Illusory causal connections and their effect on subjective probability

Sami R. Yousif (sami.yousif@yale.edu)
Department of Psychology, Yale University, New Haven, CT, 06520 USA

Frank C. Keil (frank.keil @yale.edu)
Department of Psychology, Yale University, New Haven, CT, 06520 USA

Abstract

Our world is filled with statistical information: from dice rolls to lotteries, we often act based on our *impressions* of probability. Yet the human mind is not wired to reason about truly probabilistic events, often imposing structure on data or events where no such structure exists (as in 'illusory correlations'). Here, we consider a case study in intuitive statistics: disjunctive events. For example, participants are asked to imagine a button that, when pressed, has a 1 in 100 chance of yielding a prize. They are told to imagine pressing that button 100 times. Across several paradigms, we show that people *overestimate* the probability of this disjunctive event — in stark contrast to classic demonstrations where people underestimate such probabilities (e.g., when iteratively selecting marbles from jars with replacement). These results reflect a tendency to view events as causally connected in illusory ways; implications for other domains of reasoning are discussed.

Keywords: Intuitive statistics, disjunctive events; probability; reasoning; decision-making; reservoir fallacy

Introduction

Every day, we are flooded with statistical information, whether explicitly (as when reading or watching the news) or implicitly (as when slowly learning over time when you might expect to see certain events on your way to work). For example, you may think about whether it will rain each day over the next two, which may be thought of as the probability that it will rain tomorrow and the probability that it will rain the next day. In these cases, you may implicitly form predictions about what is or is not likely to happen. In others, you may be forced to explicitly form predictions: e.g., when gambling at a roulette table, you would likely be well aware of your odds of losing money over repeated rounds. But how do you arrive at these estimates?

Probabilities about dichotomous outcomes can be difficult to discern; mental math will usually be insufficient to approximate an answer. For example, it is easy to calculate the probability of getting 'tails' three times in a row when flipping a coin. But it is much harder to predict the probability of getting exactly two "2s" when rolling a die six times. Prior work (e.g., Bar-Hillel, 1973; Brockner et al., 2002; Costello, et al., 2009). suggests that people misestimate the probability of disjunctive events in a particular way, namely by *underestimating* the probability of disjunctive events. A single (but well-cited) paper (Bar-Hillel, 1973), documents how people reason about both conjunctive events and disjunctive events in the context of retrieving marbles from jars. The paradigm is straightforward: you are playing a betting game, and you must choose between one of two jars to bet on. Let's say both jars have twenty marbles. In Jar A, there will be one red marble and nineteen blue marbles. If you choose to bet on Jar A, you can select from the jar twenty times (with replacement); you win if you get the red marble at least once. In Jar B, there will also be twenty marbles: 12 red and 8 blue. If you choose to bet on Jar B, you can select from the jar only once; you win if you get a red marble. Which jar would you choose? If you are like most of the participants, you would take the 'simple' gamble, Jar B. This would be irrational, however, because the probability of winning by betting on Jar A (~64%) is greater than the probability of winning by betting on Jar B (60%).

This finding has led to the oft repeated claim that people have a general tendency to underestimate the probability of disjunctive events (e.g., see Bazerman & Moore, 1994; Tversky & Kahneman, 1986). Is this always true? Consider the following scenario. A button, when pressed, has a 1 in 100 chance of yielding a prize. Suppose you press that button 100 times. Without using a calculator or some other tool, ask yourself: what is the probability that you will win at least one prize? Even readers with training in formal statistics may have trouble figuring out an exact answer. For most readers, we suspect that the estimate that first comes to mind is *higher* (perhaps much higher) than the correct answer: 63%. Note that this problem is formally identical to the problem of selecting marbles from jars, yet (even knowing this) you may still have an intuition to overestimate the probability of winning a prize in this case. This failure — and its seeming incongruity with classic results — is the focus of the present paper.

Impressions of structure

It is no longer surprising to assert that humans are bad 'intuitive statisticians'; a large body of work spanning nearly half a century has attempted to understand the cognitive bases of decision-making — and the many ways in which we fail to grasp our probabilistic world (e.g., Peterson & Beach, 1967; Tversky & Kahneman, 1974). Yet this literature is only one part of an even broader story. In general, we often perceive structure or causal relationships where none exist. We not only perceive regularities in our environment (Saffran et al., 1996; Turk-Browne et al., 2005), we also infer their underlying causal structure (Gopnik et al., 2004; Rottman & Keil, 2012; Steyvers et al., 2003). For example, 'illusory correlations' between features and groups may give rise to social stereotypes (Hamilton & Gifford, 1976). And it is said that our impressions of 'conceptual coherence' are governed not by similarity or features correlations, but instead by the extent to which those concepts cohere with our own prior beliefs and naïve theories (Murphy & Medin, 1985).

In light of this tendency to 'overperceive' connectedness, we ask whether subject probability estimates may be explained by illusory causal structure. Here, we seek to understand (1) whether people do reason about disjunctive events in a (discernible) context-dependent manner, and, if true, (2) whether illusory impressions of connectedness may explain this pattern. We suggest that the physical instantiations of many probability problems may matter far more than previously thought; we discuss a possible 'reservoir fallacy' that may underlie such intuitions.

Current Study

In a first set of studies, we replicate the original marbles-in-jars paradigm in several unique ways. In a second set of studies, we ask whether people *always* underestimate disjunctive events. We test participants in the button-pressing framing (explained above) and show that people have a strong tendency to *overestimate* the probability of disjunctive events (Experiment 2a). This effect is not explained by the number of events (Experiment 1a), is robust across paradigms (Experiments 2b-2c), and seems to be a part of a coherent mental model (Experiment 2c). We then discuss how these findings contribute to our understanding of reasoning about specific kinds of probabilistic events, but also how these findings advance our understanding of probabilistic reasoning generally.

Experiment 1a: Marbles — Gambles

First, we aimed to conceptually replicate classic effects documenting underestimation of disjunctive events (e.g., when selecting marbles from jars; Bar-Hillel, 1973). However, we modified this paradigm in three ways: (1) participants do not manipulate physical marbles in jars, but instead are told to imagine them; (2) the values differ from the original work; in our disjunctive bet, for example, participants are told there is a jar with 19 blue marbles and 1 red marble, from which they select 20 times with replacement; in our simple bet, three different groups of participants are told either that there is a flat 60%, 65%, or 75% chance of winning only a single selection (from a set of 20 marbles); and (3) we test adults rather than school-age students. We sought to replicate and extend these findings — to ensure that underestimation is due to the general framing of the problem and not the specific details of the paradigm.

Method

Participants Three separate groups of 100 adult participants completed a survey online through Amazon Mechanical Turk. The sample size was chosen in advance and was pre-registered. Pre-registrations for this experiment and subsequent experiments, as well as raw data and other materials, can be found at the following OSF page. All participants lived in the United States. This study was approved by the relevant Institutional Review Board.

Design & Procedure Participants are asked to choose between two bets. In Bet A, the disjunctive bet, there is a jar with 20 marbles, 19 of which are blue and one of which is red. Participants were told to imagine making 20 iterative selections from the jar, one marble at a time, with replacement. They were told that if they choose this bet, they win if they ever receive a *single* red marble. In Bet B, the simple bet, participants were told to imagine a jar with 20 marbles, x (12, 13, or 14) of which are red while the rest are blue. (Three different groups of 100 participants saw each value of x). They were told that, unlike the other bet, they could only select from this jar one time. They were told they win the bet if they receive a red marble. Thus, participants chose between a disjunctive bet of unknown probability (Bet A) and a simple bet of known probability (Bet B). To ensure that observers responses were valid, we urged participants not to use calculators or other aids and we explained that there were no tricks in the task and that we were only interested in people's intuitions. Separate groups of subjects answered the question for different values of x (12, 13, or 14). Participants simply made a selection. After this selection, they were prompted to "re-state the [question] in [their] own words, and provide a one sentence explanation for [their] answer."

Results and Discussion

Results for this experiment can be seen in Figure 1a. As is evident from the figure, participants reliably preferred the 'simple' bet in all three conditions — even when the disjunctive bet was rationally the better option (and these impressions are confirmed with analyses following). Note that the objective likelihood of winning for the disjunctive bet is always ~64%. In comparison, the objective likelihood of winning the simple bet is 60%, 65%, and 70% in each of the three conditions. Therefore, if participants are behaving rationally, they should choose the disjunctive bet more in the first conditions, about equally in the second, and less in the third. Instead, participants consistently preferred the simple bet over the disjunctive bet: across all three conditions, 217 of 300 participants chose the simple bet (p<.001; 74, 71, and 72, for each of the conditions above, respectively; all ps<.001). Thus, participants seem to prefer the simple bet even when it is objectively the poorer option. This is consistent with prior results documenting a similar effect (Bar-Hillel, 1973). Here, we conceptually replicate these effects while altering many minor experimental details — suggesting that this pattern is indeed robust and meaningful.

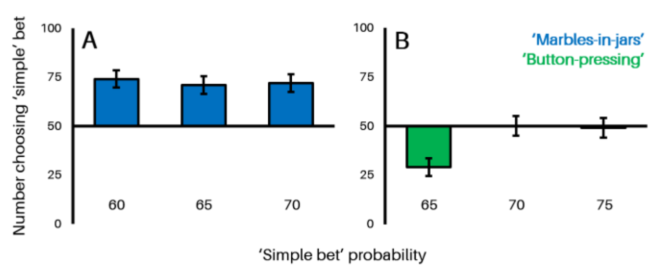

Figure 1. The number of participants who chose the 'simple' bet in (A) Experiment 1a and (B) Experiment 2a. The y-axis simply represents the number of participants (out of 100) who chose the simple bet (as opposed to the disjunctive bet). On the x-axis are the three different values of the simple bet for each experiment; note that these are not identical across experiments, by design. Error bars represent +/- 1 SE.

Experiment 1b: Marbles — Free Response

Experiment 1a shows at least one case where people do in fact underestimate the probability of disjunctive events. Nevertheless, it is still unclear how these results might vary with task demands. Here, we assess people's statistical intuitions in a different way, by having them explicitly estimate the probability of a given outcome. For example, we will again tell participants about a jar with 1 red marble (from which they select marbles one at a time, with replacement), and ask them to estimate the probability they'd receive at least one red marble if they selected x times (where x equals the number of total marbles in the jar). Instead of choosing between two gambles, they indicate their probability estimate by freely adjusting a number line from 0-100. Do people still underestimate?

Method

This experiment is identical to Experiment 1a except as otherwise noted. Three separate groups of 100 participants completed this experiment. Here, instead of choosing between two gambles, participants simply explicitly estimated the probability of receiving one or more red marble when selecting x times (20, 100, or 300) from a jar with x marbles (20, 100, or 300), only one of which was red. They made their responses by dragging and dropping a point along a number line from 0-100.

To ensure that observers responses were valid, we urged participants not to use calculators or other aids and we explained that there were no tricks in the task and that we were only interested in people's intuitions. We also noticed in pilot data that a large number of people would either answer 100% or the equivalent of (1/x)% (i.e., if there were twenty marbles, they would say 5%). To combat this, we added a note that said, e.g.: "Hint: the answer is neither 100% nor 5%." We further encouraged participants to think about this problem like a coin flip, explaining: "There is a 50% chance you'll get tails on any given flip — but that does not mean you are guaranteed to get tails in two flips. And, in general, you're more likely to get at least one tails if you flip the coin 10 times vs. 5 times."

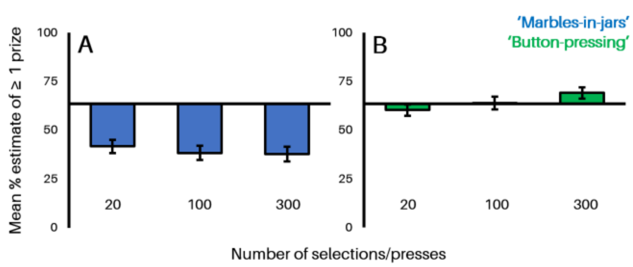

Figure 2. The mean probability estimates of receiving one or more prizes for (A) Experiment 1b and (B) Experiment 2b. The y-axis simply represents the mean response. The x-axis represents the specific gamble (i.e., the three different conditions of each experiment). The x-axis itself (set at a value of 63.5) corresponds to approximately chance performance in each case (though this differences slightly across conditions). The chance line is for the purpose of visual demonstration and is not meant to be exact. Error bars represent +/- 1 SE.

Results and Discussion

Results for this experiment can be seen in Figure 2a. As is evident from the figure, participants reliably *underestimated* the probability of the disjunctive events in all three conditions (the x-axis corresponds to approximately rational performance; these impressions are confirmed in subsequent analyses). The objective likelihood of winning for the disjunctive bet is ~63-64%; therefore, participants' responses should on average be *lower* than these values. For purposes of standardization, and to be conservative, we ask in each condition whether the observed averages differ from a value of 63%. And, indeed, that was the case: participants indicate that the estimated probability of selecting at least one red marble was 41.6%, 38.3%, and 37.8% for the 20, 100, and 300 marble conditions, respectively. All three conditions were lower than would be expected by chance if participants' answers matched the objective truth (20 selections: $t(99)=6.15$, $p<.001$, $d=.62$; 100 selections: $t(99)=6.84$, $p<.001$, $d=.68$; 300 selections: $t(99)=6.67$, $p<.001$, $d=.67$). These results offer further evidence for the tendency to underestimate the probability of disjunctive events and suggest that participants underestimate even when making explicit judgments.

Experiment 1c: Marbles — Exact Estimates

In Experiments 1a and 1b, we participants underestimated the probability of disjunctive events in the marbles-in-jars framing. But *why*? Many participants give implausible answers, as if they fail to understand the logical constraints of the problem. E.g., without thinking for very long, one may not realize that saying you have a 100% chance of getting one red marble means there is no possibility whatsoever of failing to draw that red marble. To address this, we had participants simultaneously estimate the probability of getting (a) exactly zero red marbles, (b) exactly one red marble, and (c) two or more red marbles. Participants were explicitly told the probability estimates had to add up to 100%. In this way, we force participants to reason about many possible outcomes at once — while also helping them recognize the logical constraints of the problem.

Method

This experiment is identical to Experiment 1b except as otherwise noted. A new group of 100 participants completed this experiment. Here, instead of indicating the probability of getting one or more red marble, participants simultaneously indicated the probability of getting exactly zero, exactly one, and two or more red marbles. They made their responses by dragging and dropping points along three separate number lines from 0-100. Participants were told that these probabilities constrained each other and therefore must add up to 100%; responses that did not add up to 100% were excluded and replaced.

Results and Discussion

The objective likelihood of winning zero prizes is 35.9%; the objective likelihood of winning one prize is 37.7%; and the objective likelihood of winning two or more prizes is 26.4%. (In the figure, these objective likelihoods are represented by thin grey bars.) We compare the observed values to these expected values. If participants have a general tendency to *underestimate* the probability of disjunctive events, then we should expect an *overestimated* probability for zero prizes. However, we have no specific predictions about the other two values. People could underestimate disjunctive probability because they underestimate the probability of receiving one item, or two items, or more.

Results for this experiment can be seen in Figure 3a. As is evident from the figure, participants reliably *underestimated* the probability of the disjunctive event (i.e., they overestimated the likelihood of receiving zero prizes, and thereby underestimated the probability of receiving one or more prizes). These impressions are confirmed by the following analyses. On average, participants estimated there to be a 61.5% probability of receiving zero prizes ($t(99)=6.72$, $p<.001$, $d=.67$), a 23.3% probability of receiving one prize ($t(99)=4.96$, $p<.001$, $d=.50$), and a 15.2% probability of receiving two or more prizes ($t(99)=4.69$, $p<.001$, $d=.47$). As expected, people overestimated the probability of receiving no prize at all. Interestingly, this increased probability was displaced both from the probability of receiving one prize and two or more prizes. In other words, people think the probability of receiving any number of prizes is less likely than would be expected by chance. This extra information helps us to understand *how* participants underestimate. They *are not* underestimating just because they are bad at estimating probability. Instead, they are reasoning in a consistent and coherent (albeit illogical) manner (in the sense that they view the probability of winning two or more prizes as less likely than winning one prize — which is true — and in the sense that their overestimated chance of winning exactly zero prizes detracts both from of these values).

Unlike Experiments 1a and 1b, participants in this task are explicitly forced to recognize how the probability of each outcome constraints the other. As such, if the prior results arose from a general failure to understand the problem, we might have seen an entirely different pattern of results in this paradigm. Finding a very similar pattern suggests that whatever mental model participants are engaging is highly systematic, always triggered in this context, and unlikely due to task demands.

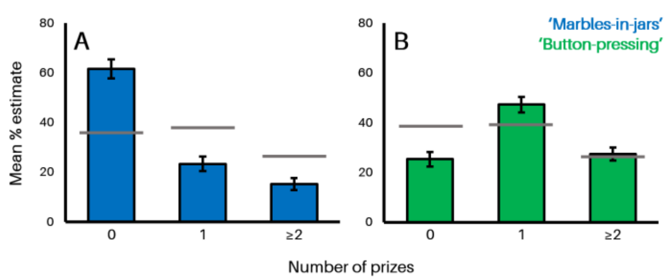

Figure 3. The mean probability estimates of receiving certain exact outcomes for (A) Experiment 1c and (B) Experiment 2c. The y-axis simply represents the mean response. The x-axis represents the specific outcome (i.e., zero, one, or more prizes). The grey bars represent the objective likelihood of each outcome in each experiment. If a value is above this bar, it would mean that participants are overestimating the probability of this specific outcome. Error bars represent +/- 1 SE.

Experiment 2a: Buttons — Gambles

Do people always underestimate the probability of disjunctive events? Previously, we gave the example of pressing a button one hundred times that has a 1% chance of yielding a prize. Intuitively, it *feels* like the chances of winning are much higher in this case. Here, we mirror the paradigm in Experiment 1a but with a slightly different framing. Participants were asked to choose between two bets involving pressing a button. In Bet A, the disjunctive bet, there is a button that can be pressed 100 times, which has a 1/100 chance to yield a prize. They were told that if they choose this bet, they win if they ever receive a single prize. In Bet B, the simple bet, participants were told that there is a button with *y*% (65, 70, 75) chance of yielding a prize. They were told that, unlike the other bet, the button could only be pressed once. Thus, participants must choose between a disjunctive bet of unknown probability (Bet A) and a simple bet of known probability (Bet B). Using this approach, we can determine whether people always underestimate disjunctive probabilities or if, instead, the intuition presented in the introduction is empirically supported.

Method

This experiment is identical to Experiment 1a except as otherwise noted. Three separate groups of 100 participants completed this experiment. Here, instead of indicating the probability of getting red marbles when selecting from a jar, the gambles were framed in terms of button presses. For example, in Bet A, the disjunctive bet, there is a button that can be pressed 100 times, which has a 1/100 chance to yield a prize. Participants were told that if they choose this bet, they win if they ever receive a single prize. In Bet B, the simple bet, participants were told that there is a button with *x*% (65, 70, 75) chance of yielding a prize. They were told that, unlike the other bet, the button can only be pressed once. Thus, the choice is between a disjunctive bet of unknown probability (Bet A) and a simple bet of known probability (Bet B). Separate groups of subjects answered the question for different values of *x* (65, 70, 75).

Results and Discussion

Results for this experiment can be seen in Figure 1b. As is evident from the figure, participants reliably preferred the 'disjunctive' bet more than they rationally should have (and these impressions are subsequently confirmed with analyses). The objective likelihood of winning for the disjunctive bet is ~63%. In comparison, the objective likelihood of winning is 65%, 70%, and 75% in each of the three conditions. Therefore, if participants are behaving rationally, we should expect that they choose the disjunctive bet less often in all three conditions. However, this is not the case. Participants overestimate the probability of the disjunctive bet in the first condition, choosing it more often than the simple bet (71 of 100; $p<.001$). In the other two conditions, participants were at chance (50/100 and 51/100, respectively) when in fact they *should* have been choosing the simple bet more often ($ps>.50$). In other words, participants seem to prefer the disjunctive bet more often than they should, even when it is objectively the poorer option. This is inconsistent with other effects of disjunctive probability estimation Experiments 1a-c here; Bar-Hillel, 1973). This experiment provides the first evidence of overestimating the probability of disjunctive events in certain contexts. Before trying to understand what accounts for this difference, we first wanted to ensure that this pattern is robust. The next 3 experiments were designed to replicate this pattern across unique paradigms.

Experiment 2b: Buttons — Free Response

Are the results in Experiment 2a generalizable? In this scenario, rather than having participants choose between two bets, they will be asked to explicitly estimate the probability of a particular outcome. Participants were told that they could press a button *x* (20, 100, 300) times which had a 1/*x* (20, 100, 300) chance to yield a prize. They were asked to indicate the probability that they would receive at least one prize. Note, crucially, that this framing is mathematically identical to the marbles-in-jars framing. Note also that we were less interested in the *absolute* answers of participants, and more in the *relative* values of their answers (compared to Experiment 1b).

Method

This experiment is identical to Experiment 1b except as otherwise noted. Three separate groups of 100 participants completed this experiment. Here, instead of indicating the probability of getting one or more red marble when selecting from a jar, participants were told that there is a button with a 1/*x* chance (20, 100, or 300) of yielding a prize when pressed. They were told to imagine pressing this button *x* times (20, 100, or 300). They made their responses by dragging and dropping a point along a number line from 0-100. (See also our discussion of exclusion criteria and how those should affect the interpretation of these results in the Method section of Experiment 1b.)

Results and Discussion

Results for this experiment can be seen in Figure 2b. As is evident from the figure, participants greatly *overestimated* the probability of the disjunctive events for all three values (relative to the marbles-in-jars framing in panel 2a). The objective likelihood of winning for the disjunctive bet is ~63-64%; therefore, we may expect that participants' responses are on average *higher* than these values; however, because of the difficulty of offering an explicit probability estimate and noise in the data, we caution against interpreting the absolute value of these estimates. Instead, we should consider the relative values (compared to the marbles-in-jars framing of Experiment 1b). That said, for purposes of standardization, and to be conservative, we ask in each condition whether the observed averages differ from a value of 64%. In fact, participants were no different from the chance value of 64% in any of the three conditions (*p*s>.05). These probability estimates differed greatly from Experiment 1b. In the 20 button condition, participants gave an average estimate that was **19 points higher** ($t(99)=4.09$, $p<.001$, $d=.58$); in the 100 button condition, participants gave an average estimate that was **25 points higher** ($t(99)=5.18$, $p<.001$, $d=.73$); and in the 300 button condition they gave an average that **was 31 points higher** ($t(99)=6.57$, $p<.001$, $d=.93$). In some situations, therefore, people *overestimate* the probability of disjunctive events: the problem posed to participants in this experiment is mathematically identical to what they encountered in Experiment 1b, yet average probability estimates differed by as much as 32%.

Experiment 2c: Buttons — Exact Estimates

We can use these same paradigms to further probe people's mental models of disjunctive events. For example, we can ask not about the probability of receiving one or more prizes, but instead about the exact probability of receiving *exactly* one prize (or exactly twice prizes, or exactly zero prizes). Mirroring the paradigm of Experiment 1c, participants simultaneously assessed the probability of receiving exactly zero, exactly one, and two or more prizes. This allows us to better understand *why* people overestimate in this context: is it because they overestimate the probability of exactly one event occurring, or because they overestimate the probability of two events occurring, or three, etc.?

Method

This experiment is identical to Experiment 1c except as otherwise noted. A separate group of 100 participants completed this experiment. Here, instead of indicating the probability of getting specific numbers of red marbles when selecting from a jar, participants were asked to indicate the probability of getting a specific number of prizes (zero, one, or more) when pressing a button 100 times with a 1/100 chance of yielding a prize.

Results and Discussion

The objective likelihood of winning zero prizes is 36.6%; the objective likelihood of winning one prize is 37.0%; and the objective likelihood of winning two or more prizes is 26.4%. We compared the observed values to these expected values. If participants have a general tendency to *overestimate* the probability of disjunctive events in this context, then we should expect an *underestimated* probability for zero prizes.

Results for this experiment can be seen in Figure 3b. Participants reliably *overestimated* the probability of the disjunctive event (i.e., they underestimated the likelihood of receiving zero prizes, and thereby overestimated the probability of receiving one or more prizes). These impressions are confirmed by the following analyses. On average, participants estimated there to be a 25.3% probability of receiving zero prizes ($t(99)=3.83$, $p<.001$, $d=.38$), a 47.2% probability of receiving one prize ($t(99)=3.28$, $p=.001$, $d=.33$), and a 26.4% probability of receiving two or more prizes ($t(99)=.40$, $p=.69$, $d=.04$). As expected, people underestimate the probability of receiving no prize at all. Interestingly, it seems that most of this probability was displaced to the probability of receiving exactly one prize (which participants consistently overestimated). This extra information helps us to understand *how* participants overestimate. As in Experiment 1c, they seem *not* to be estimating incorrectly just because they are bad at estimating probability in general; if that were the case, these results should be much more random *and* much more like those of Experiment 1c. In fact, the average for each value (zero prizes, one prizes, and two or more prizes) was significantly different from what we observed in Experiment 1c (*p*s<.001) — further confirming that participants are reasoning differently in these two contexts. Participants seem to engage different cognitive models depending on how the problem is framed. These different cognitive models cause systematic underestimation in one case (the marbles-in-jars framing; Experiments 1a-c) and overestimation in another (the button-pressing framing; Experiments 2a-c).

General Discussion

People have conflicting intuitions about disjunctive events. In some cases, people *underestimate* the probability of disjunctive events (Experiments 1a-c); but in others, they *overestimate* (Experiments 2a-c). Yet a coherent mental model may underlie these contrasting intuitions — one that is swayed by (mis)-perceptions of causal connectedness. This mental model has two key components: (1) people view probabilities as being more exact than they truly are (i.e., interpreting 1/300 to mean that we should expect almost exactly one out of every three hundred, whereas the reality is that one should expect an *average* of 1 output per 300 inputs, whether that means one receives 0, or 2, or 3, or 10 outputs for a given set of inputs; see Experiment 2d); and (2) people view disjunctive events as part of a finite-but-replenishable resource (i.e., a 'reservoir') that rapidly depletes (unless there is clear evidence of its replenishment, as when replacing marbles into a jar) — possibly reflecting a general tendency to overperceive or erroneously infer causal connections among independent events. Behavior in these tasks does not stem from arbitrary, irrational intuitions, but instead arise from an appealing (albeit incorrect) mental model of the world: that, in the absence of an explicit cue to replenishment (or independence), probabilistic outcomes are finite and must eventually occur. This is what we refer to as the 'reservoir fallacy'.

Note that the previously documented effect of *underestimation* has been explained in a very different way. Tversky and Kahneman (1986), write "…The overall probability of a disjunctive event is higher than the probability of each elementary event. As a consequence of anchoring, the overall probability… will be underestimated in disjunctive problems" (p. 47). While it may be the case that underestimation in the relevant cases may be understood as an effect of anchoring, this explanation *cannot* explain overestimation in other cases (e.g., in the button-pressing framing). Therefore, the present results call into question this domain-general explanation. Note that this same explanation is used by Tversky and Kahneman to explain why people are biased to *overestimate* conjunctive events (as in Cohen et al., 1972). If anchoring does not explain the estimation of disjunctive events, does it still explain the estimation of conjunctive events? Our results suggest that, instead, reasoning may often be guided by specific, context-dependent mental models.

That said, this 'reservoir fallacy' may be explained in other ways. Here, we have suggested that 'illusory causal connections' in certain scenarios drive differing intuitions. However, there are other ways of thinking about these results. For example, perhaps a sense of control over the marble selection (and a lack of sense of control when pressing a button) contribute to these differing intuitions. In ongoing work, we address this concern by using more carefully controlled scenarios (e.g., simultaneous vs. successive dice rolls). Even in this minimal pair, observers have different intuitions about the likelihood of certain outcomes. Here, however, a sense of control is unlikely to explain the difference: ostensibly people exert the same amount of control over dice rolls, whether or not they are successive or simultaneous. Even so, there may be other subtle ways of framing these results; 'illusory causal connections' are not the *only* possible explanation for the differences observed here. We encourage future work that addresses other possibilities.

Subjective Probability in a Causal World

We exist in a world of causes and effect; our minds are predisposed to be attuned not just to regularities in our environment (Saffran et al., 1996; Turk-Browne et al., 2005) but also to infer their causal structure (Gopnik et al., 2004; Rottman & Keil, 2012; Steyvers et al., 2003). Indeed, complex social processes like academic failure are often conceived by laypeople as stemming from a network of causes and effects (Lunt 1988, 1991). And in some cases, people *over*-interpret regularities (causal or not) — continuing to attend to where those regularities were, even after they exist no longer (Yu & Zhao, 2015). In some cases, this attention to and rapid interpretation of regularities serves a crucial functional role: to facilitate learning (Trueswell, et al., 2013). In other words: we automatically attend to the causal structure of the world and construct causal interpretations and models on highly incomplete data (as in Hamilton & Gifford, 1976).

Similarly, these findings may best be explained by a general disposition to impose or over-interpret regularities and/or causal structure when there is none. For example, when pressing a button 100 times, observers may be tempted to view those button presses as causally related in some way — each one gradually influencing the probability of the next. Yet, in the marbles-in-jars framing, people are faced with the reality that the system is 'resetting' each time, allowing them to overcome this (erroneous) disposition. This interpretation is consistent with other biases of subjective probability like the gambler's fallacy and the hot-hand fallacy, where, in both cases, people fail to properly grasp the causal independence of independent events.

Conclusion

We have shown a clear case where people *overestimate* the probability of disjunctive events. We have come to refer to this tendency as a 'reservoir fallacy', whereby disjunctive events are viewed as a being drawn from a finite-but-replenishable resource. This model explains how we reason about probability in a variety of contexts, from speculating about the probability of rain, to playing board games, to gambling — shedding light on the underlying nature, structure, and origins of our statistical intuitions.

References

Bar-Hillel, M. (1973). On the subjective probability of compound events. *Organizational Behavior and Human Performance*, *9*, 396-406.

Bazerman, M. H., & Moore, D. A. (1994). *Judgment in Managerial Decision Making*. New York: Wiley.

Brockner, J., Paruchuri, S., Idson, L. C., & Higgins, E. T. (2002). Regulatory focus and the probability estimates of conjunctive and disjunctive events. *Organizational Behavior and Human Decision Processes*, *87*, 5-24.

Cohen, J., Chesnick, E. I., & Haran, D. (1972). A confirmation of the INERTIAL-ψ effect in sequential choice and decision. *British Journal of Psychology*, *63*, 41-46.

Costello, F. J. (2009). Fallacies in probability judgments for conjunctions and disjunctions of everyday events. *Journal of Behavioral Decision Making*, *22*, 235-251.

Gigerenzer, G., & Todd, P. M. (2001). *Simple heuristics that make us smart*. New York: Oxford University Press.

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: causal maps and Bayes nets. *Psychological Review*, *111*, 3-32.

Hamilton, D. L., & Gifford, R. K. (1976). Illusory correlation in interpersonal perception: A cognitive basis of stereotypic judgments. *Journal of Experimental Social Psychology*, *12*, 392-407.

Kahneman, D. (2011). *Thinking, Fast and Slow*. London: Macmillan.

Lunt, P. K. (1988). The perceived causal structure of examination failure. *British Journal of Social Psychology*, *27*, 171-179.

Lunt, P. K. (1991). The perceived causal structure of loneliness. *Journal of Personality and Social Psychology*, *61*, 26-34.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, *92*, 289-316.

Peterson, C. R., & Beach, L. R. (1967). Man as an intuitive statistician. *Psychological Bulletin*, *68*, 29-46.

Rottman, B. M., & Keil, F. C. (2012). Causal structure learning over time: Observations and interventions. *Cognitive psychology*, *64*, 93-125.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926-1928.

Steyvers, M., Tenenbaum, J. B., Wagenmakers, E. J., & Blum, B. (2003). Inferring causal networks from observations and interventions. *Cognitive Science*, *27*, 453-489.

Trueswell, J. C., Medina, T. N., Hafri, A., & Gleitman, L. R. (2013). Propose but verify: Fast mapping meets cross-situational word learning. *Cognitive Psychology*, *66*, 126-156.

Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General*, *134*, 552.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124-1131.

Tversky, A., & Kahneman, D. (1986). Judgment under uncertainty: Heuristics and biases. In H. R. Arkes & K. R. Hammond (Eds.), *Judgment and decision making: An interdisciplinary reader* (pp. 38-55). Cambridge, England: Cambridge University Press.

Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics*, *106*, 1039-1061.

Yu, R. Q., & Zhao, J. (2015). The persistence of the attentional bias to regularities in a changing environment. *Attention, Perception, & Psychophysics*, *77*, 2217-2228.