Appalachian State University

Department of Economics Working Paper

Number 23-03| June 2023

Gambling habits and Probability Judgements in a Bayesian Task Environment

David L. Dickinson
Appalachian State University; ESI (Orange, CA); IZA (Bonn, Germany)

Parker Reid
Appalachian State University

Department of Economics
Appalachian State University
Boone, NC 28608
Phone: (828) 262-2148
Fax: (828) 262-6105
www.business.appstate.edu/economics

Gambling habits and Probability Judgements in a Bayesian Task Environment

David L. Dickinson: Department of Economics and CERPA, Appalachian State University: ESI (Orange, CA): IZA (Bonn, Germany)

Parker Reid: Appalachian State University

ABSTRACT

Little is known about how gamblers estimate probabilities from multiple information sources. This paper reports on a preregistered study that administered an incentivized Bayesian choice task to n=465 participants (self-reported gamblers and non-gamblers). The task elicits subjective probability estimates for a particular event given the base rate probability and new evidence information for that event, which allows for an assessment of one's probability assessment accuracy. Furthermore, we also estimate the degree to which both sources of information are weighted in forming subjective probability estimates. Our data failed to support our main hypotheses that experienced online gamblers would be more accurate Bayesian decision-makers compared to non-gamblers, that gamblers experienced in games of skill (e.g., poker) would be more accurate than gamblers experienced only in non-skill games (e.g., slots), or that accuracy would differ in females compared to males. Pairwise comparisons between these types of participants also failed to show any difference in decision weights placed on the two information sources. Exploratory analysis, however, revealed interesting effects related to self-reported gambling frequency. Specifically, more frequent online gamblers had lower Bayesian accuracy than infrequent gamblers. Also, those scoring higher in a cognitive reflection task were more Bayesian in weighting information sources when making belief assessments. While we report no main effect of sex on Bayesian accuracy, exploratory analysis found that the decline in accuracy linked to self-reported gambling frequency was stronger for females. Decision modeling finds a decreased weight place on new evidence (over base rate odds) in those who showed decreased accuracy, which suggests a proper incorporation of new information into one's probability assessments is important for more accurate assessment of probabilities in uncertain environments. Our results link frequency of gambling to worse performance in the critical probability assessment skills that should benefit gambling success (i.e., in skill-based games). Additional research is needed to better understand why a higher frequency of gambling is associated with lower Bayesian accuracy and why this association is greater in females compared to males.

Key Words: Gambling, Bayes Rule, Probability Judgements, Cognitive Reflection

Acknowledgments: The authors thank CERPA and the Barnes Program Student Research Grant for funding this research.

Introduction

Uncertainty can elicit diverse responses from individuals regarding their experience with high-level decision-making and metacognition. An individual's ability to accurately combine prior beliefs with newly acquired information, as suggested by Bayes Rule, can be useful in environments where an objective assessment of the uncertainty is needed. This paper examines the impact of self-reported gambling behavior on decision-making in a Bayesian choice tasks that targets effective probability comprehension skills useful in gambling environments. We followed a pre-registered design, data collection, and analysis plan in our study, and we contributed additional exploratory analysis as well.

The purpose of using an established choice task in our analysis was focused on our desire to not only assess assessment accuracy, but we also aimed to estimate the extent to which participants weighted each source of information available in the task stimulus. Our main objective was to test for differences in probability assessments by sex, whether one was a self-reported gambler or non-gambler, whether one was experienced in skilled gambling games, and whether or not individuals exhibited problem-gambling behaviors. While our preregistered hypotheses were established from previous findings, our data ultimately showed no support for hypothesized differences in assessment accuracy or information source weighting among these groups. Rather, our exploratory analysis revealed a characteristic that robustly predicted a worse ability to accurately assess probabilities in the Bayesian task: self-reported gambling frequency. We conduct analysis that further links predicted decreases in Bayesian accuracy in our sample to a decreased weight placed on new information relative to base rate odds. We later discuss implications of these findings.

Background

Bayesian updating has been extensively studied and it represents a building block decision environment of long-standing interest to psychologists (Phillips and Edwards, 1966) and economists (Grether, 1980), among others. While Bayes rule suggests a precise way to incorporate new information into updating a prior belief, cognitive short-cuts or heuristics may often be employed as an alternative (Kahneman and Tversky, 1973; Tversky and Kahneman, 1973). While some argue that individual differences in intuitive versus deliberative decision styles are not so important in risky choice environments (Steingroever et al. (2018), a larger body of literature connects Bayesian updating to one's ability to engage in more deliberative thinking (e.g., Dickinson and Drummond, 2008; Barash et al., 2019; Dickinson and McElroy, 2019).

There has been some limited attention on how gamblers evaluate probabilities, and these have focused mostly on regular or problem gambler samples (Lim et al., 2015; Cowley et al., 2015). Probability judgments are ubiquitous in the world of gambling and, while some may make money gambling, the average gambler loses money (Stetzka and Winter, 2021). Some evidence suggests that features of certain gambling games may exist to deliberately lead one to a biased judgment of the games' expected payoff (Walker et al., 2023). Another view is that gamblers fall prey to decision biases related to probability assessments (Newstead et al., 1992; Pennycook et al., 2015) in the direction consistent with a decreased reliance on System 2 (deliberative) thinking. One's ability to more accurately update probabilities should pay dividends in the world of gambling, and so the question of whether gambling experience or gambler characteristics can predict probability judgments or one's r approach to probability updating is of interest.

In the literature, regular gamblers who were more impulsive were shown to exhibit diminished use of an optimal (Bayesian-derived) probability estimate, and they also displayed slower learning rates compared to less impulsive gamblers (Lim et al., 2015). This research, however, analyzed only 87 participants from a single community (Oxford, England) and did not consider the differences of the tested subgroups. Such evidence is, however, consistent with the viewpoint that gamblers may use relatively less deliberate thought processes in updating beliefs. Ligneul et al. (2012) compared pathological gamblers and matched healthy controls using a risky choice paradigm that allowed them to estimate probability weighting function. Their finding that pathological gamblers distorted probabilities more than the healthy controls suggests that an increased overweighting of low probabilities and underweighting of high probabilities (e.g., see Kahneman and Tversky, 1984) is associated with pathological gamblers.

Further research showed that problem gamblers exhibit "illusion of control" behaviors, in which problem gamblers evaluate their gambling streaks primarily based on their largest win, rather than their largest loss (Cowley et al., 2015). The participants in this study were placed in a notable "coin-flip" testing environment, popularized by psychologists Kahneman and Tversky (1973), but the study was absent any analysis of the accuracy of participants' Bayesian estimates. There may also be important differences between males and females regarding gambling behaviors—males were observed to take more risks, to partake in riskier games, and they tended to have more problems with gambling than females (Wong et al., 2012).

We contribute to the literature by bringing new data to this question of how gambling experience and behaviors may predict performance in critical information updating environments. For our study, we recruited roughly equal sample of participants from an online platform who reported either experience or no experience with online gambling games, roughly balanced between male and female participants. Self-reported gambling frequency is assessed, as well as a validated short-screener for problem gambling behaviors. Rather than administer a risky choice task to then indirectly evaluate how probabilities are assessed as part of the analysis, we administer an incentivized Bayesian updating task that focused exclusively on one's ability to accurately assess an outcome's probability. The validated task presents participants with both base rate and sample evidence information that vary across a number of trials, which allows us to estimate the extent to which individuals value or "weight" (or distort) base rate probabilities and new information in belief updating.

Hypotheses

Our pre-registered hypotheses were based on the existing research that shows possible gambling differences among subgroups of participants. Some hypotheses focused on the accuracy of probability assessments, while others focused on how one would weight both sources of information in the Bayesian task environment—Bayes rules suggests both sources of information should be relevant in establishing one's subjective probability assessment, though there may be probability weighting biases. For example, Holt and Smith (2009) showed new information was fully weighted in accordance with Bayes rule, but probability weighting suggesting an over-weighting of low probability base rates and under-weighting of high probability base rates.

Past research indicates that males tend to overestimate the perceived odds in a gambling environment, and they exhibit different behaviors than females (Wong, et al., 2012). Thus, we anticipated that there would be a significant difference in Bayesian accuracy and information source weighting by sex in our data. Because past research connects impulsive or problem gambling behaviors with poor performance (Lim et al., 2015; Cowley et al., 2015) in Bayesian environments, we also hypothesized accuracy and information source weight differences between problem and non-problem gamblers. We also considered that experience with skill-based gambling games (e.g., poker or sports betting) would likely imply a better Bayesian decision maker as compared to a gambler who only reported experience with games of chance (e.g., slots or Pachinko). Here, we note that previous research makes a distinction between games of skill versus luck (e.g., Chantal & Vallerand, 1996; Getty et al., 2018), because games of skill involve feedback learning that is essentially a Bayesian updating exercise aimed at more accurately assessing game related probabilities. Below are the full set of our preregistered hypotheses.[1]

Accuracy Hypotheses

Hypothesis 1: Bayesian accuracy will differ by sex

Hypothesis 2: Non-problem gamblers will make more accurate probability assessments than problem gamblers

Hypothesis 3: Non-problem gamblers experienced in skilled gambling games will make more accurate assessments than those experienced only in unskilled games

Information source weight hypotheses

Hypothesis 4a: Participants will respond fully to sample evidence information

Hypothesis 4b: Participants will underweight low and overweight high probability base rates

Hypothesis 5a: Information source weights will differ between those experienced in games of skill versus those experienced only in non-skill games (or non-gamblers)

[1] Note: the numbering or our hypotheses here differs from the preregistration plan for ease of exposition, but otherwise the hypotheses are unchanged.

Hypothesis 5b: Information source weights will differ between those scoring higher on problem gambling behavior versus others (nongamblers or non-problem gamblers)

Hypothesis 6: Problem gamblers will display more sever base rate weighting bias than non-problem gamblers or non-gamblers.

Though we did not preregister a hypothesis regarding frequency of gambling and task performance, there is a basis for our exploratory analysis of the importance of self-reported gambling frequency. Frequent gamblers tend to be overconfident in their abilities to predict odds and this leads them to typically perform poorly when compared to those who do not frequently partake in gambling games (Cowley et al., 2015). And, almost by definition an impulsive or problem gambler will be a more frequent gambler. Thus, we further examine the importance of gambling frequency independent of one's problem-gambler status in the exploratory analysis we conducted.

Methods

Survey and sample screening details

The methods used were preregistered on the Open Science Framework (https://osf.io/zjsg7) to establish hypotheses, sample sizes, variable specifications, and analysis plans. When not describing pre-registered hypotheses or analysis, we will refer to our analysis as exploratory. The basic methodology was to embed a decision task within an online survey that would be administered to participants who self-reported gambling games experience and also to a sample of participants who indicated they did not play gambling games. All methods for data collection were carried out in accordance with the US Federal Policy for the Protection of

Human Subjects, and our procedures were approved by the human subjects review board at the author's academic institution.

Our sample was recruited from the Prolific subject pool (prolific.ac), which is a service tailored for researchers as an alternative to Amazon's mTurk platform for online research studies (Palan & Schitter, 2018; Peer et al., 2022). One of the benefits of Prolific is the availability of a variety of sample screening options that allow the researcher to recruit custom samples based on one or more criteria captured by Prolific in each participant's profile. Our inclusion criteria were: young adults located in the U.S. and the U.K. who were between 21 and 48 years of age who were registered to take part in research studies on the Prolific platform; those who had in a response in their profiles regarding self-reporting experience with one or more (or none) of the games from of a list of popular online gambling games—we recruited half our sample from among those reporting experience with one or more of these games, and half the sample from among those reporting they did not have experience with any of these games. We limited our study to participants between the ages of 21 and 48, as the age of 21 is the legal minimum age to gamble in most states in the United States, and research shows that cognitive decline is already evident in middle age (45-49 years) (Singh-Manoux et al., 2012). Thus, our young adult sample was chosen to eliminate any potential confound between age-related cognitive decline and task performance in our Bayesian task, which would classify as an executive function task that depends on deliberative-thinking. The Prolific recruitment platform integrates seamlessly with popular survey software platforms to allow one to conduct online studies.

Our planned sample size was partly based on available funds, but we also conducted an a priori power analysis using G*Power 3.1.9.4. Here, we found that a planned sample of n=400 would have sufficient power (power = .80 for behavioral research) to detect a small effect size (f^2 = .02) for a single regression coefficient in a multiple regression with up to 6 co-variates (e.g., age, sex, gambling experience), assuming an α = .05 error probability. A medium-small effect size (f^2 = .065) is detectible with a sample size of n=100, which means we may also conduct analysis of decision model estimates on separate subsamples (e.g., females with gambling experience) with reasonable statistical power to identify key decision weight effects, though these may also be examined in a pooled data model with interaction terms.

The Bayesian decision task

Our incentivized decision task is a modification of the Grether's design (Grether, 1980) that has been adopted by others in recent literature (Dickinson and Garbuio, 2021).[2] For the decision task, there are two boxes each populated with three balls. As shown in Figure 1, the LEFT box has two black and one white ball. Either the LEFT or RIGHT box will be selected in a trial. The participant is not told which box is selected for the current trail, but she is presented with two sources of information with which to form beliefs regarding which box was selected: the base rate or "prior odds" of either box being selected, and the results from drawing eight balls with replacement from the chosen (but hidden) box. The prior odds were represented as the chances out of ten that either box would be selected, ex ante, and this can be considered the initial information for that stimulus (trial). The results of the eight-ball sample draw can be considered the new evidence presented to the participant for that stimulus. As shown in Figure 1, the stimulus image offered a visually concise way to present the information to the participant, and the task varies the information on one or both

[2] See also Phillips and Edwards (1966) for an earlier version of a similar task referred to as the "Beads Task".

dimensions across a series of twenty trials. In the original Grether (1980) task, the response elicitation was dichotomous in the sense that, for a given set of prior odds and evidence, the participant was asked to indicate which box was thought to have been selected for that trial. Bayes rule can be used to calculate the actual posterior probability that the LEFT box was used, given the prior odds and the new sample evidence.

The task we administered differed slightly from the original task in that we elicited the participant's subjective assessment of how likely it was that the LEFT box had been selected in that trial (i.e., the “chances out of 100” that the LEFT box was used), which we call “*Left Assess*” $\in [0, 100]$. The elicitation of a precise subjective probability estimate is more in line with Grether (1992) and provide more rich data, assuming participants are incentivized to provide truthful subjective probability estimates. To this end, we followed Holt and Smith (2000) and used a Becker-DeGroot-Marshak type cross-over scoring procedure that makes it in a participant’s best interest each trial to response with her true subjective probability estimate (see Experiment Instructions in Appendix B). Because the incentivization procedure is somewhat complicated, participants are reminded at the end of the instructions that they would maximize their expected bonus payment in each trial “….by responding with your true belief of how likely you think the LEFT box was selected, given the available information!”

Table 1 shows the specific combinations of prior odds and evidence we used across the 20 total trials administered to each participant—these are highlighted cells in Table 1. Though the Bayesian posterior probabilities varied across the 20 trials used, we only selected trials where the odds and evidence favored opposite boxes. For example, in one trial the prior odds of the LEFT box were 1/10 (indicating the RIGHT box is more likely to be used) but the number of black balls drawn in the sample evidence was 7 out of 8 (i.e., a draw more likely if the LEFT box rather than the RIGHT box is used). Thus, the odds and evidence point to opposite more likely boxes and make the probability assessment task more challenging than if odds and evidence were aligned. Each subject saw the same set of stimuli, but the survey software presented the stimuli in randomized order to each participant.

Dependent Variables

The key outcome variable explored depended on the analysis of accuracy or information source decision modeling. Accuracy can be assessed both at the participant-level by averaging one's accuracy across all 20 trials, but it can also be analyzed at the trial-level for the panel data set of 20 observations per participant. The decision model analysis required use of the panel data set to examine one's subjective probability as a function of the specific odds and evidence characteristics of that trial.

Bayesian *Accuracy* of the participant at the trial level, which was also used to construct a participant-level overall *Accuracy Score*. The variable *Accuracy* was calculated using the participant's subjective estimate of the likelihood that the Left Box was used, *Left Assess* $\in [0,100]$. For each participant's trial, *Accuracy* is defined (as in Dickinson and Garbuio, 2021) by the absolute difference between *Left Assess* and the True Bayes Probability, $\in [0,1]$, given that trial's base rate odds and evidence:

$$Accuracy = 1-|(Left\ Assess/100) - \text{True Bayesian Probability}| \in [0,1]$$

After assessing *Accuracy* for each trial, the *Average* Accuracy variable was constructed to average *Accuracy* across all twenty trials. *Average Accuracy* was the key dependent variable in our participant-level analysis, and *Accuracy* was used in the trial-level analysis. Hypotheses 1-3 are examined below using both participant-level and then trial-level data.

For the decision model, the key dependent variable is the log-odds ratio of the LEFT versus RIGHT box being used, $ln\left(\frac{Belief_{LEFT}}{1-Belief_{LEFT}}\right)$. We follow the approach in Holt and Smith (2000) and consider the one-parameter specification in Wu and Gonzalez (1996) where one holds a subjective belief regarding the actual event probability under consideration, *p*. This subject belief or weighted probability, $w(p)$, is defined as follows:

$$w(p) = \frac{p^{\gamma}}{(p^{\gamma}+(1-p)^{\gamma})^{1-\gamma}} \tag{1}$$

Here, the weighting parameter, γ, equates subjective and objective (Bayesian) probabilities when $\gamma = 1$, but with $\gamma < 1$ the individual over-weights low and under-weights high probabilities. The odds ratio for one's subjective probability estimate is therefore:

$$\frac{w(p)}{1-w(p)} = \frac{p^{\gamma}}{(p^{\gamma}+(1-p)^{\gamma})^{1-\gamma}} \cdot \frac{((1-p)^{\gamma}-((1-(1-p))^{\gamma})^{1-\gamma}}{(1-p)^{\gamma}} = \left(\frac{p}{(1-p)}\right)^{\gamma} \tag{2}$$

The typical odds ratio form of Bayes Rule then writes this subjective (posterior) odds ratio of the LEFT box being used as a function of the base rate odds ratio of LEFT ($Prob_{LEFT}$) and the likelihood ratio of LEFT (i.e., the likelihood of a particular sample of evidence or new information, *S*, given one state of the world (LEFT box) versus the other (RIGHT box)). Thus, the odds ratio form of Bayes Rule in the context of our task is written as:

$$\frac{Belief_{LEFT}}{1-Belief_{LEFT}} = \left(\frac{Prob_{LEFT}}{(1-Prob_{LEFT})}\right)^{\gamma} \cdot \left(\frac{Prob(S|LEFT)}{Prob(S|RIGHT)}\right)^{\gamma} \tag{3}$$

Taking logs and generalizing such that the γ weight may differ for base rate odds versus evidence leads to the baseline specification we estimate:

$$ln\left(\frac{Belief_{LEFT}}{1-Belief_{LEFT}}\right) = \alpha + \gamma_1 ln\left(\frac{Prob_{LEFT}}{(1-Prob_{LEFT})}\right) \cdot \gamma_2 ln\left(\frac{Prob(S|LEFT)}{Prob(S|RIGHT)}\right) \quad (4)$$

In other words, the subjective odds ratio favoring the LEFT box is a function of the base rate odds of the LEFT box and the evidence that favors the LEFT box.

Independent Variables

Regarding independent variables, covariates include the participant's sex (*Female* = 0 or 1), age (in years), and country of residence (*USA* = 0 or 1). *Problem Gambling* was treated as a binary measure where we scored *Problem Gambling* = 1 to participants who answered "Yes" to any of the three questions prompted by "The NODS-CLiP* Short Problem Gambling Screen" during the survey section of the online assessment (Volberg et al., 2011).

Data was also collected on one's specific gambling experience and self-reported frequency of gambling. For specific gambling experience, participants identified which, if any, online gambling games they had participated in from a list of fourteen options: Baccarat, Black Jack, Bingo, Craps, Lottery, Pachinko, Poker, Race & Sports Book, Roulette, Slots, Video Poker, and Virtual Sports Betting. Other response options included: "Not applicable/rather not say", and "None of the ones listed above, but I do gamble." We scored *Skill Gambler* = 1 to those who participated in games of skill, such as Blackjack, Poker, Video Poker, Race & Sports Book, Virtual Sports Betting (Chantal & Vallerand, 1996; Getty et al., 2018). Gamblers who selected games other than the identified games of skill were scored *Skill Gambler* = 0. A follow-up question to those who indicated they had participated in online gambling asked their current frequency of gambling along a five-option Likert scale: "Never", "Less than once a month", One or twice a month", "Once or twice a week", "Daily". Those who were non-gamblers, and therefore they did not receive the gambling frequency question, were automatically scored as having *Gambling Frequency* = 0. We also note that many in our sample of those with self-reported gambling experience recorded a *current* gambling frequency response of "Never"). Thus, we will also distinguish in our analysis between those in our gamblers sample versus those who reported current gambling with some frequency.[3]

In addition to these covariates, we also administered a 6-item Cognitive Reflection Task (CRT) on each participant (Primi et al., 2016), which produced a 0-6 CRT score. Higher values indicate a tendency to be more reflective in one's thinking as opposed to more automatic in one's responses. A measure of self-reported sleepiness and the prior week's average sleep levels were included given previous research highlighting the link between sleepiness and Bayesian decision making (e.g., Dickinson and Drummond, 2008; Dickinson et al., 2016).

Results

The overall participant pool in Prolific was reported as n=465 individuals (n=220 with self-reported online gambling experience, n=245 without self-reported gambling experience). If we consider as *Gamblers* only those recruited from our self-reported online gambling experience sample who also reported a current gambling frequency greater than zero, then 14 participants prescreened as experienced in gambling would be considered non-gamblers for

[3] This distinction arose because the Prolific profile questions are asked at the time one registered on the system. Also, the Prolific screener question is more time non-specific for the screener by asking about the types of online gambling games they have played, whereas the *Gambling Frequency* question specifically asked how frequent is the individual's *current* gambling. Thus, our *Gambling Frequency* variable will better distinguish current gamblers from non-gamblers in our data.

practical purposes. By screening individuals who fall within our desired age range, those eligible with the desired gambling behaviors were: *Non-problem gamblers* (n=375); *Problem gamblers* (n=90); *Nongamblers* (n=259); *Gamblers* (n=206); *Women* (n=230); *Men* (n=235). Additionally, participants resided in either the *United States* (n=135) or the United Kingdom (n=330). Table 2 shows the summary statistics on key individual-specific control measures that will be used as independent variables in our analysis. From the summary statistics, on average, we can see that problem gamblers tended to be mostly male, reported gambling more frequently on average, and were roughly 32 years of age.

Hypotheses 1-3:

The first set of hypotheses focus on the accuracy of one's probability assessments, relative to Bayes rule. Table 3 shows first evidence of the *lack* of support for any of hypotheses 1-3. T-tests on the relevant pairwise subsamples highlight the lack of difference in *Average*, which summarized each participant's accuracy across the 20-trials. Recall also that tests of Hypotheses 2 and 3 require considering only the subsample of data on non-problem gamblers (Hypothesis 2) or the subsample of gamblers (Hypothesis 3), whereas the entire sample of gamblers and non-gamblers is used to evaluate Hypothesis 1. The bottom of Table 3 also shows some initial evidence from *Average Accuracy* linking more frequent gambling with reduced Bayesian accuracy. Table 4 shows estimation results from regressions of *Average Accuracy* on participant characteristics. Though Table 4 does not show a proper test of our Hypotheses 2 and 3 due to use of the entire sample across columns (1)-(3), the Table highlights that while one's identification as a *Skill Gambler* appears to *negatively* predict Bayesian accuracy, the result is spurious and due to a high correlation (rho = .710) with self-reported *Gambling Frequency*. We discuss the importance of *Gambling Frequency* later.

Tables 5-7 properly test Hypotheses 1-3 using the panel nature of the data by regressing *Accuracy*, at the trial-level, on the key indicator variables and participant characteristics. Across models (1), (2), and (3) we successively add additional control variables. Standard errors are clustered at the participant level to account for multiple observations from a given participant, and the sample size differences across Tables 5-7 reflect the need to use the full sample to test Hypothesis 1, but subsets of the data to test Hypotheses 2 and 3. We focus on the following indicator variables to test our hypotheses: *Female* in Table 5 (testing Hypothesis 1), *Skill Gambler* in Table 6 (testing Hypothesis 2), and *Problem Gambler* in Table 7 (testing Hypothesis 3). Coefficient estimates on the key indicator variables are all statistically insignificantly different from zero across all specifications, which supports rejecting Hypotheses 1-3. In fact, we find robust support in Tables 5-7 that only two variables predict one's *Accuracy* in this task: *Gambling Frequency* predicts lower *Accuracy*, while *CRT Score* predicts higher *Accuracy.*

Hypotheses 4-6:

We next turn our attention to an examination of *how* one approaches forming their subjective probability assessment. For these tests, panel estimations were performed on the trial-level data, with standard errors clustered at the participant level. Hypotheses 4a and 4b are a test of whether, in the baseline estimation specification shown in equation (4) above, $\gamma_2=1$ (Hypothesis 4a) and whether $\gamma_1<1$ (Hypothesis 4b). Table 8, column (1), shows the results for the baseline specification, while model results in columns (2) and (3) add additional control variables of interest. Across all specifications, the data reject Hypothesis 1a in favor of more *conservative* Bayesian updating (e.g., Phillips and Edwards, 1966). The data support Hypothesis 1b in the we always reject the null hypothesis test that $\gamma_1=1$ in Table 8.

Table 9 shows results of the test of Hypothesis 5a that *Skill Gamblers* will weight information sources differently than others. The preregistered hypothesis was phrased such that non-gamblers would also be compared, so Table 9 is estimated on the full panel data set. For this Hypothesis 5a test, two variables are added that interact the *Skill Gambler* indicator variable with the $ln(PriorOdds\ ratio)_{left}$ and with $ln(Likelihood\ ratio)_{left}$. The estimation results indicate that someone who self-reported experience with skill-based games places *less* weight on the sample evidence compared to one who did not report experience with skill-game gambling (or was a non-gambler). This would support Hypothesis 5a but, as noted when discussing Table 4 above, we present results in the exploratory analysis to highlight that this result is likely due to the fact that more frequent gamblers were more likely to report having played skill-based gambling games. In particular, the distinction between nongamblers and those who self-reported gambling experience was such that just having experience with online gambling meant one was quite likely to have played a skill-based game.[4] Appendix Table A2 highlights that re-estimation of the models in Table 9 to include interactions terms between *Gambling Frequency* and each of the two information sources leads to statistically insignificant coefficient estimates on the *Skill Gambler* interaction term with $ln(Likelihood\ ratio)_{left}$. In its place, the interaction between *Gambling Frequency* * $ln(Likelihood\ ratio)_{left}$ is statistically significant and negative. Exploratory analysis below will further examine the importance of *Gambling Frequency* in our data.[5]

[4] Our survey did not assess the proportion of one's gambling that involved games of chance versus skill-based games. An alternative coding of *Skill Gambler* was explored where one was considered a *Skill Gambler* if they *only* reported playing skill-based games, as opposed to some mix of skill-based games and games of chance. Unfortunately, under that alternative coding, we have only n=20 such exclusive skill-based gamblers (and only 12 of those are non-problem gambles based on the NODS-CLiP* gambling screener administered).

[5] Earlier work with this Bayesian updating task did not elicit probability estimates but participants were asked to merely indicate which box they deemed more likely, given the base rate and evidence information. This approach lends itself to non-linear probit estimation of a binary variable indicating one's subjective view of the more likely box, as a function of the odds and evidence. Here, we rescored the subjective probability estimate

Table 10 results show tests of Hypotheses 2b and 3, which focused on differences in information source weighting (2b) and base rate probability weighting (3) in the subset of *Problem Gamblers*. Interaction terms were added to the baseline specification to perform the statistical tests of these hypotheses, and it is apparent across all models (1)-(3) of Table 10 that *Problem Gamblers* weighted the information sources no differently than non-problem gamblers (or non-gamblers, as was considered in the preregistered hypothesis). Therefore, overall we find little support for our preregistered hypotheses, other at this stage we note only that our data are consistent with conservative Bayesian updating in all participant types, and these participants also show evidence of probability weighting that would over-weight low and under-weight high base rates in forming probability estimates.

Exploratory Analysis

We also report exploratory findings of hypotheses that were not preregistered, but were nevertheless of interesting in our attempts to understand our mostly null findings. In conducting our preregistered hypotheses tests on *Accuracy*, it became apparent that two characteristics of a participant in our study robustly predicted Bayesian accuracy in our incentivized task: *CRT score* (a measure of more reflective versus automatic thinking) and self-reported *Gambling Frequency*. As such, we pursued additional exploratory analysis of our Bayesian decision model specification to examine whether there was evidence of

---

data to generate the binary indicator *Left Box Likely* = 1 if one's subjective probability estimate of the trial using the LEFT box was greater than 50 (chances out of 100). Non-linear Probit models were then estimated to compare to the model estimates derived from the subjective log-odds estimates. These Probit estimations results are in Appendix A Table A3 and would compare to Appendix A Table A4. While the Probit estimations fail to fully utilize the available information in elicited responses, they nevertheless show consistency of findings. Specifically, both the odds and the evidence predict an increased probability of considering the *LEFT* box as more likely, and more frequent gamblers place a marginally lower weight on the new sample evidence (though with reduced statistical significance, $p < .10$)

differential decision weights on information sources by either or both of these two characteristics.

Table 11 shows results of this exploratory analysis, where the baseline model is modified to include interactions of the odds and evidence variables with *CRT Score* and *Gambling Frequency*. These models are estimated without the control variables that have largely been insignificant predictors (results are similar with their inclusion and are available on request). Here, models (1)-(3) differ by whether we estimate the model on the full sample or on the subsample of male or female participants. Results in Table 11 again show that participants are conservative Bayesian decision makers who engage in probability weighting as a baseline. Those with higher *CRT Scores*, which would indicate more reflective thinkers, place marginally higher weight on the sample evidence compared to those with lower *CRT Scores*, and this effect is robust in both male and female subsample estimations. *Gambling Frequency* predicts a marginally lower weight placed on sample evidence, and this result is driven by the subsample of female participants.

These Table 11 findings, in conjunction with Tables 4-7 results, suggest a mechanism connecting Bayesian accuracy to weighting the evidence more fully. That is, *CRT Score* predicts more accurate Bayesian choices and is also linked to increased weight placed on sample evidence.[6] And, *Gambling Frequency* is found to reduce Bayesian accuracy but is also linked to a reduced weight on sample evidence in the data. This result may also differ between male and female participants. A final exploratory analysis estimated the *Accuracy* model (3) from Table 3 and included interaction terms for *Female * Gambling Frequency* and *Female * CRT Score*. The results are summarized in Figure 2 and 3 and show that the

[6] See also Oechssler et al. (2009) for evidence that higher CRT improves one's accuracy in probability updating.

decline in *Accuracy* among more frequent gamblers is marginal more severe among female participants—full results behind these figures are in the Appendix Table A4. This is consistent again with the Table 11 finding that the marginally lower weight placed on sample evidence by female participants corresponds to a greater decay in *Accuracy* for more frequent female gamblers. Figure 3, highlights that the Table 11 result showing that more reflective thinkers, male or female, place marginally more weight on sample evidence is consistent with a significant increase in *Accuracy* for higher *CRT Score* participants regardless of sex.

Discussion

We set out to test a set of preregistered hypotheses that were sensibly derived from previous findings, but the data mostly failed to support those hypotheses. Rather, the exploratory analysis has directed out attention to where future research may focus efforts on the link between attention to new information, Bayesian accuracy and male-female differences in probability assessments. The most robust and consistent finding we can report from the exploratory analysis is that a more reflective style of thinking tends to pay additional attention to new sample evidence in decision making, which supports increased accuracy in making probability assessments. A second exploratory finding of note was that more frequent gamblers did surprisingly worse in the probability assessments. This result was significant only in the subset of female participants and can be linked to a decreased weight placed on new sample evidence in the estimated decision modeling.

It is worth noting that *Gambling Frequency* in our study is self-reported, and it refers specifically to *current* gambling habits. In contrast, the custom-screening of participants on Prolific was accomplished by using self-reported experience with one or more online gambling games without reference to recency of play. Indeed, some participants recruited to the sample of gamblers reported that they did not currently gamble when asked about current *Gambling Frequency*. Our intended exploration of those experienced in skill-based games was also complicated by these same data from Prolific screener questions. Namely, these data may not be current in terms of gambling habits, and many individuals reported experience with several of the listed games that included both skill-based games and games of chance, which limited the ability to identify gamblers who were more specialized in one type of game or the other in their habits.

Another limitation of the study is the cross-sectional nature of the gambling characteristic data. In other words, it is difficult to establish a causal relationship between, for example, gambling frequency and probability judgment accuracy given that gambling frequency only varies across participants in our data set. While we interpret our findings to suggest that the frequency of one's gambling impacts their judgment accuracy, we cannot say whether causation runs the other direction, or whether another unmeasured variable affects both. It is possible that those who poorly update probabilities, and are less accurate in probability judgements, do so in ways the bias one's perception of successful outcomes. This bias could then lead one to gamble more frequently, such that it is the approach to probability judgments that cause one's frequency of gambling, as opposed to vice-versa.

Notwithstanding the limits of our data, the exploratory findings reported point to an interesting association between more frequent gamblers and one's approach to probability assessments. While all participants over-weighted low and under-weighted high base rate probabilities (as others have found in this type of task. E.g., Holt and Smith, 2009), and they conservatively incorporate new information into updating beliefs (as others have found. E.g., Phillips and Edwards, 1966; Hill, 2017), more frequent gamblers were even more conservative in their incorporation of new information into their updated beliefs. This finding is noteworthy because we deliberately abstracted away from a risky choice task frequently encountered by gamblers or used in studies of gamblers to focus on a building block decision task that is of importance not only in gambling success, but also in the general domain of decision making under uncertainty.

Our results may be interpreted in light of others' work on illusion of control among gamblers (Cowley et al., 2015). While our results cannot establish causation, they are consistent with an illusion of control effect. Less accurate probability assessments do not improve one's chance of gambling success, and so the fact that those least accurate in our Bayesian probability assessment task are those who gamble more frequently could point to an illusion of control at work in their gambling habit. Our task did not provide feedback on one's accuracy across trials, and so our data show a snapshot view of how an individual approached the Bayesian inference task. In a gambling environment where feedback on success may stimulate learning, individuals may correct for faulty probability assessment efforts. Our data highlight that these more frequent gamblers may be less apt to learn from new information. We should note, however, that this speculation ignores the fact that confirmatory new information may be treated differentially compared to disconfirming information. A more complete decision environment that embeds probability judgments in a task where accuracy of those judgments also implies additional benefits (i.e., increased chance of gambling success) would help us more fully understand the implications suggested by our findings.

Conclusion

This paper reported results from a pre-registered study of self-reported gambling patterns and decision making in an online incentivized Bayesian decision task environment. Such as previous research suggests, participants weighted all available information sources in their probability assessments (Grether, 1980) However, contrary to our ex ante hypotheses, we reported no significant differences in Bayesian accuracy between male and female participants, between problem and non-problem gamblers, nor between those with experience skill-based gambling games. Consistent with this, we reported no differences in the same pairwise group comparisons in their approach to weighting base rate versus new information sources in updating probability assessments.

For our exploratory analysis, we found that those self-reporting more gambling habits were less accurate in their Bayesian updating accuracy, and those with higher scores on a cognitive reflection task were more accurate in their Bayesian accuracy. The link between frequent gambling and reduced Bayesian accuracy was significant only among females, while the link between *CRT score* and increased accuracy was true for both male and female participants (if not a bit larger in magnitude in male participants). Corresponding findings from models estimating the weights placed on base rate versus sample evidence were consistent with the hypothesis that additional weight on new information is critical for more accurate probability assessments. We leave it to future research to more systematically examine these intriguing exploratory findings.

REFERENCES

Barash, J., Brocas, I., Carrillo, J. D., & Kodaverdian, N. (2019). Heuristic to Bayesian: The evolution of reasoning from childhood to adulthood. *Journal of Economic Behavior & Organization*, *159*, 305-322.

Bar-Hillel, M. (1980). The base-rate fallacy in probability judgments. *Acta Psychologica*, *44*(3), 211-233.

Chantal, Y., & Vallerand, R. J. (1996). Skill versus luck: A motivational analysis of gambling involvement. *Journal of Gambling Studies*, *12*(4), 407-418.

Cowley, E., Briley, D. A., & Farrell, C. (2015). How do gamblers maintain an illusion of control? *Journal of Business Research*, *68*(10), 2181-2188.

Dickinson, D. L., & Drummond, S. P. (2008). The effects of total sleep deprivation on Bayesian updating. *Judgment and Decision Making*, *3*(2), 181-190.

Dickinson, D. L., Drummond, S. P., & Dyche, J. (2016). Voluntary sleep choice and its effects on Bayesian decisions. *Behavioral Sleep Medicine*, *14*(5), 501-513.

Dickinson, D. L., & Garbuio, C. (2021). The influence of dietary patterns on outcomes in a Bayesian choice task. *Health Science Reports*, *4*(3), e369.

Dickinson, D. L., & McElroy, T. (2019). Bayesian versus heuristic-based choice under sleep restriction and suboptimal times of day. *Games and Economic Behavior*, *115*, 48-59.

Hill, S. J. (2017). Learning together slowly: Bayesian learning about political facts. *The Journal of Politics*, *79*(4), 1403-1418.

Holt, C. A., & Smith, A. M. (2009). An update on Bayesian updating. *Journal of Economic Behavior & Organization*, *69*(2), 125-134.

Getty, D., Li, H., Yano, M., Gao, C., & Hosoi, A. E. (2018). Luck and the law: Quantifying chance in fantasy sports and other contests. *Siam Review*, *60*(4), 869-887.

Goodwin, P., Önkal, D., & Stekler, H. O. (2018). What if you are not Bayesian? The consequences for decisions involving risk. *European Journal of Operational Research*, *266*(1), 238-246.

Grether, D. M. (1980). Bayes rule as a descriptive model: The representativeness heuristic. *The Quarterly Journal of Economics*, *95*(3), 537-557.

Grether, D. M. (1992). Testing Bayes rule and the representativeness heuristic: Some experimental evidence. *Journal of Economic Behavior & Organization*, *17*(1), 31-57.

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, *39*(4), 341.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, *80*(4), 237.

Ligneul, R., Sescousse, G., Barbalat, G., Domenech, P., & Dreher, J. C. (2013). Shifted risk preferences in pathological gambling. *Psychological Medicine*, *43*(5), 1059-1068.

Lim, M. S., Jocham, G., Hunt, L. T., Behrens, T. E., & Rogers, R. D. (2015). Impulsivity and predictive control are associated with suboptimal action-selection and action-value learning in regular gamblers. *International Gambling Studies*, *15*(3), 489-505.

Newstead, S. E., Pollard, P., Evans, J. S. B., & Allen, J. L. (1992). The source of belief bias effects in syllogistic reasoning. *Cognition*, *45*(3), 257-284.

Oechssler, J., Roider, A., & Schmitz, P. W. (2009). Cognitive abilities and behavioral biases. *Journal of Economic Behavior & Organization*, *72*(1), 147-152.

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

Paliwal, S., Petzschner, F. H., Schmitz, A. K., Tittgemeyer, M., & Stephan, K. E. (2014). A model-based analysis of impulsivity using a slot-machine gambling paradigm. *Frontiers in Human Neuroscience*, *8*, 428.

Peer, E., Rothschild, D., Gordon, A., Evernden, Z., & Damer, E. (2022). Data quality of platforms and panels for online behavioral research. *Behavior Research Methods*, *54*(4), 1643-1662.

Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2015). What makes us think? A three-stage dual-process model of analytic engagement. *Cognitive Psychology, 80*, 34-72.

Phillips, L. D., & Edwards, W. (1966). Conservatism in a simple probability inference task. *Journal of Experimental Psychology*, *72*(3), 346.

Primi, C., Morsanyi, K., Chiesi, F., Donati, M. A., & Hamilton, J. (2016). The development and testing of a new version of the cognitive reflection test applying item response theory (IRT). *Journal of Behavioral Decision Making*, *29*(5), 453-469.

Singh-Manoux, A., Kivimaki, M., Glymour, M. M., Elbaz, A., Berr, C., Ebmeier, K. P., ... & Dugravot, A. (2012). Timing of onset of cognitive decline: results from Whitehall II prospective cohort study. *BMJ 344*.

Steingroever, H., Pachur, T., Šmíra, M., & Lee, M. D. (2018). Bayesian techniques for analyzing group differences in the Iowa Gambling Task: A case study of intuitive and deliberate decision-makers. *Psychonomic Bulletin & Review*, *25*(3), 951-970.

Stetzka, R. M., & Winter, S. (2021). How rational is gambling? *Journal of Economic Surveys*: 1-57.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*(2), 207-232.

Volberg, R. A., Munck, I. M., & Petry, N. M. (2011). A quick and simple screening method for pathological and problem gamblers in addiction programs and practices. The American Journal on Addictions, 20, 220–227.

Walker, A. C., Stange, M., Dixon, M. J., Fugelsang, J. A., & Koehler, D. J. (2022). Push outcomes bias perceptions of scratch card games. *Journal of Gambling Studies*, 1-25.

Wong, G., Zane, N., Saw, A., & Chan, A. K. K. (2013). Examining gender differences for gambling engagement and gambling problems among emerging adults. *Journal of Gambling Studies*, *29*, 171-189.

TABLE 1: Bayesian probabilities by odd-evidence combination (highlighted cells show those combinations administered to participants in the study)

	Evidence in Favor of LEFT (# black balls of 8 draws with replacement)								
Prior Odds of LEFT	0	1	2	3	4	5	6	7	8
0	0	0	0	0	0	0	0	0	0
0.1	0.0004	0.0017	0.0069	0.0269	0.1	0.3083	0.6414	0.8777	0.9664
0.2	0.0010	0.0039	0.0153	0.0587	0.2	0.5008	0.8010	0.9417	0.9848
0.3	0.0017	0.0066	0.0259	0.0965	0.3	0.6323	0.8734	0.9651	0.9911
0.4	0.0026	0.0102	0.0398	0.1425	0.4	0.7279	0.9148	0.9773	0.9942
0.5	0.0038	0.0152	0.0585	0.1995	0.5	0.8005	0.9415	0.9848	0.9962
0.6	0.0058	0.0227	0.0852	0.2721	0.6	0.8575	0.9602	0.9898	0.9974
0.7	0.0089	0.0349	0.1266	0.3677	0.7	0.9035	0.9741	0.9934	0.9983
0.8	0.0152	0.0583	0.1990	0.4992	0.8	0.9413	0.9847	0.9961	0.9990
0.9	0.0336	0.1223	0.3586	0.6917	0.9	0.9731	0.9931	0.9983	0.9996
1	1	1	1	1	1	1	1	1	1

Notes: Bayes probabilities of the LEFT box being used were calculated using Bayes rule:

$$P(Left|X) = \frac{P(X|Left)P(Left)}{P(X|Left)P(Left) + P(X|Right)P(Right)}$$

TABLE 2: Summary statistics by gambling group

Variable	Non-problem Gambler	Problem Gambler	Unskilled Gambler	Skilled Gambler
Age	Mean = 32.285 SD = 7.709	Mean = 32.022 SD = 7.929	Mean = 31.821 SD = 7.790	Mean = 32.925 SD = 7.639
Female = 1	Number (proportion) 195 (52%)	Number (proportion) 35 (39%)	Number (proportion) 154 (53%)	Number (proportion) 76 (44%)
U.K. = 1 (vs. USA)	Number (proportion) 265 (71%)	Number (proportion) 65 (72%)	Number (proportion) 221 (76%)	Number (proportion) 109 (63%)
Gambling Frequency*	Mean = 0.576 SD = 0.942	Mean = 1.800 SD = 1.104	Mean = 0.216 SD = 0.597	Mean = 1.810 SD = 0.988
CRT score	Mean = 3.325 SD = 2.121	Mean = 3.356 SD = 1.97	Mean = 3.289 SD = 2.180	Mean = 3.379 SD = 1.937
Total participants	375	90	291	174

Notes: Gambling Frequency was self-reported by participants, on a scale of 0-4: 0 (never gamble); 1 (less than once a month); 2 (once or twice a month); 3 (once or twice a week); 4 (daily). For example, a score of 1.800 would indicate that identified Problem gamblers, on average, report current gambling between less than once a month and once or twice a month.

TABLE 3: Mean *Average Accuracy* and hypothesis tests

	<u>**H1 test**</u>		<u>**H2 test**</u>		<u>**H3 test**</u>	
Test Variable: ***Average Accuracy***	Males (n=235)	Females (n=230)	Skill-Game (non-problem) Gamblers (n=99)	NonSkill-Game (non-problem) Gamblers (n=30)	Problem Gamblers (n-77)	Non-Problem Gamblers (n=129)
mean	.727	.709	.689	.690	.693	.689
t-stat (p-value)	1.2695 (p=.205)		.021 (p=.984)		-.160 (p=.873)	
Exploratory--comparison of *Average Accuracy* by current gambling frequency						
		Gfreq=0 (n=259)	*Gfreq*=1 (n=93)	*Gfreq*=2 (n=62)	*Gfreq*=3 (n=43)	*Gfreq*=4 (n=8)
	Mean	.740	.716	.674	.670	.634
	St. dev	.144	.158	.163	.166	(.211)

Notes: *SkillGame* gambler t-test used the subset of participants who reported current gambling frequency greater than zero. A similar test did not reveal any statistically significant difference in *Accuracy* between nonproblem gamblers who *exclusively* played skill-games compared to those who played a mix of skill games and games of chance, but our data set only includes n=12 current gambler participants who reported exclusively playing gambling games of skill. *Gfreq* $\in$ [0,4] describes self-reported *current* gambling frequency (0,1,2,3,4 indicates responses of "never", "less than once a month", "once or twice a month", "once or twice a week", or "daily". Nongambler participants were not shown this gambling frequency question and were scored as *Gfreq* = 0.

TABLE 4: *Average Accuracy* by sex, *Skill Gambler, Problem Gambler*

Dependent Variable: ***Average Accuracy***	(1)	(2)	(3)
Variable	Coef (st. error)	Coef (st. error)	Coef (st. error)
Constant	.747 (.012)**	.778 (.034)**	.676 (.037)**
Female (=1)	-.022 (.014)	-.025 (.015)	-.004 (.015)
Skill Gambler (=1)	-.046 (.016)**	-.045 (.016)**	-.007 (.021)
Problem Gambler (=1)	-.002 (.019)	-.003 (.020)	.010 (.020)
Age	---	-.001 (.001)	-.0003 (.001)
USA (=1)	---	-.0001 (.017)	.003 (.016)
Average Response Time	---	---	.002 (.001)**
CRT score $\in [0,6]$	---	---	.015 (.003)**
Gambling Frequency $\in [0,4]$	---	---	-.025 (.010)**
R-squared	.0248	.0268	.1004

Notes: *$p < .05$, **p < 01 for the 1-tailed test of a pre-registered one-sided hypothesis (otherwise, *p-value* is for the 2-tailed test). *N*=465 observations (participants). The significant coefficient estimates on the variable *SkillGambler* in models (1) and (2) is opposite the preregistered hypothesis. This finding is not present once controlling for one's frequency of gambling (i.e., *Gfreq* controls for those who report more frequent gambling. which spuriously relates to one being more likely to have reported playing an online gambling game of skill—the simple correlation between *GFreq* and *SkillGambler* is .710).

TABLE 5: Hypothesis 1 test (Accuracy by sex)—panel data estimates

Dependent Variable: ***Accuracy*** (trial level)	(1)	(2)	(3)
Variable	Coef (st. error)	Coef (st. error)	Coef (st. error)
Constant	.727 (.010)**	.720 (.011)**	.635 (.066)**
Female (=1)	-.018 (.014)	-.018 (.014)	-.010 (.015)
Trial #	---	.0003 (.0004)	.0003 (.0004)
Response Time	---	.0003 (.0002)	.0002 (.0002)
Age	---	---	-.0001 (.001)
Gambling Frequency $\in$ [0,4]	---	---	-.026 (.007)**
Prior Week Sleep Level	---	---	.006 (.006)
Karolinska sleepiness	---	---	.003 (.004)
CRT score $\in$ [0,6]	---	---	.016 (.003)**
R-squared	.0013	.0020	.0339

Notes: **p* < .05, ***p* < 01 for the 1-tailed test of a pre-registered one-sided hypothesis (otherwise, *p-value* is for the 2-tailed test). *N*=9300 observations (standard errors adjusted for clustering at the participant level: n=465 clusters)

TABLE 6: Hypothesis 2 test—Among non-problem gamblers (n=129), those with skill-game experience will make more accurate probability assessments than those with only non-skill-game experience.

Dependent Variable: ***Accuracy*** (trial level)	(1)	(2)	(3)
Variable	Coef (st. error)	Coef (st. error)	Coef (st. error)
Constant	.690 (.031)**	.692 (.033)**	.612 (.134)**
Skill Gambler (=1)	-.001 (.035)	-.001 (.035)	-.028 (.037)
Trial #	---	-.001 (.001)	-.001 (.001)
Response Time	---	.0005 (.0003)	.001 (.0003)
Age	---	---	-.002 (.002)
Female (=1)			-.033 (.030)
Gambling Frequency $\in$ [0,4]	---	---	-.039 (.017)*
Prior Week Sleep Level	---	---	.023 (.012)
Karolinska sleepiness	---	---	.003 (.008)
CRT score $\in$ [0,6]	---	---	.021 (.006)**
R-squared	.0000	.0016	.0593

Notes: *p < .05, **p < 01 for the 1-tailed test of a pre-registered one-sided hypothesis (otherwise, *p-value* is for the 2-tailed test). *N*=2580 observations (standard errors adjusted for clustering at the participant level: n=129 clusters). Skill-games were considered to be the following: blackjack, poker, sports betting). Non-skill-games were: slots, baccarat, craps, roulette.

TABLE 7: Hypothesis 3 test—Among gamblers (n=206), non-problem Gamblers will make more accurate probability assessments than problem gamblers.

Dependent Variable: ***Accuracy*** (trial level)	(1)	(2)	(3)
<u>Variable</u>	Coef (st. error)	Coef (st. error)	Coef (st. error)
Constant	.689 (.015)**	.686 (.017)**	.638 (.099)**
Problem Gambler (=1)	.004 (.023)	.004 (.023)	.015 (.024)
Trial #	---	.0001 (.001)	.0001 (.001)
Response Time	---	.0002 (.0002)	.0001 (.0002)
Age	---	---	-.002 (.002)
Female (=1)			-.029 (.024)
Gambling Frequency $\in [0,4]$	---	---	-.032 (.014)*
Prior Week Sleep Level	---	---	.011 (.010)
Karolinska sleepiness	---	---	.006 (.006)
CRT score $\in [0,6]$	---	---	.022 (.005)**
R-squared	.0000	.0003	.0484

Notes: **p* < .05, ***p* < 01 for the 1-tailed test of a pre-registered one-sided hypothesis (otherwise, *p-value* is for the 2-tailed test). *N*=4120 observations (standard errors adjusted for clustering at the participant level: n=206 clusters). Skill-games were considered to be the following: blackjack, poker, sports betting). Non-skill-games were: slots, baccarat, craps, roulette.

TABLE 8: Hypothesis 4a and 4b tests (Modeling subjective belief formation)

Dependent Variable: ***Ln(Subjective Odds ratio)*** $_{Left}$	(1)	(2)	(3)
<u>Variable</u>	Coef (st. error)	Coef (st. error)	Coef (st. error)
Constant	.026 (.020)	-.0003 (.042)	-.225 (.155)
Ln(Prior Odds ratio) $_{Left}$	.324 (.029)**	.324 (.029)**	.324 (.029)**
Ln(Likelihood ratio) $_{Left}$	.291 (.015)**	.291 (.015)**	.291 (.015)**
Trial #	---	.004 (.004)	.004 (.004)
Response Time	---	-.001 (.001)	-.001 (.001)
Age	---	---	.003 (.003)
Female (=1)	---	---	.022 (.040)
Gambling Frequency $\in$ [0,4]	---	---	-.018 (.019)
Prior Week Sleep Level	---	---	.030 (.015)*
Karolinska sleepiness	---	---	-.009 (.011)
CRT score $\in$ [0,6]	---	---	-.019 (.009)*
R-squared	.090	.090	.091

Notes: *p < .05, **p < 01 for the 1-tailed test of a pre-registered one-sided hypothesis (otherwise, *p-value* is for the 2-tailed test). *N*=9300 observations (standard errors adjusted for clustering at the participant level: n=465 clusters)

TABLE 9: Hypothesis 5a test—Skill-game experience and subjective belief formation

Dependent Variable: ***Ln(Subjective Odds ratio)$_{Left}$***	(1)	(2)	(3)
Variable	Coef (st. error)	Coef (st. error)	Coef (st. error)
Constant	.044 (.025)	.020 (.045)	-.202 (.155)
Ln(Prior Odds ratio)$_{Left}$	.328 (.035)**	.328 (.035)**	.328 (.035)**
Ln(Likelihood ratio)$_{Left}$	.326 (.018)**	.325 (.018)**	.325 (.018)**
Skill Gambler (=1)	-.047 (.043)	-.048 (.043)	-.045 (.051)
*Skill Gambler * Ln(Prior Odds ratio)$_{Left}$*	-.010 (.061)	-.010 (.061)	-.010 (.061)
*Skill Gambler * Ln(Likelihood ratio)$_{Left}$*	-.092 (.033)**	-.091 (.033)**	-.092 (.033)**
Trial #	---	.004 (.003)	.004 (.003)
Response Time	---	-.001 (.001)	-.001 (.001)
Age	---	---	.003 (.003)
Female (=1)	---	---	.023 (.040)
Gambling Frequency $\in$ [0,4]	---	---	-.004 (.024)
Prior Week Sleep Level	---	---	.030 (.015)*
Karolinska sleepiness	---	---	-.010 (.010)
CRT score $\in$ [0,6]	---	---	-.019 (.009)*
R-squared	.096	.096	.097

Notes: **p* < .05, ***p* < 01 for the 1-tailed test of a pre-registered one-sided hypothesis (otherwise, *p-value* is for the 2-tailed test). *N*=9300 observations (standard errors adjusted for clustering at the participant level: n=465 clusters)

TABLE 10: Hypothesis 2b and 6 tests—Problem-gamblers and subjective belief formation

Dependent Variable: ***Ln(Subjective Odds ratio)$_{Left}$***	(1)	(2)	(3)
<u>Variable</u>	Coef (st. error)	Coef (st. error)	Coef (st. error)
Constant	.029 (.022)	.002 (.043)	-.236 (.157)
Ln(Prior Odds ratio)$_{Left}$	.322 (.031)**	.322 (.031)**	.322 (.031)**
Ln(Likelihood ratio)$_{Left}$	.298 (.017)**	.298 (.017)**	.298 (.017)**
Problem Gambler (=1)	-.012 (.059)	-.010 (.058)	.026 (.058)
*Problem Gambler * Ln(Prior Odds ratio)$_{Left}$*	.009 (.079)	.011 (.079)	.011 (.079)
*Problem Gambler * Ln(Likelihood ratio)$_{Left}$*	-.037 (.042)	-.035 (.042)	-.035 (.042)
Trial #	---	.004 (.003)	.004 (.003)
Response Time	---	-.001 (.001)	-.001 (.001)
Age	---	---	.004 (.003)
Female (=1)	---	---	.024 (.040)
Gambling Frequency $\in [0,4]$	---	---	-.022 (.019)
Prior Week Sleep Level	---	---	.031 (.015)*
Karolinska sleepiness	---	---	-.008 (.011)
CRT score $\in [0,6]$	---	---	-.019 (.009)*
R-squared	.091	.091	092

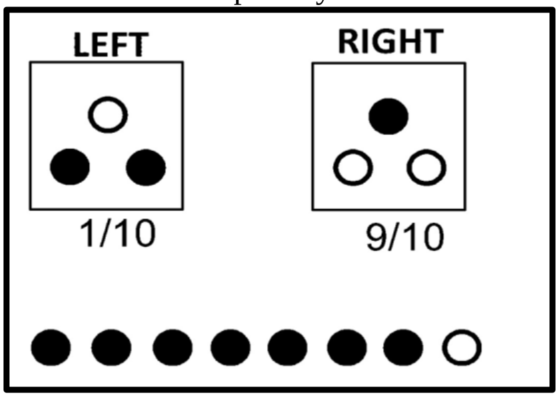
Notes: *p < .05, **p < 01 for the 1-tailed test of a pre-registered one-sided hypothesis (otherwise, *p-value* is for the 2-tailed test). *N*=9300 observations (standard errors adjusted for clustering at the participant level: n=465 clusters)

TABLE 11: Examining the important of current *Gambling Frequency* and *CRT Score* on information source weighting--Exploratory

Dependent Variable: ***Ln(Subjective Odds ratio)**$_{Left}$*	All participants (1)	Males (2)	Females (3)
Variable	Coef (st. error)	Coef (st. error)	Coef (st. error)
Constant	.105 (.041)**	.051 (.061)	.142 (.053)**
Ln(Prior Odds ratio)$_{Left}$	.305 (.057)**	.328 (.095)**	.285 (.075)**
Ln(Likelihood ratio)$_{Left}$	.211 (.038)**	.194 (.045)**	.242 (.035)**
CRT score $\in$ [0,6]	-.020 (.01)*	-.011 (.013)	-.026 (.014)
*CRT Score * Ln(Prior Odds ratio)$_{Left}$*	.001 (.013)	.002 (.020)	-.001 (.019)
*CRT Score * Ln(Likelihood ratio)$_{Left}$*	.036 (.007)**	.041 (.010)**	.027 (.009)**
Gambling Frequency $\in$ [0,4]	-.015 (.020)	-.004 (.028)	-.030 (.027)
*Gambling Frequency * Ln(Prior Odds ratio)$_{Left}$*	.020 (.030)	.003 (.038)	.043 (.050)
*Gambling Frequency * Ln(Likelihood ratio)$_{Left}$*	-.050 (.01)**	-.035 (.018)	-.077 (.018)**
Observations (clusters)	9300 (465)	4700 (235)	4600 (230)
R-squared	.1213	.1397	.1104

Notes: *p < .05, **p < 01 for the 1-tailed test of a pre-registered one-sided hypothesis (otherwise, *p-value* is for the 2-tailed test). *N*=9300 observations (standard errors adjusted for clustering at the participant level: n=465 clusters).

FIGURE 1: Sample Bayes task stimulus



Notes: Example shows trial with Prior Odds of LEFT Box = 1/10 and sample evidence of seven black balls drawn out of a sample draw (with replacement) of eight total balls

FIGURE 2:

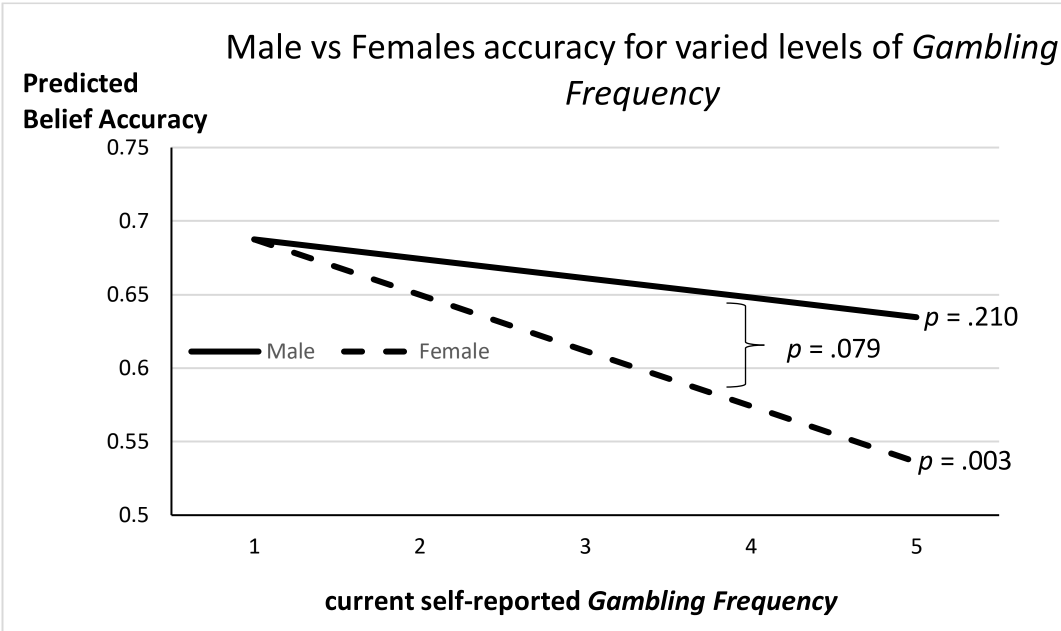
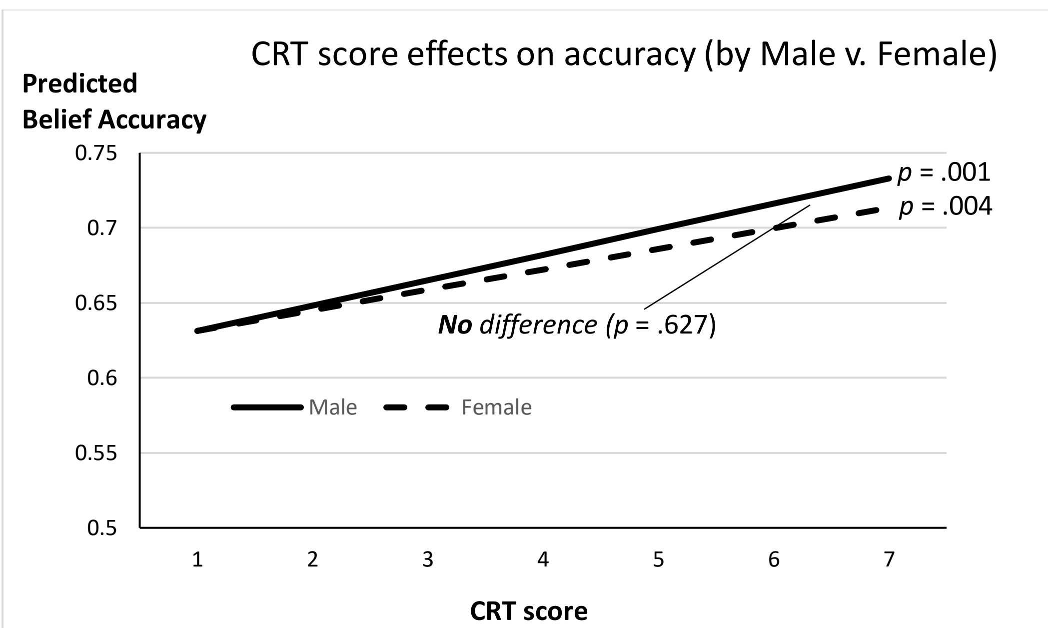


FIGURE 3:



Appendix A (additional results)

TABLE A1: Correlation matrix of key variables

	Gfreq	*ProbGamble*	*SkillGambler*	*SkillOnly*	*Female*	*Average RT*	*CRT score*
Gfreq	1.000						
ProbGamble	0.445	1.000					
SkillGambler	0.710	0.364	1.000				
SkillOnly	0.202	0.111	0.230	1.000			
Female	-0.099	-0.104	-0.090	-0.083	1.000		
Average RT	-0.035	0.068	-0.055	-0.051	-0.103	1.000	
CRT score	-0.001	0.008	0.021	0.013	-0.268	0.104	1.000

Notes: *Gfreq* $\in$ [0,4] describes self-reported *current* gambling frequency (0= never or non-gambler, higher values indicate more frequent online gambling). *ProbGamble* is an indicator = 1 if the respondent identified experienced any one or more of the 3 characteristics of problem gambling from the NODS rapid screener for adult pathological gambling. *SkillGamber* is an indicator = 1 if the individual indicated having played online gambling games of skill, while *SkillOnly* = 1 if the individual *exclusively* played games of skill (and no games of chance). *Female* = 1 that denotes sex (assigned at birth). *Average RT* is the average response time (in seconds) to the probability elicitation across the 20 trials in the Bayesian task. *CRT score* $\in$ [1,6] is one's score on the 6- item cognitive reflection task (higher scores indicating a more reflective style of thinking)

TABLE A2: Examining the important of current *Gambling Frequency* (versus Skill-game gambler)

Dependent Variable: ***Ln(Subjective Odds ratio)$_{Left}$***	(1)	(2)	(3)
Variable	Coef (st. error)	Coef (st. error)	Coef (st. error)
Constant	.044 (.025)	.023 (.045)	-.200 (.155)
Ln(Prior Odds ratio)$_{Left}$	.318 (.036)**	.318 (.036)**	.318 (.036)**
Ln(Likelihood ratio)$_{Left}$	.335 (.019)**	.334 (.019)**	.334 (.019)**
Skill Gambler (=1)	-.047 (.053)	-.049 (.053)	-.045 ((.051)
*Skill Gambler * Ln(Prior Odds ratio)$_{Left}$*	-.083 (.078)	-.084 (.078)	-.084 (.078)
*Skill Gambler * Ln(Likelihood ratio)$_{Left}$*	-.026 (.045)	-.025 (.045)	-.025 (.045)
*Gambling Frequency * Ln(Prior Odds ratio)$_{Left}$*	.046 (.039)	.047 (.039)	.047 (.039)
*Gambling Frequency * Ln(Likelihood ratio)$_{Left}$*	-.042 (.019)*	-.041 (.019)*	-.041 (.019)*
Trial #	---	.004 (.003)	.004 (.003)
Response Time	---	-.001 (.001)	-.001 (.001)
Age	---	---	.003 (.003)
Female (=1)	---	---	.022 (.040)
Gambling Frequency $\in$ [0,4]	.0003 (.025)	.0004 (.025)	-.004 (.024)
Prior Week Sleep Level	---	---	.030 (.015)*
Karolinska sleepiness	---	---	-.010 (.010)
CRT score $\in$ [0,6]	---	---	-.019 (.009)*
R-squared	.1031	.1034	.1045

Notes: *p < .05, **p < 01 for the 1-tailed test of a pre-registered one-sided hypothesis (otherwise, *p-value* is for the 2-tailed test). *N*=9300 observations (standard errors adjusted for clustering at the participant level: n=465 clusters).

TABLE A3: Non-linear Probit model estimations examining the important of current *Gambling Frequency* (versus Skill-game gambler)—Marginal effects reported

Dependent Variable: ***Left Box likely* (=1)**	(1)	(2)	(3)
Variable	Coef (st. error)	Coef (st. error)	Coef (st. error)
Ln(Prior Odds ratio)$_{Left}$	.024 (.012)*	.024 (.012)*	.024 (.012)*
Ln(Likelihood ratio)$_{Left}$	.052 (.005)**	.051 (.005)**	.051 (.005)**
Skill Gambler (=1)	-.0003 (.015)	-.001 (.015)	.003 (.015)
*Skill Gambler * Ln(Prior Odds ratio)$_{Left}$*	-.014 (.027)	-.014 (.027)	-.014 (.027)
*Skill Gambler * Ln(Likelihood ratio)$_{Left}$*	-.008 (.011)	-.008 (.011)	-.007 (.011)
*Gambling Frequency * Ln(Prior Odds ratio)$_{Left}$*	.013 (.012)	.013 (.012)	.013 (.012)
*Gambling Frequency * Ln(Likelihood ratio)$_{Left}$*	-.009 (.005)^	-.009 (.005)^	-.009 (.005)^
Trial #	---	.001 (.001)	.001 (.001)
Response Time	---	-.001 (.0003)*	-.001 (.0003)*
Age	---	---	.002 (.001)*
Female (=1)	---	---	.005 (.011)
Gambling Frequency $\in$ [0,4]	-.007 (.006)	-.007 (.006)	-.009 (.007)
Prior Week Sleep Level	---	---	.012 (.004)**
Karolinska sleepiness	---	---	.001 (.003)
CRT score $\in$ [1,6]	---	---	-.004 (.003)
Psuedo R-squared	.0380	.0386	.0397

Notes: ^*p* < .10*,* **p* < .05, ***p* < 01 for the 1-tailed test of a pre-registered one-sided hypothesis (otherwise, *p-value* is for the 2-tailed test). *N*=9074 observations (standard errors adjusted for clustering at the participant level: n=465 clusters).

TABLE A4: Bayesian *Accuracy* and the impact of *Gambling Frequency* and *CRT Score*

Dependent Variable: ***Accuracy*** (trial level)	**Full Sample** (1)	**Non-problem gamblers** (2)
Variable	Coef (st. error)	Coef (st. error)
Constant	.631 (.066)**	.597 (.132)**
Female (=1)	.020 (.030)	.060 (.083)
Trial #	.0003 (.0004)	-.001 (.001)
Response Time	.0002 (.0002)	.0005 (.0004)
Age	-.0002 (.001)	-.002 (.002)
Prior Week Sleep Level	.006 (.005)	.020 (.012)
Karolinska sleepiness	.002 (.004)	.003 (.008)
Skill Gambler (=1)	-.009 (.020)	-.034 (.036)
Gambling Frequency $\in$ [0,4]	-.013 (.011)	-.014 (.021)
CRT score $\in$ [0,6]	.017 (005)**	.032 (.022)*
*Gambling Frequency * Female*	-.025 (.014)	-.059 (.033)
*CRT score * Female*	-.003 (.627)	.001 (.013)
F-Test: *Gambling Freq + Gambling Freq * Female = 0*	*F* (1,464) = 8.81 *P* = .0032	*F* (1,128) = 8.29 *P* = .0047
F-Test: *CRT Score + CRT Score * Female = 0*	*F* (1,464) = 8.58 *P* = .0036	*F* (1,128) = 8.82 *P* = .0036
Number of observations (participant clusters)	9300 (465 participants)	2580 (129 participants)
R-squared	.0368	.0673

Notes: ^$p < .10$, *$p < .05$, **p < 01 for the 2-tailed test). Standard errors adjusted for clustering at the participant level. Figure 2 and 3 in the main text show results from the full sample model (1) coefficient estimates and their interaction terms.

Appendix B (complete survey)

<u>Informed Consent:</u> You are being asked to complete this online survey as part of a research study on decision making related dietary choice. Participation in this online survey is completely voluntary, your responses to this survey will remain completely confidential, the data will be securely stored, your name will not be recorded anywhere on this survey. The only identifier we will record will be your Prolific ID, which we as researchers cannot link to personally identifiable data of yours. This survey is estimated to take 18 minutes to complete and your payment for successful and complete survey completion will be $2.40. Additionally, the information use decision task within this survey offers **the chance of earning an *additional* $1.00 bonus payment,** depending on your choice in the task (the instructions will clearly explain how this works on that task). There are no known risks associated with this study beyond those associated with everyday life. Although this study will not benefit you personally, its results will help our understanding of how people make decisions.

For additional information related to this questionnaire, contact Dr. David Dickinson, Department of Economics, Appalachian State University, at dickinsondl@appstate.edu. Appalachian State University's Institutional Review Board (IRB) has determined this study to be exempt from review by the IRB administration**.**

○ **<u>I Consent</u>** and wish to continue with this study

○ **<u>I do not consent</u>** to participating and **do not wish to continue**

Page Break

As you do not wish to participate in this study, please return your submission on Prolific by selecting the 'Stop without completing' button

Page Break

The following questions are **screener validation questions** to make sure we get the desired sample we advertised for this survey

Page Break

What is **your current age** (in years)?

	18	26	34	43	51	59	67	75	84	92	100
Years of age ()											

What is your sex?
(i.e., what sex were you assigned at birth, such as on an original birth certificate)?

- ○ Female (1)
- ○ Male (2)

What types of online gambling / casino games have you played? Choose all that apply.

- ☐ Baccarat
- ☐ BlackJack)
- ☐ Bingo)
- ☐ Craps)
- ☐ Lottery)
- ☐ Pachinko
- ☐ Poker)
- ☐ Race & Sports Book
- ☐ Roulette1)
- ☐ Slots
- ☐ Video Poker
- ☐ Virtual Sports Betting
- ☐ None of the above
- ☐ Not applicable / rather not say

What is your current frequency of gambling (online or otherwise)?

○ Never

○ *less* than once a month

○ once or twice a month

○ once or twice a week

○ daily

Page Break

In what country do you currently reside?

○ United Kingdom

○ United States

○ Other

Page Break

Before you start, please switch off phone/ e-mail/ music so that you can focus on this study. Thank you!

Please carefully enter your Prolific ID

____

Please mark the number that best corresponds to how sleepy you feel **right now**. You may mark any number, but mark only one number.

- ○ 1. Extremely alert
- ○ 2.
- ○ 3. Alert
- ○ 4.
- ○ 5. Neither alert nor sleepy
- ○ 6.
- ○ 7. Sleepy--but no difficulty remaining awake
- ○ 8.
- ○ 9. Extremely sleepy--fighting sleep

Page Break

Over the last 7 nights, what is the average amount of sleep you obtained each night?

	0 1 2 3 4 5 6 7 8 9 10 11 12
Average nightly sleep over the LAST WEEK ()	

Page Break

Last night, how much sleep did you get?

	0 1 2 3 4 5 6 7 8 9 10 11 12
Hours of sleep LAST NIGHT ()	

Page Break

Q3.5 **What do you feel is the optimal amount of sleep for you personally to get each night**? (optimal in terms of next day alertness, performance, and functionality for you personally.)

	0 1 2 3 4 5 6 7 8 9 10 11 12
Average nightly sleep I need personally ()	

Page Break

Have there ever been periods lasting 2 weeks or longer when you spent a lot of time thinking about your gambling experiences or planning out future gambling ventures or bets?

- NO
- YES

Page Break

Have you ever tried to stop, cut down, or control your gambling?

- NO
- YES

Page Break

Have you ever lied to family members, friends, or others about how much you gamble or how much money you lost on gambling?

- NO
- YES

Page Break

As described earlier, we are interested in factors that influence the decisions you might make. In order for the results of this survey to be valid, **it is essential that you read all the instructions and questions carefully**. So we know that you have read these instructions, please place the slider below on the answer to (33+12)=? Thank you for taking the time to read these instructions.

	0 10 20 30 40 50 60 70 80 90 100
My response ()	

Page Break

INSTRUCTIONS FOR THE DECISION TASK

In each round, you will see a picture of two boxes, populated with 3 total balls each. The LEFT box contains 2 black and 1 white ball, while the RIGHT box contains 2 white and 1 black ball. One of the boxes will be selected in each round. You will not know for certain which box is selected, but **we will provide you with two pieces of information that may be helpful in how you determine which box was more likely selected**. First, we will give you the "starting chances" that either box may be selected in that round. Higher starting chances of the LEFT box means it is more likely the LEFT box will be selected in that round, for example. Secondly, we will present to you the results of having drawn 8 balls, with replacement, from whichever box was selected. Drawing with replacement means we always replace the ball after drawing so that the contents of each box are always the same when making each draw. All else equal, drawing balls from the LEFT box (which has more black than white balls) is more likely to produce more black balls in the sample set of draws, and drawing balls from the RIGHT box (which has more white than black balls) is more likely to produce more white balls in a sample of draws. Thus, both the "starting chances" and the "sample evidence" may be useful as you think of which box had more likely been selected in that round.

Winning a $0.05 bonus in each round depends on your response in that round. At the most basic level, in each round **the goal is to give your best guess about how likely *<u>the LEFT box</u>* was selected in that round. We will ask for your answer each round by asking you for your best estimate of the "*chances out of 100*" that you think the LEFT box was selected in that round**. A "0" answer means you feel there was no chance the LEFT box was selected in that round, "50" means you feel there was an equal chance the LEFT or RIGHT box was selected that round, and "100" means you feel that the LEFT box was certainly selected in that round. **Because we are asking for your response in terms of *<u>how likely you think it is the LEFT box was selected</u>*, you should indicate a response greater than "50" if you feel it is more likely the LEFT box was selected in that round, and a response less than "50" if you feel it is more likely the RIGHT box was selected (and the closer to 100 or 0 your response, the more strongly you feel the box selected was the LEFT or RIGHT, respectively).** The payment method for this task is designed so that your chances of winning a bonus that round are highest if your response is an accurate reflection of how likely you think the LEFT box was selected in that round. You will maximize your chance of the highest bonus in this decision task by being as accurate as possible in each round.

Page Break

Here's how your response generates a bonus in each round of this task. It is not the most easy to understand process, but its design actually ensures that it is in your best interest (in terms of bonus payment potential) to response with your true belief in each round.

(you can **skip these shaded details if you are not interested in the underlying process**).

In each round, the computer will draw a random number from 0 to 100. Each number from 0 to 100 is equally likely to be drawn by the computer. We'll call this number Draw 1. How you win or lose that round of the task depends on what number the computer draws for Draw 1 and your belief response (in terms of the "chances out of 100" that you think the LEFT box was selected in that round):

Payment Method 1) If Draw 1 is less than your belief response, you win if the LEFT box was selected and you do not win if the RIGHT box was selected in that round. For example, if you enter a belief response of 99, you are very likely to win a bonus if the LEFT box was selected and very likely to not win if the RIGHT box was selected. You are more likely to win the bonus in any given round where the LEFT box was selected if you give a higher belief response (i.e., chances out of 100 you feel the LEFT box was selected). Similarly, if the RIGHT box was selected in any given round, your are more likely to win the *lower* is your belief response regarding how likely the LEFT box was selected (because if indicate you feel it was less likely the LEFT box was selected, then you are also indicating that you feel it more likely the RIGHT box was selected in that round).

Payment Method 2) If Draw 1 is greater than your response, then the computer will draw a second random number from 0 to 100. As before, each number from 0 to 100 is equally likely to be drawn by the computer. We'll call this random number Draw 2. If Draw 2 is less than Draw 1, then you win the bonus for that round. If Draw 2 is greater than Draw 1, then you do not win the round. What *Payment Method 2* does is provide a way where, on average, you have a higher chance of winning the bonus than with *Payment Method 1* whenever Draw 1 is greater than your response.

The computer will therefore use your response each round to choose the whichever payoff method (Method 1 or Method 2) that offers you the best chance of earning the bonus payment in that round. **If your response represents your true beliefs about the chances the LEFT box was selected, then the computer selects the payment method that gives you the best chance of winning the bonus that round.**

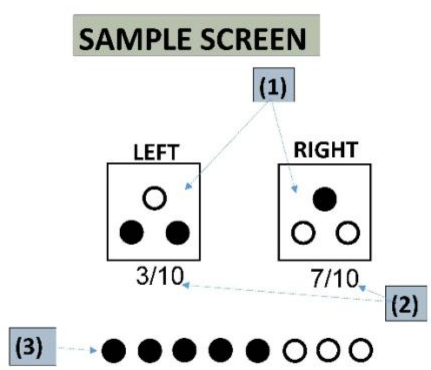
Page Break

Instructions (continued)

Again, the contest payment process is designed so that you have the best chance for earning a 5 cent bonus each round by being as accurate as possible with your response (which can earn you up to a total bonus payment of 20 rounds times 5 cents, or $1.00). The random numbers and payment calculations will happen behind the scenes after you have finished the study. As such, you will not have any feedback on your performance from one round to the next (you will only know your outcome based on the bonus you receive separately from the fixed payment for this task)

A picture of the stimulus is shown below, which succinctly reminds you of the contents of the LEFT and RIGHT boxes (this remains constant across all trials), as well as the starting-chance of selecting the LEFT versus RIGHT box and the sample evidence of the 8 ball drawn from the selected box. Across different rounds, the starting-chance and/or sample evidence may change, and so you should pay attention to these pieces of information carefully in each round because this may affect how likely you think it was that the LEFT box was selected in that round. You may use the starting-chance and sample evidence information however you like in giving your best estimate of the "chances out of 100" the LEFT box was selected in that round, and remember that you maximize your chance of the highest bonus by responding what you truly believe the chances are that the LEFT box was selected in each and every round.

Page Break



<u>The importance of each part of the stimulus image is as follows:</u>

(1) Balls inside of the box show the different contents of each box

(2) The fraction beneath the box shows the starting-chance (out of 10) that the box will be selected. ***A greater fraction below the LEFT box means the starting-chance of selecting the LEFT box is higher*** *(and a lower fraction means the starting-chance of selecting the LEFT box is lower)*. However, remember that you do **not** get to see which Box was actually selected

(3) The set of 8 balls at the bottom show the result of drawing 8 balls, with replacement, from the Box that was selected. Remember, because of the contents of each box shown in **(1)**, ***a sample draw with more black balls is more likely to come from the LEFT box*** *(and a draw with more white balls is less likely to come from the LEFT box).*

Using any of this information that seems relevant to you, **you are then asked to indicate the likelihood (chances out of 100) that you think the LEFT box had been selected** in that particular trial.

<u>Note:</u> From trial to trial, *the information in items (2) and (3) may change* (but not item (1)--the contents of each Box).

Remember, you maximize the chance of winning the bonus payment each round by responding with your true belief of how likely you think the LEFT box was selected, given the available information!

Page Break

The main assessment task starts on the next page. Please click below when ready to start.

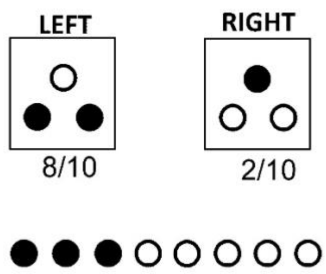
○ I'm ready to start the task

Page Break

NOTE: 2 sample trials are shown on the following pages for purposes of this Appendix. The complete survey included 20 trials where the details on each trial varied the number of black versus white balls in the 8-sample draw, and/or vary the fractions beneath the LEFT and RIGHT boxes (See Table 1, main text).

Please indicate on the scale below **how likely you think it is that the LEFT box had been selected**, given the following information below:

(remember, the fractions listed directly below each box indicate the starting-chance that the box will be selected in this trial. The row of 8 balls underneath show the result of drawing 8 balls with replacement from the box actually selected in this trial).



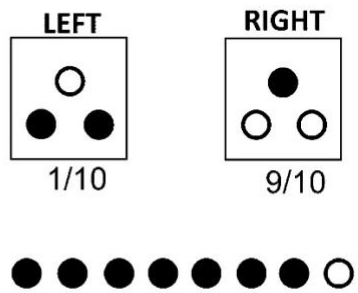
Given this information, I feel the chances out of 100 (i.e., the likelihood) that the LEFT box was selected in this trial is:

	Low chances out of 100 mean I think it is **unlikely the LEFT box was selected** (i.e., more likely the RIGHT box was selected)	Chance of 50 out of 100 means I think it is **equally likely that either the LEFT or RIGHT box were selected**	High chances out of 100 mean I think it is **more likely the LEFT box was selected** (i.e., unlikely the RIGHT box was selected)
	0 10 20 30 40 50 60 70 80 90 100		
Chances out of 100 the **LEFT** box was selected in this trial. ()			

Page Break

Please indicate on the scale below **how likely you think it is that the LEFT box had been selected**, given the following information below:

(remember, the fractions listed directly below each box indicate the starting-chance that the box will be selected in this trial. The row of 8 balls underneath show the result of drawing 8 balls with replacement from the box actually selected in this trial).

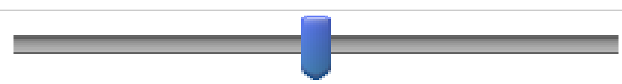


Given this information, I feel the chances out of 100 (i.e., the likelihood) that the LEFT box was selected in this trial is:

Low chances out of 100 mean I think it is **unlikely the LEFT box was selected** (i.e., more likely the RIGHT box was selected)	Chance of 50 out of 100 means I think it is **equally likely that either the LEFT or RIGHT box were selected**	High chances out of 100 mean I think it is **more likely the LEFT box was selected** (i.e., unlikely the RIGHT box was selected)

0 10 20 30 40 50 60 70 80 90 100

Chances out of 100 the **LEFT** box was selected in this trial. ()



Page Break

Finally, please answer these final questions on the next set of pages for us.

Page Break

A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost?
(please indicate your numeric answer **in cents**)

____

Page Break

If it takes 5 minutes for 5 machines to make 5 widgets, how long would it take for 100 machines to make 100 widgets?
(please indicate your numeric answer **in minutes)**

____

Page Break

If 3 elves can wrap 3 toys in 1 hour, how many elves are needed to wrap 6 toys in 2 hours?
(please give your numeric answer in **# of elves**)

____

Page Break

Jerry received both the 15th highest and the 15th lowest mark in the class. How many students are there in the class?
(please give your numeric answer in **# of students)**

____

Page Break

In an athletics team, tall members are **three** times more likely to win a medal than short members. This year the team has won 60 medals so far. How many of these have been won by short athletes?
(please give your numeric answer in **# of medals**)

____

Page Break

In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover **half** of the lake? (please indicate your numeric answer **in days**)

____