The Rationality of Frequency-Based Illusory Correlation

Abstract

When presented with two samples (a smaller sample from a Minority population and a larger sample from a Majority population) where some rare feature occurs at exactly the same rate in both samples, people reliably associate the rare feature with the Minority population and the frequent feature with the Majority population. This pattern is referred to as 'illusory correlation', reflecting the standard assumption that such associations are fundamentally irrational. In this paper we show that this assumption is incorrect, and demonstrate that this pattern of association linking rare features with the Minority and frequent features with the Majority (given a sample where those features occurred at the same proportion in both categories) is in fact normatively correct and follows a result in epistemic probability theory known as the 'Rule of Succession'. Building on this result, we present a new computational model of frequency-based illusory correlation, based on the Rule of Succession. We also discuss the implications of the Rule of Succession for our understanding of various other cognitive biases.

**keywords:** probability; rationality; biases; illusory correlation

The Rationality of Frequency-Based Illusory Correlation

Psychologists have long been interested in the apparent biases seen in people's probabilistic judgements under uncertainty: patterns that apparently represent systematic and erroneous deviations from the requirements of normative probability theory. One such pattern, referred to as 'illusory correlation', arises in probabilistic inference from samples, where people see reliable correlations between events although no such correlations exist in the observed sample data. For example: when presented with two samples (a smaller sample from a Minority population and a larger sample from a Majority population) where some rare feature occurs at exactly the same rate in both samples, people will reliably associate the rare feature with the Minority population and the frequent feature with the Majority population. Since the ability to accurately recognise associations or correlations is fundamental to intelligent behaviour, this bias has become an important research topic in the study of learning and conditioning (Alloy and Tabachnik, 1984), reasoning (Berman and Kenny, 1976; Podsakoff et al., 2003) and memory (Kao and Wasserman, 1993; Katagiri et al., 2007; Kareev, 1995b). Since illusory correlation is seen as a driving force behind stereotype formation, this is also a central topic of research in social psychology (Bar-Tal et al., 2013; Sherman et al., 2009; Kutzner and Fiedler, 2015). Finally, since illusory correlation is seen as a systematic cognitive error, its occurrence has implications for ongoing debate on the rationality or irrationality of human thought processes (see, e.g. Tversky and Kahneman, 1974; Kahneman and Tversky, 1996; Kareev, 1995a).

The name 'illusory correlation' reflects the standard assumption that such differential judgements of association, given samples in which there is no difference in rate of occurrence, are fundamentally irrational and represent "erroneous inferences about the relationship between categories of events"(Hamilton and Gifford, 1976, p.392).

Our primary aim in this paper is to show that this is assumption is incorrect. Instead, when we apply normative probability theory (and in particular, a rule from epistemic probability theory known as the 'Rule of Succession') to standard illusory correlation tasks we find that the mathematically correct and rational response is to conclude that the rare feature *is* more likely in the Minority population; that there *is* a correlation between feature occurrence and category membership.

The organisation of this paper is as follows. In the first section we give a general overview of illusory correlation research. In the second section we address the rationality of illusory correlation, describe the Rule of Succession and its consequences, and show that this rule can explain many experimental results on illusory correlation and various related biases in probabilistic reasoning. In the third section we present a new computational model of the cognitive processes behind judgements of association, based on the application of the Rule of Succession in a noisy-reasoning system. In the final section we discus the question of rationality more generally, and consider the connections between illusory correlation, the Rule of Succession, and discrimination and societal prejudice.

## Background: illusory correlation

There are two main strands of research on illusory correlation: expectancy-based (Hamilton and Rose, 1980) and frequency-based (Hamilton and Gifford, 1976). Research on expectancy-based illusory correlation investigates the relationship between people's prior expectations (their preconceptions) and their interpretation of new data. Research in this area typically finds that people's prior expectation reliably influences their perception of correlation between features in a given set of data, even when they are explicitly told to disregard all prior knowledge when assessing the data. For example,

when presented with a dataset of drawings produced by patients in the psychiatric draw-a-picture test, each paired with the patient's diagnosis, diagnosticians reported enhanced associations between certain features (e.g. head shape) and related diagnoses (e.g. worries about intelligence), even though no such association existed in the data (Chapman, 1967). Such expectancy-based correlations have been reliably observed in psychiatric and clinical diagnosis, in organisational trait perception, and in personality and self-perception (see Fiedler, 2000, for an excellent review).

Research on expectancy-based illusory correlation does not necessarily address the question of rationality in people's judgement of association: people's expectations may represent rational Bayesian priors, and the influence of such expectations on judgements of association in a given set of data may be rationally correct. A central and striking result in frequency-based illusory correlation, however, does speak directly to the question of rationality in this area. This result was first identified in an influential study by Hamilton and Gifford (1976), where participants read a set of statements describing the positive or negative behaviours of members of one of two novel categories of people: a Majority group, which occurred often in the set, and a Minority group, which occurred less often. Participants were told that these statements were sampled at random from the categories in question, and were asked to make various judgements about the Minority and Majority categories based on the sample of statements seen. Importantly, the sample of statements was specifically controlled so that negative behaviours occurred at exactly the same proportion in both categories, as did positive behaviours.

In Hamilton and Gifford's first experiment positive behaviours were frequent in the sample and negative behaviours were rare, while in their second experiment negative behaviours were frequent and positive behaviours were rare. In both experiments participants tended to reliably associate the rare feature (negative behaviour in the first

Table 1

*Proportion and count of features samples from Minority and Majority categories in Hamilton and Gifford (1976) Experiment 1, alongside participant choice probabilities and normative population probabilities.*

| Feature Sample Proportion | | Sample Count | | Participant choice probability | | Normative population probability (Rule of Succession) | |
|---|---|---|---|---|---|---|---|
| | | Majority | Minority | Majority | Minority | Majority | Minority |
| Freq. | 0.69 | 18 | 9 | 0.66 | 0.56 | 0.68 | 0.67 |
| Rare | 0.31 | 8 | 4 | 0.34 | 0.44 | 0.32 | 0.33 |
| | | 26 | 13 | 1.0 | 1.0 | 1.0 | 1.0 |

*Note.* Participant choice probabilities were calculated from choice frequency counts in Hamilton and Gifford (1976). Participants' choice probabilities show illusory correlation (regression) associating the rare feature with the minority category and the frequent feature with the majority category. This association is stronger for the smaller Minority sample. Normative population probabilities calculated from the Rule of Succession show the same pattern.

experiment, positive behaviour in the second experiment) with the Minority and the frequent feature with the Majority, with the association between rare feature and Minority being stronger than that between frequent feature and Majority (Hamilton and Gifford, 1976; Fiedler, 1991; McConnell et al., 1994; Mullen and Johnson, 1990). Table 1 shows the count of rare/frequent features in Minority and Majority samples in one of these experiments, alongside one measure of association from that experiment (this data is from Hamilton and Gifford, 1976, Experiment 1). This type of association between rare features and the Minority is shown in measures of frequency estimation (people estimate the rare feature as occurring more frequently in the Minority than the Majority); of cued recall (when asked to recall which category was associated with which feature, people recall the Minority when cued with the rare feature); and of category evaluation (when the rare feature has negative valence, people judge the Minority more negatively, while when the rare feature has positive valence, people judge the Minority more positively). This association between features and categories in such studies does not depend on prior expectations about categories (because the categories are novel) or on prior expectations about features (because the same relationship between rare features and the minority category holds irrespective of the semantic content of the features).

The observed stronger association between rare features and the Minority leads social psychologists to view illusory correlation as fundamental to stereotype formation and societal patterns of discrimination and prejudice against minorities (e.g. Smith and Alpert, 2007). This is because encounters with members of minorities are, by definition, less frequent than encounters with members of the majority for the population at large, and because there are good reasons to assume that negative behaviours are less frequent than positive behaviours. This pattern is not limited to social categories, however, but

occurs in just the same way when categories are shapes and features are colours (Primi and Agnoli, 2002), when the categories are letters and the features are shapes (Fiedler and Armbruster, 1994), or when categories are different types of coins and features are marks on those coins (Kareev, 2000).[1]

This pattern of association can be described in terms of regression towards 0.5, the center-point of the probability scale: in illusory correlation, participants' judgements are regressive (closer to 0.5) compared to the observed sample proportions, with this regression being stronger for judgements about the Minority than the Majority (e.g. Fiedler and Unkelbach, 2014). We see just this pattern in Hamilton and Gifford's data in Table 1, for example. This pattern of regression seems to hold, not just for the two-category Majority/Minority situation investigated by Hamilton and Gifford (1976), but in situations involving multiple categories. Van Rooy et al. (2013), for example, found similar patterns of illusory correlation in an experiment where participants saw statements describing 4 categories of people (A, B, C, D) with different numbers of members of each group (24 members of category A, 12 members of category B, etc.). Positive behaviours occurred more often for each category (frequent features) and negative behaviours occurred less often (rare features). Results showed a similar regressive pattern of illusory correlation, with regression effects being weaker when the size of the sample is large, and being stronger when the size of the sample is small (see Table 2).

The underlying assumption in research on frequency-based illusory correlation is that this pattern of association is fundamentally irrational, and represents a deviation

[1]Note, however, that this associative relationship only holds when participants are presented with an actual sample of items: if people are simply shown a contingency table as in Table 1, this pattern of illusory correlation does not occur.

Table 2

*Proportion and count of features in samples from categories A, B, C and D in Van Rooy et al. (2013), Experiment 1, alongside participant choice probabilities and normative probabilities.*

| Feature Sample Proportion | Sample Count | | | | Participant choice probability | | | | Normative population probability (Rule of Succession) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | A | B | C | D | A | B | C | D |
| Freq. 0.67 | 16 | 8 | 4 | 2 | 0.62 | 0.53 | 0.53 | 0.62 | 0.65 | 0.64 | 0.63 | 0.75 |
| Rare 0.33 | 8 | 4 | 2 | 0 | 0.38 | 0.47 | 0.47 | 0.38 | 0.35 | 0.36 | 0.37 | 0.25 |
| regression | | | | | 0.05 | 0.14 | 0.14 | 0.38 | 0.02 | 0.03 | 0.04 | 0.25 |

*Note.* For the smallest sample, D, the rare feature never occurs: has a sample proportion of 0. The 'Regression' row shows the difference between participants' choice probabilities for features and the actual proportion in the observed sample (first block), and similarly the difference the normative probability for features and the actual proportion in the observed sample (second block). Participants' estimates show illusory correlation (regression), with the strength of the effect increasing as sample size falls. Normative population probabilities calculated from the Rule of Succession show the same pattern.

from normatively correct patterns of probabilistic reasoning. The reasoning seems to go like this: in constructing a given sample the experimenter has explicitly selected a rate of occurrence for features, and has explicitly constructed samples so that this rate is exactly the same in the minority and majority category samples. The experimenter then assumes that this equal rate of feature occurrence across minority and majority categories in the sample should be taken, by participants, to imply an equal rate of feature occurrence across the minority and majority populations from which the samples were drawn. The assumed normative response, then, is for participants to see no association between feature occurrence and category membership. When people judge the rare feature as being more associated with the Minority and the frequent feature with the Majority, this is taken to represent a deviation from rationality. In the next section we show that this assumption is incorrect.

## Rational inference in frequency-based correlation

Studies of frequency-based illusory correlation ask participants to make probabilistic inferences: given a sample from a certain population, what relationships hold in the population overall. How should such inferences be made? More specifically, given a sample of size $N$ drawn from a given population, and given that some event $A$ occurs $K$ times in that sample, what is the optimal, normative, estimate for the population probability $P(A)$ (the probability of $A$ in the population from which the sample was drawn)? In work on illusory correlation the implicit assumption seems to be that the normatively correct probability estimate for $A$, given a sample of $N$ items containing $K$ $A$'s, is equal to the sample proportion

$$Pr(A) = \frac{K}{N} \tag{1}$$

This assumption implies that, if we have two samples (a Majority and a Minority) with the same sample proportion $Pr(A)$, then the correct inference is that those samples were drawn from populations with equal population probabilities $P(A)$: that there is no correlation between feature and sample category.

This assumption is fundamentally incorrect. To see why informally, consider an extreme case, where you are shown a sample of 2 items (neither of which are instances of $A$) that come from one population, and a sample of 20 items (none of which are instances of $A$) that come from another population. The sample proportions in both cases are $Pr(A) = 0$. Is it correct to conclude that $A$ has a probability of 0 in both populations? Obviously not: the fact that $A$ did not occur in a sample of 2 items does not allow us to conclude that $A$ will never occur in the population from which the sample was drawn. Is it correct to conclude, on the basis of these two samples, that $A$ has the same probability in both populations? Again, obviously not: for example, $P(A) = 0.25$ could reasonably hold in the first population (the probability of drawing a sample of 2 items neither of which are $A$, from a population where $P(A) = 0.25$, is $(1 - 0.25)^2 = 0.56$; a more than 50% chance), but $P(A) = 0.25$ is extremely unlikely to hold in the second population (the probability of drawing a sample of 20 items, none of which are $A$, from a population where $P(A) = 0.25$, is $(1 - 0.25)^{20} = 0.003$; a less than 1% chance).

Given a sample of $N$ items, $K$ of which are $A$, what, then, is the optimal normative estimate for the population probability of $A$? We first investigate this in a direct and simple way, via computational simulation (see Algorithm 1). For a given sample size $N$ we run this simulation by calling the function PROBABILITIES$(N)$. This function loops $10,000$ times, on each cycle randomly picking a value for the population probability $p = P(A)$ of some event $A$ ($p$ is drawn uniformly from the range $0 \ldots 1$ inclusive). On each cycle the function SAMPLE$(p, N)$ then draws a sample of $N$ items

---

**Algorithm 1** Average generating probability for samples of $N$ events containing $k$ $A$'s

---

**function** SAMPLE(p,N)  // $p$:probability of $A$; $N$:sample size

    $k \leftarrow 0$

    **for** $i \leftarrow 1$ to $N$ **do**

        $q \leftarrow$ uniform random number in $[0\ldots1]$  // generate random number.

        **if** $q < p$ **then**  // event $A$ has occurred.

            $k \leftarrow k + 1$  // increment count.

        **end if**

    **end for**

    **return** $k$  // return number of $A$'s in sample.

**end function**


**function** PROBABILITIES(N)  // $N$: sample size

    **for** $k \leftarrow 0$ to $N$ **do**  // $P_k$: list of probabilities that

        $P_k \leftarrow [\,]$  // generate samples containing $k$ $A$'s.

    **end for**

    **for** $i \leftarrow 1$ to $10,000$ **do**  // For each cycle, geforncerate a random

        $p \leftarrow$ uniform random number in $[0\ldots1]$  // probability $p$ for event $A$.

        $K \leftarrow$ SAMPLE$(p, N)$  // Draw a random sample.

        $P_K \leftarrow p$  // Add generating probability $p$

    **end for**  // to list $P_K$

    **for** $K \leftarrow 0$ to $N$ **do**

        $Pr \leftarrow K/N$  // $Pr$: Sample proportion.

        $P \leftarrow$ MEAN$(P_k)$  // $m$: mean generating probability

        PRINT$(K, Pr, P)$  // for samples containing $K$ $A$s.

    **end for**

**end function**

---

from the population with $p = P(A)$, by randomly picking $N$ values $q$, drawn uniformly from the range $0 \ldots 1$ inclusive: cases where $q < p$ are counted as an instance of event $A$ in our sample. SAMPLE$(p, N)$ then returns the number of cases which were counted as an instance of event $A$ in the drawn sample. For each $K$ from 0 to $N$ the function PROBABILITIES$(N)$ has an associated storage list $P_K$: on each cycle of our simulation where the drawn sample contains $K$ instances for event $A$, we add the probability $p = P(A)$ which generated that sample to the associated storage list $P_K$. Each list $P_K$ thus holds the set of population probabilities $p = P(A)$ which generated samples of $N$ events containing $K$ instances of $A$. After running this simulation for $10,000$ cycles, we then display the average generating probability that produced samples of size $K = 0, K = 1, \ldots K = N$. This average generating probability represents the optimal estimate for the underlying population probability $P(A)$, given an observed sample of size $N$ that contains $K$ instances of $A$. A reasoner who makes this estimate for the underlying probability given the observed sample will be, on average, closest to the true population probability that generated the observed sample.

Table 3 shows the output from this simulation for values $N = 16, 8, 4$. It is clear from this Table that, for a given sample of $N$ items containing $K$ instances of event $A$, the average probability $P$ that generated that sample differs from the sample proportion $Pr = K/N$. Specifically, the average generating probability (and so the normatively correct and optimal estimate for the population probability, given the sample in question) is regressive towards 0.5 with the degree of regression increasing as the sample size $N$ falls. For some rare feature $A$ (where $Pr(A) < 0.5$), this means that the normative population probability estimate for $A$, given a smaller sample, will be greater than the normative estimate for $A$, given a larger sample, even where $Pr(A)$ is the same across both samples: the rare feature will be more associated with the population from

Table 3

*Average generating probability P and sample proportion Pr = K/N for samples of size N containing K instances of some feature A, as generated by Algorithm 1.*

| N=16 | | | N=8 | | | N=4 | | |
|---|---|---|---|---|---|---|---|---|
| K | Pr | P | K | Pr | P | K | Pr | P |
| 0 | 0.00 | 0.06 | 0 | 0.00 | 0.10 | 0 | 0.00 | 0.17 |
| 4 | 0.25 | 0.28 | 2 | 0.25 | 0.30 | 1 | 0.25 | 0.33 |
| 8 | 0.50 | 0.50 | 4 | 0.50 | 0.50 | 2 | 0.50 | 0.50 |
| 12 | 0.75 | 0.72 | 6 | 0.75 | 0.70 | 3 | 0.75 | 0.67 |
| 16 | 1.00 | 0.94 | 8 | 1.00 | 0.90 | 4 | 1.00 | 0.83 |

*Note.* For a given sample of size $N$ containing $K$ instances of $A$, the average generating probability $P$ represents the average probability of $A$ in the population from which the sample was drawn; and thus represents the optimal normative estimate for that population probability, based on the presented sample. 0.32Note that these optimal population probability estimates precisely follow the Rule of Succession: $P = (K + 1)/(N + 2)$.

which the smaller sample was drawn. This is just the pattern of association seen in people's judgements in frequency-based illusory correlation results. We cannot, therefore, describe these patterns of association as illusory: they reflect a normatively correct inference from the presented samples to feature/category associations.[2]

## The Rule of Succession

Why do these patterns of association occur? They follow from a well-known rule in epistemic probability theory, known as Laplace's Rule of Succession. This rule states that the normative estimate for the probability of $A$ in a population, given a sample of $N$ events drawn from that population which contains $K$ occurrences of $A$ (and no further information beyond that sample) is actually given by

$$P(A) = \frac{K+1}{N+2} \tag{2}$$

This expression has been proved in various different ways, with the strongest and most general proof being given by De Finetti (1937). As Zabell (1989), in a very interesting presentation of the history and various proofs of the rule of Succession, notes, "[I]n order to attack [De Finitti's proof] one must attack the formidable edifice of epistemic probability itself. Modern philosophy continues to ignore it at its own peril". The various mean generating probabilities shown in Table 3 precisely follow this rule.

This rule undermines the standard assumption in frequency-based illusory correlation research, which is that when participants judge the rare feature to be more likely in the Minority than the Majority (and the frequent feature to be less likely), they are reasoning incorrectly and are demonstrating a systematic bias away from the requirements of normative reasoning. In fact, when we apply the Rule of Succession to

---

[2]For consistency we will continue to refer to these patterns of association as 'illusory correlations'. This naming is intended as a label rather than a description: these 'illusory correlations' are not illusory.

Hamilton and Gifford's standard illusory correlation materials we find that the normatively correct response is to conclude that the rare feature *is* more likely (more probable) in the Minority population than in the Majority population; and similarly to conclude that the frequent feature *is* less probable in the Minority population than in the Majority population (see Rule of Succession values in Table 1). Similarly, when we apply this Rule to Van Rooy et al.'s 4-category illusory correlation materials, we find that the normatively correct inferred probabilities show the same pattern of regression and illusory correlation seen in people's estimates in that experiment, with the rare feature having a higher normative population probability for categories with smaller samples than for those with larger samples, and with a pattern of increasing regression as sample size decreases (see Rule of Succession values in Table 2).

It is worth noting that, in this Rule of Succession account of illusory correlation, the effect will arise frequently for randomly selected samples from any population containing a minority and a majority group, even when there is in fact no correlation between group membership and feature occurrence. To see this consider, for example, the situation shown in Table 4 where we have two large sub-populations of some overall population $U$ such that the Minority makes up 20% of $U$ and the Majority makes up 80%, and where some rare feature occurs at exactly the same rate of 25% in both Minority population and Majority population, so there is no association or correlation between categories and features in $U$. Suppose we draw a random sample of 20 items from $U$. Since we are taking a random sample from Bernoulli-distributed variables (features are either present or absent: an item is either a member of a category or is not), the count of occurrence of categories and features in our sample will follow the Binomial distribution. In the Binomial distribution the most likely number of items in a sample is simply equal to the sample size times the population probability of that item:

so a random sample of 20 items from this universe will most likely contain $0.2 \times 20 = 4$ items from the Minority population, of which $0.25 \times 4 = 1$ will have the rare feature, and $0.8 \times 20 = 16$ items from the Majority population, of which $0.25 \times 16 = 4$ will have the rare feature (see left block in Table 4). The population probabilities inferred from these most likely samples via the normatively correct Rule of Succession, however, indicate a correlation between category membership and feature occurrence (see right block in Table 4). So the *most likely* sample from a population with no correlation between features and categories will lead a rational reasoner to infer that there *is* a correlation between features and categories in that population. Though this can appear almost paradoxical, there is no paradox here. Instead, this result is simply a reflection of the fact that, even if $S$ is the sample most likely to be drawn from a population $U$, this does not imply that $U$ is the most likely population from which sample $S$ was drawn.

**Degree of bias**

Comparing the level of regression in normative probabilities produced by the Rule of Succession against that the level of regression in participants' probability estimates (Tables 1 and 2) we see that, while participants' estimates were regressive in the normatively correct direction (that required by the Rule of Succession) producing apparent illusory correlation, the level of regression was higher in participants responses than in normatively correct population probabilities. Given this, a natural question is: to what extent do people's judgements in frequency-based correlation tasks follow the normative probability values specified by the Rule of Succession? If people are asked to estimate the probability of some feature $A$ which occurs $K$ times in a sample of size $N$, to what extent do their estimates agree with the normative population probability $P(A) = (K + 1)/(N + 2)$? The literature on frequency-based illusory correlation does

Table 4

*Most likely sample of* 20 *items from a population U in which there is no correlation between features and categories. Also shown are optimal population probability estimates inferred from this sample via the Rule of Succession.*

| Feature | Most likely sample of 20 items from $U$ | | Population probability of features inferred from this sample (Rule of Succession) | |
| --- | --- | --- | --- | --- |
| | Majority (80% of $U$) | Minority (20% of $U$) | Majority | Minority |
| Frequent (75% of $U$) | 12 | 3 | 0.72 | 0.67 |
| Rare (25% of $U$) | 4 | 1 | 0.28 | 0.33 |
| | 16 | 4 | 1.0 | 1.0 |

*Note.* In this example we assume that in population $U$ there is no association between features and categories: $P(Frequent \mid Majority) = P(Frequent \mid Minority) = P(Frequent\ feature) = 0.75$, and $P(Rare \mid Majority) = P(Rare \mid Minority) = P(Rare\ feature) = 0.25$ in $U$. The most likely the category and feature count for a random sample of 20 items from this population $U$ is most shown in the left block. The normatively correct population probabilities inferred from most likely sample of 20 items from $U$, however, show a relatively strong association between category and feature occurrence.

not, as far as we can see, give us any way of answering this question. There are two main problems. First, experiments in this literature measure correlation in various different ways (frequency estimation, cued recall, positive or negative evaluation of categories). These measures allow us to see whether participants form associations between features and categories in such frequency-based tasks (they allow us to test against the null hypothesis of no association, which researchers have assumed to be the normatively correct response in frequency-based tasks). However, they do not provide any natural way to test against the normative hypothesis, which is that judgements of association should approximate the relationship described by the Rule of Succession. This is because it is not clear how differences in frequency estimation, cued recall, or category evaluation should be related to differences in estimated population probability.

A second problem arises because the degree of association between features and categories in these experiments is typically measured in terms of averages across participants. In testing against the hypothesis that there is no association, this use of averages across participants is appropriate. In testing against the hypothesis that there is a certain degree of regression, whose extent is described by the Rule of Succession, however, this use of averages is problematic. This is because averages in general are themselves regressive, due to participant variability; and this regression, arising solely due to averaging across participants, confounds attempts to measure or estimate other forms of regression in judgement. Given these two problems, future work on people's judgement of frequency-based correlation should focus on responses at the individual level, and should use tests that allow direct comparison between participant responses and the normative requirements given by the Rule of Succession.

**Explaining illusory correlation results**

The two most general patterns seen in frequency-based illusory correlation are consistent with the Rule of Succession: that judgements are regressive (closer to 0.5) compared to observed sample proportions, with this regression being stronger for judgements about the Minority than the Majority, and with this pattern holding, not just for the two-category situation investigated by Hamilton and Gifford (1976), but in situations involving multiple categories (Van Rooy et al., 2013). Here we very briefly describe a number of other consistent results.

The degree of illusory correlation in people's inference from samples is related to sample size, with more illusory correlation (more regression) for smaller Minority samples than for larger Majority samples. This observation suggests that the degree of illusory correlation in people's responses will increase in cases where the Minority or Majority categories are 'split' into subcategories. A series of studies by Fiedler and Armbruster (1994) investigated this by examining illusory correlation for standard Majority and Minority samples and for equal-size subcategories of those samples (minority$_A$ and minority$_B$, or majority$_A$ and majority$_B$). Across these studies Fiedler and Armbruster found that illusory correlation effects were reliably stronger for these smaller subcategories than for the original categories, and that the strength of this effect was inversely related to category size. Again, this follows the Rule of Succession, in which degree of regression depends on sample size.

One aspect of the Rule of Succession is that the smaller the sample, the stronger degree of regression in probabilistic inference from that sample. In the limit, this means that regression (illusory correlation) is maximised when sample size is at its smallest: that is, with a sample size of 1. This is consistent with results on 'one-shot illusory correlation', where we find that people form strong associations between minority group

membership and rare behaviour after presentation of a single sample item of a member of the minority exhibiting that rare behaviour (Risen et al., 2007).

Illusory correlation is often presented in terms of inference from group membership to feature occurrence (inferring that members of the Minority group are more likely to have a certain rare feature). Hamilton and Gifford (1976), however, investigated the occurrence of illusory correlation in both in terms of inference from group to feature and from feature to group. In both of their experiments Hamilton and Gifford asked participants two questions, one about group membership, given behaviour (given negative behaviour, which group did the person come from?), and one about behaviour, given group membership (given a person is a member of the Minority, what is the probability of negative behaviour?). Both forms of question revealed the same pattern of illusory correlation (of regression depending on sample size). This pattern is consistent with the Rule of Succession, in which illusory correlation arises solely due to frequency of co-occurrence, and so applies equally to inference from category to feature and from feature to category.

The Rule of Succession also has implications for accounts of other biases in probability estimation or judgement, especially those which involve probability judgement from samples. One such bias is that of underconfidence in probability estimation: the finding that people's estimates for the probability of an event from an observed sample (for which the sample proportion is known) tend to be systematically biased away from the sample proportion and towards 0.5 (that is, systematically regressive, compared to the sample proportion). Erev et al. (1994), for example, found this pattern in a study where participants played a video game and then estimated the probability of different events in that game: participants reliably overestimated the probability of low-probability events and underestimated that of high-probability events.

Lichtenstein et al. (1978) found this pattern in a series of studies where participants estimated the probability of different causes of death: participants reliably overestimated low frequency causes and underestimated high-frequency causes. Teigen (1973) found this pattern in a study where participants estimated the frequency of occurrence of a given symbol in a presented sequence: participants reliably overestimated the occurrence of rare symbols and underestimated the occurrence of frequent symbols (for similar results see e.g. Poulton, 1973; Erlick, 1964). The direction of this bias follows the regressive pattern of the Rule of Succession.

## An alternative model of illusory correlation

Research on frequency-based illusory correlation has worked on the assumption that, when participants judge a rare feature (which occurs at the same rate in a smaller Minority sample and a larger Majority sample) to be correlated with membership in the Minority category, those participants are demonstrating systematic error. Our primary aim has been to show that this assumption is mistaken: when we apply the Rule of Succession to standard frequency-based illusory correlation materials we find that the correct response is to conclude that the rare feature *is* more likely in the Minority population.

In this section we turn to a secondary aim: that of presenting a new theoretical model for illusory correlation, based on the Rule of Succession. There are currently two major theoretical accounts of frequency-based illusory correlation. The first, Hamilton and Gifford's 'distinctiveness' account (e.g. Hamilton and Gifford, 1976; McConnell et al., 1994; Sanbonmatsu et al., 1987), has a privileged position because it was presented in the seminal paper by Hamilton and Gifford (1976) that first described frequency-based illusory correlation effects. In this account the observed association

between the rare feature and the Minority is assumed to arise due to the enhanced distinctiveness of the most infrequent combination: Minority members who have rare features. Such infrequent and distinctive category/feature combinations are assumed to be more salient and available to recall; so probability estimates for such combinations are enhanced due to this memory advantage, producing illusory correlation effects. While this proposal was widely accepted (leading to use of the term *distinctiveness based* illusory correlation to refer to these frequency-based results), and is still given in textbooks as the standard account of this form of illusory correlation, empirical evidence in favour of this account is somewhat weak, with a range of research showing no memory advantage for rare combinations, and with frequency-based illusory correlation arising in cases where the distinctiveness account does not seem to apply (Fiedler, 2000, 1991; Kutzner and Fiedler, 2015). To give just one illustration, Table 2, for example, shows significant regression - the characteristic feature of illusory correlation - for the rare feature in the smallest category $D$, even though no instances of that feature in that category were seen by participants and so no memory advantage could apply.

The distinctiveness account of illusory correlation is based on a particular selective bias in recall from memory. An alternative account, sometimes referred to as an 'information loss' or 'incomplete learning' account, assumes no such bias. Such accounts have been implemented computationally in various different ways (e.g. Smith, 1991; Fiedler, 2000, 1991). Speaking generally, there are two core proposals in these accounts. The first is that learning increases with the number of trials, and so the proportion of feature occurrence is better learned for the larger Majority category than for the smaller Minority category. The second proposal is that incomplete learning, or 'information loss' for the Minority category produces regression in responses, explaining illusory correlation results. This is consistent with the standard statistical idea that any error in

measurement of the relationship between two variables will produce regression (Fiedler and Unkelbach, 2014).

Both theories of these are presented as explanations for why people infer associations between rare features and minority categories from samples where, by assumption, those associations are not present. Since our results show that, in fact, it is normatively correct to infer such associations from such samples, these theories must take on a different explanatory role: rather than standing as attempts to explain a pattern of deviation from normative requirements, these theories instead stand as possible explanations for why people's judgements in illusory correlation tasks follow the normative pattern specified by the Rule of Succession (at least to a first approximation, in terms of direction of regression and increased regression with smaller sample size).

Our results lead us to propose an alternative theoretical account, where illusory correlation arises because the cognitive mechanisms of probabilistic inference are in some way designed or constructed to follow the normatively correct Rule of Succession for probabilistic inference from samples, but where these cognitive mechanisms are affected by random noise or error. We have three primary reasons for making this proposal. First, the ability to make correct probabilistic inferences (and so correct decisions) is central to successful intelligent behaviour in an fundamentally unpredictable world (Brunswik, 1955): the Rule of Succession describes normative probabilistic inference from samples, and so we would *a priori* expect people's probabilistic inference to follow this rule in some way. Second and more specifically, considering Tables 1 and 2 we see that the patterns seen in people's illusory correlation results follow the probabilities computed via the Rule of Succession, both in terms of regression relative to the sample proportion and in terms of stronger regressive effects for smaller (Minority) samples. These results are consistent with this theoretical

proposal. Third, this proposal would place illusory correlation within a theoretical framework in which a range of other systematic biases and errors in probabilistic reasoning are explained in terms of normatively correct mechanism subject to random noise (e.g. the 'probability theory plus noise' model; Costello and Watts, 2014, 2017, 2016a; Costello et al., 2018; Costello and Watts, 2018).

While the general patterns seen in frequency-based illusory correlation are consistent with the Rule of Succession, Tables 1 and 2 suggest that the degree of regression (the strength of these effects) in experimental studies of illusory correlation is larger than that produced by that Rule (subject to the qualifications about measurement described in the previous section). To account for this difference in degree of regression it is enough to note that many forms of unbiased random error will produce regression in probability estimation. We illustrate this point using a computational simulation where judgement of the probability of an event in a given category is computed by applying the rule of Succession to instances of that category recalled from memory, but there is a constant, unbiased rate $f$ of random recall failure (so that if there are $N$ instances of some category in memory, on average some random subset $(1-f)N$ will be recalled). Such random error will reduce the sample size used when applying the Rule of Succession to calculate event probability (if a sample of 10 items was seen, but the forgetting rate is 0.5, only 5 items from the sample will be recalled and used in probability calculation). As we saw earlier, regressive effects are greater for smaller sample sizes, and so random forgetting errors will increase regression levels and magnify illusory correlation effects.

A simulation of this 'Rule of Succession + random error' model is described in Algorithm 2. In this algorithm, function ESTIMATE($K, N, f$) takes as input a parameter $N$, representing sample size, a parameter $K$, representing the number of sampled items

with feature $A$, and a forgetting rate $f$ representing the chance of an sampled item being randomly forgotten. For each instance in the category sample the function draws a random number between 0 and 1: if that number is greater than $f$ then that instance is recalled, otherwise it is 'forgotten'. The function keeps track of the number, $m$, of instances which are recalled, and the number, $c$, of recalled instances which contain the feature $A$. The function applies the Rule of Succession to these counts $m$ and $c$, to produce a probability estimate for the feature $A$ in the population the category sample was drawn from. Since forgetting is random, this ESTIMATE will return randomly varying population probability estimates for $A$. The function AVERAGEESTIMATE($K, N, f$) draws $10,000$ of these estimates and returns their average, giving the mean estimate for the population probability of $A$ given a sample of $N$ events containing $K$ $A$'s, and a forgetting rate of $f$.

Table 5 shows the probabilities produced by this 'Rule of Succession + random error' model for the sample counts used in Hamilton and Gifford (1976), and Table 6 shows the probabilities produced for the sample counts used in Van Rooy et al. (2013). The random forgetting error rate was set at $f = 0.8$ for the data in Table 5, and $f = 0.6$ for the data in Table 6 (these values were selected to illustrate approximate agreement between simulated probabilities and participants' probability judgements). Comparing with the normative Rule of Succession probabilities given in Tables 1 and 2, we see that these simulated probabilities follow the overall illusory correlation pattern, but with a larger degree of regression: so giving results closer to participants estimated probabilities in those experiments. These results suggest that illusory correlation results such as those seen by, for example, Hamilton and Gifford (1976) and Van Rooy et al. (2013) can be explained by assuming that probabilistic inference is a cognitive process that in some way implements the rationally correct Rule of Succession, but that this

---

**Algorithm 2** 'Rule of Succession + random error' probability estimation model.

```
                                          // N: items in category, k items
  function ESTIMATE(k,N, f)               //  with feature, f: forgetting rate
     m ← 0                                //  m: count of recalled items
     c ← 0                                //  c: count of recalled items
     for i ← 1 to N do
        q ← uniform random number in [0...1]   // generate random number.
        if q > f then                     // This instance is recalled.
           m ← m + 1                      // increment count of items.
           if i ≤ k then                  // this instance contains feature.
              c ← c + 1                   // increment count of features.
           end if
        end if
     end for
     return (c + 1)/(m + 2)               // return probability estimate.
  end function                            // (Rule of Succession).


  function AVERAGEESTIMATE(K,N,f)
     P ← [ ]                              // holds a list of estimates
     for i ← 1 to 10,000 do
        P ← ESTIMATE(K, N, f)            // Add new estimate to list.
     end for
     return AVERAGE(P)                    // Return average estimate
  end function
```

---

process is subject to random, unbiased noise or error in recall.

Table 5

*Proportions, feature counts, and participant choice probabilities from Hamilton and Gifford's Experiment 1 (as in Table 1), alongside probabilities produced via the Rule of Succession in a simulation with a constant, unbiased 'random forgetting error rate' of $f = 0.8$.*

| Feature Sample Proportion | | Sample Count | | Participant choice probability | | Simulation: Rule of Succession + random error | |
|---|---|---|---|---|---|---|---|
| | | Majority | Minority | Majority | Minority | Majority | Minority |
| Freq. | 0.69 | 18 | 9 | 0.66 | 0.56 | 0.64 | 0.60 |
| Rare | 0.31 | 8 | 4 | 0.34 | 0.44 | 0.36 | 0.40 |

*Note.* Recall that normative probability values follow the illusory correlation pattern of association seen in people's judgements in this experiment (Table 1). 'Rule of Succession + random error' probabilities also follow this pattern, but with an increased level of regression (more closely matching participant's choice probabilities).

Table 6

*Proportions, feature counts, and participant choice probabilities for features in categories A, B C and D in Van Rooy et al. (2013), Experiment 1 (as in Table 2), alongside population probabilities produced via the Rule of Succession + random error simulation ($P_f$) with a constant, unbiased random forgetting error rate of $f = 0.6$.*

| Feature Sample Proportion | Sample Count | | | | Participant choice probability | | | | Simulation: Rule of Succession + random error | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | A | B | C | D | A | B | C | D |
| Freq. 0.67 | 16 | 8 | 4 | 2 | 0.62 | 0.53 | 0.53 | 0.62 | 0.63 | 0.60 | 0.58 | 0.61 |
| Rare 0.33 | 8 | 4 | 2 | 0 | 0.38 | 0.47 | 0.47 | 0.38 | 0.37 | 0.40 | 0.42 | 0.39 |
| regression | | | | | 0.05 | 0.14 | 0.14 | 0.38 | 0.04 | 0.07 | 0.09 | 0.39 |

*Note.* Recall that normative probability values follow the illusory correlation pattern of association seen in people's judgements in this experiment (Table 2). 'Rule of Succession + random error' probabilities also follow this pattern, but with an increased level of regression (more closely matching participant's choice probabilities).

**Implementing the Rule of Succession: working memory and noise**

Of course, the idea that people estimate probabilities in a way that explicitly follows the Rule of Succession (adding 2 to the denominator and 1 to the numerator in sample-count ratios to produce probability estimates) seems cognitively implausible. Standard models of memory (e.g. Dougherty et al., 1999; Bearden and Wallsten, 2004) typically assume that event probability is measured proportionally, in terms of the strength of response from memory-traces to a probe, with higher strength (and so higher probability) when the event is frequent in memory and so many traces respond, and lower strength (lower probability) when the event is rare and few traces respond. Such proportional responses do not give access to numerator and denominator values directly (only to their ratio), and so the Rule of Succession cannot be explicitly applied.

There are, however, cognitively plausible mechanisms that can produce or approximate the Rule of Succession for such proportional measurements. For example, consider the 'probability theory plus noise' model (Costello and Watts, 2014, 2016a, 2018), which assumes that people estimate the probability of some event $A$ by randomly sampling events (or traces) in memory and returning the proportion of instances of $A$ in the sample (the strength of response to $A$ in the sample). Since processing is subject to various forms of random error or noise, this model assumes that events, or memory traces, have some chance $d < 0.5$ of responding incorrectly to a probe: there is a chance $d$ that a $\neg A$ (*not A*) trace will incorrectly respond as $A$, and the same chance $d$ that an $A$ trace will incorrectly respond as $\neg A$. Taking $Pr(A)$ to be the sample proportion of $A$ (the proportion of traces containing $A$), with this form of noise the chance of a randomly sampled event trace responding as $A$, and so the expected average value for a

noisy probability estimate of $P(A)$, is

$$P(responds\ as\ A) = (1 - d)Pr(A) + (1 - Pr(A))d = (1 - 2d)Pr(A) + d \qquad (3)$$

with individual estimates varying independently around this expected value.

This expression is regressive towards $0.5$ relative to sample proportions $Pr(A)$. Are there values of $d$ for which this expression will produce regression equivalent to that required by the Rule of Succession? We can answer this by setting

$$(1 - 2d)Pr(A) + d = (1 - 2d)\frac{K}{N} + d = \frac{K + 1}{N + 2}$$

and solving for $d$. Some algebraic manipulation shows that this equality holds when

$$[d(N + 2) - 1](N - 2K) = 0$$

and so we see that if the rate of noise in individual memory-traces in a sample is

$$d = \frac{1}{N + 2} \qquad (4)$$

then the noisy probability estimate obtained by measuring response strength (the sample proportion) will be, on average, equal to the correct, normative population probability (calculated via the Rule of Succession). An extension of the simulation described in Algorithm 2 tested this proposal by modifying the function ESTIMATE$(K, N, f)$ so that, rather than returning the Rule of Succession value calculated from feature and category counts (after random forgetting), it returned the feature-to-category proportion after both random forgetting and noisy counting (with noise rate as in Equation 4). This simulation produced average probability estimates that were essentially the same as those produced by the Rule of Succession calculations shown in Tables 5 and 6, supporting the idea that the Rule of Succession's regressive effects can be produced by random noise. While random noise is typically seen as

detrimental to reasoning, leading to unwanted variation and error, this result suggests an interesting possibility: that random noise could play a functional and adaptive role in probabilistic reasoning, biasing responses in a way that follows the normatively correct requirements of the Rule of Succession.

**Consistency with previous results**

As this model implements the Rule of Succession, it inherits that Rule's account for various illusory correlation effects (effects of sample size, effects of category splitting, symmetry between features and categories, 'one-shot' illusory correlation, etc). In addition, this 'Rule of Succession + random error' account also predicts a relationship between cognitive load in general (and memory load in particular) and degree of illusory correlation. This is because the higher the rate of random forgetting, in this model, the smaller the recalled samples and so the higher the rate of regression due to the rule of succession. This is consistent with a range of results showing that the degree of illusory correlation in people's judgements does increase with cognitive and memory load (see Fiedler et al., 1993, for a review).

This 'Rule of Succession + random error' model may explain other biases in probability estimation or judgement, especially those which involve probability judgement from samples. One such bias is that of underconfidence in probability estimation: the finding that people's estimates for the probability of an event from an observed sample (for which the sample proportion is known) tend to be systematically biased away from the sample proportion and towards 0.5 (that is, systematically regressive, compared to the sample proportion). Erev et al. (1994), for example, found this pattern in a study where participants played a video game and then estimated the probability of different events in that game: participants reliably overestimated the

probability of low-probability events and underestimated that of high-probability events. Lichtenstein et al. (1978) found this pattern in a series of studies where participants estimated the probability of different causes of death: participants reliably overestimated low frequency causes and underestimated high-frequency causes. Teigen (1973) found this pattern in a study where participants estimated the frequency of occurrence of a given symbol in a presented sequence: participants reliably overestimated the occurrence of rare symbols and underestimated the occurrence of frequent symbols (for similar results see e.g. Poulton, 1973; Erlick, 1964). The direction of this bias follows the regressive pattern of the Rule of Succession, and so, just as above, can be explained by assuming that participants are following a rational and normatively correct process for probability estimation, but that this process is subject to random noise.

The relationship between the Rule of Succession and the regressive effects of noise may also explain another bias: people's tendency to exaggerate the degree to which small samples resemble the population from which they are drawn (referred to by Tversky and Kahneman, 1971, as the 'belief in the law of small numbers'). We know that random noise is present throughout the nervous system (e.g. Faisal et al., 2008). Given a certain base rate of unavoidable random noise in probabilistic reasoning, the above argument suggests an optimal corresponding sample size for probability estimation (and hence for optimal decision-making under uncertainty). Inverting Equation 4, we see that if the average rate of random noise in event responses is $d$, then the optimal sample size for probability estimation (the size for which the sample probability will, on average, match the normatively correct population probability given by the Rule of Succession) is

$$N = \frac{1}{d} - 2 \tag{5}$$

As this equation shows, the optimal sample size $N$ is inversely proportional to the noise

rate $d$ (as the rate of noise increases, the optimal sample size falls), and so will tend to be small for even relatively low rates of noise. For values of $d$ around $d = 0.1$ – an estimate for the median rate of noise in probability estimation from previous experimental results (Costello and Watts, 2016a) and computational simulations (Costello and Watts, 2017) – the optimal sample size $N$ from Equation 5 is around 8, for example. More generally, this equation shows that, in a noisy environment, the average sample probability obtained from a small sample size will tend to resemble the true population probability more closely than that obtained from a larger sample: a result which suggests that people's tendency to exaggerate the degree to which small samples resemble the population from which they are drawn (their 'belief in the law of small numbers') may be explained as a rational response to random noise.

It is useful to give an example illustrating this effect. Suppose we have a sample of 8 events drawn from some population, of which 2 are instances of $A$ (a sample probability of 0.25). According to the Rule of Succession, the optimal estimate for the population probability of $A$ is

$$\frac{2+1}{8+2} = 0.3$$

Suppose also that there is a rate of random noise or error $d = 0.1$ in identifying or counting these items (one time in ten we will mistakenly count an $A$ as $\neg A$ or count a $\neg A$ as $A$). Given this the average noisy sample probability (from Equation 3) is

$$(1 - 2d)0.25 + d = 0.8 \times 0.25 + 0.1 = 0.3$$

and this average sample probability equals the optimal population probability estimate. Now, by contrast, suppose we have a sample of 32 events drawn from some population, of which 8 are instances of $A$ (again, a sample probability of 0.25). According to the

Rule of Succession, the optimal estimate for the probability of $A$ in this population is

$$\frac{8+1}{32+2} = 0.26$$

while the average noisy sample probability (assuming a noise rate of $d = 0.1$) is, as before,

$$(1 - 2d)0.25 + d = 0.8 \times 0.25 + 0.1 = 0.3$$

Comparing these results we see that the average sample probability matched the optimal estimate for the population probability for the smaller sample size $N = 8$, while it significantly overestimates the optimal estimate for the larger sample size $N = 32$: sample probability estimates produced from the smaller sample better resemble the population probability of the population from which they are drawn, possibly explaining people's belief in the 'law of small numbers'.

As an aside, note that the optimal sample size for noise rates of around $d = 0.1$ is close to Miller's 'magical number $7 \pm 2$' for working memory capacity (Miller, 1956). This leads to the (very speculative) idea that the capacity of working memory may, in part, be dictated by the unavoidable presence of random noise: for decision-making under uncertainty (that is, for probabilistic reasoning) in a system with a rate of noise around $d = 0.1$, the optimal sample size for probabilistic judgement (that is, the sample size for which noise moves sample probability estimates towards normatively correct population probabilities) is somewhere in the region $7 \pm 2$. A somewhat similar argument, proposing that working-memory size around $7 \pm 2$ in various ways optimises the chances for detection of reliable correlations, is made by Kareev (2000).

### Discussion: rationality and illusory correlation

In frequency-based illusory correlation tasks people associate rare features with the Minority and frequent features with the Majority, even though these features occur at

the same rate in both categories in observed samples. We've shown that this pattern of association is consistent with normative probability theory and so is rational and correct. Readers may have concerns over our use of 'rationality' here, since this word has many different meanings (referring to utility maximisation in some contexts, internal consistency in others, agreement with normatively correct rules in still others, and so on) and has been the subject of vigorous debate for at least the last 50 years (see e.g. Sturm, 2012). This question of the 'rationality' of illusory correlation is by no means academic. Recall that frequency-based illusory correlation is seen by social psychologists as fundamental to stereotype formation, discrimination, and prejudice against minorities (e.g. Smith and Alpert, 2007). This is because illusory correlation will tend to associate negative behaviours with people who are members of a minority (since negative behaviours are typically rare, and illusory correlation causes people to associate rare features with smaller samples). This relationship between sample size and negative association is consistent with extensive social psychology research showing that discrimination or prejudice against a given group is inversely related to level of contact: the less contact you have with members of a minority group (that is, the smaller your sample size), the more likely you are to be prejudiced against that group. In a meta-analysis of 515 studies investigating prejudice and group contact, for example, Pettigrew and Tropp (2006) found that 94% of studies "show an inverse relationship between contact and prejudice".

As we have shown, this type of association between rare features and smaller samples follows directly from epistemic probability theory, purely as a consequence of differences in sample size. This suggests that prejudice against minorities may arise in rational and unbiased reasoners, simply as a consequence of reduced contact with members of a minority group. Does our argument, then, mean that such discrimination

is rational? To answer this question, it is important to be clear about the precise meaning of 'rationality' in the context of illusory correlation. The focus in illusory correlation research is on people's tendency to give (apparently) erroneous judgements of association. A rational reasoner, in this context, is one who gives answers that minimise error: in other words, answers are most likely to be correct and to reflect the true probability of association. The Rule of Succession shows that that, when a reasoner is required to make judgements about association in a population based on a sample from that population with equal feature distribution but unbalanced category membership ( fewer Minority members in the sample) and when that reasoner has no further information beyond that sample, then the mathematically correct and hence *rational* response is to judge the rare feature as more likely in the Minority than the Majority population. Importantly, however, while the Rule of Succession provides an explanation for such differential judgement of minority and majority groups (in terms of rational reasoning from unbalanced samples), it does not provide any justification for persistence in such judgement. On the contrary, the Rule of Succession tells us that differential judgements of minority and majority groups are reliably influenced by category imbalance in samples; the rational response is to be aware of this influence (a form of 'metacognitive' rationality;  Petty et al., 2007) and to therefore draw future samples in a way that is systematically biased towards the inclusion of minority members at a rate higher than the population rate.[3] A sampling process that is biased towards the inclusion of the minority is rational (again, in the sense of minimising error, or producing answers that are most likely to be correct) because this sampling process will equalise information available about minority and majority populations, and so allows a more accurate and unbiased assessment of the relationship between group

_____

[3]We'd like to thank Klaus Fiedler for this point about metacognitive rationality.

membership and feature occurrence.

Of course, this 'minimising error' view of rationality in probabilistic reasoning is only one possible perspective on human judgement. Numerous other perspectives are available, which give different views on the problem of illusory correlation. For example, a utility-maximisation analysis would require a rational reasoner to assign a cost to making errors in judgements of association, and also to assign a cost to drawing further samples from the minority and the majority population: rational behaviour, in such a perspective would depend on the relationship between these three costs. The relationship between illusory correlation and utility maximization is a important topic for future work.

### Conclusions

Our primary aim in this paper has been to show that the pattern of association seen in frequency-based illusory correlation studies, widely seen as a significant and systematic error in human reasoning, is not an error. Instead, this pattern of association follows the normatively correct Rule of Succession, both in terms of the regressive direction of judgements and in terms of the differential influence of sample size on this regression. Further, the Rule of Succession is consistent with a range of other experimental results on frequency-based illusory correlation. Conclusions about the irrationality of human judgements of association derived from this frequency-based illusory correlation research are thus called into question.

Given this, our secondary aim has been to present an alternative theoretical model of people's probabilistic inference in frequency-based illusory correlation tasks. This model sees human probabilistic reasoning as implicitly implementing the normatively correct Rule of Succession from probability theory, via the regressive effects of random

noise. This model thus represents an application of a more general 'probability theory plus noise' account, which sees human probabilistic reasoning as based on normatively correct frequentist probability theory, but subject to random noise that causes systematic biases (or in this case, produces responses that follow the Rule of Succession). In previous work we have shown that this 'probability theory plus noise' account explains various different patterns of bias in people's probability estimates for both direct and conditional probabilities, and makes a number of novel predictions about patterns of bias and agreement with probability theory for various probabilistic expressions: predictions which are supported by experimental results (see Costello and Watts, 2014, 2016b, 2017, 2016a; Costello et al., 2018; Costello and Watts, 2018). The illusory correlation results we describe here extend this account, and suggest that many aspects of human probabilistic judgement may be explained if we assume that people's reasoning is based on frequentist probability theory, but is subject to random noise or error.

It is important to stress here that we are not suggesting people are consciously aware of the equations of probability theory (or, indeed of the Rule of Succession) when judging probabilities. Instead our proposal is that people's probability judgements are derived from a 'black box' that estimates probabilities by retrieving (some analogue of) a count of instances from memory. Such a mechanism is necessarily subject to the requirements of set theory and therefore embodies the rules of probability theory, and is also subject to random errors in counting and recall, which produce systematic biases in probabilistic reasoning. We see this probability module to be unconscious, automatic, rapid, parallel, relatively undemanding of cognitive capacity and evolutionarily 'old'. Support for this view comes from that fact that people make probability judgements rapidly and easily and typically do not have access to the reasons behind their

estimations, from extensive evidence that event frequencies are stored in memory by an automatic and unconscious encoding process (Hasher and Zacks, 1984) and from evidence suggesting that infants have surprisingly sophisticated representations of probability (Cesana-Arlotti et al., 2012). Other support comes from results showing that animals effectively judge probabilities (for instance, the probability of obtaining food from a given source) and that their judged probabilities are typically close to optimal (Kheifets and Gallistel, 2012).

In terms of its application to illusory correlation, this approach is similar in some ways to Fiedler's 'information loss' account of illusory correlation (Fiedler, 1991, 2000), because both accounts give a central role to regression (caused by some form of random error) in explaining frequency-based illusory correlation. The primary theoretical difference between these two accounts lies in the functional role that regression plays. In Fiedler's information loss account, regression plays a negative role, causing bias and error in people's probabilistic reasoning and moving people's judgement away from the normatively correct responses. In our account, by contrast, regression can play a positive role, correcting proportional probability estimates and moves them towards the normatively correct 'Rule of Succession' value. We hope that an understanding of the Rule of Succession will contribute to the ongoing debate in psychology on the rationality or irrationality of human probabilistic reasoning.

References

Alloy, L. B. and Tabachnik, N. (1984). Assessment of covariation by humans and
animals: The joint influence of prior expectations and current situational information.
*Psychological Review*, 91(1):112–149.

Bar-Tal, D., Graumann, C. F., Kruglanski, A. W., and Stroebe, W. (2013). *Stereotyping
and prejudice: Changing conceptions.* Springer Science & Business Media.

Bearden, N. J. and Wallsten, T. S. (2004). Minerva-DM and subadditive frequency
judgments. *Journal of Behavioral Decision Making*, 17(5):349–363.

Berman, J. S. and Kenny, D. A. (1976). Correlational bias in observer ratings. *Journal
of Personality and Social Psychology*, 34(2):263–273.

Brunswik, E. (1955). In defense of probabilistic functionalism: A reply. *Psychological
Review*, 62:236–242.

Cesana-Arlotti, N., Téglás, E., Bonatti, L. L., et al. (2012). The probable and the
possible at 12 months: intuitive reasoning about the uncertain future. *Advances in
Child Development and Behavior*, 43:1–25.

Chapman, L. J. (1967). Illusory correlation in observational report. *Journal of Verbal
Learning and Verbal Behavior*, 6(1):151–155.

Costello, F. and Watts, P. (2014). Surprisingly rational: Probability theory plus noise
explains biases in judgment. *Psychological Review*, 121(3):463–480.

Costello, F. and Watts, P. (2016a). People's conditional probability judgments follow
probability theory (plus noise). *Cognitive Psychology*, 89:106–133.

Costello, F. and Watts, P. (2016b). Probability theory plus noise: replies to Crupi and Tentori (2015) and to Nilsson, Juslin and Winman (2015). *Psychological Review*, 123(1):112–123.

Costello, F. and Watts, P. (2017). Explaining high conjunction fallacy rates: the probability theory plus noise account. *Journal of Behavioral Decision Making*, 30(2):304–321.

Costello, F. and Watts, P. (2018). Invariants in probabilistic reasoning. *Cognitive psychology*, 100:1–16.

Costello, F., Watts, P., and Fisher, C. (2018). Surprising rationality in probability judgment: Assessing two competing models. *Cognition*, 170:280–297.

Costello, F. J. and Mathison, T. (2014). On fallacies and normative reasoning: when people's judgements follow probability theory. In *Proceedings of the 36th annual meeting of the Cognitive Science Society*, pages 361–366.

De Finetti, B. (1937). La prévision: ses lois logiques, ses sources subjectives. volume 7, pages 1–68.

Dougherty, M. R. P., Gettys, C. F., and Ogden, E. E. (1999). Minerva-DM: A memory processes model for judgments of likelihood. *Psychological Review*, 106(1):180–209.

Erev, I., Wallsten, T. S., and Budescu, D. V. (1994). Simultaneous over- and underconfidence: The role of error in judgment processes. *Psychological Review*, 101(3):519–527.

Erlick, D. E. (1964). Absolute judgments of discrete quantities randomly distributed over time. *Journal of Experimental Psychology*, 67(5):475–482.

Faisal, A. A., Selen, L. P., and Wolpert, D. M. (2008). Noise in the nervous system. *Nature Reviews neuroscience*, 9(4):292–303.

Fiedler, K. (1991). The tricky nature of skewed frequency tables: An information loss account of distinctiveness-based illusory correlations. *Journal of Personality and Social Psychology*, 60(1):24–36.

Fiedler, K. (2000). Illusory correlations: A simple associative algorithm provides a convergent account of seemingly divergent paradigms. *Review of General Psychology*, 4(1):25–58.

Fiedler, K. and Armbruster, T. (1994). Two halfs may be more than one whole: Category-split effects on frequency illusions. *Journal of Personality and Social Psychology*, 66(4):633.

Fiedler, K., Russer, S., and Gramm, K. (1993). Illusory correlations and memory performance. *Journal of Experimental Social Psychology*, 29(2):111–136.

Fiedler, K. and Unkelbach, C. (2014). Regressive judgment: Implications of a universal property of the empirical world. *Current Directions in Psychological Science*, 23(5):361–367.

Hamilton, D. L. and Gifford, R. K. (1976). Illusory correlation in interpersonal perception: A cognitive basis of stereotypic judgments. *Journal of Experimental Social Psychology*, 12(4):392–407.

Hamilton, D. L. and Rose, T. L. (1980). Illusory correlation and the maintenance of stereotypic beliefs. *Journal of Personality and Social Psychology*, 39(5):832–845.

Hasher, L. and Zacks, R. (1984). Automatic processing of fundamental information: the case of frequency of occurrence. *The American Psychologist*, 39(12):1372–1388.

Kahneman, D. and Tversky, A. (1996). On the reality of cognitive illusions.

Kao, S.-F. and Wasserman, E. A. (1993). Assessment of an information integration account of contingency judgment with examination of subjective cell importance and method of information presentation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(6):1363–1386.

Kareev, Y. (1995a). Positive bias in the perception of covariation. *Psychological review*, 102(3):490.

Kareev, Y. (1995b). Through a narrow window: Working memory capacity and the detection of covariation. *Cognition*, 56(3):263–269.

Kareev, Y. (2000). Seven (indeed, plus or minus two) and the detection of correlations. *Psychological Review*, 107(2):397–402.

Katagiri, M., Kao, S.-F., Simon, A. M., Castro, L., and Wasserman, E. A. (2007). Judgments of causal efficacy under constant and changing interevent contingencies. *Behavioural Processes*, 74(2):251–264.

Kheifets, A. and Gallistel, C. R. (2012). Mice take calculated risks. *Proceedings of the National Academy of Sciences*, in press.

Kutzner, F. L. and Fiedler, K. (2015). No correlation, no evidence for attention shift in category learning: Different mechanisms behind illusory correlations and the inverse base-rate effect. *Journal of Experimental Psychology: General*, 144(1):58–75.

Lichtenstein, S., Slovic, P., Fischhoff, B., Layman, M., and Combs, B. (1978). Judged frequency of lethal events. *Journal of Experimental Psychology: Human Learning and Memory*, 4(6):551–578.

McConnell, A. R., Sherman, S. J., and Hamilton, D. L. (1994). Illusory correlation in the perception of groups: an extension of the distinctiveness-based account. *Journal of Personality and Social Psychology*, 67(3):414–429.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 63(2):81–97.

Mullen, B. and Johnson, C. (1990). Distinctiveness-based illusory correlations and stereotyping: A meta-analytic integration. *British Journal of Social Psychology*, 29(1):11–28.

Pettigrew, T. F. and Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of personality and social psychology*, 90(5):751.

Petty, R. E., Briñol, P., Tormala, Z. L., and Wegener, D. T. (2007). The role of meta-cognition in social judgment. *Social psychology: Handbook of basic principles*, 2:254–284.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., and Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5):879–904.

Poulton, E. (1973). Unwanted range effects from using within-subject experimental designs. *Psychological Bulletin*, 80(2):113–121.

Primi, C. and Agnoli, F. (2002). Children correlate infrequent behaviors with minority groups: A case of illusory correlation. *Cognitive Development*, 17(1):1105–1131.

Risen, J. L., Gilovich, T., and Dunning, D. (2007). One-shot illusory correlations and stereotype formation. *Personality and Social Psychology Bulletin*, 33(11):1492–1502.

Sanbonmatsu, D. M., Sherman, S. J., and Hamilton, D. L. (1987). Illusory correlation in the perception of individuals and groups. *Social Cognition*, 5(1):1–25.

Sherman, J. W., Kruschke, J. K., Sherman, S. J., Percy, E. J., Petrocelli, J. V., and Conrey, F. R. (2009). Attentional processes in stereotype formation: a common model for category accentuation and illusory correlation. *Journal of Personality and Social Psychology*, 96(2):305–323.

Smith, E. R. (1991). Illusory correlation in a simulated exemplar-based memory. *Journal of Experimental Social Psychology*, 27(2):107–123.

Smith, M. R. and Alpert, G. P. (2007). Explaining police bias: A theory of social conditioning and illusory correlation. *Criminal Justice and Behavior*, 34(10):1262–1283.

Sturm, T. (2012). The ?rationality wars? in psychology: Where they are and where they could go. *Inquiry*, 55(1):66–81.

Teigen, K. H. (1973). Number and percentage estimates in sequential tasks. *Perceptual and Motor Skills*, 36(3 suppl):1035–1038.

Tversky, A. and Kahneman, D. (1971). Belief in the law of small numbers. *Psychological bulletin*, 76(2):105–110.

Tversky, A. and Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *science*, 185(4157):1124–1131.

Van Rooy, D., Vanhoomissen, T., and Van Overwalle, F. (2013). Illusory correlation, group size and memory. *Journal of Experimental Social Psychology*, 49(6):1159–1167.

Zabell, S. L. (1989). The rule of succession. *Erkenntnis*, 31(2):283–321.