**Decisions from Valuations of Unknown Payoff Distributions**

Ido Erev,

Faculty of Industrial Engineering & Management, Technion, erevido@gmail.com

Yefim Roth,

Department of Human Services, University of Haifa, rothefim@gmail.com

Doron Sonsino,

Department of Economics, Ben-Gurion University of the Negev, sonsino.doron@gmail.com

August 5, 2019

Previous studies pertaining to the effect of experience on advice-taking highlight the advantages of accurate probability estimations over direct recommendations. This investigation compares two interpretations of this finding with regard to decisions based on point estimates (valuations) of unknown distributions. It can be said that both interpretations work on the assumption that the inclination to trust reliable experts increases with experience, nevertheless it varies based on the assumed subjective measurement of reliability. As per the Mean Squared Error hypothesis, the decision makers focus on the accuracy of the predictions and work with the assumption that a proper scoring rule is being used: They trust the expert with the lower mean squared error between the prediction and the final outcomes. In line with the error-rate hypothesis, the decision makers focus on the final outcomes and trust the expert that ensures the best choice in most cases. Experimental studies favor the error-rate hypothesis and this hypothesis best fits the observed decisions based on 200 trials. Furthermore, the results indicate that when choosing between two independent payoff distributions, there is a greater likelihood for the decision makers to favor the experts that report the median, as compared to the experts that report the expected value. These findings indicate a strong underweighting of rare events in decisions based on valuations and also indicate that designers of expert systems are capable of increasing the system’s popularity by biasing the valuation toward the option that leads to the best outcome in most cases, despite the fact that it might diminish the chances of the users’ expected return.

*Key words: Decisions from experience, Black Swan, Reliance on small samples, cry wolf effect*

**1. Introduction**

New big data technologies have been successful in increasing the frequency, and the importance, of decisions based on predictions of the expected outcomes of different alternatives. For instance, while selecting the best possible route, drivers can rely on the estimated driving time provided by “exert systems” such as Google Maps and Waze. Similarly, while selecting the best possible investment options, investors can rely on earnings per share forward estimates provided by different analysts (e.g., <https://ycharts.com/indicators/sandp_500_earnings_per_share_forward_estimate>).

The decision makers in such scenarios can make use of distinct expert systems and learn how to use them based on the feedback obtained in repeated experiences.

While a majority of extant studies pertaining to the use of experts’ predictions and advice deal with the decisions that lack the influence of experience (e.g., Benjamin & Budescu, 2015; Bonaccio & Dalal, 2006; Budescu, Por, & Broomell, 2012; Cabantous, Hilton, Kunreuther, Michel-Kerjan, 2011; Erev & Cohen, 1990), recent research (Bolton & Katok, 2018) demonstrates the importance of feedback in the current context.[[1]](#footnote-1) In a particular scenario studied by Bolton and Katok (2018), the participants were asked to repeatedly choose between “take a cost of 25” and “take a risk of losing 100 with probability p.” The value of p varied from trial to trial and the participants could use two types of advice: accurate estimation of the value of p and direct recommendation (“take the risk” when the risk maximized the expected value (EV) and “take the loss” when the loss maximized EV). The results reveal that in the very first trial (prior to the feedback), the EV maximization rate was higher in response to the direct recommendation rather than in response to the accurate estimation; however, the availability of feedback reversed this pattern. Bolton and Katok (2018) show that this tendency can be described as a reflection of a “cry wolf effect” (Meyer, 2001): Experiencing cases in which the direct recommendations are wrong decreases the tendency to follow these recommendations.

It is to be noted that the study conducted by Bolton and Katok (2018) focuses on a simplified situation, wherein the decision makers are aware of all the possible results and the shape of the payoff distributions. This research extends their analysis by considering situations in which the advice, similar to the output of the expert systems described above, are point estimates (or “valuations”) of realizations from unknown distributions. This analysis starts with the initial observation that inadequate information pertaining to possible outcomes has the ability to influence the perceived capabilities and efficiency of the expert. Furthermore, different abstractions of Bolton and Katok’s cry wolf hypothesis lead to contradicting predictions in the extended settings. To elucidate further, assuming that the cry wolf effect occurs when the decision makers find the expert unreliable, and reliability is measured either by the means squared error (MSE) between the predictions and the realizations or by the error-rate (the proportion of instances in which the decision maker learns that the prediction/advice was wrong). It is to be noted that when decision makers who are aware of the shape of the payoff distributions evaluate accurate probability estimates and direct recommendations, as in Bolton and Katok’s study, both criteria (MSE and error-rate) favor the accurate estimate that minimizes the MSE, and only the direct recommendation can be wrong. In contrast, when the decision makers have limited prior knowledge of the possible outcomes as well as the shape of the payoff distribution, the two criteria can favor two different expert systems. For instance, consider the choice problem described in Table 1 (“17 with certainty” or “24 with probability .75; -4 otherwise”). The valuation of the risky prospect that minimizes the MSE is the EV (17), nevertheless this valuation is always wrong, while an expert who reports the median (24) is deemed correct in 75% of the cases. With the assumption that decision makers favor an expert with a lower error-rate, it can be said that they might prefer the experts that report the median.

**Table 1. Numerical Example of the difference between MSE and the error-rate as measures of error**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Choice task | | | Expert | |
| Option | | Payoff distributions | EV | | Median |
| R | | 24 with p = 0.75  -4 otherwise | 17 | | 24 |
| S | | 22 with certainty | 22 | | 22 |
| **Measures of Error:** | | | | | |
| Mean squared Error (MSE) | | | .**5[.75(72)+.25(212)]=73.5** | | .5[.25(282)]=98 |
| Error-rate | | | 1 | | **.25** |

*Notes.* The decision maker can rely on predictions provided by two experts. Expert EV minimizes the expected MSE, but Expert Median is less likely to be wrong (it minimizes error rate).

The present investigation analyses the impact of experience in situations in which the MSE and error-rate scoring rules favor different experts. Study 1 focuses on choices between independent distributions. As presented in Table 1, the MSE favors an expert that reports the EV, while the error-rate scoring rule favors an expert that reports the median of the relevant distributions. Figure 1 presents the distribution dependency introduced by Study 2 that affects the ranking of error rates, and both studies use the "experts’ paradigm" presented in the same figure. In each trial, the participants are asked to choose between two prospects based on their valuation provided by two "experts". One expert reports the EVs of the prospects, while the second expert reports the median. Even though the participants are given instant feedback after each trial, they have no prior information about the valuation rule.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Figure 1. The main Screens in a study that uses the “Decisions from Valuations” Paradigm** | | | | | | | | | | | |
| Instructions: In each trial of this study you will receive recommendations from two experts, and will be asked to choose between two options. Your goal is to maximize your earnings. Your final payoff will be the show-up fee (20 Shekels) plus the payoff from your selected option in one randomly selected trial (the conversion rate is 10 Points = 1 Shekel). | | | | | | | | | | | |
| Pre-choice screen (for the prospects of Table 1): | | | | | | | | | | | |
|  |  | Valuation of Expert A | | Valuation of Expert B | |  | Valuation of Expert A | | Valuation of Expert B | |  |
|  |  |  |  |
|  |  |  | 22.0 | 22.0 |  |  |  | 17.0 | 24.0 |  |  |
|  |  |  | Option Left | |  |  |  | Option Right | |  |  |
|  |  |  |  | |  | **Continue** |  |  | |  |  |
| Post-choice screen (assuming Option Right is selected): | | | | | | | | | | | |
|  |  | Valuation of Expert A | | Valuation of Expert B | |  | Valuation of Expert A | | Valuation of Expert B | |  |
|  |  |  |  |
|  |  |  | 22.0 | 22.0 |  |  |  | 17.0 | 24.0 |  |  |
|  |  |  | 22.0 | |  |  |  | **24.0** | |  |  |
|  |  |  | **You chose Right and earned 24 points** | | | | | | |  |  |
|  |  |  |  | | | **Continue** |  | | |  |  |

The experts’ paradigm can be described as a generalization of the clicking paradigm used in basic studies pertaining to decisions from experience (Erev & Haruvy, 2016). As per the basic clicking paradigm, the decision makers have no information concerning the payoff distributions and rely on their experience. An important implication of the current generalization is that it affects the strategies the decision makers are likely to consider. In addition to considering “Left” and “Right” (the strategies assumed in previous models of behavior in the clicking paradigm), the decision makers are likely to consider strategies such as “Select the option with the higher prediction by Expert A” and “Select the option with the higher average prediction.” This particular feature of the current paradigm and its implications have been analyzed in the subsequent sections.

**2. Hypotheses: MSE and Error-Rate as behavioral strategies**

The assumption that decision makers are likely to behave as if they are using the MSE scoring rule to evaluate the experts can be justified on two solid grounds. First is the observation that the MSE is a proper scoring rule, which motivates the evaluated experts to report the EV of the distributions (Brier, 1951). Consequently, it can be said that there is a high probability of the participants choosing MSE or any similar rule if they have the tendency to employ strategies that have yielded results in similar out-of-lab experiences in the past. The second justification for the use of the MSE scoring rule is that, in the current setting, trusting the expert with the lower MSE maximizes the decision makers’ expected return. Thus, effective learning will increase the use of the behavior consistent with this rule even if the participants have not learned to use this rule outside the lab.

The hypothesis that decision makers are likely to behave as if they are using the error-rate scoring rule can be justified by assuming a wide definition of the term “error.” The current analysis assumes that the decision makers consider “valuations that imply a choice of the option that led to the worst payoff in the last trial” as errors. In which case, sensitivity to the error-rate is calculated under the assumption that learning among “reaction to experts” strategies is similar to learning in basic decisions from experience tasks. Extant studies show that decisions from experience tend to rely on a small sample of past experiences, and this behavior implies a tendency to select the option that minimizes the error-rate (Hertwig & Erev, 2009, Roth, Wänke & Erev, 2016)[[2]](#footnote-2). For example, in Table 1, reliance on a sample of size one implies a selection of the low-EV gamble in 75% of the trials.

The experiments presented in the next section explore these assumptions. Study 1 shows that there is a bias to pay more attention to the expert that reports the median, as compared to the expert that reports EV. This deviation from maximization favors the hypothesis that decision makers tend to favor the expert that minimizes the error-rate over the expert with lower MSE score. Study 2 further clarifies these results by comparing the condition in which the error-rate scoring rule favors the expert that reports the median, with a condition in which the error-rate rule favors the expert that reports the EV. The results suggest the predisposition to focus on the average of the two valuations and learning to rely on the expert with the lower error-rate. Furthermore, the comparison between the two studies suggests that, in cases of a disagreement between the experts, several decision makers ignore the experts entirely. In the concluding section, the paper presents a model of sufficient conditions for the observed results and a discussion of their implications.

**3. Experimental studies**

To further clarify the relationship between "decisions from valuations" and mainstream decision research, we chose to start the current investigation by focusing in Study 1 on the problems examined in the 2015 Choice Prediction Competition (CPC15, Erev et al., 2017). This set of problems yields some of the best-known properties of both decisions under risk (decisions from description of the payoff distributions) and decisions from experience. Study 2 extends the analysis by comparing alternative interpretations of the results of Study 1.

**3.1 Study 1**

The first study analyses the manner in which people make choices based on the valuations provided by two experts and by focusing on the 120 problems presented in Table 2. As explained in the preceding section, the valuations of one expert were the EVs and the valuations of the second were the medians of the relevant distributions. The selected problems include all the choice problems studied in CPC15, except problems with ambiguity and degenerate problems in which the EV equals the median in both options. All choice problems were presented using the experts’ paradigm as illustrated in Figure 1.

***Participants***. Forty Technion students volunteered to participate in exchange for a monetary payoff, which included a show-up fee of 20 shekels and the obtained payoff in one randomly selected trial. The conversion rate was 10 points = 1 shekel (about 0.3$). The mean final payoff was about $6.

***Procedure.*** Each participant was given all 120 problems in random order. The pre-choice information was limited to two experts’ valuations of each option (see the upper part of Figure 1). The participants could neither observe the payoff distributions, nor were they informed about the experts using the EV and median rules. The names of the two experts (A or B) were determined randomly for each subject before trial 1. Each choice was followed by feedback concerning the realization of the two options as shown in Figure 1. Hence, the participants could learn to use the valuations of the two experts only from their own experience.

**3.1.1 Results**

In figure 2, the left-hand side presents the choice rates of the option with a higher valuation according to each expert and is presented in 12 blocks of 10 trials (the right side of Figure 2 presents the predictions of the post-hoc model discussed below)[[3]](#footnote-3). Table 2 presents the choice rates by problem. The mean EV-rate (the choice rate of the option with higher valuation according to Expert EV) is 0.60 (STD = 0.10), and the mean Median-rate (the choice rate of the option with higher valuations according to Expert Median) is 0.70 (STD = 0.16) and the difference is significant (t(78) = -3.31, p<0.002). Hence, it can be said that the participants paid more attention to the Expert Median. Both rates are higher than 0.5 because the two experts favor the same option in more than half (70) of the 120 problems and the mean choice rate of the options that were favored by both experts is 0.73. Thus, in more than a quarter (27%) of the trials the decision makers behaved as if they did not trust the experts and selected the “valuations dominated” option (the option with lower prediction by both experts). Analyzing the cases when the two experts favored different options reveals that participants follow Expert Median in 0.63 (STD = 0.14) while Expert Mean in only 0.37 (STD = 0.14) of the cases.

**Table 2. Study 1: Problems and Choice Rates**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Choice rates of Option 2 | | | | | | | | | | | | | | | | |
| Problem number | | Option 1 | | | Option 2 | | | Expected values | | Medians | | | CPC15 | | Here |
| Here | CPC15 | H1 | pH1 | L1 | H2 | pH2 | L2 | EV1 | EV2 | Med1 | Med2 | | fDesc | fExp | fVal |
| 1 | 1 | 3 | 1.00 | . | 4 | 0.80 | 0 | 3.00 | 3.20 | 3 | 4.0 | | 0.42 | 0.65 | 0.65 |
| 2 | 2 | 3 | 0.25 | 0 | 4 | 0.20 | 0 | 0.75 | 0.80 | 0 | 0.0 | | 0.61 | 0.62 | 0.43 |
| 3 | 5 | -3 | 1.00 | . | 0 | 0.20 | -4 | -3.00 | -3.20 | -3 | -4.0 | | 0.49 | 0.36 | 0.28 |
| 4 | 6 | 0 | 0.75 | -3 | 0 | 0.80 | -4 | -0.75 | -0.80 | 0 | 0.0 | | 0.38 | 0.41 | 0.43 |
| 5 | 7 | -1 | 1.00 | . | 0 | 0.95 | -20 | -1.00 | -1.00 | -1 | 0.0 | | 0.48 | 0.64 | 0.70 |
| 6 | 8 | 1 | 1.00 | . | 20 | 0.05 | 0 | 1.00 | 1.00 | 1 | 0.0 | | 0.39 | 0.29 | 0.23 |
| 7 | 9 | 1 | 1.00 | . | 100 | 0.01 | 0 | 1.00 | 1.00 | 1 | 0.0 | | 0.47 | 0.39 | 0.20 |
| 8 | 10 | 2 | 1.00 | . | 101 | 0.01 | 1 | 2.00 | 2.00 | 2 | 1.0 | | 0.55 | 0.42 | 0.28 |
| 9 | 11 | 19 | 1.00 | . | 20 | 0.90 | -20 | 19.00 | 16.00 | 19 | 20.0 | | 0.13 | 0.21 | 0.40 |
| 10 | 15 | 7 | 1.00 | . | 50 | 0.50 | 1 | 7.00 | 25.50 | 7 | 25.5 | | 0.78 | 0.85 | 0.70 |
| 11 | 16 | 7 | 1.00 | . | 50 | 0.50 | -1 | 7.00 | 24.50 | 7 | 24.5 | | 0.71 | 0.83 | 0.73 |
| 12 | 17 | 30 | 1.00 | . | 50 | 0.50 | 1 | 30.00 | 25.50 | 30 | 25.5 | | 0.24 | 0.29 | 0.23 |
| 13 | 18 | 30 | 1.00 | . | 50 | 0.50 | -1 | 30.00 | 24.50 | 30 | 24.5 | | 0.23 | 0.33 | 0.28 |
| 14 | 19 | 9 | 1.00 | . | 9+a8 | 1.00 | . | 9.00 | 9.00 | 9 | 3.0 | | 0.36 | 0.30 | 0.25 |
| 15 | 20 | 9 | 1.00 | . | 9+a8 | 1.00 | . | 9.00 | 9.00 | 9 | 3.0 | | 0.38 | 0.36 | 0.23 |
| 16 | 26 | 16 | 1.00 | . | 50 | 0.40 | 1 | 16.00 | 20.60 | 16 | 1.0 | | 0.50 | 0.55 | 0.28 |
| 17 | 27 | 16 | 1.00 | . | 48-a3 | 0.40 | 1 | 16.00 | 19.80 | 16 | 1.0 | | 0.50 | 0.57 | 0.50 |
| 18 | 28 | 6 | 0.50 | 0 | 9 | 0.50 | 0 | 3.00 | 4.50 | 3 | 4.5 | | 0.91 | 0.84 | 0.73 |
| 19 | 29 | 2 | 1.00 | . | 3 | 1.00 | . | 2.00 | 3.00 | 2 | 3.0 | | 0.97 | 1.00 | 0.70 |
| 20 | 30 | 6 | 0.50 | 0 | 8 | 0.50 | 0 | 3.00 | 4.00 | 3 | 4.0 | | 0.94 | 0.98 | 0.78 |
| 21 | 32 | 24 | 0.75 | -4 | 82 | 0.25 | 3 | 17.00 | 22.75 | 24 | 3.0 | | 0.68 | 0.69 | 0.33 |
| 22 | 33 | -3 | 1.00 | . | 14 | 0.40 | -22 | -3.00 | -7.60 | -3 | -22.0 | | 0.33 | 0.22 | 0.23 |
| 23 | 34 | 7 | 1.00 | . | 27+s3 | 0.10 | 4 | 7.00 | 6.30 | 7 | 4.0 | | 0.39 | 0.40 | 0.20 |
| 24 | 35 | -5 | 1.00 | . | 47 | 0.01 | -15 | -5.00 | -14.38 | -5 | -15.0 | | 0.18 | 0.05 | 0.25 |
| 25 | 36 | 28 | 1.00 | . | 88+a3 | 0.60 | -46 | 28.00 | 34.40 | 28 | 86.0 | | 0.39 | 0.57 | 0.65 |
| 26 | 37 | 23 | 0.90 | 0 | 64 | 0.40 | -7 | 20.70 | 21.40 | 23 | -7.0 | | 0.38 | 0.37 | 0.20 |
| 27 | 38 | 24 | 1.00 | . | 34 | 0.05 | 28 | 24.00 | 28.30 | 24 | 28.0 | | 0.91 | 1.00 | 0.68 |
| 28 | 39 | 29 | 1.00 | . | 33+s5 | 0.80 | 6 | 29.00 | 27.60 | 29 | 33.0 | | 0.50 | 0.66 | 0.43 |
| 29 | 40 | 3 | 0.80 | -37 | 79 | 0.40 | -46 | -5.00 | 4.00 | 3 | -46.0 | | 0.49 | 0.56 | 0.30 |
| 30 | 41 | 29 | 1.00 | . | 44+s5 | 0.40 | 21 | 29.00 | 30.20 | 29 | 21.0 | | 0.68 | 0.68 | 0.38 |
| 31 | 43 | 14 | 1.00 | . | 12 | 0.90 | 9 | 14.00 | 11.70 | 14 | 12.0 | | 0.13 | 0.00 | 0.28 |
| 32 | 44 | 23 | 1.00 | . | 24 | 0.99 | -33 | 23.00 | 23.43 | 23 | 24.0 | | 0.27 | 0.47 | 0.63 |
| 33 | 46 | 37 | 0.01 | 9 | 30 | 0.60 | -37 | 9.28 | 3.20 | 9 | 30.0 | | 0.20 | 0.30 | 0.58 |
| 34 | 47 | 11 | 1.00 | . | 57-a6 | 0.20 | -5 | 11.00 | 7.40 | 11 | -5.0 | | 0.21 | 0.15 | 0.28 |
| 35 | 51 | 42 | 0.80 | -18 | 68 | 0.20 | 23 | 30.00 | 32.00 | 42 | 23.0 | | 0.79 | 0.70 | 0.35 |
| 36 | 52 | 46 | 0.20 | 0 | 46 | 0.25 | -2 | 9.20 | 10.00 | 0 | -2.00 | | 0.36 | 0.22 | 0.35 |
| 37 | 53 | 28 | 1.00 | . | 42 | 0.75 | -22 | 28.00 | 26.00 | 28 | 42.0 | | 0.36 | 0.41 | 0.60 |
| 38 | 54 | 18 | 1.00 | . | 64 | 0.50 | -33 | 18.00 | 15.50 | 18 | 15.5 | | 0.32 | 0.29 | 0.40 |
| 39 | 55 | 43 | 0.20 | 19 | 22+s9 | 0.25 | 17 | 23.80 | 18.25 | 19 | 17.0 | | 0.20 | 0.07 | 0.35 |
| 40 | 56 | -8 | 1.00 | . | -5 | 0.99 | -34 | -8.00 | -5.29 | -8 | -5.00 | | 0.76 | 0.89 | 0.60 |
| 41 | 57 | 49 | 0.50 | -3 | 33+s9 | 0.95 | 17 | 23.00 | 32.20 | 23 | 33.0 | | 0.77 | 0.75 | 0.78 |
| 42 | 58 | 85 | 0.40 | -7 | 40 | 0.25 | 24 | 29.80 | 28.00 | -7 | 24.0 | | 0.60 | 0.56 | 0.73 |
| 43 | 59 | 17 | 0.25 | 16 | 43 | 0.40 | 2 | 16.25 | 18.40 | 16 | 2.00 | | 0.51 | 0.49 | 0.30 |
| 44 | 60 | 51 | 0.10 | 21 | 38 | 0.60 | 1 | 24.00 | 23.20 | 21 | 38.0 | | 0.37 | 0.30 | 0.70 |
| 45 | 61 | 26 | 0.25 | 25 | 29+a7 | 0.05 | 24 | 25.25 | 24.25 | 25 | 24.0 | | 0.67 | 0.56 | 0.23 |
| 46 | 62 | 25 | 1.00 | . | 45 | 0.20 | 17 | 25.00 | 22.60 | 25 | 17.0 | | 0.32 | 0.34 | 0.30 |
| 47 | 63 | 17 | 1.00 | . | 60+s5 | 0.10 | 15 | 17.00 | 19.50 | 17 | 15.0 | | 0.68 | 0.69 | 0.55 |
| 48 | 65 | 12 | 0.40 | -16 | -5 | 1.00 | . | -4.80 | -5.00 | -16 | -5.0 | | 0.33 | 0.45 | 0.75 |
| 49 | 66 | 45 | 0.60 | 2 | 54-a5 | 0.10 | 20 | 27.80 | 23.40 | 45 | 20.0 | | 0.43 | 0.43 | 0.40 |
| 50 | 67 | 85 | 0.25 | 4 | 54 | 0.25 | 11 | 24.25 | 21.75 | 4 | 11.0 | | 0.45 | 0.43 | 0.50 |
| 51 | 68 | 12 | 1.00 | . | 102 | 0.20 | -14 | 12.00 | 9.20 | 12 | -14.0 | | 0.39 | 0.31 | 0.25 |
| 52 | 70 | 18 | 1.00 | . | 35 | 0.75 | -19 | 18.00 | 21.50 | 18 | 35.0 | | 0.38 | 0.58 | 0.63 |
| 53 | 71 | 13 | 0.60 | -20 | 76 | 0.20 | -26 | -0.20 | -5.60 | 13 | -26.0 | | 0.38 | 0.28 | 0.18 |
| 54 | 72 | -9 | 1.00 | . | 13 | 0.25 | -8 | -9.00 | -2.75 | -9 | -8.0 | | 0.82 | 1.00 | 0.65 |
| 55 | 73 | 2 | 1.00 | . | 51+s7 | 0.05 | 0 | 2.00 | 2.55 | 2 | 0.00 | | 0.37 | 0.41 | 0.40 |
| 56 | 75 | 13 | 1.00 | . | 50 | 0.60 | -45 | 13.00 | 12.00 | 13 | 50.0 | | 0.35 | 0.50 | 0.75 |
| 57 | 77 | 1 | 1.00 | . | 38 | 0.40 | -9 | 1.00 | 9.80 | 1 | -9.0 | | 0.65 | 0.63 | 0.40 |
| 58 | 78 | 19 | 1.00 | . | 44 | 0.05 | 9 | 19.00 | 10.75 | 19 | 9.00 | | 0.11 | 0.12 | 0.23 |
| 59 | 79 | 32 | 0.01 | 19 | 65 | 0.01 | 9 | 19.13 | 9.56 | 19 | 9.00 | | 0.14 | 0.02 | 0.20 |
| 60 | 80 | 3 | 1.00 | . | 50 | 0.40 | -36 | 3.00 | -1.60 | 3 | -36.0 | | 0.47 | 0.43 | 0.28 |
| 61 | 83 | 9 | 1.00 | . | 64 | 0.01 | 9 | 9.00 | 9.55 | 9 | 9.00 | | 0.87 | 0.99 | 0.58 |
| 62 | 84 | 27 | 1.00 | . | 22 | 0.99 | -7 | 27.00 | 21.71 | 27 | 22.0 | | 0.08 | 0.00 | 0.23 |
| 63 | 85 | 20 | 1.00 | . | 70 | 0.25 | 6 | 20.00 | 22.00 | 20 | 6.00 | | 0.43 | 0.44 | 0.15 |
| 64 | 87 | -2 | 1.00 | . | 4+s7 | 0.99 | -34 | -2.00 | 3.62 | -2 | 4.00 | | 0.81 | 0.98 | 0.78 |
| 65 | 89 | 17 | 1.00 | . | 44 | 0.10 | 17 | 17.00 | 19.70 | 17 | 17.0 | | 0.88 | 1.00 | 0.70 |
| 66 | 90 | 10 | 1.00 | . | 31 | 0.75 | -49 | 10.00 | 11.00 | 10 | 31.0 | | 0.42 | 0.55 | 0.68 |
| 67 | 91 | 7 | 1.00 | . | 16 | 0.10 | 10 | 7.00 | 10.60 | 7 | 10.0 | | 0.92 | 1.00 | 0.70 |
| 68 | 92 | 8 | 0.80 | -37 | 102 | 0.20 | -29 | -1.00 | -2.80 | 8 | -29.0 | | 0.39 | 0.32 | 0.25 |
| 69 | 94 | 7 | 1.00 | . | 6 | 0.75 | 1 | 7.00 | 4.75 | 7 | 6.00 | | 0.10 | 0.02 | 0.25 |
| 70 | 95 | -3 | 0.05 | -9 | 42-a7 | 0.40 | -24 | -8.70 | 2.40 | -9 | -24.0 | | 0.72 | 0.57 | 0.43 |
| 71 | 96 | 35 | 0.50 | -47 | -10 | 0.75 | -15 | -6.00 | -11.25 | -6 | -10.0 | | 0.30 | 0.31 | 0.23 |
| 72 | 97 | 10 | 1.00 | . | 45 | 0.20 | -5 | 10.00 | 5.00 | 10 | -5.0 | | 0.22 | 0.21 | 0.20 |
| 73 | 98 | 94 | 0.50 | -40 | 36+s7 | 0.75 | -21 | 27.00 | 21.75 | 27 | 35.0 | | 0.51 | 0.45 | 0.60 |
| 74 | 99 | 22 | 1.00 | . | 44+s5 | 0.40 | 15 | 22.00 | 26.60 | 22 | 15.0 | | 0.65 | 0.71 | 0.45 |
| 75 | 100 | 18 | 0.60 | -29 | -1 | 1.00 | . | -0.80 | -1.00 | 18 | -1.0 | | 0.54 | 0.53 | 0.33 |
| 76 | 101 | 28 | 1.00 | . | 73+s3 | 0.05 | 27 | 28.00 | 29.30 | 28 | 27.0 | | 0.83 | 0.76 | 0.40 |
| 77 | 103 | 27 | 0.80 | -4 | 77+a6 | 0.10 | 22 | 20.80 | 27.50 | 27 | 22.0 | | 0.83 | 0.77 | 0.53 |
| 78 | 104 | -6 | 1.00 | . | 3 | 0.99 | -27 | -6.00 | 2.70 | -6 | 3.0 | | 0.85 | 0.98 | 0.78 |
| 79 | 105 | 30 | 1.00 | . | 90 | 0.01 | 36 | 30.00 | 36.54 | 30 | 36.0 | | 0.90 | 1.00 | 0.68 |
| 80 | 106 | 2 | 1.00 | . | 34+s5 | 0.05 | -5 | 2.00 | -3.05 | 2 | -5.0 | | 0.20 | 0.12 | 0.23 |
| 81 | 107 | 25 | 1.00 | . | 65+s5 | 0.25 | 9 | 25.00 | 23.00 | 25 | 9.0 | | 0.39 | 0.32 | 0.28 |
| 82 | 108 | 16 | 1.00 | . | 91 | 0.20 | -11 | 16.00 | 9.40 | 16 | -11.0 | | 0.27 | 0.18 | 0.28 |
| 83 | 109 | 11 | 1.00 | . | 26 | 0.50 | -9 | 11.00 | 8.50 | 11 | 8.5 | | 0.33 | 0.40 | 0.25 |
| 84 | 110 | 12 | 1.00 | . | 29-a2 | 0.80 | -35 | 12.00 | 16.20 | 12 | 28.0 | | 0.56 | 0.77 | 0.60 |
| 85 | 111 | 28 | 1.00 | . | 47 | 0.60 | -13 | 28.00 | 23.00 | 28 | 47.0 | | 0.25 | 0.39 | 0.50 |
| 86 | 112 | -7 | 1.00 | . | 28+s7 | 0.20 | -18 | -7.00 | -8.80 | -7 | -18.0 | | 0.51 | 0.26 | 0.20 |
| 87 | 113 | 9 | 0.95 | 0 | 37+a6 | 0.25 | -3 | 8.55 | 7.00 | 9 | -3.0 | | 0.35 | 0.37 | 0.20 |
| 88 | 114 | 72 | 0.01 | -2 | 112 | 0.25 | -33 | -1.26 | 3.25 | -2 | -33.0 | | 0.44 | 0.32 | 0.38 |
| 89 | 115 | 50 | 0.40 | 5 | 20+s7 | 0.80 | -17 | 23.00 | 12.60 | 5 | 20.0 | | 0.17 | 0.20 | 0.60 |
| 90 | 116 | 2 | 1.00 | . | 45+s5 | 0.05 | 3 | 2.00 | 5.10 | 2 | 3.0 | | 0.95 | 1.00 | 0.55 |
| 91 | 117 | -6 | 1.00 | . | 7 | 0.50 | -30 | -6.00 | -11.50 | -6 | -11.5 | | 0.33 | 0.34 | 0.18 |
| 92 | 118 | 26 | 1.00 | . | 46-a6 | 0.50 | 10 | 26.00 | 28.00 | 26 | 10.0 | | 0.47 | 0.57 | 0.23 |
| 93 | 119 | 19 | 0.40 | 12 | 100+a2 | 0.25 | -12 | 14.80 | 16.00 | 12 | -12.0 | | 0.33 | 0.35 | 0.15 |
| 94 | 120 | -9 | 0.95 | -26 | -1 | 0.10 | -11 | -9.85 | -10.00 | -9 | -11.0 | | 0.57 | 0.42 | 0.23 |
| 95 | 121 | -8 | 1.00 | . | 21+s3 | 0.01 | 0 | -8.00 | 0.21 | -8 | 0.0 | | 0.79 | 0.99 | 0.73 |
| 96 | 122 | 68 | 0.05 | -14 | -11 | 0.90 | -36 | -9.90 | -13.50 | -14 | -11.0 | | 0.36 | 0.40 | 0.48 |
| 97 | 123 | 28 | 0.75 | -13 | 57 | 0.10 | 16 | 17.75 | 20.10 | 28 | 16.0 | | 0.74 | 0.63 | 0.23 |
| 98 | 124 | 15 | 0.95 | 7 | 42 | 0.01 | 19 | 14.60 | 19.23 | 15 | 19.0 | | 0.85 | 0.98 | 0.85 |
| 99 | 125 | 28 | 1.00 | . | 41 | 0.40 | 12 | 28.00 | 23.60 | 28 | 12.0 | | 0.29 | 0.36 | 0.23 |
| 100 | 126 | -8 | 1.00 | . | 80+s7 | 0.20 | -18 | -8.00 | 1.60 | -8 | -18.0 | | 0.53 | 0.52 | 0.53 |
| 101 | 128 | -3 | 1.00 | . | 32 | 0.40 | -16 | -3.00 | 3.20 | -3 | -16.0 | | 0.64 | 0.56 | 0.38 |
| 102 | 130 | 72 | 0.40 | -41 | 16 | 0.01 | 1 | 4.20 | 1.15 | -41 | 1.00 | | 0.61 | 0.54 | 0.78 |
| 103 | 131 | 18 | 1.00 | . | 45 | 0.01 | 11 | 18.00 | 11.34 | 18 | 11.0 | | 0.19 | 0.06 | 0.28 |
| 104 | 132 | 11 | 1.00 | . | 20+s7 | 0.99 | 4 | 11.00 | 19.84 | 11 | 20.0 | | 0.81 | 0.98 | 0.75 |
| 105 | 133 | 3 | 1.00 | . | 8+s9 | 0.99 | -17 | 3.00 | 7.75 | 3 | 8.0 | | 0.71 | 0.94 | 0.70 |
| 106 | 134 | 27 | 0.05 | 24 | 31+s3 | 0.50 | 10 | 24.15 | 20.50 | 24 | 20.0 | | 0.34 | 0.37 | 0.25 |
| 107 | 135 | 6 | 1.00 | . | 8 | 0.50 | -1 | 6.00 | 3.50 | 6 | 3.5 | | 0.25 | 0.29 | 0.23 |
| 108 | 136 | 4 | 1.00 | . | 25 | 0.01 | -5 | 4.00 | -4.70 | 4 | -5.0 | | 0.16 | 0.05 | 0.23 |
| 109 | 138 | 23 | 1.00 | . | 21 | 0.80 | 16 | 23.00 | 20.00 | 23 | 21.0 | | 0.13 | 0.02 | 0.25 |
| 110 | 139 | 14 | 1.00 | . | 35+s7 | 0.60 | -9 | 14.00 | 17.40 | 14 | 34.0 | | 0.48 | 0.65 | 0.70 |
| 111 | 140 | -2 | 1.00 | . | 9 | 0.25 | 8 | -2.00 | 8.25 | -2 | 8.0 | | 0.91 | 0.98 | 0.78 |
| 112 | 141 | 28 | 0.80 | -26 | 22 | 0.75 | 2 | 17.20 | 17.00 | 28 | 22.0 | | 0.77 | 0.62 | 0.33 |
| 113 | 142 | 23 | 1.00 | . | 29 | 0.80 | -8 | 23.00 | 21.60 | 23 | 29.0 | | 0.30 | 0.54 | 0.73 |
| 114 | 143 | 67 | 0.50 | -39 | 93 | 0.25 | -15 | 14.00 | 12.00 | 14 | -15.0 | | 0.53 | 0.63 | 0.18 |
| 115 | 144 | 16 | 0.80 | 12 | 15+s9 | 1.00 | . | 15.20 | 15.00 | 16 | 15.0 | | 0.42 | 0.42 | 0.40 |
| 116 | 145 | 17 | 0.50 | -27 | 3+s7 | 0.75 | -35 | -5.00 | -6.50 | -5 | 2.0 | | 0.42 | 0.34 | 0.60 |
| 117 | 146 | 45 | 0.20 | 3 | 75+s5 | 0.05 | 13 | 11.40 | 16.10 | 3 | 13.0 | | 0.79 | 0.84 | 0.80 |
| 118 | 147 | 29 | 1.00 | . | 36+s7 | 0.10 | 32 | 29.00 | 32.40 | 29 | 32.0 | | 0.88 | 0.97 | 0.60 |
| 119 | 149 | 12 | 1.00 | . | 31+s3 | 0.10 | 12 | 12.00 | 13.90 | 12 | 12.0 | | 0.86 | 0.94 | 0.65 |
| 120 | 150 | 16 | 1.00 | . | 24-a3 | 0.05 | 12 | 16.00 | 12.60 | 16 | 12.0 | | 0.35 | 0.17 | 0.23 |
|  | | | | | | | | | | | | Choice rates of Option 2  **[Choice rates of Option 2 consistent with EV maximization]** | | | |
|  |  |  |  |  |  |  |  | Expected values | | Medians | | | CPC15 | | Here |
|  |  |  |  |  |  |  |  | EV1 | EV2 | Med1 | Med2 | | fDesc | fExp | fVal |
| Min |  |  |  |  |  |  |  | -9.9 | -14.38 | -41 | -46 | | 0.08  **[0.23]** | 0  **[0.22]** | 0.15  **[0.15]** |
| Max |  |  |  |  |  |  |  | 30 | 36.54 | 45 | 86 | | 0.97  **[0.97]** | 1.00  **[1.00]** | 0.85  **[0.85]** |
| Mean |  |  |  |  |  |  |  | 11.36 | 11.75 | 10.96 | 8.39 | | 0.5  **[0.66]** | 0.51  **[0.69]** | 0.44  **[0.58]** |
| STD |  |  |  |  |  |  |  | 11.76 | 11.86 | 13.36 | 18.96 | | 0.24  **[0.20]** | 0.28  **[0.18]** | 0.21  **[0.20]** |

*Notes*. The left-hand side presents the 120 problems (from CPC15) used in Study 1. Option 1 yields H1 with probability pH1, L1 otherwise. Option 2 yields H2 with probability pH2, L2 otherwise. In some of the problems, Hb is a lottery that provides a basic payoff (the number in the relevant cell) plus additional terms with a mean of 0. The distribution of the additional terms is described below (a3,a5, e.t.c.,). EV1 (EV2) and Med1 (Med2) presents the expected values and medians of Option 1 (Option 2). The right hand side presents the observed choice rates of Option 2 in CPC15: Column fDesc (from description) presents the first 5 trials that were made from description. Column fExp (from experience) shows the choice rate in the last 5 trials after 15 to 19 trials with experience (feedback). Column fVal (from valuations) shows the choice rate in Study 1 here.

a3: a draw from (-2, 0.5; 0, 0.25; 4)

a5: a draw from (-4, 0.5; -2, 0.25; 2, 0.125; 10, 0.00625; 26)

a6: a draw from (-5, 0.5; -3, 0.25; 1, 0.125, 9, 0.0625, 25, 0.03125; 57)

a7: a draw from (-6, 0.5; -4, 0.25; 0, 0.125, 8, 0.0625, 24, 0.03125; 56, 0.03125, 120)

a8: a draw from (-7, .5; -5, .25 -1, .125; +7, .0625; +23, .03125; 55, 0.01563; 119, 0.0078; 247)

s2: a draw from (-1, 0.50; 1)

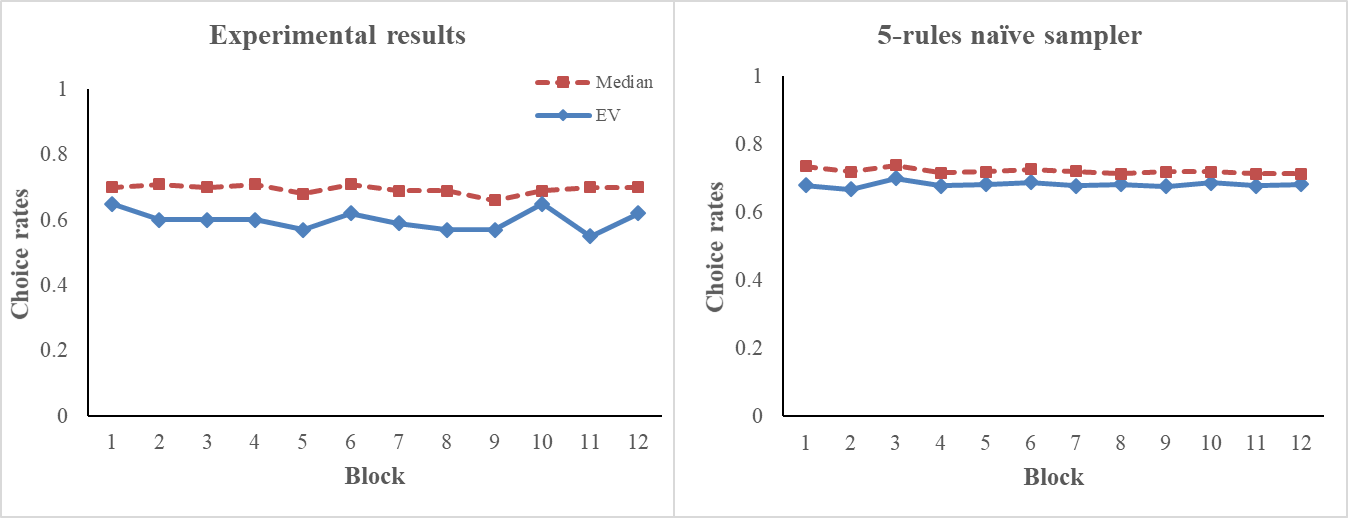
s3: a draw from (-1, 0.25; 0, 0.5; 1)

s5: a draw from (-2, 0.0625; -1, 0.25; 0, 0.375; 1, 0.25; 2)

s7: a draw from (-3, 0.0156; -2; 0.09375; -1, 0.23438; 0, 0.31250; 1, 0.23438; 2, 0.09375, 3)

s9: a draw from (-4, 0.0039; -3, 0.03125, -2, .10938, -1, 0.21875, 0, 0.27344, 1, 0.21875; 2, 0.10938; 3, 0.031, 4)

**Figure 2. Study 1: Mean Choice Rates and Model Predictions**



*Notes.* Mean choice rates of the option with the higher valuation according to each of the two experts (EV and Median) in 12 blocks of 10 trails (left), and the predictions of the “5-rules naïve sampler” model described below (Right).

***Individual differences.*** Table 3 presents the choice rates consistent with five decision rules by participants. The five rules include: EV (select the option suggested by Expert EV), Median (select the option suggested by Expert Median), Average (select the option favored by the average of the two experts (Hastie & Kameda, 2005; Larrick & Soll, 2006; Winkler, 1981)), Safe (select the option with the smaller difference between the two experts), and Risk (select the option with the larger difference between the two experts (Harries, Yaniv, & Harvey, 2004)). The decision to consider rule Average was made due to the fact that it implies the rational choice under the assumption that both experts are equally useful. Rules Safe and Risk were considered to capture the “valuations dominated” (i.e., contrary to the recommendation of both experts) choices. The results show large individual differences: While only one of the 40 participants was best fitted by the EV rule, 11 were best fitted by the Median rule, 12 by the Average, 10 by the minimum difference (Safe), and 6 by the maximum difference (Risk). The difference between the Median- and the EV-rates was significant for 14 of the 40 subjects (Sign test, p <.05), and in all 14 cases, the Median-rate was higher. For 4 of the subjects the difference between the Median- and the Average-rates was significant (Sign test, p <.05) and the Average-rate was higher in all four cases.

Furthermore, 16 of the 40 participants (40%) were more sensitive towards the absolute difference between the two valuations than to their values. Ten of them showed an inclination for the safer option (smaller difference between the two valuations) and the other six preferred the riskier option (larger difference between the two valuations). Sign test reveals that in 9 out of the 16 cases, one of the “Safe or Risk” rules fits the choice rate significantly better (p < .05) than any of the other three rules. Furthermore, these results suggest that at least some of the choices of the “valuations dominated” option are not merely random choices; we will further discuss the return to the individual differences in a subsequent section “Implications to Descriptive Models.”

**Table 3. Study 1: Choices Proportions Consistent with Five Distinct Rules by Subject**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| X  (rule name) | Higher  Expected Value  (EV) | Higher  Median  (Median) | Higher average of the two experts  (Average) | Smaller difference between the two experts  (Safe) | Larger difference between the two experts  (Risk) | Against both experts |
| Participant |  |
| 1 | 0.68 | **0.73** | 0.72 | 0.61 | 0.39 | 0.19 |
| 2 | 0.73 | 0.77 | **0.78** | 0.60 | 0.40 | 0.11 |
| 3 | 0.56 | 0.52 | 0.53 | 0.28 | **0.72** | 0.43 |
| 4 | 0.49 | 0.58 | 0.58 | **0.72** | 0.28 | 0.44 |
| 5 | 0.66 | 0.76 | **0.77** | 0.69 | 0.31 | 0.17 |
| 6 | 0.68 | **0.96** | 0.93 | 0.63 | 0.37 | 0.01 |
| 7 | 0.55 | 0.66 | 0.63 | **0.67** | 0.33 | 0.33 |
| 8 | 0.69 | 0.81 | **0.82** | 0.74 | 0.26 | 0.11 |
| 9 | 0.56 | **0.79** | 0.74 | 0.56 | 0.44 | 0.21 |
| 10 | 0.52 | 0.60 | 0.58 | **0.79** | 0.21 | 0.41 |
| 11 | 0.53 | 0.56 | 0.55 | 0.41 | **0.59** | 0.43 |
| 12 | 0.68 | **0.94** | 0.91 | 0.56 | 0.44 | 0.03 |
| 13 | **0.63** | 0.57 | 0.61 | 0.46 | 0.54 | 0.36 |
| 14 | 0.76 | 0.83 | **0.84** | 0.56 | 0.44 | 0.04 |
| 15 | 0.52 | 0.47 | 0.50 | 0.34 | **0.66** | 0.53 |
| 16 | 0.50 | 0.56 | **0.56** | 0.49 | 0.51 | 0.44 |
| 17 | 0.61 | 0.72 | 0.70 | **0.74** | 0.26 | 0.24 |
| 18 | 0.64 | 0.87 | **0.88** | 0.59 | 0.41 | 0.10 |
| 19 | 0.51 | 0.57 | 0.54 | **0.74** | 0.26 | 0.44 |
| 20 | 0.46 | 0.66 | 0.63 | **0.94** | 0.06 | 0.40 |
| 21 | 0.72 | 0.90 | **0.95** | 0.56 | 0.44 | 0.03 |
| 22 | 0.46 | **0.57** | 0.56 | 0.49 | 0.51 | 0.50 |
| 23 | 0.61 | **0.72** | 0.72 | 0.69 | 0.31 | 0.26 |
| 24 | 0.60 | 0.72 | 0.69 | **0.87** | 0.13 | 0.24 |
| 25 | 0.47 | **0.69** | 0.65 | 0.63 | 0.37 | 0.37 |
| 26 | 0.55 | **0.86** | 0.81 | 0.56 | 0.44 | 0.16 |
| 27 | 0.69 | **0.90** | 0.89 | 0.53 | 0.47 | 0.04 |
| 28 | 0.61 | **0.66** | 0.65 | 0.51 | 0.49 | 0.29 |
| 29 | 0.77 | 0.88 | **0.92** | 0.56 | 0.44 | 0.00 |
| 30 | 0.79 | 0.78 | **0.83** | 0.44 | 0.56 | 0.07 |
| 31 | 0.45 | 0.42 | 0.42 | 0.32 | **0.68** | 0.61 |
| 32 | 0.51 | 0.43 | 0.46 | 0.22 | **0.78** | 0.54 |
| 33 | 0.50 | 0.55 | 0.55 | **0.62** | 0.38 | 0.44 |
| 34 | 0.54 | 0.57 | 0.56 | **0.78** | 0.22 | 0.41 |
| 35 | 0.64 | **0.83** | 0.81 | 0.56 | 0.44 | 0.13 |
| 36 | 0.47 | 0.38 | 0.41 | 0.15 | **0.85** | 0.61 |
| 37 | 0.74 | 0.89 | **0.92** | 0.58 | 0.42 | 0.01 |
| 38 | 0.56 | 0.56 | 0.55 | **0.60** | 0.40 | 0.41 |
| 39 | 0.60 | 0.64 | **0.66** | 0.37 | 0.63 | 0.31 |
| 40 | 0.71 | 0.90 | **0.91** | 0.57 | 0.43 | 0.04 |
| # Best fitted | **1** | **11** | **12** | **10** | **6** |  |
| Min | 0.45 | 0.38 | 0.41 | 0.15 | 0.06 | 0.00 |
| Max | 0.79 | 0.96 | 0.95 | 0.94 | 0.85 | 0.61 |
| Mean | 0.60 | 0.70 | 0.69 | 0.57 | 0.43 | 0.27 |
| SDT | 0.1 | 0.15 | 0.15 | 0.17 | 0.17 | 0.18 |

*Notes.* 1. All the rules are of the type “choose the option with property X. The top row presents property X under each rule (and the rule’s name).

2. All the within-subject comparisons of EV with Med are significant (Sign test, p<.05), when the difference between the compared rate is .11 or larger. The best fitted Safe or Risk strategy is significantly higher than the best fitted alternative strategy when the difference is larger than 0.16.

3. Column “Against both experts” presents the individual choice rates *against* the suggestion of both experts, in 30% cases when both experts favored same alternative.

***Relationship to description-experience gap.*** The comparison between the decisions from (experts’) valuations that are examined here and the decisions from description and experience, as confronted in Erev et al. (2017) CPC15 study[[4]](#footnote-4), is presented in table 4. The table focuses on six of the problems examined in the CPC15 study. The “**Description”** column presents the choice rate of the riskier option in the first 5 trails in the CPC study. These first 5 decisions were made based on a complete description of the payoff distributions, prior to the subjects receiving any feedback for the outcome of their previous choices. The **“Description & Experience”** column presents the choice rate of the riskier option in the last 5 trials of the CPC study. Feedback was provided after each choice starting in trail 6 and while making the last 5 decisions the participants could rely on feedback from at least 15 past experiences with such problems. A comparison of the two CPC15 columns shows that experience decreases the weighing of the rare events and increases sensitivity to the win rate. The “**Valuation”** column shows that the choice rates in the experts’ paradigm analyzed here are similar and even more extreme than those at the “**Description & Experience”** column. In the current setting, decisions from valuation leads to a reversed Allais paradox/certainty effect, reversed reflection effect, and even stronger underweighting of rare events than the ones observed in choices from experience.

**Table 4. Three Classical Choice Anomalies**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | EV | Median | Description  (Block 1) | Description& Experience  (Block 5) | Valuations |
| Allais paradox/ Certainty effect (Kahneman & Tversky, 1979, following Allais, 1953) | A: 3 with certainty  B: 4, 0.80; 0 otherwise  Aʹ: 3, 0.25; 0 otherwise  Bʹ: 4, 0.20; 0 otherwise | 3  3.2  0.75  0.8 | 3  4  0  0 | 0.42  0.61 | 0.65  0.62 | 0.65  0.43 |
| Reflection effect (Kahneman & Tversky, 1979) | A: 3 with certainty  B: 4, 0.80; 0 otherwise  Aʹ: −3 with certainty  Bʹ: −4, 0.80; 0 otherwise | 3  3.2  -3  -3.2 | 3  4  -3  -4 | 0.42  0.49 | 0.65  0.36 | 0.65  0.28 |
| Over- and under-weighting of rare events (Kahneman & Tversky, 1979) | A: 2 with certainty  B: 101, 0.01; 1 otherwise  A: -1 with certainty  B: -20, 0.05; 0 otherwise | 2  2  -1  -1 | 2  1  -1  0 | 0.55  0.48 | 0.42  0.64 | 0.28  0.70 |

*Notes.* The results obtained in six problems that demonstrate classical choice anomalies in decisions under risk.

The singular explanation pertaining to the intensified underweighting of rare events, observed in decisions from valuations, reflects the joint impact of two related factors: first, the tendency to pay more attention to the error-rates (rather than to the MSE or the average earnings) leads participants to favor Expert Median over Expert EV. Second, the median estimates mask the rare events.

**3.2 Study 2**

The main goal of Study 2 is to compare the error-rate hypothesis, presented in the preceding section, with two alternative explanations of the results of Study 1. Under the first alternative explanation, the results of Study 1 reflect the use of absolute difference score to evaluate the experts. This rule justifies the choice of Expert Median (the Median minimizes the absolute error) and in Study 1 it agrees with the prescription of the error-rate hypothesis. To address this “absolute difference hypothesis,” Study 2 presents a detailed analysis of the four classes of 50 problems summarized in Table 5 and Appendix 2 (the values of S and C that distinguish between the different problems in each class were selected from the set {-2,-1, 0, 1, 2} and the uniform distribution U(2, 20) respectively). In study 2, the realization of the options are connected in a way that guarantees that the option with higher median minimizes the error-rate in Classes 1 and 2 (nevertheless, the expected return can be reduced by an average of 1.2 points by selecting the option of the higher median ), while the option with higher EV minimizes the error-rate in Classes 3 and 4. As per this design, it can be concluded that, in the present scenario, the predictions of the error-rate hypothesis differ from the predictions of the absolute difference hypothesis (that always favors Expert Median).

A second alternative explanation to the results of Study 1 is suggested by a comparison of column Median and Average in Table 3. This comparison suggests that in Study 1 the two rules provide very similar predictions and very similar fit of the results. Hence it can be concluded that the descriptive value of the error-rate hypothesis (that predicts high Median rate) depends entirely on the tendency of the participants to abide by the average rule. Study 2 addresses this alternative explanation by focusing on problems where in the Expert Median is less “decisive” (i.e., has smaller absolute difference between predictions) than the Expert EV (and for that reason only the implications of Expert EV agrees with the implications of the average of the two valuations). This condition makes way for inequality between the average rule hypotheses and the error-rate.

Furthermore, the main aim of Study 2 is to further simplify the nature of the errors that affect decisions in the current setting. In Study 1, reliance on Expert Median played the key role in minimizing the rate of the two types of errors that are of great importance for the decision makers (Cabantous et al., 2011): incorrect predictions (disagreement between the predicted and the observed outcomes), and implied errors (choosing the option that provides lower payoff). In Study 2, each option involves multiple outcomes to ensure that the two experts could not be compared based on their prediction error (the probability of prediction error for each expert was above 0.999). This design allows us to assess the effect of the implied errors.

**Table 5. Study 2: Four Problems Types**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| General format | | | | | | |
| Payoff from Safe | | S + C + εs | | | | |
| Payoff from Risk | | (S, 0.9; V08, 0.08; V02, 0.02) + C + εR | | | | |
| The values of S and C were selected before the start of the experiment and presented in Appendix 2.  The values of u1 in each trial was drawn from the uniform distribution U(-5,+5), and the value of u2 was set using the following rule: u2 = u1+1 if u1 < 4; and u1-9 otherwise. Thus, the realization of u2 is the sum of u1 and the outcome of a binary lottery that pays +1 if u1 < 4; and -9 otherwise. | | | | | | |
| Class parameters: Class 1 2 3 4 | | | | | | |
| εs | | | u1 | u2 | u2 | u1 |
| V08 | | | 10 | -10 | 10 | -10 |
| V02 | | | -100 | +100 | -100 | +100 |
| εR | | | u2 | u1 | u1 | u2 |
| Valuations: | | | | | | |
| Safe | EV | | C + S | C + S | C + S | C + S |
|  | Median | | C + S | C + S | C + S | C + S |
| Risk | EVa | | C + 0.9S -1.2 | C + 0.9S + 1.2 | C + 0.9S - 1.2 | C + 0.9S + 1.2 |
|  | Median | | C + S + 0.33 | C + S - 0.33 | C + S + 0.33 | C + S - 0.33 |
| Implications of the valuations: | | | | | | |
| EV favors Risk | | | No | Yes | No | Yes |
| Median favors Risk | | | Yes | No | Yes | No |
| Median more decisive | | | No | No | No | No |
| P(Payoff from Risk = Median’s valuation) | | | 0.001 | 0.001 | 0.001 | 0.001 |
| Median’s win-rate | | | 0.83 | 0.83 | 0.11 | 0.11 |
| Main results. Choice rate of the options with higher valuation according to the Expert Median (median-rate): | | | | | | |
|  | | | 0.55 | 0.54 | 0.40 | 0.40 |

*Notes.* aThe computation of the mean and the median of the risky prospect is explained in Appendix 3.

***Participant.*** Just as in study 1, eighty Technion student volunteers participated in this experiment in exchange for a financial payoff (the mean payoff was approximately 6$).

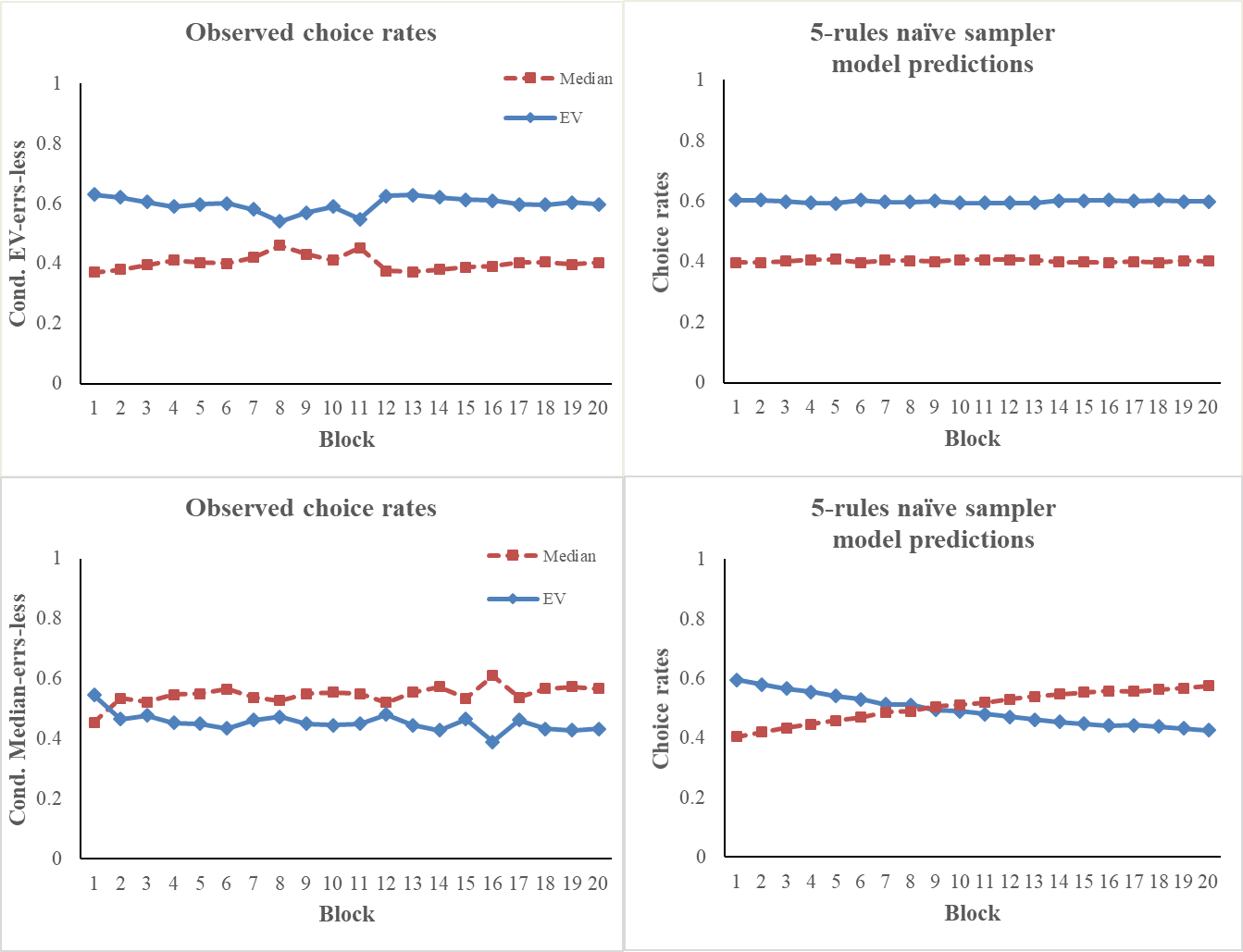
***Design and Procedure.*** The study used a between-subject design, where in forty participants were assigned to Condition Median-errs-less and confronted the problems in Class 1 and 2 twice. Each of the 100 problems were presented first in the first block and then in the second block of 100 trials. The order in each block was randomly determined and the experts in the first block were named A and B, while those in the second block were named C and D. The remaining forty participants were assigned to Condition EV-errs-less that employed the same procedure with the problems from Class 3 and 4. The instructions and screens were the same as in Study 1 (recall Figure 1).

**3.2.1 Results**

In table 4, the lower row presents the aggregate choice rates (and Appendix 2 presents the choice rates by problem). As per the predictions of the error-rate hypothesis (and in contradiction to the alternative explanations of the results of Study 1), the choice rate of the option suggested by Expert Median (the option with higher Median) is higher than 50% in Condition Median-errs-less (M = 0.55, STD = 0.15, t(39) = 1.99, p = 0.054) and significantly lower than 50% in Condition EV-errs-less (M = 0.40, STD = 0.14, t(39) = 4.29, p<0.001). The difference between the two conditions is significant (t(78) = 4.42, p<0.001).

In Figure 3, the left-hand graph presents that the mean choice rates in 20 blocks of 10 trials and the results reveal a relatively flat curve in Condition EV-errs-less, along with a fast initial adjustment in Condition Median-errs-less. A single explanation of the difference between the two conditions reflects the inclination to prefer the option suggested by the average of the two experts and an increase in the chances to follow the expert that wins more over time. The flat learning curves in Condition EV-wins more and the similar pattern in Study 1, can be the product of the fact that in these conditions the prescription of the average rule was similar to the prescription of the error-rate hypothesis. In the scenario when the expert with the least percentage of errors disagrees with the average (only in Condition Median-errs-less of Study 2), the results depict an evident increase in the inclination to follow the expert that errs less. The following right-hand graph presents the model predictions.

**Figure 3. Study 2: Mean Choice Rates and Model’s Predictions**



*Note*. Mean choice rates of the option with the higher valuation according to each of the two experts (EV and Median) (Left) and the “5-rules naïve sampler” model predictions described below (Right).

***Individual differences.*** Table 6 presents the choice rates that are consistent with four of the rules described in Table 3 (the average rule was omitted as its prescription, in the current study, is identical to the prescription of the EV rule). Each choice rate is based on 200 observations and differences of 0.15 or more are significant in a sign test (p < .05). Two major differences from the pattern documented in Study 1 are revealed in the results: larger individual differences and more subjects that are best described by rules Safe or Risk. Specifically, the proportion of subjects best described by rules Safe or Risk is 0.40 (16/40) in Study 1 and 0.64 (51/80) in the current study. Further analysis reveals that this difference between the two studies increases with time and the proportion of subjects best described by “Safe or Risk” increased with time in the Study 2 (from 0.58 in the first 50 trials to 0.68 in the last 50 trials), and decreased with time in Study 1 (from 0.425 in the first 50 trials to 0.325 in the last 50 trials). Due to a single explanation of the current increase, many participants in Study 2 stopped trusting the experts and chose a side (see related observation in Ecken & Pibernik, 2015). There are at least two good reasons to stop trusting the experts in Study 2: They always contradict each other and they are almost never accurate (the final outcome almost always differs from the valuation (Du et al., 2011; Yaniv & Kleinberger, 2000)).

**Table 6. The proportion of choices consistent with four distinct rules by condition and participant**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| EV-wins-more | | | | |  | Median-wins-more | | | | |
| Participant number | EV | Med | Safe | Risk |  | Participant number | EV | Med | Safe | Risk |
| 1 | **0.84** | 0.17 | 0.49 | 0.52 |  | 41 | 0.47 | 0.54 | **0.80** | 0.20 |
| 2 | **0.93** | 0.08 | 0.52 | 0.48 |  | 42 | 0.50 | 0.50 | **0.88** | 0.13 |
| 3 | **0.76** | 0.25 | 0.54 | 0.46 |  | 43 | 0.51 | 0.49 | 0.02 | **0.99** |
| 4 | 0.59 | 0.42 | 0.35 | **0.65** |  | 44 | 0.46 | 0.55 | 0.40 | **0.61** |
| 5 | **0.86** | 0.14 | 0.51 | 0.50 |  | 45 | 0.51 | 0.50 | **1.00** | 0.01 |
| 6 | 0.63 | 0.38 | 0.14 | **0.86** |  | 46 | 0.49 | 0.51 | **0.73** | 0.27 |
| 7 | **0.72** | 0.28 | 0.47 | 0.53 |  | 47 | 0.41 | 0.59 | 0.38 | **0.63** |
| 8 | 0.52 | 0.49 | 0.39 | **0.61** |  | 48 | 0.50 | 0.50 | **0.53** | 0.47 |
| 9 | **0.78** | 0.22 | 0.38 | 0.63 |  | 49 | **0.53** | 0.47 | 0.51 | 0.49 |
| 10 | 0.51 | 0.50 | **1.00** | 0.01 |  | 50 | 0.51 | 0.49 | **0.79** | 0.21 |
| 11 | 0.46 | 0.55 | 0.32 | **0.69** |  | 51 | 0.45 | **0.56** | 0.46 | 0.55 |
| 12 | 0.68 | 0.32 | **0.79** | 0.22 |  | 52 | 0.17 | **0.83** | 0.45 | 0.56 |
| 13 | 0.52 | 0.48 | **0.54** | 0.46 |  | 53 | 0.23 | **0.77** | 0.36 | 0.64 |
| 14 | 0.49 | 0.51 | **0.83** | 0.17 |  | 54 | **0.54** | 0.47 | 0.51 | 0.50 |
| 15 | 0.56 | 0.44 | **0.71** | 0.29 |  | 55 | 0.57 | 0.44 | **0.69** | 0.32 |
| 16 | **0.89** | 0.11 | 0.46 | 0.55 |  | 56 | 0.56 | 0.45 | 0.42 | **0.58** |
| 17 | **0.57** | 0.43 | 0.49 | 0.51 |  | 57 | 0.50 | 0.51 | 0.01 | **1.00** |
| 18 | 0.50 | 0.51 | **0.62** | 0.38 |  | 58 | 0.07 | **0.94** | 0.47 | 0.54 |
| 19 | 0.45 | 0.55 | 0.42 | **0.58** |  | 59 | 0.49 | 0.51 | 0.34 | **0.67** |
| 20 | **0.53** | 0.47 | 0.50 | 0.50 |  | 60 | 0.55 | 0.45 | **0.58** | 0.43 |
| 21 | 0.47 | 0.53 | **0.64** | 0.36 |  | 61 | 0.13 | **0.87** | 0.51 | 0.50 |
| 22 | 0.57 | 0.43 | 0.40 | **0.60** |  | 62 | 0.51 | 0.49 | **0.94** | 0.06 |
| 23 | **0.92** | 0.08 | 0.44 | 0.57 |  | 63 | 0.50 | 0.51 | 0.10 | **0.90** |
| 24 | 0.49 | 0.51 | **0.53** | 0.48 |  | 64 | **0.58** | 0.42 | 0.50 | 0.51 |
| 25 | 0.46 | 0.54 | 0.42 | **0.59** |  | 65 | 0.50 | 0.50 | 0.49 | **0.52** |
| 26 | 0.49 | 0.51 | **1.00** | 0.00 |  | 66 | 0.62 | 0.39 | 0.38 | **0.63** |
| 27 | 0.51 | 0.50 | **0.71** | 0.29 |  | 67 | 0.57 | 0.44 | **0.70** | 0.30 |
| 28 | 0.48 | 0.53 | **0.54** | 0.46 |  | 68 | 0.52 | 0.49 | **0.61** | 0.39 |
| 29 | **0.71** | 0.29 | 0.38 | 0.63 |  | 69 | 0.07 | **0.94** | 0.49 | 0.51 |
| 30 | **0.55** | 0.45 | 0.52 | 0.48 |  | 70 | 0.58 | 0.43 | 0.42 | **0.58** |
| 31 | 0.52 | 0.49 | 0.28 | **0.73** |  | 71 | 0.47 | 0.54 | **0.72** | 0.28 |
| 32 | 0.52 | 0.49 | 0.28 | **0.73** |  | 72 | 0.54 | 0.46 | **0.86** | 0.15 |
| 33 | 0.55 | 0.45 | 0.22 | **0.79** |  | 73 | 0.08 | **0.92** | 0.50 | 0.50 |
| 34 | 0.45 | 0.56 | 0.09 | **0.91** |  | 74 | 0.48 | **0.53** | 0.50 | 0.51 |
| 35 | 0.51 | 0.50 | 0.01 | **0.99** |  | 75 | 0.51 | 0.49 | 0.28 | **0.73** |
| 36 | 0.49 | 0.52 | **0.56** | 0.44 |  | 76 | 0.53 | 0.47 | 0.21 | **0.79** |
| 37 | 0.55 | 0.46 | **0.89** | 0.12 |  | 77 | 0.45 | 0.56 | **0.80** | 0.20 |
| 38 | **0.80** | 0.20 | 0.44 | 0.57 |  | 78 | 0.49 | 0.51 | 0.22 | **0.78** |
| 39 | **0.68** | 0.33 | 0.37 | 0.64 |  | 79 | 0.58 | 0.43 | 0.20 | **0.81** |
| 40 | 0.50 | 0.51 | **1.00** | 0.01 |  | 80 | 0.50 | 0.50 | 0.23 | **0.78** |
| Min | 0.45 | 0.08 | 0.01 | 0 |  |  | 0.07 | 0.39 | 0.01 | 0.01 |
| Max | 0.93 | 0.56 | 1 | 0.99 |  |  | 0.62 | 0.94 | 1 | 1 |
| Mean | 0.60 | 0.40 | 0.50 | 0.50 |  |  | 0.44 | 0.56 | 0.51 | 0.49 |
| STD | 0.14 | 0.14 | 0.23 | 0.23 |  |  | 0.15 | 0.15 | 0.24 | 0.24 |
| # best fitted | **14** | 0 | 14 | 12 |  |  | 3 | 8 | 14 | **15** |

*Notes*. 1. The rules are described in Table 2.

2. 0.5 was rounded upward.

3. All the within-subject comparisons are significant (Sign test, p<.05) when the difference between the compared rate is .15 or larger. When the compared rates are not mutually exclusive, smaller differences can be significant too.

**4. Implications descriptive models**

Previous research shows a tendency to maximize the win-rate (and minimize the error-rate), in the case of basic decisions from experience tasks. Comparison of alternative abstractions of this pattern highlights the descriptive value of the assumption that people tend to rely on small samples of past experiences (see a review in Erev & Haruvy (2016)). Specifically, the main findings documented in studies of decisions from experience among unmarked prospects can be captured by the naïve sampler model (Erev & Roth, 2014). This model assumes two decision modes, namely exploration (random choice) and exploitation. In exploitation trials, the decision maker j considers a small sample consisting of kj past trials and selects the option that provided the highest average payoff in these trials. The value of kj is a property of the decision maker and assumed to be distributed uniformly in the population between 1 and K (where K is a free parameter). To further clarify the relationship between the decisions from valuations to pure decisions from experience, this section questions the assumptions that should be added to the naïve sampler model to capture the current findings.

The first generalization we considered, referred to as the “5-rules naïve sampler” model, assumes that the decision maker considers the five rules summarized in Table 2. Additionally, it assumes random choice among these five strategies in the “early trials.” The exact probability of a random choice in trial t in an experiment with T trials is set to be equal to:

P(random choice at t) =

where α is a free parameter that captures the exploration rate. Thus, the model has only two free parameters: K that denotes the maximal sample size (the actual sample size is between 1 and K, with the mean at ), and α. Appendix 3 presents a numerical example that clarifies the assumed processes.

The right-hand curves in Figures 2 and 3 present the predictions of the 5-rules naïve sampler model for the current experiments with the parameters that best fit the data (K = 12, α = 0.4). The fit was quantified with the Mean Squared Distance between the observed and the predicted (reproduced) EV-rates and Median-rates over the three conditions (the 104 blocks X conditions over the two experiments). The MSD score obtained was 0.0025.

In addition, the model captures two non-trivial qualitative features of the observed results. First is the difference between the initial behavior in the two studies: The participants tend to follow Expert Median in the early trials of Study 1 and Expert EV in the early trials of Study 2. The model captures this pattern because Strategy Average tends to favor Expert Median in Study 1 and Expert EV in Study 2. The other four strategies imply equal EV- and Median-rates and for that reason the predicted initial rates reflect the bias of Strategy Average. The second interesting qualitative feature is the observation of a preference reversal with time in Condition Median-errs-less in Study 2 and relatively flat curves in the other cases. The model captures this pattern because in Condition Median-errs-less, the less decisive expert (the one that affects the average less), becomes more influential with experience.

Further analyses reveal that it is not easy to find a simpler model that provides a similar fit. For example, elimination of either one of the five rules increase the MSD score by more than 100%.

An analysis of the individual choice rates reveals the shortcomings of the 5-rules naïve sampler model: it under-predicts the large individual differences. For example, the model with the parameters that fits the aggregate choice rate predicts that the Safe-rate in Condition Median-wins-more of Study 2 will vary between 33% to 62%. Nevertheless, the observed range is between 1% to 100% and this shortcoming can be addressed with generalizations of the 5-rules model that allows for positive serial dependencies in the sampling of past experiences. That is, if a past experience is sampled at trial t, it is also likely to be used again in trial t+1 with high probability (with probability φ, the agent uses the previous sample). With large φ, the model predicts that each agent will use the same small samples for many trials. Since each small sample has the ability to favor a different rule, this assumption implies large individual differences. Also, with large φ the model predicts higher inertia (repetition) rates by agents that tend to ignore the experts and are best described by rules Safe and Risk. This prediction is descriptive; for example, the inertia rates in trials 151-200 of Study 2 are 79% for subjects whose choices in trails 1-150 are best described by “Safe or Risk,” and only 51% for subjects whose choices in trails 1-150 are best described by “EV or median.”

Under an alternative explanation of the large individual differences, the participants are more sensitive to the observed outcomes predicted by our model. This explanation implies that the participants that follow the EV rule are those who experience large advantages of this rule. Nevertheless, this prediction is not supported by the data: the correlation between the EV-rate in trial 151-to 200 of Study 2 and the observed benefit from using this rule is insignificant (r = 0.08, p > 0.1, ns).

Figure 4 presents the proportion of choices consistent with Expert EV, after trial 50 in Study 2, as a function of the number of trials since the last observation of extreme outcome (the rare V02 event). Notice that Expert Median ignores these events and their occurrence clarifies the advantage of following Expert EV. Thus, the assumption that the deviations from maximization reflect cognitive limitations implies higher EV-rate immediately after the V02 event and a decline with time. The results do not reflect this pattern; rather, they reveal a wavy recency pattern of the type documented by Plonsky et al. (2015). For example, the EV rate was lower 3 trials after the V02 event (53%) than 10 trials after a V02 event (58%). Plonsky et al. explain this pattern with the assertion that one of the contributors to the tendency to rely on small samples is an effort to select the strategy that leads to the best outcomes in similar situations in the past and the participants have less experiences after observing rare events. This is especially surprising given the task complexity.

**Figure 4. Study 2 percent of choices consistent with Expert EV prescription as a function of the number of trials since last observed extreme outcome**

**5. General Discussion**

Previous studies pertaining to the effect of experience on advice taking highlights the advantage of accurate probability estimates over direct recommendations (Bolton & Katok, 2018). The current investigation compares two interpretations of this finding in the context of decisions from valuations of unknown payoff distributions. Both interpretations assume that experience increases the tendency to trust experts that appear to be reliable (Engelmann et al., 2009), but differ with respect to the assumed subjective measurement of reliability. As per the MSE hypothesis, the decision makers concentrate on how accurate the predictions are and behave as if they use a proper scoring rule: They learn to trust the expert with the lower mean squared error between the prediction and the final outcomes. Under the error-rate hypothesis, the decision makers focus on the final outcomes and learn to trust the expert that guarantees the best choice in most cases. The results favor the error-rate hypothesis and this hypothesis best fits the observed decisions even after 200 trials of experience when the behavior, it predicts, impairs expected return. Furthermore, the results suggest that given a choice among two independent payoff distributions, the decision makers are more likely to rely on experts that report the median over experts that report the expected value.

These findings have two clear practical implications. First, they suggest that designers of expert systems can increase the system’s popularity by biasing the valuation in favor of the option that leads to the best outcome in most cases, even if this bias impairs the users’ expected return. This bias is predicted to be particularly effective when the competing expert systems favor these “counterproductive but less likely to err” options. In these settings the bias has two “positive” effects: it increases the weighing of the biased expert relative to the alternative experts and it also reduces the tendency to ignore both experts. In other words, competition among experts does not guarantee an increase in accuracy. Second, in most settings, the expected bias is likely to trigger insufficient sensitivity to rare events, which will in turn enhance the “Black Swan” effect (Taleb, 2007).

The theoretical implications of the current analysis involve the generalization of basic decision-making research to situations in which the decision makers must learn how to use the available information (in our settings, the valuations and obtained outcomes). Our results suggest that the abstraction of the decision processes in this setting requires two sub-models. The first summarizes the decision rules considered by the decision makers. Our results can be summarized with the assumption that the decision makers considered at least five rules: follow Expert A, follow Expert B, follow the average of the two experts, select the safer option (the one on which the two experts agree), and select the riskier option (the one on which the two experts disagree).

The second sub-model summarizes the manner in which the decision makers make an informed selection among these rules. The bias toward the expert that wins in more cases can be captured with the assumption that this choice is likely to be made based on small samples of past experiences. In that respect, decisions from valuations are similar to basic decisions from experience. Furthermore, the effort to capture the large individual differences suggests high inertia rates. When the two experts disagree and their accuracy rates are low, many decision makers behave as if their experience leads them to ignore the experts’ valuations altogether and repeat their last choice.

The error-rate hypothesis and model supported here leads to problematic predictions concerning the impact of “fake news,” which we hope to explore in future research. Specifically, the prediction that decision makers tend to trust experts with a higher success rate, even when the valuations provided by these experts impair expected earning and do not minimize the mean squared error, implies a tendency to prefer certain sources of fake news over sources that provide well-calibrated valuations (see related assertions in Radzevick & Moore, 2011; Greenstein & Zhu, 2012, 2014). In addition, the error-rate hypothesis and the current model can be used to suggest interventions that reduce the impact of fake news. For example, the current analysis suggests that the presentation of the aggregate gain (the average payoff from following the option favored by each expert over several trials) that increases the proportion of cases in which the unbiased expert wins, will increase the tendency to follow the accurate valuations.

To conclude, the current investigation of decisions from valuations of unknown payoff distributions highlights four main observations: (1) Experience can decrease the trust in well-calibrated sources of information. (2) Experience increases the tendency to trust the expert that minimizes the error-rate (implies the best choice) in most cases. (3) Consensus among experts increases trust, but does not guarantee acceptance of their advice. To be more specific, in the current study, in more than 25% of the trials where in the two experts favored the same option, the decision makers chose the alternative option. (4) The effect of experience on the weighting of experts, in the current setting, can be captured by a simple generalization of the naïve sampling model that assumes reliance on small sample of past experiences and the generalization suggests learning among five simple rules and high inertia rate.

**Appendix 1. Study 2: Problems and Choice Rates**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  | |  | | Risk choice rate | |
|  |  |  |  |  | Expert EV valuations | | Expert Median valuations | | Median wins more | EV wins more |
| Observation | C | S | V08 | V02 | Safe | Risk | Safe | Risk | εS = u1  εR = u2  Class1 | εS = u2  εR = u1  Class3 |
| 1 | 2.26 | 2 | 10 | -100 | 4.26 | 2.86 | 4.26 | 4.59 | 0.55 | 0.40 |
| 2 | 2.29 | 1 | 10 | -100 | 3.29 | 1.99 | 3.29 | 3.62 | 0.57 | 0.35 |
| 3 | 2.3 | 0 | 10 | -100 | 2.30 | 1.10 | 2.30 | 2.63 | 0.59 | 0.36 |
| 4 | 2.33 | -1 | 10 | -100 | 1.33 | 0.23 | 1.33 | 1.66 | 0.58 | 0.53 |
| 5 | 2.85 | -2 | 10 | -100 | 0.85 | -0.15 | 0.85 | 1.18 | 0.50 | 0.28 |
| 6 | 3 | 1 | 10 | -100 | 4.00 | 2.70 | 4.00 | 4.33 | 0.53 | 0.37 |
| 7 | 3.24 | 0 | 10 | -100 | 3.24 | 2.04 | 3.24 | 3.57 | 0.58 | 0.40 |
| 8 | 3.36 | 2 | 10 | -100 | 5.36 | 3.96 | 5.36 | 5.69 | 0.55 | 0.51 |
| 9 | 3.73 | -1 | 10 | -100 | 2.73 | 1.63 | 2.73 | 3.06 | 0.55 | 0.40 |
| 10 | 3.96 | -2 | 10 | -100 | 1.96 | 0.96 | 1.96 | 2.29 | 0.53 | 0.40 |
| 11 | 4.26 | 2 | 10 | -100 | 6.26 | 4.86 | 6.26 | 6.59 | 0.60 | 0.30 |
| 12 | 4.42 | 0 | 10 | -100 | 4.42 | 3.22 | 4.42 | 4.75 | 0.56 | 0.41 |
| 13 | 4.57 | -1 | 10 | -100 | 3.57 | 2.47 | 3.57 | 3.9 | 0.56 | 0.45 |
| 14 | 4.8 | 1 | 10 | -100 | 5.8 | 4.5 | 5.8 | 6.13 | 0.59 | 0.48 |
| 15 | 4.83 | -2 | 10 | -100 | 2.83 | 1.83 | 2.83 | 3.16 | 0.56 | 0.41 |
| 16 | 5.13 | 2 | 10 | -100 | 7.13 | 5.73 | 7.13 | 7.46 | 0.63 | 0.34 |
| 17 | 5.17 | -1 | 10 | -100 | 4.17 | 3.07 | 4.17 | 4.5 | 0.52 | 0.36 |
| 18 | 5.81 | 0 | 10 | -100 | 5.81 | 4.61 | 5.81 | 6.14 | 0.55 | 0.38 |
| 19 | 5.87 | -2 | 10 | -100 | 3.87 | 2.87 | 3.87 | 4.2 | 0.60 | 0.42 |
| 20 | 5.93 | 1 | 10 | -100 | 6.93 | 5.63 | 6.93 | 7.26 | 0.50 | 0.48 |
| 21 | 6.24 | -1 | 10 | -100 | 5.24 | 4.14 | 5.24 | 5.57 | 0.54 | 0.39 |
| 22 | 6.26 | -2 | 10 | -100 | 4.26 | 3.26 | 4.26 | 4.59 | 0.58 | 0.36 |
| 23 | 6.46 | 1 | 10 | -100 | 7.46 | 6.16 | 7.46 | 7.79 | 0.51 | 0.37 |
| 24 | 6.49 | 2 | 10 | -100 | 8.49 | 7.09 | 8.49 | 8.82 | 0.60 | 0.39 |
| 25 | 6.61 | 0 | 10 | -100 | 6.61 | 5.41 | 6.61 | 6.94 | 0.60 | 0.46 |
| 26 | 7.08 | -1 | 10 | -100 | 6.08 | 4.98 | 6.08 | 6.41 | 0.53 | 0.43 |
| 27 | 7.4 | 2 | 10 | -100 | 9.4 | 8.00 | 9.4 | 9.73 | 0.52 | 0.38 |
| 28 | 7.41 | 0 | 10 | -100 | 7.41 | 6.21 | 7.41 | 7.74 | 0.60 | 0.49 |
| 29 | 7.45 | -2 | 10 | -100 | 5.45 | 4.45 | 5.45 | 5.78 | 0.58 | 0.41 |
| 30 | 7.56 | 1 | 10 | -100 | 8.56 | 7.26 | 8.56 | 8.89 | 0.53 | 0.39 |
| 31 | 8.06 | 1 | 10 | -100 | 9.06 | 7.76 | 9.06 | 9.39 | 0.54 | 0.45 |
| 32 | 8.18 | -2 | 10 | -100 | 6.18 | 5.18 | 6.18 | 6.51 | 0.55 | 0.30 |
| 33 | 8.36 | 0 | 10 | -100 | 8.36 | 7.16 | 8.36 | 8.69 | 0.52 | 0.46 |
| 34 | 8.9 | -1 | 10 | -100 | 7.90 | 6.80 | 7.90 | 8.23 | 0.51 | 0.41 |
| 35 | 8.94 | 2 | 10 | -100 | 10.94 | 9.54 | 10.94 | 11.27 | 0.52 | 0.42 |
| 36 | 9.06 | 2 | 10 | -100 | 11.06 | 9.66 | 11.06 | 11.39 | 0.52 | 0.32 |
| 37 | 9.41 | -2 | 10 | -100 | 7.41 | 6.41 | 7.41 | 7.74 | 0.57 | 0.40 |
| 38 | 9.42 | -1 | 10 | -100 | 8.42 | 7.32 | 8.42 | 8.75 | 0.48 | 0.35 |
| 39 | 9.46 | 1 | 10 | -100 | 10.46 | 9.16 | 10.46 | 10.79 | 0.48 | 0.41 |
| 40 | 9.52 | 0 | 10 | -100 | 9.52 | 8.32 | 9.52 | 9.85 | 0.65 | 0.40 |
| 41 | 10.07 | 0 | 10 | -100 | 10.07 | 8.87 | 10.07 | 10.4 | 0.56 | 0.32 |
| 42 | 10.17 | -2 | 10 | -100 | 8.17 | 7.17 | 8.17 | 8.5 | 0.52 | 0.38 |
| 43 | 10.56 | 1 | 10 | -100 | 11.56 | 10.26 | 11.56 | 11.89 | 0.50 | 0.41 |
| 44 | 10.88 | -1 | 10 | -100 | 9.88 | 8.78 | 9.88 | 10.21 | 0.50 | 0.48 |
| 45 | 10.96 | 2 | 10 | -100 | 12.96 | 11.56 | 12.96 | 13.29 | 0.62 | 0.49 |
| 46 | 11.18 | 2 | 10 | -100 | 13.18 | 11.78 | 13.18 | 13.51 | 0.51 | 0.35 |
| 47 | 11.44 | -1 | 10 | -100 | 10.44 | 9.34 | 10.44 | 10.77 | 0.55 | 0.37 |
| 48 | 11.61 | 0 | 10 | -100 | 11.61 | 10.41 | 11.61 | 11.94 | 0.63 | 0.41 |
| 49 | 11.76 | 1 | 10 | -100 | 12.76 | 11.46 | 12.76 | 13.09 | 0.56 | 0.42 |
| 50 | 11.81 | -2 | 10 | -100 | 9.81 | 8.81 | 9.81 | 10.14 | 0.55 | 0.40 |
| Min  Max  Mean |  |  |  |  |  |  |  |  | 0.48  0.65  0.55 | 0.28  0.53  0.40 |
| Observation | C | S | V08 | V02 | Safe | Risk | Safe | Risk | εS = u2  εR = u1  Class2 | εS = u1  εR = u2  Class4 |
| 51 | 2.26 | 2 | -10 | 100 | 4.26 | 5.26 | 4.26 | 3.93 | 0.42 | 0.48 |
| 52 | 2.29 | 1 | -10 | 100 | 3.29 | 4.39 | 3.29 | 2.96 | 0.45 | 0.68 |
| 53 | 2.3 | 0 | -10 | 100 | 2.30 | 3.50 | 2.30 | 1.97 | 0.42 | 0.63 |
| 54 | 2.33 | -1 | -10 | 100 | 1.33 | 2.63 | 1.33 | 1.00 | 0.41 | 0.61 |
| 55 | 2.85 | -2 | -10 | 100 | 0.85 | 2.25 | 0.85 | 0.52 | 0.50 | 0.59 |
| 56 | 3 | 1 | -10 | 100 | 4.00 | 5.10 | 4.00 | 3.67 | 0.44 | 0.50 |
| 57 | 3.24 | 0 | -10 | 100 | 3.24 | 4.44 | 3.24 | 2.91 | 0.49 | 0.62 |
| 58 | 3.36 | 2 | -10 | 100 | 5.36 | 6.36 | 5.36 | 5.03 | 0.41 | 0.64 |
| 59 | 3.73 | -1 | -10 | 100 | 2.73 | 4.03 | 2.73 | 2.40 | 0.47 | 0.56 |
| 60 | 3.96 | -2 | -10 | 100 | 1.96 | 3.36 | 1.96 | 1.63 | 0.37 | 0.62 |
| 61 | 4.26 | 2 | -10 | 100 | 6.26 | 7.26 | 6.26 | 5.93 | 0.43 | 0.62 |
| 62 | 4.42 | 0 | -10 | 100 | 4.42 | 5.62 | 4.42 | 4.09 | 0.47 | 0.66 |
| 63 | 4.57 | -1 | -10 | 100 | 3.57 | 4.87 | 3.57 | 3.24 | 0.41 | 0.54 |
| 64 | 4.8 | 1 | -10 | 100 | 5.80 | 6.90 | 5.8 | 5.47 | 0.40 | 0.56 |
| 65 | 4.83 | -2 | -10 | 100 | 2.83 | 4.23 | 2.83 | 2.50 | 0.45 | 0.61 |
| 66 | 5.13 | 2 | -10 | 100 | 7.13 | 8.13 | 7.13 | 6.80 | 0.44 | 0.60 |
| 67 | 5.17 | -1 | -10 | 100 | 4.17 | 5.47 | 4.17 | 3.84 | 0.50 | 0.66 |
| 68 | 5.81 | 0 | -10 | 100 | 5.81 | 7.01 | 5.81 | 5.48 | 0.52 | 0.55 |
| 69 | 5.87 | -2 | -10 | 100 | 3.87 | 5.27 | 3.87 | 3.54 | 0.55 | 0.62 |
| 70 | 5.93 | 1 | -10 | 100 | 6.93 | 8.03 | 6.93 | 6.60 | 0.44 | 0.52 |
| 71 | 6.24 | -1 | -10 | 100 | 5.24 | 6.54 | 5.24 | 4.91 | 0.41 | 0.65 |
| 72 | 6.26 | -2 | -10 | 100 | 4.26 | 5.66 | 4.26 | 3.93 | 0.51 | 0.50 |
| 73 | 6.46 | 1 | -10 | 100 | 7.46 | 8.56 | 7.46 | 7.13 | 0.42 | 0.59 |
| 74 | 6.49 | 2 | -10 | 100 | 8.49 | 9.49 | 8.49 | 8.16 | 0.47 | 0.57 |
| 75 | 6.61 | 0 | -10 | 100 | 6.61 | 7.81 | 6.61 | 6.28 | 0.52 | 0.62 |
| 76 | 7.08 | -1 | -10 | 100 | 6.08 | 7.38 | 6.08 | 5.75 | 0.50 | 0.55 |
| 77 | 7.4 | 2 | -10 | 100 | 9.4 | 10.4 | 9.4 | 9.07 | 0.54 | 0.67 |
| 78 | 7.41 | 0 | -10 | 100 | 7.41 | 8.61 | 7.41 | 7.08 | 0.48 | 0.62 |
| 79 | 7.45 | -2 | -10 | 100 | 5.45 | 6.85 | 5.45 | 5.12 | 0.44 | 0.61 |
| 80 | 7.56 | 1 | -10 | 100 | 8.56 | 9.66 | 8.56 | 8.23 | 0.51 | 0.60 |
| 81 | 8.06 | 1 | -10 | 100 | 9.06 | 10.16 | 9.06 | 8.73 | 0.45 | 0.53 |
| 82 | 8.18 | -2 | -10 | 100 | 6.18 | 7.58 | 6.18 | 5.85 | 0.45 | 0.60 |
| 83 | 8.36 | 0 | -10 | 100 | 8.36 | 9.56 | 8.36 | 8.03 | 0.36 | 0.59 |
| 84 | 8.9 | -1 | -10 | 100 | 7.9 | 9.2 | 7.9 | 7.57 | 0.46 | 0.58 |
| 85 | 8.94 | 2 | -10 | 100 | 10.94 | 11.94 | 10.94 | 10.61 | 0.45 | 0.58 |
| 86 | 9.06 | 2 | -10 | 100 | 11.06 | 12.06 | 11.06 | 10.73 | 0.44 | 0.60 |
| 87 | 9.41 | -2 | -10 | 100 | 7.41 | 8.81 | 7.41 | 7.08 | 0.38 | 0.60 |
| 88 | 9.42 | -1 | -10 | 100 | 8.42 | 9.72 | 8.42 | 8.09 | 0.51 | 0.67 |
| 89 | 9.46 | 1 | -10 | 100 | 10.46 | 11.56 | 10.46 | 10.13 | 0.49 | 0.64 |
| 90 | 9.52 | 0 | -10 | 100 | 9.52 | 10.72 | 9.52 | 9.19 | 0.45 | 0.65 |
| 91 | 10.07 | 0 | -10 | 100 | 10.07 | 11.27 | 10.07 | 9.74 | 0.43 | 0.62 |
| 92 | 10.17 | -2 | -10 | 100 | 8.17 | 9.57 | 8.17 | 7.84 | 0.46 | 0.54 |
| 93 | 10.56 | 1 | -10 | 100 | 11.56 | 12.66 | 11.56 | 11.23 | 0.42 | 0.57 |
| 94 | 10.88 | -1 | -10 | 100 | 9.88 | 11.18 | 9.88 | 9.55 | 0.51 | 0.60 |
| 95 | 10.96 | 2 | -10 | 100 | 12.96 | 13.96 | 12.96 | 12.63 | 0.48 | 0.64 |
| 96 | 11.18 | 2 | -10 | 100 | 13.18 | 14.18 | 13.18 | 12.85 | 0.53 | 0.58 |
| 97 | 11.44 | -1 | -10 | 100 | 10.44 | 11.74 | 10.44 | 10.11 | 0.51 | 0.58 |
| 98 | 11.61 | 0 | -10 | 100 | 11.61 | 12.81 | 11.61 | 11.28 | 0.48 | 0.58 |
| 99 | 11.76 | 1 | -10 | 100 | 12.76 | 13.86 | 12.76 | 12.43 | 0.42 | 0.64 |
| 100 | 11.81 | -2 | -10 | 100 | 9.81 | 11.21 | 9.81 | 9.48 | 0.43 | 0.64 |
| Min  Max  Mean |  |  |  |  |  |  |  |  | 0.36  0.55  0.46 | 0.48  0.68  0.60 |

*Notes*. Study 2 problems set. All the problems involved a choice between Option Safe that provides S + C + εS and Option Risk that yields (S, 0.9; V08, 0.08; V02, 0.02) + C + εR. Each pair of S & C were used in each of the four classes of problems that differ with respect to the value of the rare outcomes (V02 and V08) and the error terms combinations (εR and εS). The right-hand columns present the Risk choice rates. Note that the prospect with error term = u2, is better most of the time.

**Appendix 2. The Computation of the EVs and the Medians of the Risky Prospects in Study 2:**

3.1 The risky prospect “(S, 0.9; +10, 0.08; -100, 0.02) + C + εR”

EV:

Since the EV of εR = 0, the EV of the prospect is: 0.9(S) + 0.08(10) + 0.02(-100) + C = 0.9S + C – 1.2

Median:

In the worst 2% of the cases the payoff is between -105 + C to -95 + C

In the middle 90% of the cases the payoff is between -5 + S + C to +5 + S + C

In the best 8% of the cases the payoff is between 5 + C to 15 + C

Thus: when δ is a very small number (δ convergence to 0):

The (2+ δ)th percentile (the value better than (2 + δ)% of the payoffs) is -5 + S + C

The (92- δ)th percentile (the value better than (92- δ )% of the payoffs) is 5 + S + C

Thus, in the range between the 2nd and the 92nd percentile, the payoff increases by 10/90 with every percentile. Therefore, the 50th percentile (the median) is:

-5 + S + C + 48(10/90) = S + C + 1/3

3.2 The risky prospect “(S, 0.9; -10, 0.08; +100, 0.02) + C + εR”

EV:

Since the EV of εR =0, the EV of the prospect is: 0.9(S) + 0.08(-10) + 0.02(100) + C = 0.9S + C + 1.2

Median:

In the worst 8% of the cases the payoff is between -15 + C to -5 + C

In the middle 90% of the cases the payoff is between -5 + S + C to +5 + S + C

In the best 2% of the cases the payoff is between 95 + C to 105 + C

Thus: when δ is a very small number (δ convergence to 0):

The (8+ δ)th percentile (the value better than (8 + δ)% of the payoffs) is -5 + S + C

The (98- δ)th percentile (the value better than (98- δ )% of the payoffs) is + 5 + S + C

Thus, in the range between the 8nd and the 98nd percentile, the payoff increases by 10/90 with every percentile. Therefore, the 50th percentile (the median) is:

-5 + S + C+ 42(10/90) = S + C - 1/3

**Appendix 3. Numerical example of the computations assumed by the 5-startegy model:**

The current example focuses on the choice in trial 7, of an agent with ki=4, and α = 0.4, after experiencing the following 6 trials (from Study 1):

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Trial | Left  Expert EV | Left  Expert Median | Right  Expert EV | Right  Expert Median | Left  Observed | Right  Observed | Agent’s Choices |
| 1 | 22 | 22 | 17 | 24 | 22 | 24 | Left |
| 2 | 2 | 2 | 2 | 1 | 2 | 1 | Right |
| 3 | -3 | -3 | -7.6 | -22 | -3 | 15 | Right |
| 4 | 14 | 14 | 11.7 | 12 | 14 | 12 | Left |
| 5 | 25 | 25 | 22.6 | 17 | 25 | 17 | Left |
| 6 | 16 | 16 | 20.6 | 1 | 16 | 50 | Right |
| 7 | -5 | 3 | 4 | -46 |  |  | Right |

The probability of a random choice at trial 7 is: P(random choice at t) = 0.97

Let’s assume that the agent makes a deliberate choice at trial 7 (the 0.03 event occurred):

Let’s further assume that the agent recalls trials {2,3,5 and 5} (notice that the sample size is ki=4).

Now, the agent calculates the payoffs based on each of the five rules in her recalled sample:

* EV = [1 + (-3) + 25 + 25]/4 = 12

*Explanation:* In Trial 2 (the first sampled trial), Expert EV is indecisive (Left EV = Right EV = 2). In this case the computation selects randomly one of the sides. Here, the computation used Right and the observed payoff is “1”.

* Median = [2 + (-3) + 25 + 25]/4 = 12.25
* Average = [2 + (-3) + 25 + 25]/4 = 12.25

*Explanation:* The Average rule calculates the average suggestion for each strategy. For example, in Trial 2, the Average Left is while the Average Right is. Thus the computation choose Left (and the obtained payoff from Left is “2”).

* Safe = [2 + (-3) + 25 +25]/4 = 12.25

*Explanation:* Rule Safe choose the strategy with lower difference between the two experts. For example, in Trial 2, the difference between the experts for Left alternative valuations is 0, (2 – 2 = 0) while the difference between Right alternative valuations is 1 (2-1 = 1). Thus, based on Rule Safe, in Trial 2, the chosen alternative is Left (smaller difference), with obtained payoff of 2.

* Risk = [1 + 15 + 17 + 17]/4 = 12.50

In the recalled sample, the highest mean payoff is from Rule Risk (12.5), thus the choice in trial 7 is based on Rule Risk: The chosen alternative is the one with higher (absolute) difference between the expert valuations, which is Right (-5 – 3 vs. 4 – (-46)).

**References**

Abdellaoui M, Klibanoff P, Placido L (2015) Experiments on compound risk in relation to simple risk and to ambiguity. *Management Science* 61(6):1306-1322.

Allais M (1953) Le comportement de l’homme rationnel devant le risque: critique des postulats et axiomes de l’école américaine. *Econometrica: Journal of the Econometric Society* 21(4):503–546. JOUR. https://doi.org/10.2307/1907921

Brier GW (1950) Verification of forecasts expressed in terms of probability. *Monthey Weather Review* 78(1):1–3.

Bolton, G, & Katok, E (2018). Cry wolf or equivocate? Credible forecast guidance in a cost-loss game. *Management Science*, 64(3): 1440-1457. https://doi.org/10.1287/mnsc.2016.2645

Bonaccio, S., & Dalal, R. S. (2006). Advice taking and decision-making: An integrative literature review, and implications for the organizational sciences. *Organizational behavior and human decision processes*, 101(2): 127-151.

Budescu, DV, Por, HH, & Broomell, SB (2012). Effective communication of uncertainty in the IPCC reports. *Climatic change*, 113(2): 181-200.

Cabantous, L, Hilton, D, Kunreuther, H, & Michel-Kerjan, E (2011). Is imprecise knowledge better than conflicting expertise? Evidence from insurers’ decisions in the United States. *Journal of Risk and Uncertainty*, 42(3): 211-232.

de Palma, A, Abdellaoui M, Attanasi G, Ben-Akiva M, Erev I, Fehr-Duda H, Fok D, Hertwig R, Picard N, Wakker PP, Walker JL, Weber M (2014) Beware of black swans: Taking stock of the description–experience gap in decision under uncertainty. *Marketing Letters* 25(3):269-280.

Du, N, Budescu, DВ , Shelly, MK, & Omer, TC (2011). The appeal of vague financial forecasts. *Organizational Behavior and Human Decision Processes*, 114(2): 179-189.

Ecken P, Pibernik R (2015) Hit or miss: what leads experts to take advice for long-term judgments? *Management Science* 62(7): 2002–2021.

Erev, I., & Cohen, B. L. (1990). Verbal versus numerical probabilities: Efficiency, biases, and the preference paradox. *Organizational behavior and human decision processes*, 45(1): 1-18.

Erev I, Ert E, Plonsky O, Cohen D, Cohen O (2017) From anomalies to forecasts: Toward a descriptive model of decisions under risk, under ambiguity, and from experience. *Psychological Review*, 124(4): 369-409.

Erev I, Haruvy E (2016) Learning and the economics of small decisions. *The Handbook of Experimental Economics*, *2*, 9781400883172--011.

Erev I, Roth AE (2014) Maximization, learning, and economic behavior. *Proceedings of the National Academy of Sciences* 111(Supplement 3): 10818–10825.

Fox CR, Tversky A (1998) A belief-based account of decision under uncertainty. *Management science* 44(7): 879-895.

Greenstein S, Zhu F (2012) Is Wikipedia Biased? *American Economic Review* 102(3): 343–348.

Greenstein S, Zhu F (2014) *Do Experts Or Collective Intelligence Write with More Bias?: Evidence from Encyclopædia Britannica and Wikipedia*. Harvard Business School.

Harries C, Yaniv I, Harvey N (2004) Combining advice: The weight of a dissenting opinion in the consensus. *Journal of Behavioral Decision Making* 17(5): 333–348.

Hastie R, Kameda T (2005) The robust beauty of majority rules in group decisions. *Psychological Review* 112(2): 494-508.

Hertwig R, Erev I (2009) The description--experience gap in risky choice. *Trends in Cognitive Sciences* 13(12): 517–523.

Han, Y, & Budescu, D (2019). A universal method for evaluating the quality of aggregators. *Judgment and Decision Making*, 14(4): 395-411.

Kahneman D, Tversky A (1979) Prospect Theory: An Analysis of Decision under Risk. *Econometrica* 47(2): 263–292.

Larrick RP, & Soll, JB (2006) Intuitions about combining opinions: Misappreciation of the averaging principle. *Management Science* 52(1): 111–127.

Lejarraga, T., & Gonzalez, C (2011). Effects of feedback and complexity on repeated decisions from description. *Organizational Behavior and Human Decision Processes*, 116(2): 286-295.

Nevo I, Erev I (2012) On surprise, change, and the effect of recent outcomes. *Frontiers in Psychology*, 3, 24.

Plonsky O, Teodorescu K, Erev I (2015) Reliance on small samples, the wavy recency effect, and similarity-based learning. *Psychological review* 122(4): 621-647.

Radzevick JR, Moore DA (2011) Competing to be certain (but wrong): Social pressure and overprecision in judgment. Management Science 57(1): 93–106.

Roth Y, Wänke M, Erev I (2016) Click or Skip: The Role of Experience in Easy-Click Checking Decisions. *Journal of Consumer Research* 43(4): 583–597.

Taleb, NN (2007) *The black swan: The impact of the highly improbable* (Vol. 2). Random house.

Wakker PP (2010) *Prospect theory: For risk and ambiguity*. Cambridge university press.

Winkler RL (1981) Combining probability distributions from dependent information sources. *Management Science* 27(4): 479–488.

Yaniv, I., & Kleinberger, E (2000). Advice taking in decision making: Egocentric discounting and reputation formation. *Organizational behavior and human decision processes*, 83(2): 260-281.

1. Benjamin and Budescu (2015) examined situations in which the advisers learned the distribution from experience, and the decision makers gained experience with different types of advisers. However, the paper does not examine the impact of experience on the outcomes from the decisions made based on the advice. [↑](#footnote-ref-1)
2. In pure decisions from experience tasks, the decision makers do not receive prior information of the incentive structure and have to rely on their personal experience. Experience was found to increase sensitivity to the winning rate even when the decision makers are presented with a full description of the incentive structure (Erev et al., 2017). Increased sensitivity to the winning rate implies underweighting rare events and a reversal of some of the classical deviations from the prediction of expected utility theory. The best known exceptions to this effect of experience come from studies that focus on one-shot choices in which the possible outcomes are known, and the decision makers have to use their experience to estimate the probabilities (Fox & Tversky, 1998; Abdellaoui, Klibanoff, & Placido, 2015; de Palma et al., 2014). [↑](#footnote-ref-2)
3. We ignore the 5% (4.16%) of trials where the mean (median) advisor provides the same evaluations for the two options. [↑](#footnote-ref-3)
4. Table 4’s problems are five of the 30 problems examined in the “replication” experiment in CPC15. Each of 125 participants faced each of the 30 problems for 25 trials. The final payoff was determined by the sum of the show-up fee and one randomly selected choice. [↑](#footnote-ref-4)