

Naturalistic Decision Framing in Computer Mediated Scientific Exploration

Lelyn D. Saner

Dynamic Decision Making Laboratory
Carnegie Mellon University
lsaner@andrew.cmu.edu

Cleotilde Gonzalez

Dynamic Decision Making Laboratory
Carnegie Mellon University
coty@cmu.edu

ABSTRACT

Motivation – This study was conducted to determine how people frame decisions naturally. **Research Approach** - Decision framing was analyzed in the computer mediated Mars Exploration Rover (MER) mission planning process. Outcomes of proposed actions were coded as positively or negatively framed, with rates of framing compared relative to action specificity and expertise. **Findings** - It was found that positive framing was preferred in general, and that increased expertise led to more balanced consideration of positive and negative outcomes. **Research Limitations/Implications** – This study was exploratory and experimentation is necessary to formalize the observed patterns further. **Originality/Value** - The findings suggest that providing decision-makers additional information about the costs of using technology may more rapidly lead to better mental representations in technology-supported decision-making. **Take away message** – Different kinds of expert behavior promote different patterns of decision framing relative to how technology is used for decision-making.

Keywords

Framing, Decision Making

INTRODUCTION

The concept of a decision frame was originally introduced by Tversky and Kahneman (1981), who defined it as “the decision-maker’s conception of the acts, outcomes, or contingencies associated with a particular choice.” They also pointed out that, although a decision frame is influenced partly by the way the decision problem is formulated, it is also subject to the “norms, habits, and personal characteristics of the decision-maker.” The possible influences on framing therefore span multiple levels of psychology (e.g. cognitive, affective, personality, etc.), and different studies invariably focus on different components of the decision representation (Tversky & Kahneman, 1981; Kuhberger, 1998; Levin, Schneider, & Gaeth, 1998; Levin, Gaeth, Schrieber, & Lauriola, 2002). Furthermore, although prospect theory (Kahneman & Tversky, 1979) assumes that people respond passively to presented framed information (Thaler, Tversky, Kahneman, & Schwartz, 1997), research in cognitive psychology has shown that mental representations of a problem do not always coincide exactly with the representations people are presented (Hayes & Simon, 1977). People draw selectively from the available information to construct mental representations such as decision frames. Therefore, in order to predict decision-making behavior in a broad range of situations, it is important to be able to predict what types of decision frames people will adopt to motivate their choices.

One of the ways in which natural situations constrain decision-making is that they differ on the availability of tools and resources. For example, a person who owns a vehicle has one more transportation option than a person who does not have one. However, if that second person has a large amount of money, he or she has a better chance of purchasing a car (i.e. adding car as an option) than someone who has very little money. As such, an important consideration in situations where goals are implemented by way of technology is that the technology itself, and the resources that are needed to use it, can become a prominent component of the decision frame. Using the same example, when a person is trying to decide how to get somewhere, there are the general attributes of the decision that are not particular to individual options, such as how far away the destination is and how quickly one wants to get there. However, when considering the car as an option, there are attributes that are specific to the car that can limit its value as an option at any given point in time (e.g. engine running properly, sufficient amount of fuel for the distance to be driven, etc.) If one considered a bicycle instead of a car, fuel level would not be an issue. In other words, technology often adds attributes to be framed and its impact needs to be accounted for when predicting the way people will frame decisions.

In the literature on framing effects, most manipulations of decision frames are reducible to changing the positive or negative valence of the information given (Levin, Schneider, & Gaeth, 1998). Rather than looking further at the effects of framing on choice, however, this study was conducted to identify factors that might predict peoples’ decision frames themselves (i.e. predict mental representations of decisions). We ran an exploratory study of decision framing in a naturalistic setting in order to determine whether people are more likely to select positive or negative information for

expressing their options. In this paper we report some of the findings from that study that are related to the role of technology in decision framing.

For our sample, we conducted an analysis of the planning process for the U.S. National Aeronautics and Space Administration’s (NASA) Mars Exploration Rover (MER) Mission. In this situation, technology played a dual role. First, all of the mission goals were conducted via the operations of the robot rovers, such that all of the alternatives that were considered were a function of what the rovers were able to do with the objects of interest that were at hand, as well as what data they were able to process with the resources available at the time. Second, because the rovers were millions of miles away from the people who were controlling them, all communications between the rovers and the planners were mediated by technology. Computers had to be used to sequence conceptual plans into instruction programs and the programs had to be transmitted by radio signals. If a rover malfunctioned in any way, and if it could not be sent a program to repair itself, that functionality was permanently lost. In this paper, we examine two technology related factors and their impact on frame selection.

First, in naturalistic situations, isolating a small set of clear alternatives to choose from is not always an easy task, and different situations often require that different decision-making strategies be used (Payne, Bettman, and Johnson, 1992, 1993) Alternatives rarely come pre-packaged in such environments and often need to be generated by decision makers with information about the circumstances. When particular technology is used to carry out actions, however, there is often a small set of clearly defined possible operations available to be utilized. Different combinations of these operations constitute different alternatives, each of which will have advantages and disadvantages. It may be that two alternatives will differ by only one or two low-level operations or parameter settings, so that they are distinct from each other but very closely related. Distinguishing unique alternatives can still be a challenge from an observer’s standpoint, however, so one factor to consider in predicting frames is how specifically the use of technology is described when possible actions are proposed as alternatives. The question, therefore, is whether the level of action specification will predict positive or negative expression of outcomes.

Expertise is also a factor that has been shown to be related to mental representations (Chi, Feltovich, & Glaser, 1981; Schunn, McGregor, & Saner, 2005). Particularly when it comes to selecting strategic actions to take (Schunn, et al., 2005), people with a great deal of experience in a domain have a greater number of strategies in their repertoire than those who are novices in the same domain. In addition, those strategies are usually of better quality and experts usually apply them more appropriately than novices do. One measure of expertise in this context is the general increase in experience that scientists gained between the beginning of the mission and the end. Another advantage of using computer technology to implement plans is that the programs that have been written for one activity sequence can be saved and used again later when the same sequence is needed in a similar situation. With documentation of how it was done before, parameters and targets can be changed and the rest of the program left unchanged, eliminating the need to rewrite similar sequences from scratch. Reusing previous work is therefore indicative of expert behavior, and a difference in references to positive and negative outcomes related to patterns of reuse would suggest that the expertise is affecting mental representations as well.

In the following sections, we will first describe the features of the MER context in more detail. We will then outline the methods used for coding the predictors and decision frames. We will report our findings regarding the effects of action formulation and expertise on frame selection. Finally, we will discuss the importance of these factors when technology is used for decision-making.

METHODS

Overview of the Mars Exploration Mission: In the summer of 2003, NASA launched two robot probes to Mars to begin a major planet surface exploration mission. These “Mars Exploration Rovers” (MERs) landed on the surface in February of 2004. Each rover was intended to be fully operational for at least 90 Mars days, called “sols,” and this time period was referred to as the *nominal mission*. The rovers have stayed functional beyond that period, however, and the mission is ongoing. The goal of the mission is to collect as much data as possible about the surface conditions on the planet. The aspects of most interest are the weather patterns, the characteristics of the terrain, and the geological composition of surface material. The rovers are identical in construction. Each is equipped with solar panels for recharging power cells, three different communication antennae, and data is stored in a flash-memory device where it is buffered for transmission back to Earth. To perform its data collection functions, each rover has six specialized scientific instruments, which are mounted on the rover in locations appropriate to the uses.

The planning of rover operations was done by teams of scientists from five areas of science that stayed onsite at the NASA Jet Propulsion Lab at the California Institute of Technology during the nominal mission. There were geologists, geochemists, mineralogists, atmospheric scientists, and NASA’s long-term



planning researchers. Every science group was represented by 8-10 scientists who ranged in experience from being graduate students in training to being tenured professors who were leaders in their fields. Many scientists worked on both MERs over the course of the mission. On any given sol, however, the scientists in each area were divided into sub-groups and exclusive interdisciplinary teams ran operations for each MER. All rover operations, including driving, scientific analysis, communications, and maintenance were managed through a daily regime of large group planning meetings and periods of breakout discussion. The scientists followed the same general schedule of events each day and kept to a Mars sol schedule (Local Sol Time, LST).

The first meeting on each sol was held to review what was returned from the previous sol's activities. After this initial meeting scientists met in their topic area groups to analyze and interpret the data that was pertinent to their interests. Not every group got new data each day, but each group was represented at the meeting in order to be updated on the status of the rover and to determine if there were opportunities available that their group would want to pursue. The second large group meeting of the day was where scientists presented their requests for what activities to conduct on the following sol. As such the first meeting was crucial to decision-making in terms of specifying the preconditions for the decisions, but the second meeting is where the major choices were made about what activities to pursue. If the planners learned in the first meeting that a spectrometer that they wanted to use was malfunctioning, they would have to use their discussion time to decide whether a different instrument would suffice or whether to abandon that alternative altogether. Across these two meetings, the situation status was monitored, advantages and disadvantages of different courses of action were weighed, constraints were considered, and trade-offs were negotiated until a rough conceptual plan was derived for what activities the rover would engage in during the following sol. Then, after a period of some finer grain tuning of the plan by the individual science groups, there was a meeting to finalize the sequence of implementation for the selected operations. This sequence of meetings and planning tasks was conducted each sol, for each rover, throughout the entire mission.

The Mars exploration mission was not a situation of simple choices between given alternatives. The primary task, in fact, was to generate the alternatives and outline the contingencies associated with them. The scientists and engineers had to have a plan, a sequence of prioritized operations, as developed as possible by the end of each day that would fill the available time in the following sol. Final decisions were not often reached within one meeting because the specific choices that led to the plan were distributed across all of the scientists' interactions throughout the day. However, because the focus of analysis in this study was on the representation of alternatives and not on the choice between them, it was not necessary to locate the final moment of choice on each action. In addition, although single individuals suggested possible alternatives, no single person made final decisions on what the plan would include. Decisions were reached by way of a consensus building process among all of the participating planners.

Design: Because they were in different locations on the planet, operations for each MER were planned separately and treated as separate missions (MER-A and MER-B). The large group meetings were video-recorded for a subset of the 90 sols of the nominal mission by a team of human factors researchers whose visits to the site were spaced throughout the life span of the missions. Cameras were inconspicuously placed around the room and were not salient to the scientists while they worked. The cameras remained fixed in those locations and did not move to follow who was speaking in the large group meetings.

Procedure: A sample of 18 recorded meetings was selected to be analyzed in this study, nine from each MER, and all were the second large-group meeting of the day. Selection was based on the audibility of the dialogue in the meetings and their distribution across the mission lifetime. The selected planning meetings were transcribed from the video data and segmented into individual statements. Segments were divided whenever a new proposition was made, but not all statements were related to decision-making. Only statements pertaining to the plan for the next sol were included in the analysis. Comments about meeting logistics and general information were excluded as off task. Of the material that was devoted to rover activity planning, segmenting was further verified to create new segments whenever a new action was proposed or a unique outcome for an action was mentioned. Each segment represented a single decision proposition that, once stated, could be considered in the deliberation process as an individual alternative. A verbal protocol analysis was performed on the segmented transcripts focusing on the level of planning in the suggested actions and the framing of the outcomes (Ericsson & Simon, 1993). Codes for actions and outcomes were pre-defined, as described below, and coded by two trained coders independently. For further details about the design, procedure, or the coding criteria described next, please contact the first author.

Coding of Outcomes: The main determinant of positive or negative frame is the gain or loss in some resource. There was a relatively well-defined set of resources (or attributes) in the MER context that could be gained or lost in the course of conducting operations. Some of these attributes were moment-to-moment opportunities, including the opportunity to perform an operation, to gather scientific data, and to drive some distance. Other resources were electric power, time, data volume (i.e. flash drive memory storage), and the ability to fit activities into the overall plan. An outcome was framed positively when a statement referred to gaining or saving some amount of a resource (e.g. "we'll get a picture of that type of soil" or "we'll save 20 minutes by going in that direction"), resolving an existing problem, or making the overall situation more advantageous (e.g. "It will probably be a smoother drive if we go that route.").

Negative framing was the opposite; coded when a loss of some resource or some cost was identified, or when doing an operation might lead to a new problem (e.g. "taking the picture at that resolution might lead to a flash issue.").

Coding of Actions: The rover operation proposed in each statement was the core of an action. Proposed actions were coded as being on one of three levels of action specificity, based on the level of detail that was explicitly used to describe the suggested activity and on the scope of the goal that would be accomplished by doing the operation. There was also a standing set of mission milestones that were imposed by NASA to measure overall mission success and which served as large umbrella goals (e.g. a minimum distance to be driven by each rover). The criteria for distinguishing levels of activity were specific to the Mars exploration mission context. They were determined with input from researchers at the NASA Ames research laboratory.

The most general level of activity specification, *high-level* planning, included statements about the accomplishment of the mission success criteria (e.g. "We need to drive another 250 meters to reach that milestone"), references to particular long-range scientific goals (e.g. "We should go to Victoria crater"), and explicitly mentioned multi-sol plans (e.g. a four-day exploration of a large rock formation or a three-day super-drive with several stops). Some of these statements did not even specify what tools would be used for the activity. The next level of planning, however, required them to do so. Because the planners were most concerned, on a given sol, with sequencing an activity plan for the following sol, references to specific instruments, either singly or as a set, were the primary examples of *basic-level* planning, (e.g. "... so we're either going to do Microscopic Imaging or Mini-TES"). Basic-level planning was also often a matter of referring to a category of activity. For example, because there were several instruments mounted on the Instrument Deployment Device (IDD) arm, planners often referred to "IDD-ing" as an activity. In those cases, it was understood that one or more of the instruments would be used for an operation, but that the particular combination and targets would be specified later. Finally, the most detailed level of planning, *low-level* planning, involved the identification of specific instruments to be used on specific targets, in some cases with specific parameters set. This planning was coded if an instrument-target pairing was identified or if a setting was proposed for a specific operation (e.g. what target to take a picture of and what filter settings to use, what resolution to do spectroscopy at, how deep to scratch a rock, etc.). Any discussion that was devoted to the specific ordering, timing, and prioritization of operations was also classified as low-level planning.

Coding of Expertise: The expertise of the planners was not measured directly, but rather as a function of mission progress and sol meetings. The meetings that occurred early in the mission (prior to MER-A Sol 52) were compared to those that occurred later in the mission (after MER-A Sol 52), with the assumption that expertise was lower during the early phase and increased as they progressed to the later phase. The other measure of expertise was focused on the reuse of previously generated sequences of operations. The plans formulated in each meeting were coded for the relative amount of novel material. A decrease in plan novelty, therefore, would be an indicator of increased expertise. The highest proportion of novelty observed in any meeting was 90%, so a split-half division was performed on the sample of meetings. Based on the proportion of novel planning within it, each meeting was coded as having low novelty (i.e. less than 45% novel planning) or high novelty (i.e. more than 45% novel planning).

RESULTS

When all of the transcripts had been segmented, there were 4992 individual segments. Of these, 597 statements (12%) were devoted to planning with actions explicitly connected to outcomes. Most operations were specified at the basic-level of detail, with basic-level planning statements accounting for a significantly higher proportion of planning (Basic=50%, S.E.= 2%, n=597) than both of the other two categories combined, $X^2(1, N=597)=73.88, p<.005$. There was no relationship between planning level and framing, however (see Figure 1). The preference was for positive framing at all levels of action detail. The high-level action proposals exhibited the highest proportion of positive framing, 81% (S.E.= 7%, n=26). This was followed by basic-level actions, of which 69% (S.E.= 3%, n=205) were framed positively. And at a slightly smaller rate, low-level proposed actions were framed positively 64% (S.E.= 3%, n=172) of the time. Even though the proportion of positive framing decreased as planning became more detailed, none of the pair-wise differences between these proportions were significant.

To assess the effect of expertise on framing, a 2*2 ANOVA was conducted with both general experience (early vs. late mission phase) and plan reuse (high vs. low novelty sol plan) as predictors of decision frame. As shown on Figure 2 there was a main effect of mission phase, with a drop in the proportion of positive framing from early to late mission occurring at both levels of plan novelty, $F(1, 507)=5.48, p<0.05$. There was also a significant main effect of plan novelty, $F(1, 507)=7.71, p=0.01$, such that there was a higher rate of positive framing on the low novelty sols than on the high novelty sols within each phase of the mission. Because decreased plan novelty is associated with increased expertise, the effect of novelty on framing was actually the opposite that of the mission phase expertise measure. There was no interaction between these two measures of expertise.

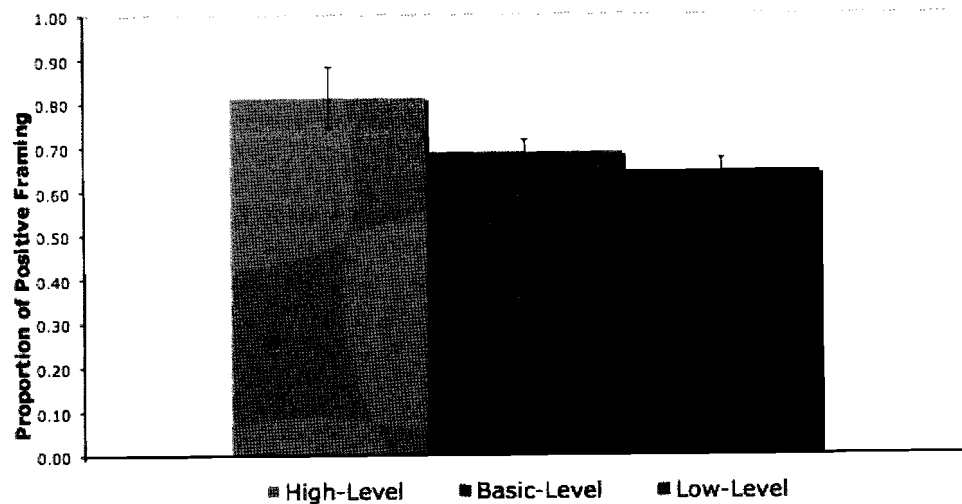


Figure 1. Proportion of Positive Framing by Planning Level

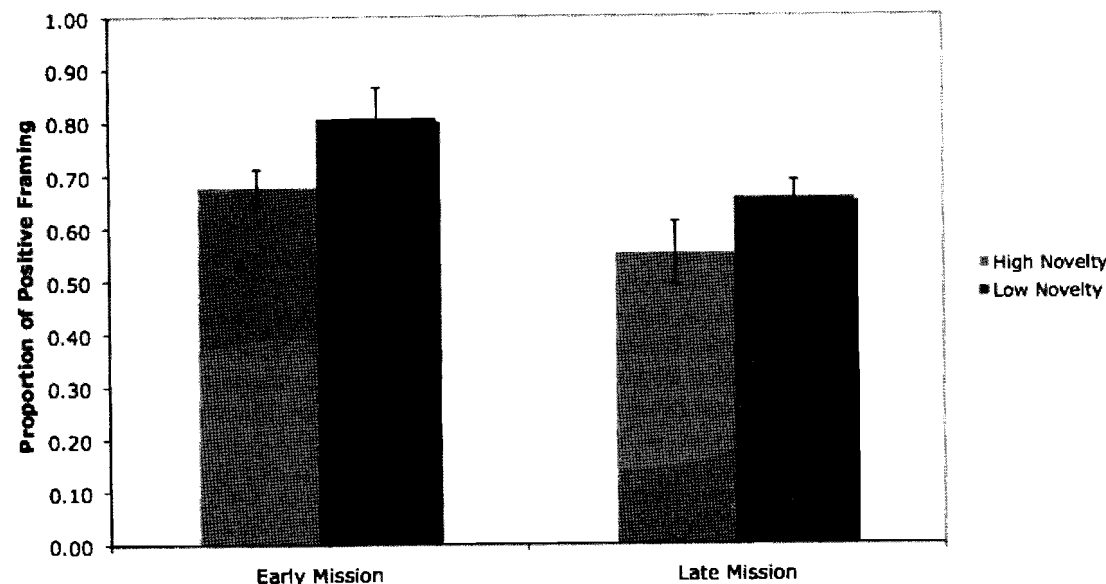


Figure 2. Proportion of Positive Framing by Expertise

CONCLUSION

In this study, we examined the question of how peoples' selection of decision frames is affected when technology is used to plan and carry out chosen courses of action. We observed that people generally prefer positive framing regardless of the level of detail at which they specify proposed actions. This suggests that, when decision makers consider what tools to use for a task, they are more likely to focus on what they will gain with each tool than on what it will cost them to use it. At the same time, although the effect was not statistically significant, the trend in frame selection across planning levels indicates that the balance of positive and negative framing improves with more highly specified actions. One implication of this finding is that people may need to be more directly reminded of what it costs to use a particular piece of technology for a task in order for them to make rational decisions. Less costly tools may be available that will do the same job in a more cost-effective way, and awareness of such options could save resources.

We also observed that, as decision makers gain general experience with the technology they are using, their bias toward positive framing decreases. This is congruent with prior research showing that experts develop more complete and balanced representations of problems (Chi, et al., 1981), and it indicates that experience leads to better cost-benefit

accounting practices. At the same time, regardless of general experience, when plans had a higher proportion of reused material, which is also indicative of expert thinking, the rate of positive framing was also higher.

One possible explanation for why the two measures of expertise had opposite effects is that the specificity of experience represented by each measure may have been different. It may be that a general increase in experience made planners more cognizant of what negative outcomes could potentially be encountered at any time in the process of exploring Mars, which might explain the observed increase in negative frames during late mission. At the same time, specific experience with particular activities would have given the planners specific actual knowledge of what would be gained by doing those activities in different conditions. Reuse of developed plan material is a sign of expertise because it is more efficient than developing a plan from the ground up and it can only be done when a new situation is recognized to be similar to one already experienced. The known, real gains might have been used to motivate the reuse of operational procedures in favor of untried procedures where the outcomes were less certain, which might explain why framing was more positive on the low novelty sols. Furthermore, because the outcomes of used procedures became matters of record immediately, the planners' general level of experience would not influence this process, which would explain the observed consistency of the difference in both early and late mission phases.

One implication of these findings is that improved training, perhaps in the form of high-fidelity simulations and scenario-based problem solving, is likely to accelerate decision-makers' development of rich representations that make them more responsible with their resources. In addition, technology should be designed to support detailed documentation of methods that have been proven effective. Ideally, that documentation should include statistics regarding the resource costs associated with using the methods under different conditions, as well as the payoffs.

Given the exploratory nature of this research, it is impossible to draw strong inferences about what causes people to prefer one frame to another. Further research is needed to better characterize the representation selection patterns observed in our study. Our next steps with this research include further analysis of sequences of statements related to the same proposed actions, as well as exploring the link between frame selection and action implementation. Based on this study, however, there is evidence that technology does impact mental representations, and a better understanding of the process by which it does so could lead to improved performance in any number of domains.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Christian Schunn (University of Pittsburgh, PA) and members of the NASA-Ames Human Factors research lab for consulting on this research. Thanks also to Zarna Shah and Jamie Kammer for assistance with data coding.

REFERENCES

- Chi, M. T. H., Feltovich, P., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5, 121-152.
- Hayes, J. R., & Simon, H. A. (1977). Psychological differences among problem isomorphs. In N. J. Castellan, D. B. Pisoni, and G. R. Potts (Eds.), *Cognitive Theory*. Hillsdale, NJ: Erlbaum
- Kahneman, D. & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263-292.
- Kuhberger, A. (1998). The influence of framing on risky decisions: A meta-analysis. *Organizational Behavior and Human Decision Processes*, 75(1), 23-55.
- Levin, I. P., Gaeth, G. J., Schreiber, J., & Lauriola, M. (2002). A new look at framing effects: Distribution of effect sizes, individual differences, and independence of types of effects. *Organizational Behavior and Human Decision Processes*, 88(1), 411-429.
- Levin, I. P., Schneider, S. L., & Gaeth, G. J. (1998). All frames are not created equal: A typology and critical analysis of framing effects. *Organizational Behavior and Human Decision Processes*, 76(2), 149-188.
- Payne, J. W., Bettman, J. R. & Johnson, E. J. (1992). Behavioral decision research: A constructive processing perspective. *Annual Review of Psychology*, 43, 87-131.
- Payne, J. W., Bettman, J. R. & Johnson, E. J. (1993). *The Adaptive Decision Maker*. New York: Cambridge University Press
- Schunn, C.D., McGregor, M. U., & Saner, L. D. (2005). Expertise in ill-defined problem solving domains as effective strategy use. *Memory and Cognition*, 33(8), 1377-1387.
- Thaler, R. H., Tversky, A., Kahneman, D., & Schwartz, A. (1997). The effect of myopia and loss aversion on risk taking: An experimental test. *The Quarterly Journal of Economics*, 112, 647-661.
- Tversky, A. & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211(4481), 453-458.