifications are uncommon and occur with protocols on the border between the two strategies. To estimate this, we performed a second set of analyses on the 46 cases where the clustering gave two groups. In these cases, cluster assignment varied: some participants might be assigned to the *problem-focused* group on one occasion and the *solution-focused group* on another. We call the cases where a participant was classified in a different cluster than in the 48-group cluster a misclassification. We estimated how frequent the misclassifications were and discovered that of the 2160 (46x48-48) classification events, only 141 were cases where a protocol was classified differently than in the full sample. The protocols were classified correctly 92.9% of the time. Additionally, six participants accounted for 90% of the misclassifications (127 of 141); these participants were classified around 50% of the time as *problem-focused* and 50% of the time as *solution-focused*. We learned that misclassifications are uncommon, happening around 6.6% of the time and that they are over-represented among a few participants (six out of 48) who are at the border of being classified as *problem-focused* or *solution-focused.* This finding, together with the finding that two clusters are a robust separation of the think-aloud protocols, led us to see the results of the full-sample clustering as robust— in terms of both the number of clusters, and the classifications of each participant to each cluster.

8.3 Research design of Study 2

Study two followed a mixed factorial experimental design, as it mixed a between-subject design and a within-subject design (Oehlert 2010; Anderson and McLean 2018). This is because the experiment had three experimental conditions (between-subject design), and each participant performed two tasks—one before the manipulation and one afterward (within-subject design). We follow this procedure to better pinpoint the causality of the attention-focus mechanism by reducing the amount of unexplained variance in our analyses and studying only the behavioral change induced by the manipulations.

We hosted the online experiment using the Qualtrics platform. Qualtrics provides multiple tools for creating tasks, such as Likert scales, multiple-choice quizzes, and even drag-and-drop interfaces. However, in our experiment, we wrote our own interface in JavaScript. We did this to capture detailed timings of every movement the participant made while solving the task—every click, drag, and drop. We recorded and timed each of these events, as well as the time it took the participant to read the task, and to leave the task once their ranking was finished. The motivation for this level of detail was to understand the thinking processes of the participants beyond reaction time and solution performance (Yu et al. 2012). We followed the methods of studies that have used mouse actions to provide a "*more direct measure of the evolution of a particular response*" (Freeman 2018). It has been found that mouse movements are more representative of cognitive processing than other subjective measures, such as self-reports (Fedor et al. 2015), and have proven a reliable means for examining continuous cognitive processes in real-time (Yu et al. 2012), providing a good proxy for the analysis of thinking processes (Ollinger et al. 2013).

We recruited participants for our experiment through the online platform Prolific and paid them based on their performance. Prolific is a dedicated research-subject pool and recruiting platform employed in multiple studies in recent years. We ask the platform to direct participants who are native English speakers, have at least a Bachelor's degree, and are between 25 and 55 years old (both inclusive). For comparisons between Prolific and other online participant recruitment platforms, see Palan and Schitter (2018) and Peer et al. (2017). The top quartile of participants received twice the base rate of the platform for the 30 minutes of solving the experiment: a total of £5 (British pounds sterling; GBP). The middle half received 1.5 times the base rate of the platform: £3.75. The bottom quartile was paid £2.50: the base rate for the half an hour the experiment required of the participants.

Figure A.4 presents a depiction of the experimental procedure of Study 2. An experimental session unfolded according to the following sequence. First, Prolific referred us a participant. The participant arrived at the experiment and was welcomed. We requested their compliance with the experiment's conditions and presented the incentives scheme, and the participant performed an attention check regarding the incentive scheme. After passing the attention check, the participant performed the first task: the "winter survival" problem. From this task, we saved the final ranking; the reading, processing, and total time spent on the task; the number of moves; and the time each move took.

Insert Figure A.4 about here.

After the participant confirmed the solution, they were asked to write down the motivation for how they built the ranking, and afterward to explain to us the process of how they built the ranking. Specifically, we asked them to "imagine what a tape recorder would play if it had recorded what you thought about when ranking the items." After they finished writing, we asked the participants to select, from a list of prototypical processes, the one that most closely resembled their own thinking process. Given that they had just spent time writing about their thinking process, they had a deliberate comparison from which to self-select. Additionally, participants needed to answer an attention check at this point.

Up to this point, the experience of every participant in the study had been the same (i.e., the three conditions were hitherto indistinguishable). The next step was the manipulation, which was shown as a transition page between the two tasks. We created three pages, one per manipulation, to incentivize all participants similarly except for the focus of their attention.

In the manipulation, we wanted participants to focus their attention on the specific phases of problem-solving in which the strategies of Study 1 differed. To do this, we took the coding scheme the raters used to code the problem-solving phases in Study 1 and created recommendations for the participants directly out of the coding scheme. By using the same language as the coding scheme, we can be closer to the original difference in attention focus found in Study 1, even if this difference is not linked to a theoretical or cognitive mechanism.

The manipulations had three parts. Figures A.5, A.6, and A.7 show screen captures of the participants' graphical interface on each of the three conditions. The first part explained that experts recommend individuals should spend most of their time thinking about either: a) "the problem in the way that feels most natural to them" in the case of the *control* condition, b) "the framing of the problem" in the case of the *framing disposition* condition, or c) "the implementation of the solution" in the case of the *implementation disposition* condition. The literature on naturalistic decision-making (Klein 2017), planning (Steiner 2010), and forecasting (Tetlock and Gardner 2016) give recommendations that could be interpreted as suggestions to focus on ways that follow either of the three recommendations. So, given that there is no scientific consensus, and we can find research findings backing all three manipulations, we did not deceive the participants. The three recommendations could lead to either better performance or no performance difference; in this study, we give support to the latter.

Insert Figure A.5 about here.

The second part of the manipulation was present only in the treatment conditions (i.e., the *framing disposition* and *implementation disposition* conditions). This part explained what thinking about the "framing of the problem" or "implementation of the solution" means. The explanation was given in three bullet points. The bullet points were précis of the coding schemes for the phases used most differently by the two strategies identified in Study 1. Namely, the three bullet points of the *framing-disposition* condition came from summarizing the *frame stating* and *frame assuming* phases. In the case of the *implementation disposition* condition, the bullet points were summaries of the coding schemes of the *implementation* and *implementation evaluation* phases.

Insert Figure A.6 about here.

The third and final part of the manipulation slide was a reminder for the participants to focus their attention, as shown in Figures A.5, A.6, and A.7. After reading the manipulations, the participants started the "NASA survival" problem. The interface for the three conditions was the same, except for the sentence shown in red (see Figure A.3). In this sentence, we reminded participants to "Please direct your attention and effort" in either: a) "ways that feel natural to you"; "to think about the framing of the problem"; or "to think about the implementation of the solution." The sentence is the last part of the manipulation we gave the participants.

Insert Figure A.7 about here.

We recorded the same variables in this task as in the "winter survival" problem (e.g., rankings, times, number of moves, process self-selection). After the participant finished the "NASA survival" problem, we asked them to answer a set of control questions. In between the control questions, we included an attention check. The control questions had two aims: first, to understand whether the participants had understood the task and paid attention, and second, to collect demographic information about the participant. The main questions on this second theme were *age*, *gender*, *education level*, *reading habits*, and *experience in survival training*. These questions were important during the piloting process and helped us to improve the experimental design.

We removed participants with experience in survival training because they might not see the novelty in the problems—or, at least, less than other participants did. A participant who sees the problem as less novel will also see it as less strategic. Study 2 aims at replicating the conditions of Study 1, where the problem was novel to all participants. Thus, removing the participants with experience in the contexts presented was a way of maintaining the similarity between the studies. After the control questions were answered, we thanked the participants and referred them back to Prolific, where they could claim remuneration for partaking in our experiment.

From the pilot studies, we estimated an effect size f of around 0.2 for this study. With this, we were able to estimate the sample size based with G*Power 3.1 (Faul et al. 2009). We specify a three condition by two task ANOVA with effect size *f* of 0.2, α of 0.05, and power of 0.95. The control condition was larger to understand better the baseline behavior and the correlational results we found while piloting this experiment.

The experiment was held at the end of May 2018. The experiment included 523 participants, from which 51 were removed as they had participated in survival training. Out of the 472 remaining participants, 276 were in the control group, 97 in the *framing disposition* condition, and 99 in the *implementation disposition* condition.

8.3.1 Piloting of Study 2

Four pretests and two pilot studies preceded Study 2. The first four pretests were mainly aimed at selecting the tasks, refining the instructions, and polishing the computer interface. The two pilot studies were mainly aimed at refining the manipulation and incentive scheme. After each study, we held debriefing sessions.

For the initial four pretests, we included three tasks" the "NASA survival" task, the "winter survival" task, and a case analysis regarding the LEGO Corporation circa 2003 obtained from Wellian (2010), which we later eliminated. We had initially attempted to capture participants' thinking process by asking them to describe it briefly. The four pretests took place from March–August 2017 and were carried out with samples of 21, 19, 60, and 61 participants. The pilots were run on the Prolific platform, and participants were incentivized with a flat base rate. From the pretests, in addition to our main goal of polishing the interface, we learned that participants found it very difficult to abstract their thinking processes cleanly. They mixed up the description of their thinking process with justifications about the motivation for their choices. In later studies, we asked participants first to explain their motivation, later explain their process, and finally match their thinking processes to the prototypical quotes we wrote to resemble *problem-focused* or *solution-focused* problem-solving processes. This change seemed to help, but still, participants acknowledged that they were frequently unaware of their own thinking process, meaning that the self-classification measures were of limited use. As a result, we dropped the case analysis regarding the LEGO Corporation after the first pilot. We dropped the problem because we needed each participant to solve two problems for our experiment, and solving two lengthy cases could easily lead to cognitive depletion—an interesting topic for future studies, but not our purpose.

After the pretests, we explored different ways to manipulate the participants' attention. Our first attempt at creating a manipulation that could be at the root cause of the different strategies observed in Study 1 involved using episodic future thinking. Episodic future thinking is the ability of the brain to imagine future events in detail (Schacter et al. 2017). We attempted to recreate the solutions observed in the *solution-focused* cluster by inducing episodic future thinking. In October 2017, we ran the first pilot study, the first with enough statistical power to refute hypotheses (Faul et al. 2009). The first pilot study was run in Prolific and involved 158 participants. Our assumption for this pilot study was that the *implementation disposition* condition would use more episodic future thinking and imagine how the solution could be implemented in the future at the expense of focusing on the problem in the present. In contrast, the *framing disposition* condition would be more present-focused. To test this, we ran a pilot study with two conditions (*high episodic future thinking* and *control*), but the manipulation was ineffective. We did not find an effect on the manipulation checks regarding the use of episodic future thinking and temporal focus, nor did we find an effect on performance or behavior. We found some correlational results on the participants' self-selection, but nothing causal.

After the first pilot study, we decided to remove our interpretation from the design of the experiment and strip down the manipulation to the core of the actual differences we found in Study 1 (i.e., the attention paid to different phases of the problem). We then created manipulations to incentivize participants to focus their attention on either *implementation* and *implementation evaluation* or *frame stating* and *frame assuming.* We tested the manipulations in April 2018 with our second pilot study, which included 46 participants. The key learning from this pilot was that participants forgot the incentive scheme at the end of the experiment. We needed to incentivize participants' performance more directly and check whether they had understood the incentive scheme. We found out that an attention check at the end was too late. Better results were obtained by inserting an attention check on the incentive scheme at the start of the experimental procedure. With this amendment, we ran the experiment shown in Study 2 in May 2018.

Finally, we ran a small scale, follow-up study with a group of 51 student volunteers during June 2018. This study helped us understand individuals' actual thinking processes, and allowed for an in-depth qualitative debriefing of the final experimental design presented in the paper. We validated our expectations, but more importantly, we learned two points. First, we gained a qualitative understanding of the quantitative measures we obtained from the past behavioral experiments we had run. For example, participants elaborated on what was going on in their heads when they were taking a long time before a drag-and-drop move (i.e., they were thinking carefully about how the solution in their heads might unfold through time), or what was going on when they were making many moves (i.e., they were trying to connect different elements of the problem, and thinking about their use and possibilities for different sub-goals). Second, we checked for the ecological validity of the task by asking the students—most of whom had years of experience as managers—to draw parallels between the experiment tasks and some real work situations. While not all were able to perceive parallels straight away, some made immediate connections with personal challenges they had faced as managers, when they had ensured the survival of their startup or their business unit, and the jobs of their employees under stressful conditions. Overall, while time-consuming, the various studies were a fundamental source for building the final research design for Study 2 and for understanding its robustness.

8.3.2 List of variables used in Study 2

In Study 2, we collected detailed measures of the participants' problem-solving processes. In the paper, we discuss four behavioral change variables named *total time*, *time per move*, *number of moves*, and *performance*. To build each of these variables, we need one variable from the "winter survival" problem (before the manipulation) and one from the "NASA survival" problem (after the manipulation). In addition to these variables, we collected the age, gender, education level, and reading habits of the participants, referred to as a *reader*. Finally, we collected two variables that relate to the choice participants made when they self-selected their thinking process into texts representing prototypical *problem-focused* or *solution-focused* problem-solving strategies. The variable *solution-focused* is a dichotomous variable with a value of 1 if the participant said the text closest to their thinking process was a prototypical *solution-focused* process, and -1 if they chose a *problem-focused* process. The variables and brief descriptions are shown in Table A.4.

Insert Table A.4 about here.

Table A.5 shows descriptive values and zero-order correlations of the variables collected for Study 2. Note that for the exception of the variables related to the condition the participants were assigned, all dichotomous variables are presented in contrast coding (-1 and 1) instead of binary coding (Davis, 2010). It is important to note that our data collection method allowed us to record the step-by-step process by which participants built their solutions. This allowed us to create analyses similar to the move-and-time analysis from Fedor et al. (2015). This data is not included in this document, as here we are focused on the effects of the manipulation on behavioral change.

Insert Table A.5 about here.

Table A.6 presents the equivalent robust regressions of the ordinary linear squares regressions shown in Table 9 of the main manuscript.

Insert Table A.6 about here.

8.4 Replication studies

As part of the first revision to the paper, we ran two new experiments. For the second revision, we ran one more experiment. The three experiments are preregistered in the Open Science Framework and publicly available. You can find them on the following three links:

- **Pandemic pilot study:** https://osf.io/a7sm5
- **Study 3:** https://osf.io/nvfdc
- **Post-hoc manipulation check study:** https://osf.io/abg8v

The first data collection is what we refer to as the pandemic pilot study. The pandemic pilot study is a preregistered experiment where we first introduced the pandemic control. We included the pandemic control variables at the end of the study and increased participants' payments above the minimum levels required by the hosting platform, Prolific.ac. In this study, we found that the participants behaved differently to Study 2, spending much more time than before on each task.

We learned from this mistake, and in the second preregistered experiment, we collected the pandemic control variables in a follow-up study. The second preregistration is an exact replication of Study 2, but with a larger sample size and a set of pandemic control variables collected in a follow-up survey to keep the replication as close as possible to the original experiment. The data collected for this preregistration is what we refer to as Study 3.

Lastly, for the second revision we collected ran a preregistered experiment. The experiment was a direct replication of Study 2 and Study 3. However, after the second task is finished, there is an added question to perform a manipulation check. In the prior studies, we had open questions for manipulation checks. Yet, in this study, we added a further question to have a quantitative view of how the manipulations affect the mental activities the participants employ while solving the tasks.

In the next three sections, we present the three preregistered experiments in more detail.

8.5 Pandemic pilot study

As part of the first revision of this paper, we preregistered an experiment where we aimed to test whether the task order could have biased the results of Study 2. In doing this, we followed the same experimental design as in Study 2—a three-condition by two-task design. The sample size was chosen to test effects of size f larger than 0.2, thus requiring 98 participants per condition; 294 in total. The experiment employed counterbalancing, and thus we also had two different task orders, giving a total of six experimental blocks, each with 49 participants.

The experiment was held in late April 2020, in the midst of the COVID-19 pandemic. During the pandemic, survival is more present in people’s minds. This is problematic for our study, as we aim to follow participant’s problem-solving as they solve a novel task. In Study 2, we excluded people who had survival training from the study to ensure the novelty of the tasks. We kept that filter in this study, but given that the pandemic brings survival to mind, we also include a set of pandemic control variables. These control measures are taken from psychology, which has studied how the threat of disease affects people’s decision-making (Taylor 2019).

We employed six different pandemic control scales in total. Three of them are taken from Taylor (2019): the Monitoring/Blunting coping style measure (Miller 1987; Steptoe 1989), Perceived Vulnerability to Disease (Duncan et al. 2009), and the Intolerance to Uncertainty scale (Carleton et al. 2007). These scales gave us a comprehensive view of how the participants would react to a disease or threat. However, they did not control whether the participants saw the COVID-19 pandemic as a threat or not. For that, we employed three other scales. First, the Fear of COVID-19 scale from Ahorsu (2020) and two sets of financial questions of our own making. One of the scales asks participants if they or someone else in their household has lost their job or a significant part of their income in the crisis resulting from the COVID-19 pandemic. The second set of questions asks whether the participants find their work in Prolific to be more important now, during the COVID-19 pandemic, than before. We asked this both in general and also financially. Table A.7 summarizes the pandemic control variables added to the pandemic pilot study.

Insert Table A.7 about here.

The combination of the six scales helps us obtain a more nuanced view of how COVID-19 has affected the participants. In all cases, the higher the scales, the more a participant has been affected, or the more they would react to perceived diseases or threats. Therefore, in using the six scales as control variables, we can account for part of the latent change in experimental conditions that COVID-19 brings to our data collection.

8.5.1 Changes to research design

The use of counterbalancing and the inclusion of the pandemic control variables are the two changes made in this pilot study compared to Study 2. The use of a counterbalanced design leads to a three-condition by two-task-order factorial design with six blocks. We follow Study 2 and use effect size f of 0.2 and α of 5% and power of 0.95 to determine the sample size. We require 49 participants per block, giving a total of 98 participants per condition and 294 participants in the full study.

Adding the pandemic control variables increases the time required to finish the experiment. In Prolific.ac, the platform that provided the participants, participants were shown the list of open studies, the base payment available from the study, and the expected time to finish it. Including the pandemic control variables led to an increase in the stated duration of the study. In Study 2, we told participants that the study would take 30 minutes. In the pandemic control pilot study, we told participants they would require 45 minutes to finish. This increase in time led to a change in the participants' behavior: they not only spent longer on the entire study but allocated more time to solve the "winter survival" and "NASA survival" problems. We explain these differences in the next section.

8.5.2 Comparison of Study 2 and Pandemic Pilot

In this study, we give participants performance-based incentives in the form of a bonus. The incentive scheme is the same as in Study 2; top performers earn up to twice as much as the lower-performing participants (Hertwig and Ortmann 2001). This performance-based payment is calculated from the solutions given to the "NASA survival" and "winter survival" problems. These tasks are shown at the beginning of the study, and participants know that only these two tasks affect their performance.

In the pandemic pilot study, we found that the participants' baseline behavior changed when compared to Study 2. Table A.8 shows the mean and standard deviation of the *total time, time per move, number of moves,* and *performance* of all participants in the pandemic pilot study—a total of eight variables[1]. Table A.8 compares the behavior of the participants in Study 2 and the pandemic pilot study. We calculate the t-test on each of the eight variables in Table A.8 between all the participants in Study 2 (N=472) and all the participants of the pandemic control study (N=294). We find that the *total time* spent on each problem by participants in the pandemic pilot study is much higher than the time spent in Study 2. The same is true for the *time per move,* although only marginally in the "winter survival" problem. Finally, the *number of moves* in the "winter survival" problem is higher in the pandemic pilot. In general, participants deliberate more in the pandemic pilot study than in Study 2.

Insert Table A.8 about here.

We argue that participants expended more effort because we told them that they would require 50% longer to finish the experiment and because they know that their payment depends only on their

[1] Note that the table mixes the order in which participants see the tasks. The large deviations are also present if we show instead the only the baseline of people who saw each task

performance in the "winter survival" and "NASA survival" problems. We changed this in Study 3. The second preregistered experiment we performed during the review process.

8.5.3 Task-order bias

The mixed-factorial design of this experiment and Studies 2 and 3 (presented below), should not be affected by standard task-order bias. In our studies, we use the participants as their own baselines, given that they perform two tasks. A simple within-subject design experiment would lead to a task-order bias. This does not happen in the mixed factorial design because the average treatment effect is calculated between experimental conditions (i.e., between participants). Thus, the task-order effect is canceled out. The effect is canceled because we compare experimental conditions, and the task order can be assumed to affect all conditions the same.

In the pandemic pilot study, we employed a counterbalance research design to test the assumption that our results were not biased by the order in which the participants solved the tasks. We find support for this expectation by failing to reject the null-hypotheses of zero difference between participants' behavior if they solve the "NASA survival" problem first or the "winter survival" problem first. The results are shown in Table A.9. This table presents the three standardized change variables, the main variables of this study. It compares the mean of participants who solve the "winter survival" problem first with the mean of the participants who solve the "NASA survival" problem first. The three t-tests are non-significant. The lowest p-value is 0.189, with the highest absolute t-statistic being 1.317.

Insert Table A.9 about here.

The addition of the pandemic control variables increased the effort participants made to solve the tasks during the pandemic pilot study. This deviation makes the study a failed replication of Study 2. However, the pandemic pilot study did allow us to test our expectation that the order of the tasks should not bias the mixed factorial design. Having found support for this expectation, we removed the counterbalancing of task order in Study 3 (shown below).

8.6 Study 3

From the pandemic pilot study, we learned that to replicate Study 2, we needed to keep the instrument used for collecting the responses as close as possible to the one used in Study 2. To do this in Study 3, we used the exact survey employed in Study 2, with the same remuneration scheme, variable definition, and data analysis scripts. The direct replication allowed us to gain trust in the validity of the results and in the way, the variables were defined and hypotheses tested in Study 2. The experiment was held in late May 2020.

8.6.1 Changes to the research design

To guarantee that the data collection in Study 3 was as similar as possible to Study 2, we used the same survey as in Study 2. However, several days after the experiment was finished and all participants had given their responses, we opened up a new study in which only respondents of Study 3 were allowed to participate. This second study included the pandemic control variables. There was a small dropout rate, as some participants did not respond to both surveys, but over 92.8% responded to the second survey.

In addition to using a second survey, we employed a larger sample in Study 3 than in Study 2. We did this because ex-post effect-size analyses in Study 2 showed us that the average effect size f of the results on *total time, time per move,* and *number of moves* in Study 2 (Table 9) was 0.118 (Cohen 1969). To account for the smaller effect sizes, we estimated the sample size for Study 3. We found that an experiment with three conditions and two tasks that can detect an effect size f of at least 0.125 with α of 5% and power of 0.95, would require 747 participants, or 249 per condition. We use this value in Study 3 to replicate Study 2 soundly.

8.6.2 Comparison of Study 2 and Study 3

We discuss this in more detail in the main paper. In sum, Study 3 differs slightly from Study 2, but the deviations are less than half the size of what the experiment is meant to measure, and thus we argue that Study 3 is a good replication of Study 2.

Table A.10 shows descriptive values and zero-order correlations of the variables collected for Study 3.

--

Insert Table A.10 about here.

8.6.3 Hypothesis Testing

In Table 9 of the manuscript, we compare the hypothesis testing of Studies 2 and 3 together. In Table 9, the results of Study 3 include both the demographic and pandemic controls. In Table A.11 we separate the control variables to show the results for Study 3 with each set separately. Table A.12 presents the same estimates but with robust regressions instead of OLS. Table A.13, presents robust regressions for Study 3 with both sets of control variables (equivalent to Table 9 Models 3 and 4 but with robust regressions).

Insert Table A.11, A.12, and A.13 about here.

8.7 Post-hoc manipulation check study

In the second revision, we carried out a new data collection in late February 2021. The experiment was preregistered and followed the structure of Studies 2 and 3. Yet, experiment's goal was not to replicate the findings of the prior studies but to perform a qualitative manipulation check. In Studies 2 and 3, we used a qualitative question to check the process people espouse when solving the tasks. Yet, we could not uncover direct differences between the espoused problem-solving processes of the participants. We assumed there was a difference, as we found causal behavioral traces of the manipulation, and these traces were broadly replicated. Causality aside, without a manipulation check, we were not able to identify the mechanism that drove the behavioral change. We could infer the reason but not identify it. In this experiment, we added a question to identify the mechanism.

8.7.1 Changes to the research design

This experiment's aim differs from the prior studies' aim. In this study, the hypotheses relate to how the manipulation changes the "espoused decision policy" followed by the participants on the different conditions and show how these deviations align to the "in-use decision policies" we inferred from Studies 2 and 3 (Grégoire et al. 2019).

To test this relationship, we employ a dichotomous question and test it through contingency tables and χ^2 tests. The question is shown in Figure A.8, and it is placed directly after participants give the rank-ordered solution to the “NASA survival” problem. The question directly asks for the mental activities the participants employ more while solving the task—specifically, whether the participants employ more mental activities related to *framing* of the problem or to *implementation* of the solution. The use of these activities was the focus of the manipulation, and thus by testing this, we can identify the causal effect we discovered in Study 2 and replicated in Study 3.

Insert Figure A.8 about here.

We expected that the changes in frequency would be relatively large and thus related to a larger effect size than the results of Studies 2 and 3. We follow the same procedures as in the prior studies to estimate the sample size. Yet, the hypotheses test is done through contingency tables and thus employs a different estimation process. From G*Power 3.1, we find an estimated sample size of 252 participants with 84 participants per condition when using an effect size f of 0.25, α of 0.05, and power of 0.95.

8.7.2 Comparison of Study 2 and Study 3

As the focus of this study was not to test the hypotheses of Studies 2 and 3, we did not delve into the variables used in the prior studies. The raw variables were collected but not analyzed as the sample size was smaller and we expected to get unclear results as in the pandemic pilot study. The focus of this study is the question shown in Figure A.8 and discussed below.

8.7.3 Hypothesis Testing

The main results of this manipulation check are shown in Table 10 in the manuscript. The first aspect to notice is that the control condition has a small bias towards employing mental activities related to the framing of the problem (i.e., top row). Namely, of the 84 participants in the control condition, 10 more people espoused the use of framing than implementation activities. Compared to the control condition, the *framing disposition* and *implementation disposition* conditions espouse mental activities that deviate significantly from the *control* condition. These results are shown in Table A.14, where we present the four different χ^2. The top row presents the test for the 3-factor condition variable, the results are quite conclusive and present a high effect size (f ~ 0.5, Cohen 1969). The other three comparisons present comparisons between two of the conditions only. In all cases the results are significant. However, the comparison between the *framing disposition* condition and the *control* condition has a smaller effect size. This is due to the bias in the *control* condition where the participants espoused activities related to the framing of the problem more often than the implementation of the solution.

Insert Table A.14 about here.

REFERENCES

Ahorsu, D. K., Lin, C. Y., Imani, V., Saffari, M., Griffiths, M. D., & Pakpour, A. H. (2020). The fear of COVID-19 scale: development and initial validation. *International journal of mental health and addiction*.

Anderson, V. L., & McLean, R. A. (2018). *Design of experiments: a realistic approach*. Routledge.

Carleton, R. N., Norton, M. P. J., & Asmundson, G. J. (2007). Fearing the unknown: A short version of the Intolerance of Uncertainty Scale. *Journal of anxiety disorders*, 21(1), 105-117.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and psychological measurement,* 20(1), 37-46.

Cohen, J. (1969*). Statistical power analysis for the behavioral sciences.* Academic press.

Davis, M. J. (2010). Contrast coding in multiple regression analysis: Strengths, weaknesses, and utility of popular coding structures. *Journal of Data Science*, 8(1), 61-73.

Duncan, L. A., Schaller, M., & Park, J. H. (2009). Perceived vulnerability to disease: Development and validation of a 15-item self-report instrument. *Personality and Individual differences*, 47(6), 541-546.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods,* 41(4), 1149-1160.

Fedor, A., Szathmáry, E., & Öllinger, M. (2015). Problem solving stages in the five square problem. Frontiers in psychology, 6, 1050.

Gibbs, A. J., Dale, M. B., Kinns, H. R., & MacKenzie, H. G. (1971). The transition matrix method for comparing sequences; its use in describing and classifying proteins by their amino acid sequences. *Systematic Biology*, 20(4): 417-425.

Grégoire, D. A., Binder, J. K., & Rauch, A. (2019). Navigating the validity tradeoffs of entrepreneurship research experiments: A systematic review and best-practice suggestions. *Journal of Business Venturing*, 34(2), 284-310.

Greve, H. R. (2018). Show me the data! Improving evidence presentation for publication. *Management and Organization Review*, *14*(2), 423-432.

Hall, J., & Watson, W. H. (1970). The effects of a normative intervention on group decision-making performance. *Human Relations*, 23(4), 299-317.

Hennig, C. (2015). *fpc: Flexible Procedures For Clustering*. R Package Version 2.1-6.

Hertwig, R., & Ortmann, A. (2001). Experimental practices in economics: A methodological challenge for psychologists?. *Behavioral and Brain Sciences*, 24(3), 383-403.

Johnson, D. W., & Johnson, F. P. (1982). *Joining together* (2nd ed.). Prentice Hall. 111-116

Kaufman L, Rousseeuw P. J. (1990). Partitioning around medoids (program pam). *Finding groups in data: an introduction to cluster analysis*: 68-125.

Klein, G. A. (2017). *Sources of power: How people make decisions*. MIT press.

Krippendorff K. (2012). *Content analysis: An introduction to its methodology.* Sage.

Laureiro-Martínez, D., & Brusoni, S. (2018). Cognitive flexibility and adaptive decision-making: Evidence from a laboratory study of expert decision makers. *Strategic Management Journal*, 39(4), 1031-1058.

Lipshitz, R., & Bar-Ilan, O. (1996). How problems are solved: Reconsidering the phase theorem. *Organizational Behavior and Human Decision Processes*, 65(1): 48-60.

Lombard, M., Snyder-Duch, J., & Bracken, C. C. (2002). Content analysis in mass communication: Assessment and reporting of intercoder reliability. *Human communication research*, *28*(4), 587-604.

McLellan, E., MacQueen, K. M., & Neidig, J. L. (2003). Beyond the qualitative interview: Data preparation and transcription. *Field methods*, 15(1), 63-84.

Neuendorf, K. A. (2002). *The content analysis guidebook*. Sage.

Ocasio, W., & Joseph, J. (2018). The attention-based view of great strategies. *Strategy Science*, 3(1), 289-294.

Oehlert, G. W. (2010). *A first course in design and analysis of experiments*. W. H. Freeman

Öllinger, M., Jones, G., Faber, A. H., & Knoblich, G. (2013). Cognitive mechanisms of insight: the role of heuristics and representational change in solving the eight-coin problem. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(3), 931.

Palan, S., & Schitter, C. (2018). Prolific. ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, 17, 22-27.

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. J*ournal of Experimental Social Psychology*, 70, 153-163.

Pentland, B. T. (2003). Sequential variety in work processes. *Organization Science,* 14(5), 528-540.

Poland, B. D. (1995). Transcription quality as an aspect of rigor in qualitative research. *Qualitative inquiry*, 1(3), 290-310.

Salvato, C. (2009). Capabilities unveiled: The role of ordinary activities in the evolution of product development processes. *Organization science*, 20(2), 384-409.

Schacter, D. L., Benoit, R. G., & Szpunar, K. K. (2017). Episodic future thinking: Mechanisms and functions. *Current opinion in behavioral sciences*, 17, 41-50.

Steiner, G. A. (2010). *Strategic planning*. Simon and Schuster.

Steptoe, A. (1989). An abbreviated version of the Miller behavioral style scale. *British Journal of Clinical Psychology*, 28(2), 183-184.

Taylor, S. (2019). *The Psychology of Pandemics: Preparing for the Next Global Outbreak of Infectious Disease*. Cambridge Scholars Publishing.

Tetlock, P. E., & Gardner, D. (2016). *Superforecasting: The art and science of prediction*. Random House.

Wellian, E. (2010), *LEGO: Consolidating distribution*, Case Study, IMD International

Yu, Z., Wang, F., Wang, D., & Bastin, M. (2012). Beyond reaction times: Incorporating mouse-tracking measures into the implicit association test to examine its underlying process. *Social Cognition*, 30(3), 289-306.

FIGURES

Figure A.1: Text and layout of the problem used in Study 1

Imagine that you are the leader of the Karabayos, a small tribe (22 women and 26 men) in the Amazon rainforest. There are hundreds of tribes in the world that, like yours, have never had any contact with other peoples and are scattered in the vast jungles of South America, New Guinea and the Indian Ocean.

You all have one characteristic in common: you are the most vulnerable people in the world and you want to be left in peace. And for a good reason! The history of contact between indigenous tribes and the rest of the world has always been particularly unfortunate.

Contact with other people is almost always a disaster for these types of tribes, who have lived according to a lifestyle largely intact for more than 10,000 years. Your lifestyle does not include the use of television, microwaves, or cars. You never feel the need of any of these. Most of these tribes live in hidden places inside the forest. However, many of these hidden places are getting closer to the areas under the control of rubber producers, loggers, settlers and drug traffickers, which endanger the survival of the tribes.

Even when the loggers do not kill any tribe members directly (which often happens), after contact with other peoples the tribes are decimated within a year or two from many diseases (such as influenza, measles, chicken pox) against which these tribes have no immunity.

The peace and harmony of your tribe, and the abundance of your lands, which for centuries have allowed you to live in balance with nature is constantly endangered by the approach of "civilization".

Your group lives in an area that for years has been full of fruit trees and animals of all kinds. Traditionally, men hunt with bows and blowguns, while women stay at home to take care of children. You are aware that some parts of the area you live in are bordering areas of the "whites" and, for years, your people have avoided coming into contact with them. In recent years, you have realized that the trees no longer produce as many fruits as they used to, and that many of the animals you used to hunt have disappeared. You ask yourself what you should do to save your tribe.

Figure A.2: User interface and layout of the "winter survival" problem

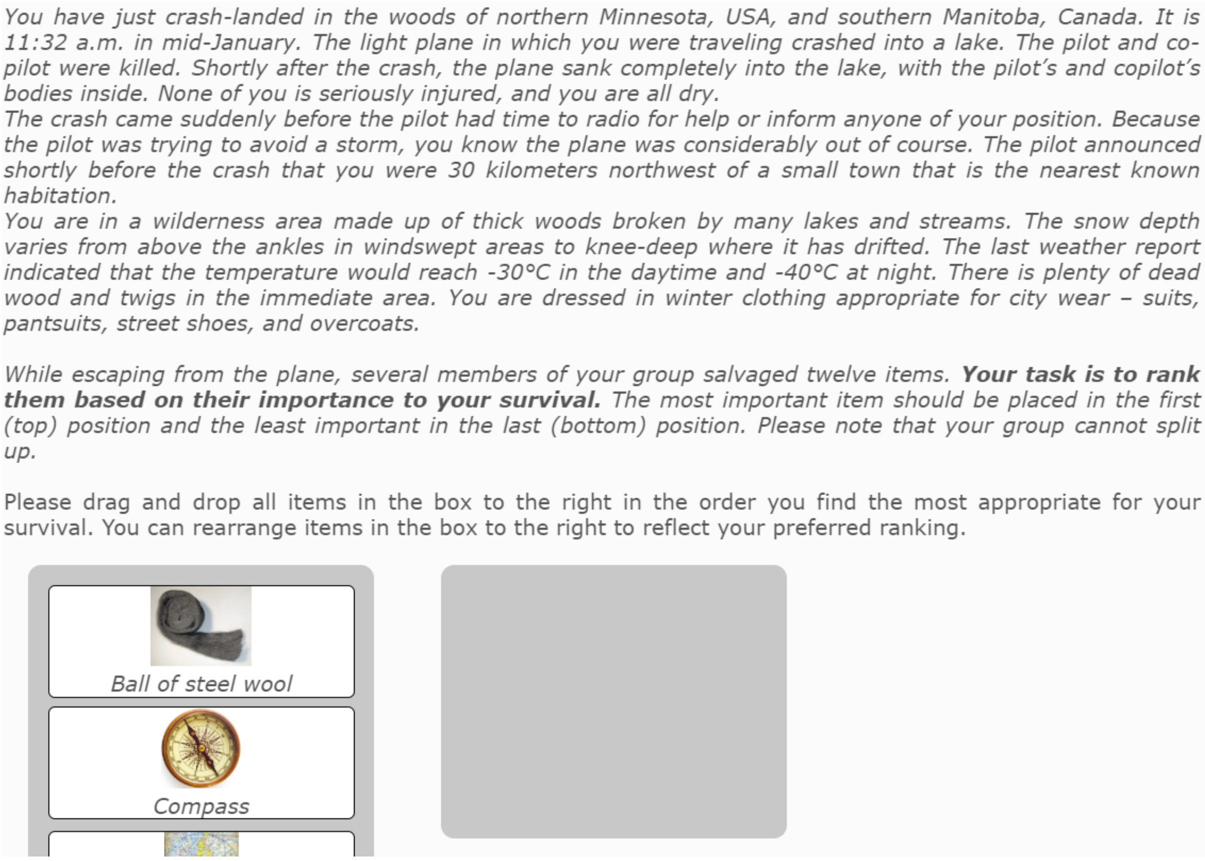

Figure A.3: User interface and layout of the "NASA survival" problem, implementation disposition condition

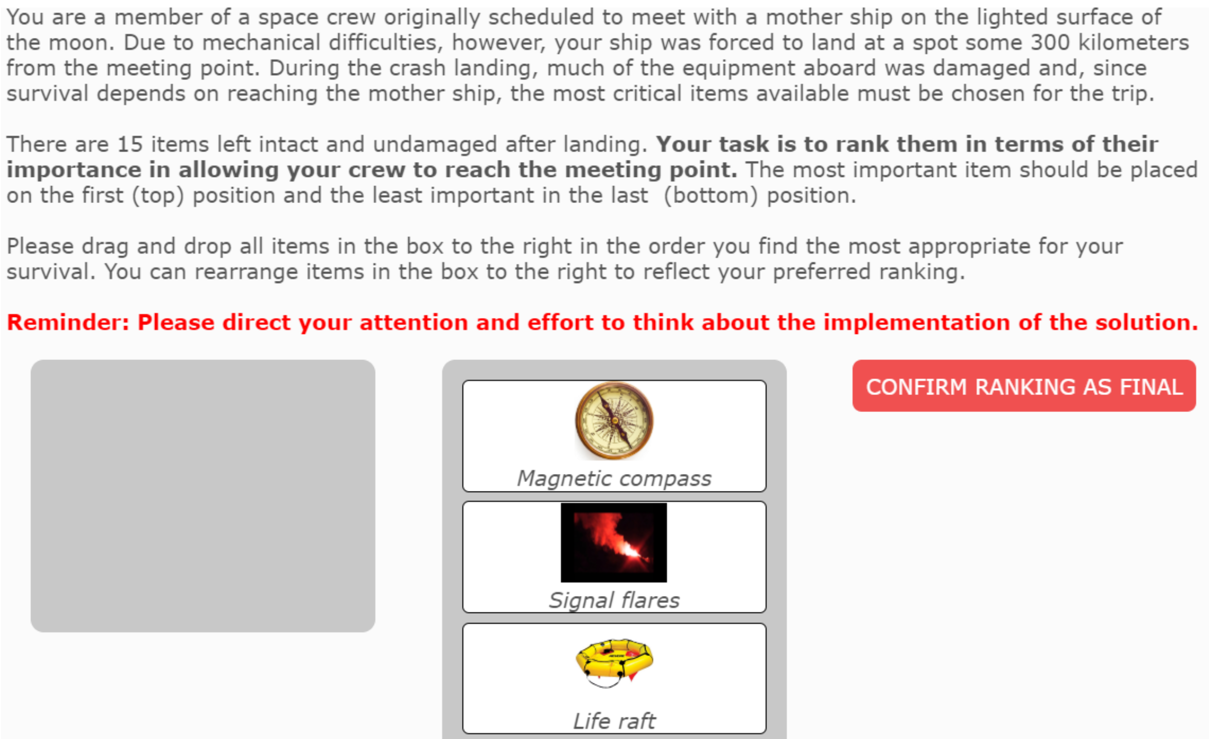

Figure A.4. Visualization of the research procedure

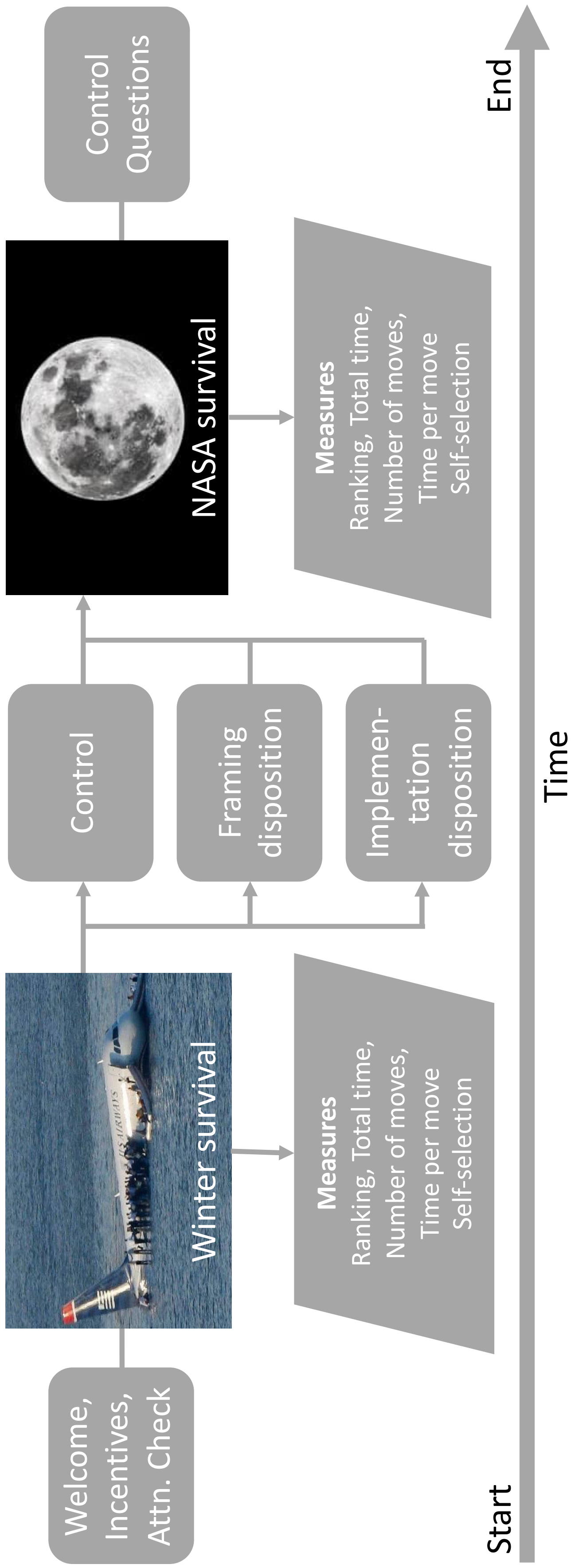

Figure A.5: Manipulation for the control condition of Study 2

Welcome to the second task.

Your solution to this problem will affect your performance evaluation. There is no time limit.

Research and experts in problem-solving recommend that in order to achieve higher performance, individuals should spend most of their time thinking about the problem in the way that feels more natural to them.

It is crucial that while solving the next task you direct your attention and effort to think about the problem in the way that seems most natural to you.

On the next page, please carefully read the problem description and answer accordingly.

Figure A.6: Manipulation for the framing disposition condition of Study 2

Welcome to the second task.

Your solution to this problem will affect your performance evaluation. There is no time limit.

Research and experts in problem-solving recommend that in order to achieve higher performance, individuals should spend most of their time thinking about the framing of the problem.

The framing of the problem involves important mental activities linked to higher performance, such as:

- **Analyze the problem by recalling the available information**
- **Empathizing to identify with the situation**
- **Develop hypotheses/assumptions to gain an understanding of the problem**

It is crucial that while solving the task you direct your attention and effort to think about the framing of the problem.

On the next page, please carefully read the problem description and answer accordingly.

Figure A.7: Manipulation for the implementation disposition condition of Study 2

Welcome to the second task.

Your solution to this problem will affect your performance evaluation. There is no time limit.

Research and experts in problem-solving recommend that in order to achieve higher performance, individuals should spend most of their time thinking about the implementation of the solution.

The Implementation of the solution involves important mental activities linked to higher performance, such as:

- **Design the sequence of actions that could unfold during the solution of the problem**
- **Anticipate how events will carry out**
- **Evaluate the feasibility of the solutions**

It is crucial that while solving the task you direct your attention and effort to think about the implementation of the solution.

On the next page, please carefully read the problem description and answer accordingly.

Figure A.8: Manipulation check question

When solving the problem, on what mental activities did you spend more time (Please select one of the two options):

○ - Analyzing the problem by recalling the available information

- Empathizing to identify with the situation

- Developing hypotheses/assumptions to gain an understanding of the problem.

○ - Designing the sequence of actions that could unfold during the solution of the problem

- Anticipating how events will play out

- Evaluating the feasibility of the solutions.

TABLES

Table A.1: List of items used in the "winter survival" problem

Order Shown	Correct Ranking	Item
1	2	Ball of steel wool
2	12	Compass
3	1	Cigarette lighter (without fluid)
4	11	Sectional air map made of plastic
5	3	Extra shirt and pants for each survivor
6	10	Quart of 100-proof whiskey
7	8	Newspaper (one per person)
8	6	Hand ax
9	9	Loaded 0.45 caliber pistol
10	5	6x6 meter of heavy-duty canvas
11	4	Can of shortening (margarine)
12	7	Family size chocolate bar (one per person)

Table A.2: List of items used in the "NASA survival" problem

Order Shown	Correct Ranking	Item
1	15	Box of matches
2	6	15 meters of nylon rope
3	13	Portable heating unit
4	12	One case of dehydrated pet milk
5	3	Stellar map (of how constellations look on the moon)
6	14	Magnetic compass
7	10	Signal flares
8	5	Solar-powered FM receiver transmitter
9	4	Food concentrate
10	8	Parachute silk
11	11	Two 0.45 caliber pistols
12	1	Two fifty-kilo tanks of oxygen
13	9	Life raft
14	2	20 liters of water
15	7	First aid kit containing injection needles

Table A.3: Comparison of the "winter survival" and "NASA survival" problems

	Winter survival	NASA survival
Source	Johnson and Johnson (1982, p. 111)	Hall and Watson (1970)
Novel Situation	Plane crash on a Canadian lake during winter	Space module crashed on the moon
Confounding Uncertainty	Closest town is 30 km away Survival in low temperature	Meeting point is 300 km away Survival in the Moon surface
Process complexity	Rank-order 12 elements	Rank-order 15 elements
High stakes goal	Group survival	Crew survival
Span of control	Over group	Over crew
Optimal Solution	Stay near crash location until rescue	Walk to meeting point

Table A.4. List of variables in Study 2

	Variables	**Definition**
Experimental condition	1. Control	Member of the *control* condition (1 if yes, 0 if no)
	2. *Framing* disposition	Member of the *framing disposition condition* (1 if yes, 0 if no)
	3. *Implementation disposition*	Member of the *implementation disposition* condition n (1 if yes, 0 if no)
Demographics	4. Gender	Gender of the participant (1 if female, -1 if male)
	5. Age	Age of the participant in years
	6. Postgraduate	The education level of the participant (1 if holds a masters degree or higher, -1 if lower)
	7. Reader	Reading habits of the participant (1 if more than twice a week, -1 if less)
Behavioral change	8. Performance	Change in standardized performance between tasks; negative value implies improvement [s.d.]
	9. Total Time	Change in standardized total time between tasks [s.d.]
	10. Time per Move	Change in standardized time per move between tasks [sd.]
	11. # of Moves	Change in standardized total time between tasks [s.d.]
"NASA survival" problem	12. Performance	Performance in the "NASA survival" problem; lower is better, 0 is perfect
	13. Total Time	Total time spent in the "NASA survival" problem [s]
	14. Time per Move	Time per move in the "NASA survival" problem [s]
	15. # of Moves	Number of moves in the "NASA survival" problem
	16. Self-selection *Solution-focused*	Self-selected a prototypical *solution-focused* problem-solving text (1 if yes, -1 if no) in the "NASA survival" problem
"Winter survival" problem	17. Performance	Performance in the "winter survival" problem; lower is better, 0 is perfect
	18. Total Time	Total time spent in the "winter survival" problem [s]
	19. Time per Move	Time per move in the "winter survival" problem [s]
	20. # of Moves	Number of moves in the "winter survival" problem
	21. Self-selection *Solution-focused*	Self-selected a prototypical *solution-focused* problem-solving text (1 if yes, -1 if no) in the "NASA survival" problem

Table A.5: Descriptive and zero-order correlation table of the variables of Study 2 (part 1 of 3)

	Variables	1.	2.	3.	4.	5.	6.	7.
Experimental condition	1. *Control*	1						
	2. *Framing disposition*	-0.604 (0.000)	1					
	3. *Implementation disp.*	-0.611 (0.000)	-0.262 (0.000)	1				
Demographics	4. Gender	0.006 (0.895)	-0.024 (0.600)	0.017 (0.719)	1			
	5. Age	-0.001 (0.989)	-0.012 (0.795)	0.013 (0.783)	0.027 (0.556)	1		
	6. Postgraduate	-0.038 (0.415)	0.010 (0.827)	0.035 (0.442)	-0.070 (0.129)	-0.034 (0.460)	1	
	7. Reader	0.034 (0.462)	-0.095 (0.040)	0.053 (0.252)	0.122 (0.008)	0.173 (0.000)	0.200 (0.000)	1
Behavioral change	8. Performance	0.028 (0.546)	-0.017 (0.71)	-0.017 (0.718)	-0.017 (0.717)	0.080 (0.082)	0.024 (0.601)	-0.020 (0.667)
	9. Total Time	-0.131 (0.005)	0.083 (0.071)	0.075 (0.102)	-0.102 (0.027)	0.021 (0.652)	-0.109 (0.018)	-0.055 (0.230)
	10. Time per Move	-0.097 (0.035)	0.004 (0.931)	0.114 (0.014)	-0.070 (0.130)	0.006 (0.898)	-0.059 (0.198)	-0.087 (0.060)
	11. # of Moves	-0.073 (0.114)	0.104 (0.023)	-0.015 (0.738)	-0.011 (0.817)	-0.048 (0.293)	-0.005 (0.921)	0.073 (0.112)
"NASA survival" problem	12. Performance	0.046 (0.316)	-0.025 (0.587)	-0.031 (0.500)	-0.050 (0.274)	-0.016 (0.723)	-0.032 (0.492)	-0.014 (0.769)
	13. Total Time	-0.036 (0.430)	0.013 (0.778)	0.031 (0.499)	-0.005 (0.920)	0.144 (0.002)	-0.079 (0.087)	-0.007 (0.878)
	14. Time per Move	-0.030 (0.511)	-0.038 (0.406)	0.075 (0.105)	0.008 (0.861)	0.100 (0.030)	-0.048 (0.300)	-0.033 (0.469)
	15. # of Steps	-0.041 (0.372)	0.103 (0.025)	-0.053 (0.255)	0.013 (0.782)	-0.022 (0.628)	-0.051 (0.273)	0.044 (0.343)
	16. Self-sel. *Sol. focus*	-0.017 (0.712)	-0.067 (0.147)	0.087 (0.059)	0.042 (0.368)	-0.013 (0.779)	0.027 (0.552)	-0.006 (0.900)
"Winter survival" problem	17. Performance	0.011 (0.816)	-0.003 (0.945)	-0.010 (0.831)	-0.029 (0.531)	-0.118 (0.011)	-0.062 (0.179)	0.012 (0.800)
	18. Total Time	0.097 (0.034)	-0.072 (0.119)	-0.047 (0.313)	0.099 (0.032)	0.118 (0.011)	0.035 (0.449)	0.049 (0.285)
	19. Time per move	0.077 (0.096)	-0.042 (0.367)	-0.051 (0.264)	0.084 (0.068)	0.091 (0.049)	0.018 (0.689)	0.062 (0.178)
	20. # of Steps	0.039 (0.400)	-0.013 (0.773)	-0.034 (0.464)	0.024 (0.607)	0.031 (0.506)	-0.044 (0.345)	-0.037 (0.426)
	21. Self-sel. *Sol. focus*	-0.041 (0.377)	0.052 (0.263)	-0.002 (0.967)	-0.030 (0.517)	0.010 (0.820)	-0.001 (0.981)	0.006 (0.905)
	M **SD**	276 of 472	97 of 472	99 of 472	269 of 472	34.88 8.51	161 of 472	219 of 472

Note: p-value of the pairwise correlations shown in parenthesis

The mean value and standard deviation of the behavioral change variables (8-11) are not 0 and 1 respectively, because we use only the data of the participants in the control condition to standardize the "NASA survival" problem data.

Table A.5: Descriptive and zero-order correlation table of the variables of Study 2 (continuation 2 of 3)

	Variables	8.	9.	10.	11.	12.	13.	14.
Behavioral change	8. Performance	1						
	9. Total Time	-0.065 (0.159)	1					
	10. Time per Move	-0.032 (0.489)	0.662 (0.000)	1				
	11. # of Moves	-0.073 (0.112)	0.263 (0.000)	-0.139 (0.002)	1			
"NASA survival" problem	12. Performance	0.631 (0.000)	-0.114 (0.014)	-0.119 (0.010)	-0.095 (0.038)	1		
	13. Total Time	-0.043 (0.350)	0.489 (0.000)	0.353 (0.000)	0.072 (0.119)	-0.259 (0.000)	1	
	14. Time per Move	-0.039 (0.396)	0.283 (0.000)	0.536 (0.000)	-0.151 (0.001)	-0.208 (0.000)	0.695 (0.000)	1
	15. # of Steps	-0.045 (0.327)	0.227 (0.000)	-0.117 (0.011)	0.523 (0.000)	-0.112 (0.015)	0.359 (0.000)	-0.162 (0.000)
	16. Self-sel. *Sol. focus*	-0.083 (0.073)	0.052 (0.256)	0.002 (0.959)	0.047 (0.313)	-0.123 (0.007)	0.029 (0.533)	-0.001 (0.990)
"Winter survival" problem	17. Performance	-0.638 (0.000)	-0.031 (0.507)	-0.078 (0.09)	-0.002 (0.964)	0.194 (0.000)	-0.202 (0.000)	-0.157 (0.001)
	18. Total Time	0.025 (0.595)	-0.544 (0.000)	-0.332 (0.000)	-0.197 (0.000)	-0.133 (0.004)	0.466 (0.000)	0.381 (0.000)
	19. Time per move	-0.003 (0.946)	-0.449 (0.000)	-0.572 (0.000)	0.006 (0.899)	-0.072 (0.118)	0.289 (0.000)	0.386 (0.000)
	20. # of Steps	0.035 (0.442)	-0.064 (0.163)	0.037 (0.422)	-0.574 (0.000)	-0.005 (0.918)	0.267 (0.000)	0.007 (0.884)
	21. Self-sel. *Sol. focus*	-0.033 (0.477)	0.067 (0.149)	0.039 (0.392)	0.024 (0.609)	0.060 (0.191)	-0.013 (0.779)	-0.011 (0.814)
	M	-0.039	0.029	0.025	0.033	49.34	363.0	8.853
	SD	1.265	1.015	1.093	1.077	15.48	208.8	5.921

Table A.5: Descriptive and zero-order correlation table of the variables of Study 2 (continuation 3 of 3)

	Variables	15.	16.	17.	18.	19.	20.	21.
"NASA survival" problem	15. # of Steps	1						
	16. Self-sel. *Sol. focus*	-0.007 (0.878)	1					
"Winter survival" problem	17. Performance	-0.054 (0.242)	-0.018 (0.697)	1				
	18. Total Time	0.114 (0.013)	-0.025 (0.581)	-0.163 (0.000)	1			
	19. Time per move	-0.029 (0.533)	-0.003 (0.946)	-0.068 (0.142)	0.734 (0.000)	1		
	20. # of Steps	0.398 (0.000)	-0.057 (0.217)	-0.050 (0.282)	0.322 (0.000)	-0.034 (0.462)	1	
	21. Self-sel. *Sol. focus*	-0.003 (0.955)	0.031 (0.501)	0.101 (0.028)	-0.080 (0.083)	-0.054 (0.244)	-0.028 (0.546)	1
	M	28.92	176 of 472	45.13	324.5	10.17	20.05	53 of 472
	SD	12.31		9.54	287.2	11.37	7.37	

Note: p-value of the pairwise correlations shown in parenthesis

The mean value and standard deviation of the behavioral change variables (8-11) are not 0 and 1 respectively, because we use only the data of the participants in the control condition to standardize the "NASA survival" problem data.

Table A.6: Robust linear regression of performance and behavioral change of Study 2

	Dependent Variable	
	Change in # of moves [s.d.] (1)	**Change in time per move [s.d.] (1)**
Framing disposition	0.334	0.027
	(0.154, 0.513)	(-0.095, 0.149)
Implementation disposition	0.029	0.181
	(-0.149, 0.206)	(0.060, 0.301)
Gender	0.007	-0.013
	(-0.064, 0.079)	(-0.061, 0.036)
Age	-0.007	0.005
	(-0.016, 0.001)	(-0.0002, 0.011)
Postgraduate	-0.050	-0.002
	(-0.126, 0.025)	(-0.054, 0.049)
Reader	0.043	-0.069
	(-0.030, 0.117)	(-0.119 -0.019)
Constant	0.204	-0.253
	(-0.101, 0.509)	(-0.460, -0.045)
Observations	472	472
Residual Std. Error	0.679 (df=465)	0.465 (df=465)

Note: 95% confidence intervals shown in parentheses.

Table A.7. List of variables added in the replication studies

	Variables	**Definition**
Pandemic controls	22. Fear of COVID-19	Measured responses to the Fear of COVID-19 scale by Ahorsu et al. (2020)
	23. Monitor/Blunt Scale	Abbreviated version of the Monitoring/Blunting scale for psychological coping with threats (Steptoe, 1989)
	24. Perceived Vulnerability to Disease	Perceived Vulnerability to Disease 15-item scale (Duncan et al., 2009)
	25. Intolerance to Uncertainty	Twelve-item Intolerance to Uncertainty Scale by Carleton et al. (2007)
	26. Job Loss	Two yes/no questions asking whether the participant or someone in their household has lost their job due to the COVID-19 crisis
	27. Prolific Importance	Two five-level Likert scale questions relating to the relative importance of the work in Prolific in general and financially for the participant

Table A.8: Descriptive statistics and comparison of Study 2 and the pandemic pilot study

Task	Measure	Study 2		Pandemic pilot study		t-test	
		Mean	s.d.	Mean	s.d.	p-value	t-statistic
"Winter survival" problem	Total time [s]	324.5	287.2	384.3	229.3	0.002	3.181
	# of moves	20.05	7.37	21.38	9.21	0.037	2.091
	Time per move [s]	10.17	11.37	11.40	6.78	0.061	1.874
"NASA survival" problem	Total time [s]	363.0	208.8	454.8	361.6	0.000	3.959
	# of moves	28.92	12.31	29.08	11.34	0.852	0.186
	Time per move [s]	8.85	5.92	11.01	9.68	0.001	3.448

Table A.9: Testing the effect of task order

Measure	Winter then NASA [s.d.]	NASA then Winter [s.d.]	t-test	
			p-value	t-statistic
Std. Change in Total Time	0.020	-0.073	0.377	0.885
Std. Change in Time per move	-0.053	-0.088	0.763	0.302
Std. Change in Number of Moves	0.060	-0.089	0.189	1.317

Table A.10 Descriptive and zero-order correlation table of the variables of Study 3 (part 1 of 4)

	Variables	1.	2.	3.	4.	5.	6.	7.
Experimental condition	1. *Control*	1						
	2. *Framing disposition*	-0.500 (0.000)	1					
	3. *Implementation disposition*	-0.500 (0.000)	-0.500 (0.000)	1				
Demographics	4. Gender	-0.006 (0.876)	-0.080 (0.028)	0.086 (0.019)	1			
	5. Age	-0.018 (0.629)	-0.050 (0.170)	0.068 (0.064)	0.014 (0.693)	1		
	6. Postgraduate	0.034 (0.355)	-0.014 (0.703)	-0.020 (0.586)	0.034 (0.348)	-0.037 (0.307)	1	
	7. Reader	-0.059 (0.107)	0.010 (0.795)	0.049 (0.177)	0.050 (0.174)	0.201 (0.000)	0.111 (0.002)	1
Behavioral change	8. Performance	-0.016 (0.660)	0.008 (0.827)	0.008 (0.825)	-0.023 (0.531)	0.043 (0.235)	0.043 (0.242)	-0.018 (0.619)
	9. Total Time	-0.143 (0.000)	0.131 (0.000)	0.013 (0.729)	-0.013 (0.716)	0.080 (0.028)	-0.003 (0.926)	0.031 (0.395)
	10. Time per Move	-0.072 (0.051)	0.036 (0.320)	0.035 (0.338)	0.007 (0.854)	0.046 (0.207)	-0.008 (0.828)	0.045 (0.218)
	11. # of Moves	-0.092 (0.012)	0.111 (0.002)	-0.019 (0.605)	-0.030 (0.411)	-0.004 (0.921)	-0.071 (0.054)	0.053 (0.149)
"NASA survival" problem	12. Performance	-0.003 (0.927)	-0.035 (0.343)	0.038 (0.298)	-0.011 (0.773)	-0.041 (0.268)	-0.020 (0.576)	-0.061 (0.094)
	13. Total Time	-0.152 (0.000)	0.064 (0.078)	0.087 (0.017)	0.006 (0.877)	0.170 (0.000)	-0.023 (0.531)	0.084 (0.022)
	14. Time per Move	-0.101 (0.006)	-0.008 (0.824)	0.109 (0.003)	0.026 (0.472)	0.133 (0.000)	-0.029 (0.426)	0.104 (0.004)
	15. # of Steps	-0.037 (0.311)	0.092 (0.012)	-0.055 (0.136)	-0.044 (0.229)	0.007 (0.848)	-0.015 (0.672)	0.036 (0.329)
	16. Self-sel. *Impl. focus*	-0.049 (0.18)	-0.037 (0.308)	0.086 (0.018)	-0.001 (0.981)	0.027 (0.459)	0.006 (0.861)	-0.015 (0.691)
"Winter survival" problem	17. Performance	0.015 (0.679)	-0.046 (0.209)	0.031 (0.400)	0.016 (0.671)	-0.093 (0.011)	-0.072 (0.051)	-0.044 (0.234)
	18. Total Time	-0.024 (0.511)	-0.088 (0.017)	0.112 (0.002)	0.027 (0.460)	0.140 (0.000)	-0.029 (0.424)	0.081 (0.028)
	19. Time per move	-0.042 (0.253)	-0.066 (0.071)	0.108 (0.003)	0.029 (0.432)	0.128 (0.000)	-0.031 (0.395)	0.087 (0.018)
	20. # of Steps	0.070 (0.057)	-0.034 (0.353)	-0.036 (0.330)	-0.011 (0.763)	0.012 (0.748)	0.068 (0.065)	-0.025 (0.498)
	21. Self-sel. *Impl. focus*	-0.044 (0.227)	0.000 (1.000)	0.044 (0.227)	0.015 (0.690)	0.096 (0.009)	0.037 (0.315)	-0.010 (0.775)
Pandemic controls	22. Fear of COVID-19	-0.014 (0.705)	-0.009 (0.819)	0.023 (0.540)	0.243 (0.000)	-0.076 (0.043)	0.012 (0.752)	0.001 (0.985)
	23. Monitor/Blunt Scale	-0.051 (0.174)	0.022 (0.567)	0.030 (0.430)	0.028 (0.454)	-0.073 (0.052)	0.050 (0.183)	-0.007 (0.860)
	24. Perceived Vulnerability	0.057 (0.126)	-0.018 (0.624)	-0.039 (0.296)	-0.081 (0.031)	-0.123 (0.001)	-0.085 (0.023)	-0.032 (0.400)
	25. Intolerance to Uncertainty	-0.005 (0.897)	-0.028 (0.455)	0.033 (0.376)	0.126 (0.001)	-0.095 (0.011)	-0.063 (0.093)	-0.068 (0.070)
	26. Job Loss	0.025 (0.506)	-0.019 (0.605)	-0.006 (0.883)	0.045 (0.232)	-0.024 (0.519)	-0.028 (0.464)	0.028 (0.449)
	27. Prolific Importance	0.050 (0.185)	-0.040 (0.289)	-0.010 (0.792)	0.079 (0.036)	-0.056 (0.135)	-0.013 (0.722)	-0.040 (0.289)
	M **SD**	249 of 747	249 of 747	249 of 747	424 of 747	34.21 7.60	256 of 747	337 of 747

Note: p-value of the pairwise correlations shown in parenthesis

The mean value and standard deviation of the behavioral change variables (8-11) are not zero and one respectively, because we use only the data of the participants in the control condition to standardize the "NASA survival" problem data.

Table A.10 Descriptive and zero-order correlation table of the variables of Study 3 (continuation 2 of 4)

	Variables	8.	9.	10.	11.	12.	13.	14.
Behavioral change	8. Performance	1						
	9. Total Time	-0.087 (0.017)	1					
	10. Time per Move	-0.037 (0.316)	0.640 (0.000)	1				
	11. # of Moves	-0.136 (0.000)	0.255 (0.000)	-0.146 (0.000)	1			
"NASA survival" problem	12. Performance	0.598 (0.000)	-0.076 (0.039)	-0.062 (0.091)	-0.049 (0.185)	1		
	13. Total Time	-0.056 (0.129)	0.761 (0.000)	0.468 (0.000)	0.147 (0.000)	-0.142 (0.000)	1	
	14. Time per Move	0.003 (0.926)	0.451 (0.000)	0.772 (0.000)	-0.132 (0.000)	-0.114 (0.002)	0.671 (0.000)	1
	15. # of Steps	-0.101 (0.006)	0.229 (0.000)	-0.141 (0.000)	0.612 (0.000)	-0.111 (0.002)	0.272 (0.000)	-0.160 (0.000)
	16. Self-sel. *Sol. focus*	0.021 (0.576)	0.005 (0.881)	0.025 (0.497)	0.039 (0.286)	0.039 (0.288)	-0.010 (0.781)	-0.003 (0.943)
"Winter survival" problem	17. Performance	-0.532 (0.000)	0.022 (0.550)	-0.023 (0.535)	0.107 (0.003)	0.361 (0.000)	-0.086 (0.019)	-0.124 (0.001)
	18. Total Time	0.040 (0.273)	-0.273 (0.000)	-0.203 (0.000)	-0.139 (0.000)	-0.105 (0.004)	0.417 (0.000)	0.364 (0.000)
	19. Time per move	0.060 (0.103)	-0.288 (0.000)	-0.349 (0.000)	0.022 (0.556)	-0.076 (0.038)	0.293 (0.000)	0.326 (0.000)
	20. # of Steps	0.055 (0.131)	-0.061 (0.098)	0.023 (0.529)	-0.539 (0.000)	-0.061 (0.096)	0.114 (0.002)	-0.013 (0.722)
	21. Self-sel. *Sol. focus*	0.034 (0.347)	-0.045 (0.224)	-0.051 (0.164)	-0.002 (0.951)	0.026 (0.485)	-0.016 (0.659)	-0.035 (0.334)
Pandemic controls	22. Fear of COVID-19	-0.015 (0.688)	0.002 (0.966)	-0.022 (0.553)	-0.002 (0.958)	0.079 (0.034)	-0.015 (0.688)	-0.044 (0.238)
	23. Monitor/Blunt Scale	0.026 (0.494)	-0.039 (0.302)	0.008 (0.829)	-0.015 (0.693)	0.013 (0.738)	-0.026 (0.492)	0.032 (0.398)
	24. Perceived Vulnerability	-0.020 (0.600)	-0.025 (0.504)	-0.012 (0.752)	0.015 (0.690)	0.020 (0.598)	-0.051 (0.171)	-0.017 (0.657)
	25. Intolerance to Uncertainty	-0.019 (0.618)	-0.005 (0.892)	0.052 (0.170)	-0.075 (0.044)	0.032 (0.399)	0.017 (0.657)	0.069 (0.068)
	26. Job Loss	-0.038 (0.317)	0.034 (0.371)	0.050 (0.183)	0.045 (0.236)	0.028 (0.451)	0.051 (0.173)	0.061 (0.105)
	27. Prolific Importance	0.004 (0.917)	0.022 (0.567)	-0.046 (0.220)	0.048 (0.201)	0.054 (0.153)	0.072 (0.055)	-0.002 (0.947)
	M	0.005	0.318	0.210	0.056	48.82	383.8	9.551
	SD	1.163	1.401	1.487	1.191	15.76	260.7	6.976

Note: p-value of the pairwise correlations shown in parenthesis.
The mean value and standard deviation of the behavioral change variables (8-11) are not 0 and 1 respectively, because we use only the data of the participants in the control condition to standardize the "NASA survival" problem data.

Table A.10 Descriptive and zero-order correlation table of the variables of Study 3 (continuation 3 of 4)

	Variables	15.	16.	17.	18.	19.	20.	21.
"NASA survival" problem	15. # of Steps	1						
	16. Self-sel. *Sol. focus*	0.014 (0.7)	1					
"Winter survival" problem	17. Performance	-0.001 (0.987)	0.017 (0.638)	1				
	18. Total Time	0.083 (0.024)	-0.023 (0.533)	-0.158 (0.000)	1			
	19. Time per move	-0.026 (0.476)	-0.041 (0.264)	-0.150 (0.000)	0.839 (0.000)	1		
	20. # of Steps	0.336 (0.000)	-0.032 (0.389)	-0.129 (0.000)	0.254 (0.000)	-0.054 (0.144)	1	
	21. Self-sel. *Sol. focus*	0.058 (0.114)	0.061 (0.094)	-0.013 (0.721)	0.038 (0.295)	0.024 (0.519)	0.064 (0.079)	1
Pandemic controls	22. Fear of COVID-19	-0.010 (0.796)	-0.009 (0.812)	0.100 (0.008)	-0.026 (0.494)	-0.036 (0.34)	-0.008 (0.830)	-0.007 (0.860)
	23. Monitor/Blunt Scale	-0.023 (0.535)	-0.038 (0.307)	-0.017 (0.659)	0.017 (0.661)	0.038 (0.31)	-0.007 (0.847)	0.006 (0.873)
	24. Perceived Vulnerability	-0.071 (0.058)	-0.008 (0.825)	0.043 (0.248)	-0.043 (0.251)	-0.008 (0.830)	-0.094 (0.012)	-0.047 (0.212)
	25. Intolerance to Uncertainty	-0.072 (0.056)	-0.020 (0.589)	0.055 (0.146)	0.033 (0.376)	0.029 (0.446)	0.013 (0.721)	-0.005 (0.897)
	26. Job Loss	0.013 (0.721)	0.003 (0.932)	0.073 (0.052)	0.030 (0.419)	0.019 (0.618)	-0.039 (0.302)	0.017 (0.654)
	27. Prolific Importance	0.047 (0.212)	-0.028 (0.450)	0.051 (0.172)	0.080 (0.032)	0.069 (0.067)	-0.007 (0.85)	0.008 (0.825)
	M	28.19	274 of 747	44.23	337.2	10.21	20.91	87 of 747
	SD	10.38		11.98	233.7	8.25	7.04	

Table A.10 Descriptive and zero-order correlation table of the variables of Study 3 (continuation 4 of 4)

	Variables	22.	23.	24.	25.	26.	27.
Pandemic controls	22. Fear of COVID-19	1					
	23. Monitor/Blunt Scale	0.088 (0.020)	1				
	24. Perceived Vulnerability to Disease	0.061 (0.106)	0.089 (0.018)	1			
	25. Intolerance to Uncertainty	0.374 (0.000)	0.086 (0.022)	0.086 (0.022)	1		
	26. Job Loss	0.106 (0.005)	0.102 (0.006)	0.020 (0.602)	0.075 (0.045)	1	
	27. Prolific Importance	0.163 (0.000)	0.125 (0.001)	0.024 (0.528)	0.188 (0.000)	0.309 (0.000)	1
	M	15.17	8.769	58.32	35.43	232 of 747	6.897
	SD	6.14	2.167	6.13	8.62		1.287

Note: p-value of the pairwise correlations shown in parenthesis.

Table A.11: OLS regressions of Study 3 with demographic or pandemic controls

	Dependent Variable			
	Change in # of moves [s.d.] (1)	**Change in time per move [s.d.]** (2)	**Change in # of moves [s.d.]** (3)	**Change in time per move [s.d.]** (4)
Framing disposition	0.324 (0.116, 0.532)	0.225 (-0.037, 0.487)	0.339 (0.125, 0.553)	0.245 (-0.017, 0.506)
Implementation disposition	0.110 (-0.098, 0.319)	0.208 (-0.055, 0.470)	0.125 (-0.091, 0.342)	0.297 (0.033, 0.562)
Gender	-0.029 (-0.115, 0.057)	0.007 (-0.101, 0.116)		
Age	-0.002 (-0.014, 0.009)	0.007 (-0.007, 0.022)		
Postgraduate	-0.093 (-0.184, -0.003)	-0.013 (-0.126, 0.101)		
Reader	0.074 (-0.014, 0.162)	0.051 (-0.059, 0.162)		
Fear of COVID-19			0.004 (-0.011, 0.020)	-0.011 (-0.030, 0.007)
Monitor/Blunt Scale			-0.015 (-0.056, 0.027)	0.002 (-0.048, 0.053)
Perceived Vulnerability			0.005 (-0.009, 0.020)	-0.003 (-0.020, 0.015)
Intolerance to Uncertainty			-0.014 (-0.025, -0.002)	0.013 (-0.001, 0.027)
Job Loss			0.046 (-0.053, 0.146)	0.115 (-0.006, 0.236)
Prolific Importance			0.056 (-0.017, 0.129)	-0.081 (-0.171, 0.009)
Constant	-0.026 (-0.445, 0.392)	-0.189 (-0.716, 0.337)	-0.229 (-1.266, 0.807)	0.503 (-0.764, 1.769)
Observations	747	747	710	710
R2	0.023	0.008	0.026	0.020
Adjusted R2	0.015	0.0003	0.015	0.009
Residual Std. Error	1.182 (df = 740)	1.487 (df = 740)	1.191 (df = 701)	1.455 (df = 701)
F Statistic	2.860 (p=0.009, df=6)	1.033 (p=0.402, df=6)	2.339 (p=0.175, df =8)	1.807 (p=0.073, df=8)

Note: 95% confidence intervals shown in parentheses.

Table A.12: Robust linear regressions of Study 3 with demographic controls or pandemic control

	Dependent Variable			
	Change in # of moves [s.d.] (1)	**Change in time per move [s.d.]** (2)	**Change in # of moves [s.d.]** (3)	**Change in time per move [s.d.]** (4)
Framing disposition	0.199 (0.029, 0.369)	-0.009 (-0.128, 0.110)	0.232 (0.058, 0.406)	0.001 (-0.121, 0.122)
Implementation disposition	0.060 (-0.111, 0.230)	0.044 (-0.075, 0.163)	0.073 (-0.103, 0.249)	0.078 (-0.045, 0.201)
Gender	-0.035 (-0.106, 0.035)	0.009 (-0.040, 0.058)		
Age	-0.007 (-0.016, 0.003)	0.008 (0.002, 0.015)		
Postgraduate	-0.089 (-0.163, -0.015)	0.016 (-0.035, 0.068)		
Reader	0.066 (-0.006, 0.138)	0.028 (-0.022, 0.079)		
Fear of COVID-19			-0.001 (-0.013, 0.012)	-0.008 (-0.017, 0.0003)
Monitor/Blunt Scale			-0.018 (-0.051, 0.016)	0.002 (-0.022, 0.025)
Perceived Vulnerability			0.0001 (-0.012, 0.012)	-0.002 (-0.006, 0.010)
Intolerance to Uncertainty			-0.012 (-0.021, -0.003)	0.001 (-0.005, 0.008)
Job Loss			0.065 (-0.016, 0.145)	0.014 (-0.042, 0.071)
Prolific Importance			0.018 (-0.041, 0.078)	-0.018 (-0.060, 0.023)
Constant	0.178 (-0.164, 0.520)	-0.219 (-0.458, 0.021)	0.418 (-0.425, 1.262)	0.147 (-0.442, 0.735)
Observations	747	747	710	710
Residual Std. Error	0.837 (df = 740)	0.601 (df = 740)	0.874 (df = 697)	0.588 (df = 697)

Note: 95% confidence intervals shown in parentheses.

Table A.13: Robust regressions of Study 3 with demographic and pandemic controls

	Dependent Variable	
	Change in # of moves [s.d.] (1)	**Change in time per move [s.d.]** (2)
Framing disposition	0.197 (0.022, 0.372)	-0.002 (-0.123, 0.120)
Implementation disposition	0.057 (-0.120, 0.234)	0.069 (-0.054, 0.192)
Gender	-0.029 (-0.103, 0.046)	0.020 (-0.032, 0.072)
Age	-0.008 (-0.018, 0.002)	0.008 (0.001, 0.014)
Postgraduate	-0.097 (-0.173, -0.021)	0.010 (-0.043, 0.063)
Reader	0.060 (-0.014, 0.134)	0.027 (-0.025, 0.078)
Fear of COVID-19	0.0005 (-0.012, 0.013)	-0.009 (-0.018, 0.0002)
Monitor/Blunt Scale	-0.016 (-0.049, 0.018)	0.003 (-0.021, 0.026)
Perceived Vulnerability	-0.003 (-0.015, 0.009)	0.003 (-0.005, 0.012)
Intolerance to Uncertainty	-0.013 (-0.022, -0.004)	0.002 (-0.004, 0.008)
Job Loss	0.060 (-0.020, 0.141)	0.012 (-0.044, 0.068)
Prolific Importance	0.017 (-0.043, 0.076)	-0.015 (-0.056, 0.027)
Constant	0.855 (-0.105, 1.814)	-0.231 (-0.898, 0.436)
Observations	710	710
Residual Std. Error	0.871 (df = 697)	0.588 (df = 697)

Note: 95% confidence intervals shown in parentheses.

Table A.14: χ^2 tests for the espoused manipulation checks

Subset of conditions compared	χ^2	**df**	**p-value**	ϕ	ϕ **95% C.I.**
Three conditions	56.89	2	0.000	0.475	[0.347, 0.595]
Implementation vs. Framing	52.86	1	0.000	0.561	[0.410, 0.712]
Implementation vs. Control	24.57	1	0.000	0.382	[0.231, 0.534]
Framing vs. Control	5.92	1	0.015	0.188	[0.034, 0.339]