Benford's Law: Testing the Effects of Distributions and Anchors on Number Estimation

Duyi Chi (dchi6051@uni.sydney.edu.au)
Bruce D. Burns (bruce.burns@sydney.edu.au)
University of Sydney, School of Psychology, Brennan MacCallum Building (A18)
Camperdown, NSW 2006, Australia

Abstract

Recent research (e.g., Burns & Krygier, 2015; Chi & Burns, 2022) demonstrated that people could exhibit a strong bias towards the smaller first digits, which is consistent with the pattern predicted by Benford's law. However, this psychological phenomenon was predominantly observed when generating meaningful numbers for decision-making. We investigated explanations rooted in the statistical acquisition of distributional information and the impact of anchoring during number estimation. Undergraduate students were asked to estimate the weight, lifespan and group-size of animals after learning different distributions of these variables, supplied with an anchored value, either an average or a starting point, for reference. The Benford bias reasonably emerged regardless of the variable distribution, yet was strongly influenced by the anchored information. Notably, showing average values significantly suppressed Benford bias. These findings offered insights into the cognitive process of number estimation in the presence of statistical evidence and anchored information.

Keywords: Benford's law, number estimation, anchoring, statistical learning

Introduction

Decision-making and judgement under uncertainty have been investigated traditionally by looking for heuristics and biases (Tversky & Kahneman, 1974). However other research has demonstrated that people's behaviour could follow statistical regularities. For example, the numerical prediction of everyday events often aligns with accurate prior probabilities (e.g., Griffith & Tenenbaum, 2006). Recent psychological research (Burns & Krygier, 2015) has found a bias that is consistent with a well-established regularity regarding the first digits of data, Benford's law (BL). This law proposes that first-digit frequencies follow a log distribution where digit-1 occurs 30% of the time while digit-9 has no more than 5% occurrence (Benford, 1938). Such a distribution has been widely discovered to apply to datasets from numerous domains in naturally occurring settings, like finance indicators, mathematical topics, and physical observations. This principle regarding the first digits posits that in any numerical domain covering a wide range of values without confined limits, the leading digits of its data are expected to follow a logarithmic distribution that decreases monotonically. Hence, it has been incorporated into the auditing and accounting process and used for detecting falsified data (Miller, 2015).

Recent studies (e.g., Burns & Krygier, 2015; Diekmann, 2007) found that people could spontaneously generate a first-digit bias that approximates Benford's law when producing unknown numbers, such as the length of a river or national debt of a country. They did not find a perfect fit of human data to BL, with deviations like spikes at digit-5 sometimes, but its pattern accounts for a large amount of variance in human first-digit data. Such findings were extended to other tasks using visual stimuli, for example, estimating the quantities of jellybeans in a jar or dots in displays (Chi & Burns, 2022). In these cases, a similar trend was observed: a strong preference for smaller leading digits (i.e., 1, 2, 3) and a monotonic decline as the first digit increases, suggesting a Benford bias in human number production.

Therefore, this behavioural alignment with BL, a phenomenon typically observed in naturally occurring data, has raised interesting questions. We do not expect that people will perfectly fit to Benford's law, instead, we can try to measure the size of a Benford bias towards the first digit distribution suggested by Benford's law. The existence of this bias leads to two questions: 1) What are the processes leading to Benford bias in human behaviour; 2) Can Benford bias help explain other phenomena of number estimation?

Possible Explanations of Benford's Bias

To understand why people could have a Benford bias, several studies have sought to provide explanations based on theories of learning, speculating that individuals might have been exposed to this statistical pattern during their lifetime, given its ubiquitous presence in various datasets.

Tests of the Recognition Hypothesis. Theories of statistical learning, well-established in the laboratory, suggest that people can quickly grasp statistical relationships, like conditional probability (Fiser & Aslin, 2002), even with minimal exposure. This forms the basis for the Recognition Hypothesis, which posits that regular exposure to data consistent with Benford's Law (BL) might lead to an unconscious sensitivity towards the statistical pattern of *first digits*. If the bias for smaller first digits is indeed a result of unconscious acquisition due to long-term exposure in the environment, this preference should extend beyond tasks of number generation to also include tasks where participants shall *recognize* and *select* the numerical answers.

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

To test the Recognition Hypothesis, Burns and Krygier (2015) introduced a selection task into one of their studies. This task presented numerical questions similar to those used in their generation tasks (e.g., national debt, region population, etc.). However, instead of generating a number as a response, participants were asked to choose an answer from nine numerical options, each with a different first digit. If people have Benford bias due to long-term exposure in their environment, the options with lower first digits should be selected more frequently than those with higher first digits. Contrary to this prediction, except for a slight elevation for digit-1, the relative frequencies of the first digits chosen by participants were close to a flat distribution. Unlike their number generation tasks, people appeared to have no strong preference for any first digits when selecting answers. However, it was argued that providing too many choices might demotivate individuals to make a rational judgment by exhibiting undesired behaviour (Iyengar & Lepper, 2000)

Chi and Burns (2022) argued that a better approach might be to directly contrast the first digits. So, they tested the Recognition Hypothesis with what was considered a much more sensitive paradigm for a selection task by directly pitting a smaller and a larger FSD of a number against each other. For example, a forced-choice item was offered between 2xx and 8xx where x is a random digit. If people implicitly learned that the lower first digits occurred more frequently than the higher ones due to exposure to this pattern in the environment, the forced-choice selected responses should have strongly favoured the number with smaller first digits. However, again their data did not show a systematic difference in preferences for any first digit.

The Benford bias, characterized by generating numbers with smaller leading digits, has been notably observed in tasks where participants generate responses to unfamiliar questions, such as estimating electricity consumption and jellybeans. In contrast, recognition tasks all show no such bias, with participants not displaying any differential first-digit patterns when selecting answers. This suggests that the Benford bias may be a product of the process by which people generate responses (Burns, 2009; Chi & Burns, 2022) rather than the sensitivity to the first-digit patterns. Thus we believe the Recognition Hypothesis has to be rejected. Consequently, it becomes essential to explore alternative explanations for the Benford bias, focusing on the cognitive processing involved in producing estimates.

An important implication of the lack of a Benford bias in selection tasks is that it shows that Benford bias arises only when people generate numbers. This puts important constraints on how it could be explained.

Optimal Statistical Inferences in Everyday Prediction. Although previous examinations failed to support the Recognition hypothesis, the idea that Benford bias is due to sensitivity to the first-digit distribution emergent from long-term exposure to this real-world relationship, remains intuitively appealing due to its alignment with other research on how learning statistical relationships in the environment can influence decision-making. A study by Griffith and Tenenbaum (2006) strongly supported the idea that people's cognitive judgment can be explained in terms of optimal statistical inferences that follow prior probabilities with evidence from estimation in real-life scenarios. Their experiments captured our attention not only because the estimated values were collected as outcomes, but also because of the identical domains of the testing items used, as those found in Benford bias. Their questions were selected from familiar activities in daily life, such as the movie gross and run time, poem lengths, lifespan, lengths of marriages, and baking time for cakes.

Griffith and Tenenbaum's (2006) studies examined how humans made predictions about events by asking them to predict numerical outcomes based on a pre-existing reference, such as estimating the lifespan of a man with a current age of 65. The findings showed that people's cognitive judgments align with optimal statistical inferences informed by accurate prior probabilities. It appears that people are aware that data come from different distributions, and their responses agree with the appropriate prior distribution depending on the context. For instance, when estimating a person's total lifetime, the responses were consistent with the use of the appropriate Gaussian prior; but, when predicting the total gross of a movie, the responses tended to follow a power-law prior.

The Distribution Hypothesis. Despite the failure of the Recognition Hypothesis to explain the observation of Benford bias in first-digit distributions, people also seem aware of how each variable should be distributed. Such conjecture, combined with the mathematical proposal that BL shall emerge from the numbers aggregated from the variables following the power law (Berger & Hill, 2015; 2020), directly build up the basis for establishing an alternative explanation, the Distributional Hypothesis.

Assuming that individuals can produce estimates consistent with a power law, Benford bias should emerge as anticipated. Given the mathematical evidence that power laws tend to follow Benford's law, while normally distributed data should not, the Distribution Hypothesis posits that the Benford bias arises from generating values from underlying distributions that follow a power law. Thus, assuming people can learn the distribution of variables, this makes a testable prediction that varying what people think the distribution of numbers is will affect the degree to which they fit BL.

Using Benford Bias to Examine Anchoring

One of the few behavioural phenomena regarding number estimation that has been well studied are anchoring effect (Kahneman and Tversky, 1974). In this phenomenon, the first piece of information encountered serves as an anchor or reference point and influences the subsequent estimates of a value, even if they are not related. For instance, recalling the last two digits of a social security number impacted the estimated cost for an item with an unknown value, in the group with higher social security numbers placed higher bids than those with lower digits (Ariely et al., 2003). The effect has been observed across various domains, including legal judgment (e.g., Enough & Mussweiler, 2001), purchasing decisions (e.g., Mussweiler, Strack & Pfeiffer, 2000), predicting and negotiation (e.g., Galinsky & Mussweiler, 2001).

When the anchoring effect was introduced by Kahneman and Tversky (1974) they described it as a result of insufficient adjustments from the initial value offered. When individuals are expected to generate an estimate with limited information, they tend to adjust the value from the initially encountered anchor. However, this effortless modification is inadequate, leading to estimates closer to the anchor than they should be. While some studies using self-generated anchors supported this theory (e.g., Epley & Gilovich, 2001), others argued that it only applies when the anchor is outside a reasonable range. For instance, most would agree that 153 is not the likely answer for Mahatma Gandhi's life expectancy, so people would adjust away from that number. However, when an anchor is within a plausible range or provided externally, the *Anchoring-and-Adjusting* model might not fully explain the outcome.

An alternative theory, *Selective Accessibili*ty, derived from confirmatory hypothesis testing, has emerged as another dominant explanation (e.g., Mussweiler & Strack, 1999). Specifically, it posits that people evaluate the target against the anchor value by actively generating information that aligns with the idea that the anchor is a plausible targeted value (referred to as the Selectivity Hypothesis). When it's not, they move to the next guess, yet not without examining all attributes of the anchor itself in the first place. This process of generating such information enhances its accessibility, thereby influencing its utilization in forming the ultimate absolute judgment (referred to as the Accessibility Hypothesis). Traditionally, it was believed that a more extreme anchor would produce a stronger anchoring effect, which is supported by both the *Anchoring-and-Adjusting* and *Selective Accessibility* models. However, Wegener and colleagues (2001) found that extremely high or low anchors did not produce a stronger effect than moderate anchors. While no single theory may fully capture the nuances of the anchoring effect, they collectively provide valuable insights into the cognitive processes associated with this bias in estimation.

As discussed above, evidence of a Benford bias has been found in several tasks when people have to generate estimated numerical answers to questions they are not sure of. Therefore, the presence of Benford bias could be used as an indicator of a number generation process. This could be important for understanding anchoring effects because the Anchoring-and-Adjusting and Selective Accessibility models propose different processes. Selective Accessibility appears to suggest that anchoring is no different to any other number generation process, just one that is biased by the information associated with the anchor. Therefore, selective accessibility predicts that the results of anchoring tasks should show a strong Benford bias. In contrast, Anchoring-and-Adjusting proposes that anchoring tasks involve a different process to that used when people normally generate numerical estimates, a process that is tightly constrained by a specific number. In general, evidence for Benford's law is less likely to be found when the possible answers are tightly constrained. Therefore, the *Anchoring-and-Adjusting* model predicts that the results of anchoring tasks should show a weak Benford bias.

The Present Study

The study presented here aimed to both test the Distribution Hypothesis as an explanation for Beford bias and, assuming that we continue to find evidence of Benford bias, to examine the effects of anchors on Benford bias. We tested the distribution hypothesis by varying the underlying distribution of variables participants tried to estimate, and we examined anchoring effects by giving near or far anchors.

To follow the paradigm of previous studies showing Benford bias, our tasks continued to ask for estimates of factual knowledge. When assessing the reliability of the Distribution Hypothesis, it was essential to offer the variables characterised by different shapes of distributions for contrast to see if individuals are capable of producing and following them accordingly. Hence, the attributes of animals were chosen for their distinct distribution characteristics: animal weight typically follows a normal distribution (Uchmański, 1983), lifespan is often left-skewed (Muradian, 1989), and group-size generally adheres to a power law (Griesser et al., 2011).

Similar to the standard paradigm in statistical learning research, prior to the testing phase, the exposure stage introduced the statistical data illustrating the distribution of each variable estimated, using human-related examples for clarity. A quiz was also supplied to ensure a basic understanding of the materials. As proposed by the distribution hypothesis, if individuals are sensitive to the underlying distribution of a variable, we expect to observe Benford bias when participants generate answers in response to questions characterised by a power law such as group-size, but not from Gaussian variables like weight.

To test the effects of anchoring on Benford bias two types of anchors were introduced: Average and Juvenile. For the *Average* type, when participants had to make an estimate, a number about the average value (e.g., the mean weight of a Megabat) was given. We anticipated participants generating answers close to the average, so it was usually effectively a near anchor which should suppress the Benford bias. For the *Juvenile* type of anchor, we gave participants juvenile information (e.g., the current age of a young sea otter). Given that participants were asked about adult animals, this effectively usually was a far anchor.

The effects of both anchored value and statistical regularities to answer questions do not necessarily have to be mutually exclusive. Each direction of tests simply offers a unique viewpoint that highlights a possible mechanism when making reasonable estimates. The task of Animal Estimation in this experiment was a 3 (variables: weight vs. lifespan vs. group-size) × 2 (anchors: average vs. juvenile) mixed design.

Methods

Participants

236 first-year psychology students with an average age of 20.17 (*SD* = 2.335) provided valid responses. Most participants identified as female (65.68%), while males accounted for about 33.05%. The two most common native languages were English (46.61%) and Chinese (38.14%).

Procedure and materials

Each participant completed 27 estimation items in the task of Animal Estimation, which involves two stages. During the exposure phase, participants were introduced to visual examples representing three different statistical distributions, accompanied by thorough explanations to ensure comprehension of distribution concepts. These examples utilised human-related data, illustrating a normal distribution through human birth weight, a left-skewed distribution via female life expectancy, and a power law distribution through the size of WhatsApp groups. This was followed by a manipulation check with six quiz questions to ensure a basic level of understanding.

Here is a series of pictures of a Sea Otter. Please generate the estimates to Picture A only:

If the average weight of the Sea Otter is 25,000 g, please estimate the weight for this Sea Otter. ________ (g)

If the average lifespan of the Sea Otter is 4,500 days, please estimate the lifespan for this Sea Otter. ______(days)

If the average group size of the Sea Otters in the wild is 250, please estimate the group size for this Sea Otter. ________

Figure 1: Example of questions including average anchors for the sea otter in Picture A about its weight, lifespan, and group-size. Also shown are the four images depicting animals of various sizes and age groups used for this specific animal. Nine wild animals were selected: sea otters, Atlantic herrings, cape buffalo, red-winged blackbirds, bottlenose dolphins, plains garter snakes, Atlantic horseshoe crabs, Megabats, and greater flamingos.

In the subsequent testing phase, participants were presented with images of nine animals and asked to estimate for each their weight, lifespan, and group-size. However, they were randomly assigned to a question version containing an anchor related to either the average or the juvenile data. In the average anchor group, four distinct images were prepared for each animal, each depicting the animal at different sizes and ages. Among these, one image (labelled as A, B, C, or D) was randomly assigned to be presented to a participant (see Figure 1) showing average values. Those in the juvenile group were asked to predict the animal's mature weight, total lifespan, and final group-size based on a starting point given for each variable (see Figure 2). The first digits of anchored data for each variable were uniformly distributed.

It is a picture of a juvenile Megabat. Please give an estimate to the questions below:

If the weight of this young Megabat is 750 g, please estimate its weight once it matures.________ (g)

If this young Megabat has been living for 4,500 days, please estimate its total lifespan. ________ (days)

If 8,500 Megabats were observed at a point of time during their gathering in the wild, please estimate the final group size.

Figure 2: This is an example of questions including juvenile information regarding a young megabat's weight, lifespan, and group-size.

Results

Standard methods for testing goodness of fit, such as the chi-square (χ^2) test, provided a way to determine whether the datasets conform to a hypothetical model or not. However, in our study, we aimed to quantify the degree of fit rather than just obtaining a binary answer of yes or no. This was achieved by calculating the eta-squared (η^2) for the digit liner contrast weighted according to the proportions specified by BL. Having an η^2 measures how much variances can be explained by Benford's proportions, with a larger size indicating a stronger fit to the expected model.

A three-way interaction was found between the first digits, variables asked and anchor types, as indicated by the contrast analysis weighted by the proportions of BL, $F(1,169)= 7.462$, $p= .007$, $\eta^2 =.042$. Additionally, a main effect of anchor type on the first-digit distribution was reported, $F(1,169) = 198.728$, $p< .001$, $\eta^2 = .54$. Therefore, the following sections separately discussed the first-digit pattern observed for average and juvenile conditions.

First-digit Patterns When Given Juvenile Anchors

In the group where participants were given juvenile information, the first-digit distribution of the estimates for three variables asked visually displayed a reasonable monotonic decline (see Figure 3). The test of within-subject contrasts weighted by the proportions of BL indicated that the variances of estimates in three rounds were all reasonably explained by Benford's proportions (Weight: $\eta^2 = .767$; Lifespan: $\eta^2 = .734$; Group-size 3: $\eta^2 = .584$). It also reported a significant interaction between the variable questioned and the first digits, $F(1,64) = 4.482$, $p = .038$, $\eta^2 = .065$. Fluctuations were observed at digit-1, 2 and 8. Digit-5 was consistently elevated. Nevertheless, a strong Benford bias was consistently present across all three variables. Contrary to expectations, Benford bias was not disrupted as a result of introducing a Gaussian variable.

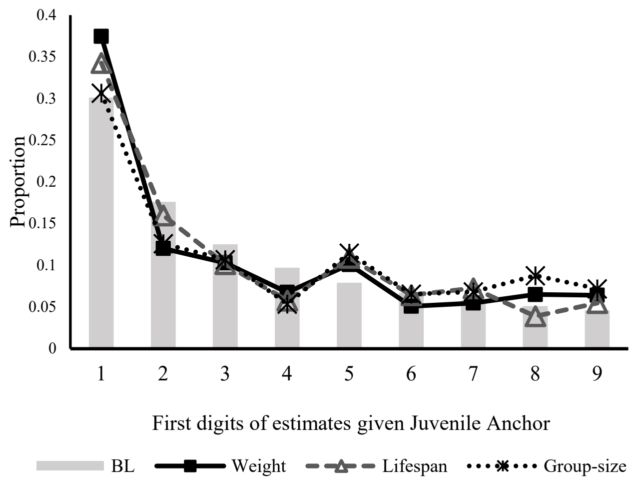

Figure 3: The first-digit proportions of the estimates in response to three types of variables questioned with Juvenile anchors, compared to BL.

First-digit Patterns When Given Average Anchors

Contrary to the pattern of data for juvenile anchors, when an average anchor was offered the first-digit distribution largely deviated from a monotonic decline (see Figure 4). The contrast analysis showed that the BL accounted for only a small portion of the variances for the three variables (weight: η^2= .081, lifespan: η^2= .297, group-size: η^2= .18). The first-digit patterns were not significantly different from each other reported by the within-subject contrast analysis weighted by the proportions of BL, $F(1,105) = .203$, p =.653. Unlike the juvenile group, it lacked clear evidence to support the presence of a Benford bias when showing an average anchor.

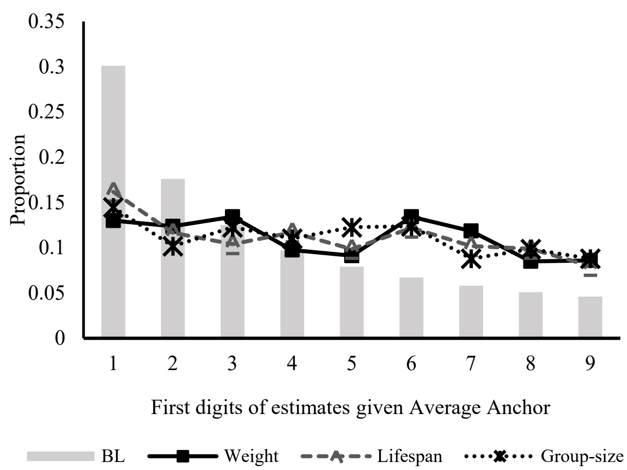

Figure 4: The first-digit proportions of the estimates in response to three types of variables questioned with Average anchors, compared to BL.

Analysis of Estimated Values

The distribution of the estimated values was assessed to see if people were capable of following the underlying distributions of the variables. In the case where participants provided estimates based on juvenile anchors, none of the distributions for the estimated answers conformed to normality. All twenty-seven distributions (3 variables × 9 animals) exhibited strong positive skewness based on a normality test. In contrast, normal distributions were more frequently detected in the group provided with average anchors. Five specific patterns of estimates, the group-size of herrings, the lifespan of blackbirds, the weight of dolphins, the lifespan of snakes, and both the age and group-size of crabs, were found to be closely normally distributed. Their skewness values fell within the -1 to 1 range, meeting the criteria for approximate normal distributions defined by Hair et al. (2022). Some estimates were even found to be negatively skewed, such as the estimates for the group-size of Garter snakes. It seems that the type of variable being estimated did not systematically affect the underlying distribution of the numerical responses; however, the nature of the anchored data provided did have an impact.

We also assessed the influence of anchors by measuring the differences between the mean estimated values and reference values provided. The group that received the average anchors showed that most estimated weights were not statistically different from the average data in the question. The data suggested that approximately 55.4% of the weight estimates, 69.9% of the lifespan estimates, and 65.9% of the group-size estimates were generated within 10% differences of the average anchor values provided in the questions. This implies that people are more likely to make minor adjustments when the average value is known. However, the group with juvenile anchors produced estimates that largely deviated from the starting point values offered in the questions, demonstrating a completely different estimation pattern. Our data additionally showed that, in the juvenile group, about 56.3% of the weight estimates, 64.7% of the lifespan estimates, and 47.8% of the group-size estimates were more than double the starting point value. More specifically, over a third of the estimates were a direct result of multiplying the starting value (e.g., 750) or the first digit of the starting value (e.g., 700) provided in the question by an integer, such as 2, 3, 4 or 10.

Discussion

Testing the Distribution Hypothesis

Unlike what we expected due to previous studies (e.g., Griffith & Tenenbaum, 2006; Lewandowsky, Griffiths & Kalish, 2009) which indicated that predictions are informed by accurate prior probabilities, our data did not show significant differences in the first-digit pattern across different variables with different distributions. The results indicated that even when participants were explicitly informed about the underlying distributions, they might not be influenced by those statistical relationships when solving unknown questions. For example, most estimated weights in our sample deviated strongly from a typical bell curve when they were supposed to be a Gaussian variable. Thus, the evidence collected so far challenged the utility of the Distribution Hypothesis's emphasis on sensitivity to the distributions of the estimated variables for explaining Benford bias, at least at the aggregated level.

Nevertheless, the absence of adherence to certain regularities does not necessarily mean participants had failed to utilize the knowledge they were given. There might be a gap between knowing and applying, an area still under-explored in Psychology. Examples of probability learning, focusing on behavioural changes over time, suggested that the different strategies used by adults and children in transitioning from probability matching to maximising rewards were a result of choices rather than the knowledge of associations (Plate, Shutts, Green & Pollak, 2018). Sensitivity to underlying distributions was often reported from responses estimated for everyday activities, like baking times (Griffiths & Tenenbaum, 2006). However, anomalies were also detected in highly unfamiliar situations, such as in predictions about pharaohs' reigns, where people's predictions failed to follow Bayesian calculations based on optimal statistical judgment. The questions asked in our study such as the weight or lifespan of wild animals may not be easily answered based on daily experiences, hence such deviation from the expected distribution might additionally offer insights into understanding the constraints of the utility of people's sensitivity to accurate prior probabilities.

The Effects of Anchors on Benford bias

The animal estimation task involved asking participants factual questions about variables like the weight and group-size of animals. While learning about the distributions of these variables showed limited impact, the type of anchor used significantly influenced estimation behaviours. From examining the first-digit patterns of estimated numbers of participants given juvenile anchors we found evidence of a strong Benford bias, whereas for the group given average anchors we did not. This suggests that the type of anchor can disrupt the Benford bias by altering the cognitive process of number generation.

In line with the expectations proposed in *Anchoring and Adjusting* (Kahneman & Tversky, 1974), the use of average anchors as references was supposed to constrain the range of plausible answers, which ultimately suppresses the emergence of Benford bias. In contrast, showing juvenile anchors appears to lead to the same sort of number generation process we saw in Burns & Krygier (2015) when no anchor was presented.

A previous analysis of this data set that examined full-number responses further revealed that most mean estimates closely aligned with the average anchors, while estimates in the juvenile group often exceeded the starting point by significant multiples. This pattern suggests that individuals adjust their estimation strategies based on the available information, often using multiplication for projections from a known starting point and addition or subtraction for average-based estimations. Such behavioural shifts indicate that cognitive judgment heavily relies on anchored information, with distribution expectations being primed by the type of anchors presented. For instance, an average score is often associated with a normal distribution, while a starting point typically suggests a growth pattern following a power law. Stereotypical reasoning about measures of central tendency is widespread among students from primary to tertiary education (Ismail & Chan, 2015). These stereotypes often align with everyday assumptions, such as the notion that an average score is its midpoint (Mokros & Russell, 1995), which may not always be true.

This reliance on anchored information in reasoning implies that people tend to favour mental shortcuts over newly acquired factual evidence, especially in uncertain situations. Even participants who passed our quiz and demonstrated a basic understanding of variable distributions often resorted to solutions based on assumptions drawn from specific anchor types. Whether consciously or unconsciously, they disregarded the factual information presented, showing a preference for a simple adjustment based on the anchors for judgment. This suggests that overriding established norms in response to general knowledge questions is challenging, even when explicit factual information is presented. Such an approach might be a practical means of navigating uncertainty, akin to representative bias or base-rate neglect (Kahneman & Tversky, 1972).

It could be pointed out that many studies of anchoring effects use arbitrary anchors, which our study did not. Whereas arbitrary anchors can produce impressive anchoring effects, their usefulness in explaining anchoring depends on the assumption that they invoke the same cognitive processes as nonarbitrary anchors. Thus both arbitrary and nonarbitrary anchors have a role in exploring the phenomenon.

Conclusions

In summary, the unexpected results in the test of sensitivity to the underlying distribution of variables pointed out that further research is needed to fully capture the constraints of statistical learning. Our study also found that specific anchored information could disrupt the Benford bias by altering the cognitive process in number estimation. This finding potentially extends our understanding of the anchoring effect and its underlying mechanisms regarding the adjustment. To the extent that a strong Benford bias indicates a number estimation process, our results suggest that far anchors induce general number estimation processes whereas near anchors induce a different process, plausibly adjustment. Future studies could attempt to offer questions related to everyday activities to examine the Distribution hypothesis in familiar situations. Nonarbitrary anchors could also be supplied for contrast to see if it makes a difference when examining the impact of the anchoring effect on Benford bias.

References

Ariely, D., Loewenstein, G., & Prelec, D. (2003). "Coherent arbitrariness": Stable demand curves without stable preferences. *Quarterly Journal of Economics*, *118*, 73-105.

Berger, A., & Hill, T. P. (2015). A Short Introduction to the Mathematical Theory of Benford's Law. In: S.J. Miller (ed.) *Benford's Law: Theory and Applications* (pp. 23-6)., Princeton University Press: Princeton, NJ.

Berger, A., & Hill, T. P. (2020). *The mathematics of Benford's law: a primer.* Statistical Methods & Applications.

Benford, F. (1938). The law of anomalous numbers, *Proceedings of the American Philosophical Society, 78,* 551-572.

Burns, B. D. (2009). Sensitivity to Statistical Regularities: People (largely) Follow Benford's Law. In N. A. Taatgen & H. van Rijn (Eds.), *Proceedings of the Thirty-First Annual Conference of the Cognitive Science Society* (pp. 2872-2877). Austin, TX: Cognitive Science Society.

Burns, B. D., & Krygier, J. (2015). Psychology and Benford's Law. In S. J. Miller (Ed.), The Theory and Applications of Benford's Law (pp. 259-268). Princeton University Press.

Chi, D., & Burns, B. D. (2022). Why do people fit to Benford's Law? – A test of the recognition hypothesis. In J. Culbertson, A. Perfors, H. Rabagliati, & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society* (pp. 3648-3654).

Diekmann, A. (2007). Not the First Digit! Using Benford's Law to Detect Fraudulent Scientific Data. *Journal of Applied Statistics*, *34(3)*, 321-329.

Enough, B., & Mussweiler, T. (2001). Sentencing under uncertainty: Anchoring effects in the courtroom 1. *Journal of Applied Social Psychology*, *31*(7), 1535-1551.

Epley, N., & Gilovich, T. (2001). Putting adjustment back in the anchoring and adjustment heuristic: *Differential processing of self-generated and experimenter-provided anchors. Psychological science*, *12*(5), 391-396.

Fiser, J. Z., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28,* 458-467.

Galinsky, A. D., & Mussweiler, T. (2001). First offers as anchors: The role of perspective-taking and negotiator focus. *Journal of personality and social psychology*, *81*(4), 657.

Griesser, M., Ma, Q., Webber, S., Bowgen, K., & Sumpter, D. J. (2011). Understanding animal group-size distributions. *PloS One*, *6*(8).

Griffith, T., & Tenenbaum, J. (2006). Optimal Predictions in Everyday Cognition. Psychological Science, 17(9), 767-73.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A primer on partial leastsquares structural equation modeling (PLS-SEM)* (3rd ed.). Thousand Oaks, CA: Sage.

Ismail, Z., & Chan, S. W. (2015). Malaysian students' misconceptions about measures of central tendency: An error analysis. *AIP Conference Proceedings*, *1643*, 93-100.

Iyengar, S. S., & Lepper, M. R. (2000). When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology*, *79*(6), 995.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, *3(3)*, 430-454.

Kreuzer, M., Jordan, D., Antkowiak, B., Drexler, B., Kochs, E. F., & Schneider, G. (2014). Brain electrical activity obeys Benford's law. *Anesthesia & Analgesia*, *118*, 183-191.

Kubovy, M. (1977). Response availability and the apparent spontaneity of numerical choices. *Journal of Experimental Psychology: Human Performance and Performance*, *2*, 359-364.

Lewandowsky, S., Griffiths, T. L., & Kalish, M. L. (2009). The wisdom of individuals: Exploring people's knowledge about everyday events using iterated learning. *Cognitive science*, *33*(6), 969-998.

Miller. S. (Eds.). (2015). *The Theory and Applications of Benford's Law*. Princeton, NJ, Princeton University Press.

Mokros, J., & Russell, S. J. (1995). Children's concepts of average and representativeness. *Journal for Research in Mathematics Education*, *26*(1), 20–39.

Muradian, K. H. (1989). Distribution of the species-specific lifespan in various taxonomic animal groups. *Zhurnal Obshchei Biologii*, *50*(1), 22–26.

Mussweiler, T., & Strack, F. (1999). Hypothesis-consistent testing and semantic priming in the anchoring paradigm: A selective accessibility model. *Journal of Experimental Social Psychology*, *35*(2), 136-164.

Mussweiler, T., Strack, F., & Pfeiffer, T. (2000). Overcoming the inevitable anchoring effect: Considering the opposite compensates for selective accessibility. *Personality and Social Psychology Bulletin*, *26*(9), 1142-1150.

Plate, R. C., Fulvio, J. M., Shutts, K., Green, C. S., & Pollak, S. D. (2018). Probability learning: Changes in behavior across time and development. *Child development*, *89*(1), 205-218.

Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General*, *134*(4), 552.

Uchmański, J. (1985). Differentiation and frequency distributions of body weights in plants and animals. Philosophical Transactions of the Royal Society of London. Series B, *Biological Sciences*, *310*(1142), 1–75.

Wegener, D. T., Petty, R. E., Detweiler-Bedell, B. T., & Jarvis, W. B. G. (2001). Implications of attitude change theories for numerical anchoring: Anchor plausibility and the limits of anchor effectiveness. *Journal of Experimental Social Psychology*, *37*(1), 62-69.