Effect of Intention on Outcome Bias in Decision Making—Implications for Safety Management

Atsuo Murata*, Tomoko Nakamura

Department of Intelligent Mechanical Systems, Graduate School of Natural Science and Technology, Okayama University, Okayama, Japan
Email: murata@iims.sys.okayama-u.ac.jp

Received 5 November 2015; accepted 28 November 2015; published 2 December 2015

Copyright © 2015 by authors and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY).
http://creativecommons.org/licenses/by/4.0/

Abstract

The evidence of outcome bias was explored in a two-player (Player 1: allocator and Player 2: evaluator) economic game experiment where the reward allocation was made between two players. The experimental factors were the intention of an allocator (Player 1), the type of chosen dice (selfish, fair, and generous), and the outcome (selfish, fair, and generous). The outcome bias occurred when the type of dice chosen by the allocator (Player 1) was not only a selfish one but also a generous one. The comparison between the two conditions (intentional and no-intentional conditions) definitely showed that Player 2 punished Player 1 to a larger extent when the outcome was disadvantageous for Player 2 (selfish outcome) and Player 2 rewarded Player 1 when the outcome was advantageous (generous outcome) irrespective of whether the die was chosen out of the three types intentionally or not. Moreover, the outcome bias was not observed when the outcome was fair. Thus, we could verify the hypothesis that we are readily got trapped in the outcome bias. Some implications were given for safety management that put more emphasis on the process than on the outcome.

Keywords
Decision Making, Outcome, Process, Intention, Outcome Bias, Cognitive Bias, Safety Management

1. Introduction

Reason [1] stated the judgmental heuristics, the irrationality and the cognitive “backlash” were potential risk

*Corresponding author.

http://dx.doi.org/10.4236/jbbs.2015.513053
factors of human errors leading to incidents, crashes, collisions, or disasters. However, he has not demonstrated a systematic model on how such cognitive biases are related to the distorted decision making, and become a trigger of critical incidents such as crashes, collisions, and disasters. Dekker [2] pointed out a situation of developing a vicious circle of repeated occurrences of similar incidents, crashes, collisions, or disasters, and suggested that the cause of such a vicious circle was the hindsight bias, pointed out by Fischhoff [3] or the outcome bias [4] that put more emphasis on outcomes than processes. In other words, he suggested that the analysis of incidents, crashes, collisions, or disasters should be conducted not with hindsight or by considering outcomes but with foresight or in consideration of processes.

Several studies explored the outcome bias in a variety of areas [5]-[16]. Lipe [5] investigated the effects of outcomes (information ex post to the decision) on the variance investigation decision by a manager, and showed that outcomes apparently affected the investigation decision. Marshall and Mowen [6] carried out an experimental investigation of the outcome bias in evaluation of salesperson’s achievements. He showed that the outcome bias affected a salesperson’s achievement evaluation, and a salesperson’s achievement was not evaluated based on the process of the sales activities. A higher outcome led to a higher assessment of salesperson’s achievements. Reimers and Butler [7] examined the effect of outcome knowledge on auditors’ judgmental evaluations, and showed that the evaluation of auditor’s judgment was considerably influenced by the outcome. Dekker [8] and Dekker [9] showed that the worse the consequence of the action was, the more any preceding acts were seen as blameworthy. Baron and Hershey [17] explored an outcome bias in evaluating a decision maker on either medical matters or monetary gambles. In medical matters, the participants rated the decision maker as more competent when the outcome was more favorable than when it was unfavorable. In monetary gambles, the participants tended to rate the thinking of a player as better when the outcome not chosen turned out to be inferior than when it turned out to be superior. In spite of being instructed not to put an emphasis on outcomes in evaluating a decision maker, the participants tended to evaluate the decision maker based on outcomes. Thus, it was proven that an outcome played a direct and major role when evaluating the quality or right and wrong of decision making. In this manner, outcome biases are ubiquitous in a variety of activities.

Recently, collisions, crasher, or disasters occur frequently due to the violation of safety rules or regulations. These include a disaster such as the unintended critical mass accident at JCO Tokai Works Test Facility uranium processing plant [18]. In spite of increasing opportunities of accessing the data base of the detailed analysis of crashes, collisions, or disasters, similar and undesirable events repeatedly occur. We generally tend to make a thing of not the process but the outcome in decision making. Such a cognitive property hinders appropriate safety management since the process that potentially induces crashes, collisions, or disasters is ignored as long as our activities result in proper outcomes on the surface.

In this manner, our propensity to attach a great deal of importance to the outcome becomes a potential risk in a variety of activities for safety management. In the critical mass disaster at JCO Tokai Works Test Facility uranium processing plant, illegal alternations of the manufacturing process of uranium solution were conducted six times. Although the manufacturer recognized that the manufacturing process must not be altered to assure plant safety, the successful alternation (outcome) induced an optimistic bias (an optimistic feeling that a minor alternation was not problematic because this induced no major problems) and became a trigger of escalating the alternation of the manufacturing process by apparently placing a more emphasis on efficiency than on safety.

In behavioral economics, the following game is frequently used to explore the role of outcome and intention in moral (ethical) judgment [19]-[21]. One player makes an allocation of money and another player rewards or punishes an allocator by adding or subtracting money depending on the allocation. However, in such a game situation, it is impossible to know whether the response to an allocation is based on the actual allocation, the intended allocation, or an interaction between these two factors. To cope with such a problem, Cushman et al. [22] examined how we responded to others’ accidental behavior. They investigated the balance of outcomes and intentions in a two-player economic game where the intentions and the outcomes in monetary allocations were sometimes mismatched. Player 1 was required to allocate $10 between oneself and Player 2 by rolling one of three types of dice (The first die had a high probability of a selfish outcome that was favorable for Player 1, the second die had a high probability of a fair outcome, and the third die had a high probability of generous outcome that was favorable for Player 2). Player 2 could infer Player 1’s intention on the basis of the die choice of Player 1. Any of the three dice can yield any of three possible outcomes (selfish, fair, and generous outcomes). Consequently, they found that Player 2’s response was affected not by the intention of Player 1 but by the accidental outcome of Player 1’s roll of the die. This corresponds to the outcome bias.
However, the method and the technique for quantifying and measuring the amount of outcome bias have not yet been established in Cushman et al. [22]. In Cushman et al. [22], the experiment was conducted under the condition where Player 1 can freely choose one of the three types of dice. In order to separate the effects of the outcome and the intention and accurately measure the amount of outcome bias, it is necessary to carry out a similar experiment even under the condition where the type of dice cannot be chosen intentionally and compare the results (allocation of money) when the die could be chosen intentionally with those when the die could not be chosen intentionally. Few studies seem to explore how the intention of choosing the actions affected the outcome bias. The difference of punishment or reward between the two conditions must be taken into account so that we can get insights into the outcome bias. Taking into account such an experimental condition where the type of dice that yields one of the three possible outcomes (selfish, fair, and generous outcomes) cannot be chosen intentionally, we conducted an experiment to further explore the outcome bias.

Using a two-player (Player 1: an allocator and Player 2: an evaluator of Player 1’s allocation) economic game experiment where reward allocation was made between two players, it was examined how the outcome bias was affected by the intention of the allocator (whether the type of dice can be chosen intentionally or not). It was examined how the outcome of the game (selfish, fair, or generous), the intention of the allocator (intentionally, or non-intentionally), and the actually chosen die (selfish, fair, or generous) affected the situation of outcome bias. Based on such investigations, some implications for safety management were given from the perspective of outcome bias.

2. Method

The study was conducted on November, 2014. The investigation was approved by the Ethics Committee, Dept. of Intelligent Mechanical Systems, Okayama University.

2.1. Participants

Twenty-one undergraduate or graduate students (eleven male and ten female) aged from 20 to 23 years old participated in the experiment. Participants without knowledge or skill in psychology or behavioral economics were recruited. All participants agreed with participating in the experiment after receiving a brief explanation of the contents of the experiment.

2.2. Task

Player 1 was required to allocate $10 between himself or herself and Player 2 by rolling one of the following three types of dice (see Figure 1). The dice were the same with those designed and used in Cushman et al. [22].

Die A (selfish): Player 1 receives $10 by rolling a 1, 2, 3, or 4; If Player 1 rolls a 5, the reward is divided evenly ($5/$5); If Player 1 rolls a 6, Player 2 receives $10.

Die B (fair): Rolling a 1 provides Player 1 with $10; Rolling a 6 provide Player 2 with $10; If Player 1 rolls a 2, 3, 4, or 5, the reward is divided evenly ($5/$5).

Die C (generous): Player 1 receives $10 by rolling a 1; If Player 1 rolls a 2, the reward is divided evenly...
($5/$5); If Player 1 rolls a 3, 4, 5, or 6, Player 2 receives $10.

It must be noted that Player 1 does not have a perfect control over the allocation of $10. Even if Player 1 chooses the die with the highest probability of aiming at a desired outcome according his or her intention, any die can result in any of three possible outcomes. For example, even if Player 1 chooses the die A (selfish intention), Player 1 does not necessarily get a desired outcome (in this case, rolling a 1, 2, 3, or 4). Therefore, Player 1 has partial but imperfect control over the allocation of $10.

2.3. Design and Procedure

The game was carried out both when the type of dice can be chosen with own intention (intentional choice condition) and when it cannot be chosen intentionally (non-intentional choice condition). The intentional choice condition means that Player 1 can intentionally choose the type of dice. The non-intentional choice condition means that the type of dice was assigned by the experimenter (experimental program) and Player 1 cannot choose the type of dice intentionally.

The experimental factors were whether the die type could be chosen intentionally or not, the type of dice (selfish, fair, and generous) and the type of the outcome (selfish outcome (Player 1; $10), fair outcome ($5 for each), and generous outcome (Player 2; $10)). All factors were within-subject factors. The experimental procedure was simulated on a PC. It must be noted that the participants only played the role of Player 2. The computer agent acted the role of Player 1. Player 2 was permitted to assign punishment or reward to Player 1 at his or her disposal according to the outcome, the type of die, and whether the die type was chosen intentionally or not, and the interaction of three factors (outcome by type of die by whether the die type was chosen intentionally or not). Player 2 was given an opportunity to respond to Player 1’s allocation outcome by increasing or decreasing Player 1’s payoff by any amount up to $10.

Cushman et al. [22] used a human (Player 1) vs. human (Player 2) experiment. The choice of Player 1 is inevitably limited because of the fear of punishment by Player 2, and it is expected that Player 1 does not choose the selfish intention (die A). In such a situation, it is actually impossible to observe the reaction (decision) of Player 2 for a variety of chosen dice (selfish, fair, and generous) and outcomes (selfish, fair, and generous). Therefore, Player 2’s attempt to reward or punish Player 1 was enforced by the experimenter in only 1/10 of all trials. Only 1/10 of all trials were chosen randomly by the experimenter. Player 1 was able to get away with the selfish allocation (choice) most of the time without fear of punishment in case of the human (Player 1) vs. human (Player 2) game situation.

In Cushman et al. [22], the experimenter has an imperfect control over the intention of Player 1 due to the reason mentioned above. The aim of this study was to observe the reaction (decision) of Player 2 for a variety of dice (selfish, fair, and generous), outcomes (selfish, fair, and generous), and the intention of the allocator (Player 1). If we make a computer program so that the role of Player 1 is played by a computer agent, we can more readily observe the reaction (decision) of Player 2 for a variety of the type of dice, intentions of Player 1, and outcomes of the game. Therefore, this study is different from Cushman et al. [22] in that a computer (Player 1) vs. human (Player 2) experimental paradigm was used.

There were eighteen combinations of intention (two levels: intentionally chosen, and non-intentionally chosen), the type of dice (three levels: selfish, fair, and generous), and outcome (three levels: selfish, fair, and generous). For each of eighteen combinations, the participants assigned punishment or reward to the reaction of Player 1 10 times according to the outcome, the type of dice, whether the die type was intentionally chosen or not, and the interaction of three factors. Each participant carried out a total of 180 assignments (trials) of punishment or reward. The order of performance of 180 trials was randomized across the participants.

3. Results

When the die could be chosen intentionally (see Figure 2), it tended that Player 2 punished Player 1 for the selfish outcome irrespective of the type of dice (selfish die, fair die, or generous die). Even when the die could not be chosen intentionally (see Figure 3), a similar result was obtained. These results indicate that Player 2’s reaction to the allocation mediated by Player 1 is affected not by the intention of Player 1 but by the outcome of rolling a die. In other words, independently of whether the die could be chosen intentionally or not, Player 2 punished the selfish outcome.

When the die could be chosen intentionally, Player 2 tended to reward the generous outcome for three types
of dice (selfish, fair, or generous die choice by Player 1). Even when the die could not be chosen intentionally, Player 1 tended to react similarly. This also indicates that Player 2 puts emphasis not on the Player 1’s intention but on the outcome of Player 1.

As for the fair outcome, the situation of punishment/reward by Player 2 differed between the intentional and non-intentional conditions (see Figure 2 and Figure 3). When the die could be chosen intentionally, Player 2 tended to punish Player 1 irrespective of the type of dice that Player 1 chose for the allocation of $10 (selfish, fair, or generous intention). When the die could not be chosen intentionally, it tended that Player 2 provided Player 1 with a small quantity of reward. The data were statistically analyzed using analysis of variance (ANOVA).

A three-way (outcome by type of dice chosen by Player 1 by intention of Player 1) ANOVA conducted on the punishment/reward revealed significant main effects of outcome ($F(2,369) = 48.259, p < 0.01$) and intention ($F(1,369) = 4.143, p < 0.05$). A significant type of dice by intention interaction ($F(2,369) = 6.103, p < 0.05$) was detected, which indicates a different punishment/reward situation between the intentional and the non-intentional conditions.

A two-way (outcome by type of die chosen by Player 1) ANOVA (Analysis of Variance) conducted on the punishment/reward revealed a significant main effect of outcome ($F(2,369) = 28.609, p < 0.01$) and type of die ($F(2,369) = 3.619, p < 0.05$) for the intentional choice condition. No significant interaction between the two factors was detected. For the non-intentional condition, as a result of a similar two-way ANOVA conducted on the punishment/reward, only a significant main effect of outcome ($F(2,369) = 34.286, p < 0.01$) was detected.

4. Discussion
4.1. Effect of Intention on Outcome Bias
As shown in Figure 2, Player 2 punished Player 1 for the selfish outcome irrespective of the type of dice (selfish
The outcome bias is evident even from the following result. As for the intentional condition (Figure 2), the choice of die A must be punished irrespective of the results of rolling the die A from the rational viewpoint. If the choice of die A is rewarded when the outcome is generous for the evaluator (Player 2), this corresponds to an outcome bias. Although the choice of die B (fair die) or C (generous die) must not be punished from the rational viewpoint, a great deal of importance is attached not to the intentional choice of the die but to the outcome. Therefore, the choice of die B or C was punished when the outcome was selfish. This is also indicative of an outcome bias. Such tendencies were also observed for the non-intentional condition (see Figure 3).

The evaluations by Player 2 for the fair outcome were different between the intentional and non-intentional conditions. Here, it must be noted that the amount of reward or punishment of the fair outcome was smaller than that of the selfish and generous outcomes for both intentional and non-intentional conditions. When it was assumed that Player 1 intentionally chose one of the three dice A-C and the outcome was fair, Player 2 tended to punish Player 1 by a small amount. On the other hand, for the non-intentional condition, Player 2 rewarded the fair outcome irrespective of the chosen die. Especially for the combination of the fair outcome and the selfish die, the evaluation of Player 2 was apparently different between the intentional and non-intentional conditions. This indicates that the participant placed emphasis on not the outcome but the intention when the die could not be chosen intentionally. Different from the extreme outcomes (selfish and the generous outcomes), the evaluation seemed to depend on the intention of Player 1 when the outcome was neutral (fair). The tendency above can be further verified by the statistically significant interaction between the type of die chosen by Player 1 (Dice A, B, or C) and the intention of Player 1 (intentional or non-intentional choice conditions).

The outcome of the game seems to have more strong effects on the punishment/reward decision of Player 2 than the intention of Player 1 when the outcome is extreme (selfish or generous outcome). On the other hand, the evaluation of Player 1 by Player 2 seems to be more strongly affected not by the outcome of the game but by the intention of Player 1 when the outcome is neutral (in this case, fair). Thus, we could verify the hypothesis that we are readily led to the outcome bias irrespective of whether the die could be chosen intentionally or non-intentionally when the outcome is extreme (selfish or generous outcome). In the range of this experiment, the outcome bias does not readily occur when the outcome is neutral (fair). Rather, when the outcome was neutral (fair), the evaluation of Player 2 was affected by the intention of Player 1.

Fishhoff [3] empirically showed that a hindsight bias existed in our estimation of the occurrence probability of some events. In his experiment, participants were given scenarios and required to estimate occurrence probabilities for different outcomes. When the outcome was told in advance, the occurrence probability tended to be estimated high owing to a hindsight bias. The hindsight seems to lead to biased evaluations of decision quality. Although this bias is similar to an outcome bias, the outcome in a hindsight bias was not evaluated by controlling the intent (selfish, fair, and generous) and the outcome (selfish, fair, and generous) like an outcome bias. The outcome in a hindsight bias was evaluated by using occurrence probabilities. A hindsight bias might be included in an outcome bias, and distort our evaluation of decision (punishment/reward) together with the outcome bias that ignores the intention of allocators and places emphasis on the outcome.

Zakay [23] showed that managers counted good or excellent outcomes as one of the criteria for evaluating decisions made by other managers due to an outcome bias. Mitchell and Kalb [24] showed how the knowledge on the outcome affected both responsibility for the outcome and the probability of the outcome. The participants were required to read descriptions of poor performances by nurses that resulted in poor achievements. The knowledge on the outcome was found to affect both responsibility for the outcome and the probability of the outcome. It seems certain that outcome biases are ubiquitous in our daily activities.

Cushman et al. [22], using only the condition where the dice were chosen intentionally by the participant, empirically demonstrated the outcome bias on the basis of the finding that Player 2 tended to provide Player 1 with more punishments when the outcome was selfish one irrespective of the intention of Player 1 (intentional choice of Die A-C). To demonstrate the outcome bias more systematically, the data under this condition must be compared with the data when the die was not intentionally chosen by Player 1. The comparison between the two conditions (intentional and no-intentional conditions) above definitely shows that Player 2 punished Player 1 to a larger extent when the outcome was disadvantageous for Player 2 (selfish outcome) irrespective of whether the die was chosen out of three types intentionally or not (see Figure 2 and Figure 3). In this manner, the outcome strongly affected the decision of punishment/reward by Player 2 irrespective of the intention of Player 1 when
the outcome was selfish or generous. Player 2 seems to place more emphasis on the outcome than on the intention of the allocator (Player 1) when the outcome is selfish or generous.

Cushman et al. [22] found that Player 2 rewarded fair outcomes irrespective of the intention of Player 1. This corresponds to an outcome bias. On the other hand, it tended that Player 2 in this study punished the selfish intention for the fair outcome (see Figure 2). This shows that Player 2 (the participant) in this study did not behave according to the outcome. Such a difference might be explained by the inference that the participants (Player 2) in this study were aware of Player 1’s intention (choice of a selfish die) more strongly than those in Cushman et al. [22]. It is speculated that incorporating the intentional condition of Player 1 into the experimental factors made the awareness of Player 1’s intention by Player 2 (the participant) stronger. This must affect the behavior of the participant (Player 2) for the fair outcome, and eventually lead to the punishment of selfish die for the fair outcome. The different behavior of Player 2 between the intentional and the non-intentional conditions for the fair outcome should be further explored to confirm the validity of this study.

4.2. Implication for Safety Management

It is possible that the outcome bias potentially becomes a root cause of such an unintended disaster as the critical mass disaster at JCO Tokai Works Test Facility uranium processing plant [18]. Figure 4 illustrates a model that shows how an outcome bias induces a hindsight bias and an optimistic bias and leads to such a disaster. As suggested by Baron and Hershey [18], both outcome and judged probability of an outcome affect an evaluation of decision. As a result of an illegal alteration of the manufacturing process of uranium solution, it happened that no abnormality occurred. The outcome must affect the judged probability of outcome. The probability of leading to an abnormal state must be underestimated on condition that the illegal alteration induced no abnormality. To put it another way, the probability of maintaining a normal state by an illegal alteration must be overestimated. The information on the outcome must bias the evaluation of illegal decision together with the judged probability of outcome (hindsight). Thus, the illegal alteration must be optimistically evaluated. Consequently, illegal and minor alterations of the manufacturing process were readily repeated six times until the critical mass disaster occurred at JCO Tokai Works test facility uranium processing plant in Japan. As suggested in this study, the outcome bias does not readily occur when the outcome was neutral (fair) as compared with the extreme (selfish or generous) outcome. This might indicate that the neutral outcome or the neutral viewing of the outcome is potentially effective for avoiding incidents, collisions, crashes, or disasters caused by an outcome bias.

An outcome bias gives us a lesson that we tend to take a serious view of not the process or the intention of actors but the outcome itself. Therefore, we should not get trapped in an outcome bias especially when we engage in the safety management activities, and must take a serious view of not the outcome but the process or the intention of actors. More concretely, we should be very careful so that we should have more priority on the process of actions or the intention of actors so that we should not be keen on the outcome itself. Future research should also explore how a hindsight bias and an outcome bias are linked with each other.

![Figure 4](image-url)
Although the importance of considering the judgmental heuristics and biases is pointed out in the analysis of human errors (for example, Reason [1]), it has not been demonstrated systematically that such biases are related to the distorted decision making which becomes a trigger of crashes, collisions, or disasters. In the framework of traditional analysis of human errors, crashes, collisions, or disasters, it has not been focused on how the distortion of decision making due to cognitive biases such as an outcome bias is related to such unfavorable events. Cognitive biases cause the distortion of decision making, lead to human errors in judgment, and presumably trigger incidents, crashes, collisions, or disasters if the commitment to the biased judgment, decision making, and behavior is escalated [25].

Cognitive biases hinder the progress of safety management and technology. A promising addressing to such an issue might lie in the consideration and the steady elimination of cognitive biases that is unexpectedly and unconsciously interfering with large-scale man-machine systems. Introducing appropriate safety interventions that will ensure that cognitive biases such as an outcome bias do not eventually manifest themselves as causal factors of incidents, crashes, collisions, or disasters would enable us to address a cognitive-bias-related safety issue appropriately.

5. Conclusions

The evidence of an outcome bias was explored in a two-player economic game experiment where the reward allocation was made between two players by taking into account the intention of the allocator (Player 1) when choosing a type of a die. The outcome bias occurred not only when the outcome was selfish but also when it was generous. Moreover, the intention of the allocator did not affect the decision by the evaluator (Player 2). Different from the selfish and the generous outcome, the evaluation of Player 2 was less affected by the outcome of the game when the outcome was fair. Rather, the evaluation of Player 2 tended to be affected by the intention of Player 1.

It can be concluded that the outcome of the game more strongly affected the punishment/reward reaction (decision) by the evaluator (Player 2) than the intention of the allocator (Player1) when the outcome was extreme (selfish or generous). In this manner, we could verify the hypothesis that we are readily get trapped in the outcome bias when the outcome is selfish or generous.

At present, it is not possible to show quantitatively how cognitive biases hinder objectively acquiring a lesson from crashes, collisions, disasters, or incidents. Future research should explore and clarify this process in more detail and get further insights into the implications for safety management. As this study also did not present an experimental evidence on how outcome and hindsight biases were related and led to an optimistic bias as shown in a qualitative model in Figure 4, we should systematically investigate how the three cognitive biases are interrelated each other. Moreover, a check sheet of cognitive biases within the system or the procedure must be developed, and it should be examined whether this is effective as safety interventions that ensure that cognitive biases do not eventually manifest themselves as causal factors of incidents, crashes, collisions, or disasters.

Acknowledgements

This work was partly supported by Grant-in Aid for Scientific Research (B) (26282095), Japan Society for the Promotion of Science (JSPS).

References


