UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
Sensitivity to statistical regularities: People (largely) follow Benford's law

Permalink
https://escholarship.org/uc/item/5vb810w9

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 31(31)

ISSN
1069-7977

Author
Burns, Bruce

Publication Date
2009

Peer reviewed

eScholarship.org
Powered by the California Digital Library
University of California

Sensitivity to statistical regularities: People (largely) follow Benford's law

Bruce D. Burns (bburns@psych.usyd.edu.au)
School of Psychology, Brennan MacCallum Bldg, A18
University of Sydney, NSW 2006, Australia

Abstract

Recent decision making research has emphasized people's sensitivity to statistical relationships in the environment. A little-known relationship is Benford's law, that the first digits of numbers representing many natural and human phenomena have a logarithmic distribution (Benford, 1938). Benford's law is being used to help detect fraudulent financial data, but this assumes that people will not follow Benford's law when generating data. In two studies I examined whether people follow Benford's law. In Study 1 participants were given nine questions (e.g., "Length of the Indus river: km") chosen to have a flat distribution of first-digits for correct answers. The generated distribution was close to Benford's law. In Study 2 the results for generated data were replicated with new questions, and a selection task was also given in which participants selected from nine possible answers. Selected answers were a poor fit to Benford's law. Taken together the results suggest that Benford's law is a product of the way people generate responses, rather than sensitivity to the relationship itself.

Keywords: Benford's law; randomness; decision making; forensic accountancy.

Introduction

A common theme in recent research into reasoning and decision making has been that people are influenced by statistical relationships in the environment. This is a key part of adaptive approaches to decision making such as that of Gigerenzer, et al, (1999) and it underpins apparent automaticities in everyday life (Bargh & Ferguson (2000). Key to these approaches is that people follow statistical relationship of which they have little awareness. Rarely though is it possible to test if people are truly acting precisely in accord with an unknown pure statistical relationship, rather than just following heuristics broadly consistent with measurable relationships. There are usually too many issues around conditions and sampling that exactly what the statistical relationship a person may have experienced cannot be stated precisely. Therefore *Benford's law* (also called the *first-digit law*) offers an interesting test case, because it is a precise statistical relationship that is both universal and little known.

Table 1: Percentage frequency of each first digit from theory (Newcomb, 1881) and data (Benford, 1938).

	1	2	3	4	5	6	7	8	9
Newcomb (BND)	30.1	17.6	12.5	9.7	7.9	6.7	5.8	5.1	4.6
Benford's data	30.6	18.5	12.4	9.4	8.0	6.4	5.1	4.9	4.7

History. Benford's law is named after the physicist Frank Benford, who gathered data supporting a proposal by Newcomb (1881). Newcomb noted that the first pages of logarithmic tables seemed to wear out faster, and deduced that the frequency of the first digit in numbers from nature are logarithmic. In its most general form as stated by Hill (1995), the leading digit d ($d \in \{1, \ldots, b-1\}$) in base b ($b \geq 2$) occurs with probability $P(d)=\log_b(d+1)-\log_b d = \log_b((d+1)/d)$. For a base 10 number system this gives the distribution Newcomb first proposed and which I will refer to as the *Benford / Newcomb distribution* (BND).

Benford (1938) empirically demonstrated the validity of the law by collecting first-digit distributions for a number of quantities, both natural and human. For 20 different quantities with an average of 1011 data points each he found the mean distribution shown in Table 1, which for each digit was within the margin of error of Newcomb's (1881) prediction. These quantities included rivers, physical constants, cost data, populations and newspaper circulations. Because logarithms yield the only frequency distribution that is invariant under transformation, it does not matter what units these quantities are expressed in or even if they are expressed as reciprocals. Benford noted that the poorest fitting data was data such as physical data (e.g., weights) and the best was for numbers generated by no known relationship (such as river lengths), so he called it the *law of anomalous data*. Since Benford's demonstration this logarithmic relationship has been shown to apply to many types of naturally occurring data such that its validity is not disputed, though its scope is a more open question. Not all data would be expected to follow Benford's law, for example telephone numbers, lottery numbers, and populations of villages under 1000 would not.

Benford's law long remained just a curiosity for mathematicians who argued over how it could be derived from more general mathematical principles (see Raimi, 1976). Hill (1995) seems to have resolved the mathematical issue somewhat by deriving Benford's law from the assumption of scale invariance. Further, Hill (1998) argued that because a distribution seems to fit better the *more* it arises from completely unrelated data (e.g, batting averages, areas of rivers); the critical point may be that the data is a combination of different distributions. So he proposed a theorem, "If distributions are selected at random (in any 'unbiased' way) and random samples are taken from each of these distributions, then the significant-digit frequencies of the combined sample will converge to Benford's distribution, even though the individual distributions selected may not closely follow the law." (p. 6).

Interest in Benford's law rose when it was demonstrated that it could be used to detect fraudulent data, such as in tax returns. Nigrini's (1992) PhD thesis advanced the idea that Benford's law could be used to detect fraud on the basis that fraudulent data would not fit the law. The *Wall Street Journal* (Berton, 1995) reported that the chief financial officer for the district attorney's office in Brooklyn had detected company check fraud using Nigrini's analysis. Nigrini (1996) analyzed a sample of over 100,000 unaudited US tax returns and found that digit-1 occurred slightly higher than expected for interest received and slightly lower for interest paid, consistent with under-reporting of income and over-reporting of expenses. Nye and Moul (2007) found that GDP number follow Benford's law unless they are from Africa, where economic data had already been considered suspect. Mlodinow (2008, pp. 83-84) reports the case of Kevin Lawrence, an entrepreneur jailed after an investigation sparked by a forensic accountant who found that his company's various checks and wire transfers did not fit Benford's law. Benford's law has now become a standard tool of forensic accountancy (Zikmund, 2008). Its uses though go beyond fraud detection. For example, Hassan (2003) suggests that Benford's law could be used to help with the problem of detecting inaccuracies in databases.

Previous empirical research. The usefulness of Benford's law as a tool for fraud detection partly rests on the assumption that people are poor at deliberately generating numbers that conform to it, just as they are generally poor at generating random numbers (Rapoport & Budescu, 1992). As Bolton & Hand (2002) point out in their review of statistical fraud detection, "The premise behind fraud detection using tools such as Benford's law is that fabricating data which conform to Benford's law is difficult.**"** (p.238) So several attempts have been made to test whether people generate Benford's law (see Table 2).

Hsü (1948) asked 1044 participants to "write a 4-digit number that must be original, i.e., created in your own mind" but found no relationship between the first-digit and Benford's law. Hill (1988) asked mathematics students to generate a 6-digit number "out of their heads" and found very similar data to Hsü. The two studies differed in that the largest frequency in Hsü was for digit-4 (i.e., the digit "4") and for Hill it was for digit-6. This could be explained by Kubovy's (1977) results for priming. In his Experiment 3 participants were asked to generate a 4 digit number and produced the biggest peak for digit-4, whereas when asked to generate a number between 1000 and 9999 they produced the biggest peak for digit-1. All four of the above samples produced a peak for digit-1, although largest for Kubovy (perhaps because asking for the "first" number that came to mind produced more priming), but none were a fit to Benford's law. Thus the consensus has been that just as people are poor at producing random data, they are poor at producing data that fits to Benford's law.

However, Diekmann's (2007) data challenged this conclusion. He first showed that unstandardized regression coefficients reported in journals were a good fit to Benford's law. He then asked students in sociology or economics to fabricate multiple four-digit "plausible values" of regression coefficients that would support a hypothesis. He found a reasonable fit to Benford's law (see Table 2). Interesting there was no evidence of a priming effect for digit-4 despite participants being asked to generate a 4-digit coefficient. However the samples were small (10 or 13 participants) and the pattern could be due to knowledge about regression coefficients, such that they tend to be low for data in social science. Thus in two studies I further investigated people's ability consistency with Benford's law.

Understanding whether (or when) people follow Benford's law is important for both practical and theoretical reasons. Practically, the value of Benford's law as a detector of fraud or error is a product of being able to predict when invalid data will nevertheless fit it. Theoretically, it is valuable because it is a precise distribution that every person has had exposure to over their lives. Thus it could be a useful test case for how sensitive people are to a statistical relationship that they are not consciously aware of.

Table 2. Percentage frequency of each first digit reported in previous studies, indicating the question that was asked. Kubovy (1977) data is from his Experiment 3 and were estimated from measurements of his graph for all digits except 1, 4 and 5.

	1	2	3	4	5	6	7	8	9
Hsü (1948): 4-digit "created in your own mind" (n=1044)	13.3	9.2	14.3	15.5	6.6	9.3	12.6	9.1	10.5
Hill (1988): 6-digit "out of their heads" (n=742)	14.7	10.0	10.4	13.3	9.7	15.7	12.0	8.4	5.8
Kubovy (1977): first 4-digit that comes to mind (n=190)	24.2	12.1	11.1	27.4	2.6	6.4	7.8	5.7	2.7
Kubovy (1977): first number between 1000 and 9999 that comes to mind (n=116)	51.7	5.3	11.7	4.3	10.3	0.8	6.1	5.3	4.4
Diekmann (2007, Exp. 1): 4-digit regression co-efficients (n=10x10)	37	21	10	11	9	2	3	6	1
Diekmann (2007, Exp. 2): 4-digit regression co-efficients (n=13x10)	26.2	19.2	10.8	5.4	12.3	5.4	5.4	10.8	4.7

Study 1

Perhaps the critical difference between earlier studies and Diekmann (2007) was that the latter asked about something meaningful rather than a random number. So to test Benford's law I asked students to estimate the types of quantities that Benford had gathered data on, such as river drainage areas and newspaper circulations. If people are sensitive to statistical relationships in the environment then the best test of consistency with BND would be asking them to generate data that is known to fit Benford's law, rather than data that is know to not fit it such as random numbers.

Method

Participants. The task was given to 127 students (101 female, 26 male) in a psychology class as part of a set of tasks designed to illustrate reasoning phenomena.

Procedure & Materials. Using the types of quantities used by Benford (1938), a set of nine questions was constructed. Answers were drawn from entries in Wikipedia on 28/5/2008. The nine questions were selected so that one had a correct answer with each of the first-digits 1 through 9. Thus either correct or random answers would yield a flat distribution of first-digits. Questions were not selected completely randomly as I tried to avoid well known items or the largest or smallest examples of a category. The selected questions were as follows, with correct answers (not shown to participants) displayed in squared brackets.

1. US gross national debt: $ [9] trillion
2. The number 2 raised to the power of 33: [8,589,934,592]
3. The peak summer electricity consumption of Melbourne: [7000] MV
4. Atomic weight of zinc: [65.39]
5. Population of the urban area of Philadelphia, USA: [5,330,000]
6. Area drained by the Pearl (Xi Jiang) river: [437,000] km^2
7. Length of the Indus river: [3,180] km
8. Daily circulation of UK newspaper *The Daily Mail*: [2,340,255]
9. Infant mortality rate of Afghanistan: [157.43] deaths per 1000 live births

Participants were presented with the nine questions on a computer screen with an answer box next to each. They were asked to "Please try to estimate the following values. Even if you have no idea, just guess." They could not proceed to the next task until a legitimate number had been entered for each question.

Results

The first digit of each participant's answer was extracted and the percentage of their nine answers which used each digit was calculated. Figure 1 shows the mean percentages for each first-digit together with lines representing the distribution for Benford's law and for correct answers.

As can be seen from Figure 1 the data was a reasonable fit to Benford's law, except for digit-5. The data is a much better fit to Benford's law than the flat distribution that would represent random or correct responding. To test fit we calculated for each participant a root-mean square (RMS) relative to Benford's law (RMS-Benford) and relative to the flat distribution (RMS-flat). RMS-Benford ($M = 11.9$, $sd = 3.5$) was lower than RMS-flat ($M = 13.2$, $sd = 3.5$), $t(126) = 4.19$, $p < .001$.

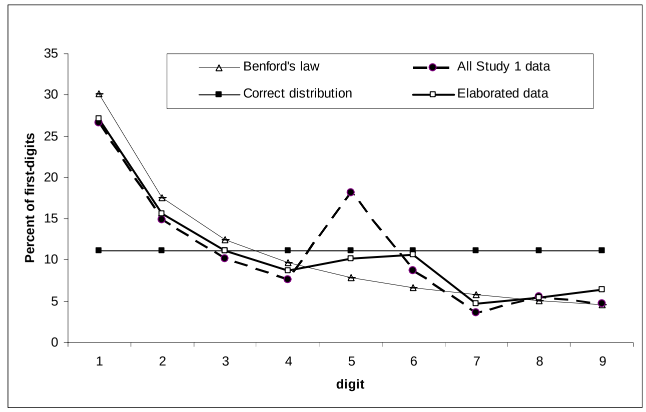

Figure 1: Distribution of first-digits for all Study 1 data, elaborated only data, Benford's law, and the distribution for the correct answers.

With 127 participants choosing 10 numbers there is enough power for the first-digit distribution to be statistically significantly different from BND. Descriptively though the data looks to be a good fit to BND, except for digit-5. An oversupply of digit-5 was also present in the fraudulent financial data reported by Hill (1998), but it is unknown if that is generally characteristic of fraudulent data. Another data set in which there are peaks at digit-5 is that of Scott, Barnard & May (2001). They asked participants to provide numbers within various constraints, and like the studies described in Table 1 they did not find fits to BND (which was not their focus), but often they found a peak at digit-5. However they made an interesting distinction between *unelaborated* numbers (numbers consisting of only one digit that was not zero) and *elaborated* numbers (those containing at least two nonzero digits). They found that the peak at digit-5 was largely due to unelaborated responses, which they argued was because elaborated responses arose from more use of executive processes. So that greater processing should yield distributions less likely to be the result of a single bias.

Therefore I separated elaborated from unelaborated numbers. The distribution of the elaborated numbers is shown in Figure 1 (representing 37% of the data). As can be seen the peak at digit-5 is reduced. The elaborated distribution is still significant different from BND, $X^2(7) = 19.6$, $p < .05$, but the data appears to be a better fit. The equivalent test against the correct (flat) distribution yields, $X^2(7) = 144.9$, $p < .05$.

Diekmann’s (2007) data also produced a slight peak at digit-5 but more in line with that found in my elaborated data than the whole data set. This could be because it is a reasonable assumption that most of his data was elaborated. He asked participants to fabricate four-digit regression coefficients and they were probably aware that such values rarely containing three zeros.

It is possible that people just happen to better know the low-digits answers and that produced the observed distribution. To check for this I removed answers with the correct first-digit, but there was little change to the distribution.

Discussion

Study 1 suggests that although people are not aware of Benford’s law their estimates fit reasonably well to the BND. So it represents a regularity of the environment that people are not aware of, yet largely conform to. The main discrepancy is for digit-5. Scott, et al. (2001) found various biases in how people generate numbers, one of which was a focus on magnitude. To the extent that participants in Study 1 were focused on magnitude it would increase digit-1 responses but also it might increase digit-5 responses because they would represent half magnitudes. A participant deciding between two different magnitudes may have split the difference and thus have digit-5 as the first digit. This is however purely speculation.

Why then do people appear to produce reasonable fits to the BND? There seem to be at least two possibilities. First, it could be that due to participants’ lifelong exposure to data that fits to Benford’s law they may have become sensitive to this distribution, even if not conscious of it. Second, it could be due to some fundamental property of how people generate numbers. That there are biases in number generation is well know from research into people’s failure to generate truly random numbers (see Rapoport & Budecsu, 1992). One way to try to address whether conforming to Benford’s law is due to learning or is a property of number generation itself is to see if it holds for a selection task as well as a generation task.

Study 2

The second study presented a selection task as well as a generation task. The generation task was intended to deal with a possible weakness of Study 1, which was that it presented everyone with the exact same nine questions. Therefore it is possible that the apparent fit of the data in Figure 1 was an artefact of those particular questions. So in Study 2 participants did a generation task with better randomization of questions.

The selection task gave participants the same types of questions as the generation task and asked them to choose from amongst possible answers. If people fit to the BND because they are sensitive to that distribution, then their chosen answers to difficult questions should also fit to this.

Method

Participants. The task was given to 335 students (243 female, 92 male) in a psychology class as part of a set of tasks designed to illustrate decision making phenomena.

Materials. For the generation task open-ended questions were asked about nine different domains (e.g., river lengths in km) and nine different targets (e.g., Indus). Targets were selected so that one each started with each digit 1 through 9. Targets for the nine domains questions were not widely spaced, did not include the largest or smallest case for a domain, and the highest value target for each question was a different first-digit. All data was drawn from Wikipedia pages on 24/8/2008. These constraints were designed to allow more complete randomisation of the first-digit of correct answers. The nine questions used were similar to Study 1 but some had to be replaced to fit to new constraints. The new set (indicating location of targets) was:

1. Infant mortality rate (deaths/1000 live births) for [target1]? __________
2. Atomic weight of [target2]? __________
3. In square kilometers, the area drained by the river [target3]? __________
4. In MILES, length of the river [target4]? __________
5. Total energy consumption per capita (kg of oil equivalent) for [target5]? __________
6. External debt per capita ($US) for [target6]? __________
7. Expenditure on US TV advertising (US$millons) by [target7]? __________
8. Population of metropolitan area of [target8]? __________
9. Daily newspaper circulation of [target9]? __________

All participants received all nine domain questions but for each the targets were randomized with the constraint that for each participant they would receive a target with each of the number 1-9 as first digit of the correct answer. Thus if a participant knew all the correct answers they would produce a flat distribution.

For the selection task the same nine domain questions were asked but different targets were selected (still with the constraint that each had a different first-digit). However instead of having to type in a number as their answer, the correct quantities for all nine potential targets were presented as possible answers. For each question participants had to chose one of the nine quantities.

Procedure. In class participants completed on a computer a set of tasks designed to illustrate well known heuristics and biases in decision making. Early in the set they did the selection task (one question per screen), and then after a further set of tasks they did the generation task. Although the two tasks asked about the same nine domains, for each individual participant the tasks asked about different targets.

Results

Generation task. Some participants did not get to the end of the set of tasks, so only 290 completed the generation task that came late in the sequence. As shown in Table 3, the distribution of first-digits in the generation task was very similar to Study 1, despite utilizing new questions and proper randomization.

To test fit I again calculated each participant's root-mean square (RMS) relative to Benford's law (RMS-Benford) and the flat distribution (RMS-flat). RMS-Benford (M = 12.2, sd = 4.1) was lower than RMS-flat (M = 13.1, sd = 4.4), $t(289)$ = 4.16, p < .001. Again digit-5 deviated the most from BND, so I examined the distribution of elaborated responses (46% of all responses). As shown in Table 3, elaborated responses again had a greatly reduced peak for digit-5 and appear to more closely approximate the BND. As would be expected with this much statistical power, the elaborated distribution is still significant different from BND, $X^2(7)$ = 51.9, $p < .05$, but the equivalent test against the correct (flat) distribution yields, $X^2(7) = 403.2$, $p < .05$.

Table 3: First-digit distributions for Study 2 generation (plus elaborated responses only) and selection tasks (plus correct answers only). Values are mean percentages with standard deviations, where appropriate.

.

digit	Study 2 (generation)	Study 2 (generation: elaborated)	Study 2 (selection)	Study 2 (selection: correct)
1	23.9 (18.6)	26.6	14.3 (11.2)	22.7
2	15.1 (13.4)	16.3	13.4 (11.3)	14.6
3	9.7 (10.6)	9.1	9.7 (8.7)	9.1
4	7.6 (9.8)	9.6	10.1 (9.9)	7.8
5	19.0 (16.4)	11.7	10.6 (10.7)	8.8
6	9.1 (10.8)	8.4	9.9 (9.8)	9.1
7	5.5 (9.3)	7.2	10.5 (10.3)	9.6
8	5.8 (8.3)	5.7	8.7 (9.2)	7.1
9	4.2 (7.9)	5.2	12.9 (11.4)	11.1

Selection task. More participants (335) completed the selection task because it was earlier in the sequence. As can be seen in Table 3, the selection task yielded a much flatter distribution than the generation task. RMS calculated in the same way as for generation found that RMS-flat (M = 10.1, sd = 2.7) was now lower than RMS-Benford (M = 11.9, sd = 3.1), $t(334)$ = 12.27, p < .001. This indicates that the flat distribution was a better fit to the data than the BND. There was no correlation between the RMS-Benford scores calculated for the generation and selection tasks, $r(290)$ = .062, p = .29.

One reason why the selection task might lead to a flatter distribution than the generation task could be that multiple choice responses are more vulnerable to various strategies, such as choosing the first option. The materials were designed so that any such strategy would produce a flat distribution of first-digits, so to the extent that participants employed such strategies they would flatten the overall distribution. Giving all the correct answers would also lead to a flat distribution but if Benford's law is a statistical relationship that people are sensitive to, maybe they are more likely to remember correct answers that fit to it. So post-hoc I looked at the first-digit distribution of all answers that were correct, which was 13.3% of all answers. The proportions displayed in Table 3 are somewhat closer to the BND, but note that each participant is not making an equal contribution to the data.

Discussion

The replication of the generation data from Study 1 with a better randomized set of questions and answers suggests that people's fit to Benford's law is a robust phenomenon for meaningful questions. The finding that the selection task does not fit to BND raises question over what is the basis of the fit in the generation task. It also argues against some possible alternative explanations of the generation task fit to the BND. For example, it could be argued that people just are expressing a belief that low numbers are more likely than high numbers for any quantity and this has consequences for the first-digit. However if this was the case it should be just as strong in a selection task.

General Discussion

These results only scratch the surface of understanding how people fit to Benford's law, but by expanding on Diekmann's (2007) finding they show it was premature to assume that the first-digits people generate have flat distribution for unknown quantities. Overall people's generated distributions were close to Benford's law, largely deviating at a specific point, digit-5. However this distribution did not emerge when participants had to select from nine answers. This argues against Benford's law arising due to knowledge about the distribution of first-digits and instead being a product of the way people generate quantities.

The results have implications for both the practical and theoretical aspects of Benford's law. The practical implications are that applications of Benford's law to search for fraud need to be more nuanced, otherwise they may at best be a waste of time and at worst increase the confidence in data that is in fact invalid. Although previous research suggested that first-digits have a flat distribution in generated data, this appears to have been a product of asking for context-free numbers. Understanding what distributions people produce and under what conditions would allow Benford's law to be more effectively applied to detecting fraud. For example, my data and Diekmann's consistently find an elevated frequency of digit-5, and the fraudulent data reported by Hill (1998) also showed a peak at digit-5 (though in that case it was 61.2%). So it could be that financial data showing elevation at digit-5 should be flagged even if the overall fit to the BND is good.

From a theoretical perspective Benford's law offers a way to examine people's sensitivity to a statistical relationship in the environment. To the extent people follow Benford's law, why do they? One possibility is sensitivity to a statistical relationship present in the environment, so they have learned that low digits tend to start the numbers representing the length of rivers (as Gigerenzer et al, 1999, suggests they have become sensitive to the relationship between newspapers and size of cities). The finding that the selection task did not yield the BND argues against this explanation for Benford's law, but perhaps there are other ways to present a selection task that would yield a better fit to it.

A second possibility is that fit to the BND tells us something about generation, that people are Benford's law generators. Hill's (1995) "Random samples from random distributions theorem" proposes that Benford's law is what you get when you take random samples from random distributions. Thus people may have a greater ability to act randomly than has been claimed. Rapoport and Budescu (1992) found that people can produce sequences that pass test of randomness when not asked explicitly to generate random numbers, but instead played a game. Similarly tests of Benford's law that asked participants to generate random numbers found no evidence, but when asked to generate something meaningful then they fit better to randomness. Thus these results add to the picture that people can generate random numbers under suitable conditions.

These two explanations are not necessarily incompatible. Scott, et al. (2001) used number generation to investigate executive functioning and argued that in particular their elaborated data was a result of more complex processes due to reconciliation of multiple biases. If such complexity is applied generating meaningful numbers then these biases should themselves have meaningful distributions. Thus if Benford's law arises from the random selection from random distributions, as Hill (1995) argued, then when cognitive processes sample from random distributions Benford's law could emerge. Thus it may not be that Benford's law itself is a statistical regularity of the environment that people are aware of, instead it could just be the inevitable result of people's sensitivity to many statistical regularity drawn on when estimating unknown quantities. However this is a purely speculative proposal.

References

Bargh, J.A., & Ferguson, M. J. (2000). Beyond behaviorism: On the automaticity of higher mental processes. *Psychological Bulletin,126*, 925-945.

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

Berton, L. (1995). He's Got Their Number: Scholar Uses Math to Foil Financial Fraud. *The Wall Street Journal*, July 10.

Bolton, R.J., & Hand, D.J. (1999). Statistical fraud detection: A review. *Statistical Science, 17*, 235-255.

Diekmann, A. (2007). Not the first digit! Using Benford's law to detect fraudulent scientific data. *Journal of Applied Statistics, 34*, 321-329.

Gigerenzer, G., Todd, P. M., & the ABC Research Group (1999) *Simple heuristics that make us smart.* London: Oxford University Press.

Hassan, B. (2003). Examining data accuracy and authenticity with leading digit frequency analysis. *Industrial Management & Data Systems, 103*, 121-125.

Hill, T.P. (1988). Random-number guessing and the first digit phenomenon. *Psychological Reports, 6*, 967-971.

Hill, T.P. (1995). Base-invariance implies Benford's law. *Proceedings of the American Mathematical Society, 123*, 887-895.

Hill, T.P. (1998). The first digit phenomenon. *The American Scientist, 10*, 354-363.

Hsü, E. H. (1948). An experimental study on "mental numbers" and a new application. *Journal of General Psychology, 38*, 57-67.

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

Mlodinow, L. (2008). *The drunkard's walk: How randomness rules our lives.* New York: Patheon books.

Newcomb, S. (1881). Note on the frequency of use of the different digits in natural numbers. *American Journal of Mathematics,* 4, 39-40.

Nigrini, M. J. (1992). *The Detection of Income Tax Evasion Through an Analysis of Digital Frequencies*. Ph.D. thesis. Cincinnati, OH: University of Cincinnati.

Nigrini, M. J. (1996). A taxpayer compliance application of Benford's law. *Journal of the American Taxation Association, 1*, 72-91.

Nye, J., & Moul, C. (2007). The Political Economy of Numbers: On the application of Benford's law to international macroeconomic statistics. *The B.E. Journal of Macroeconomics, 7(1),* Article 17.

Raimi, R.A. (1976). The first digit problem. *American Mathematical Monthly, 83*, 521-538.

Rapoport, A., & Budescu, D. V. (1992). Generation of random series in two-person strictly competitive games. *Journal of Experimental Psychology: General, 121*, 352-363.

Scott, S.K., Barnard, P.J., & May, J. (2001). Specifying executive representations and processes in number generation tasks. *The Quarterly Journal of ExperimentalPsychology: Section A, 54*, 641-664.

Zikmund, P. E. (2008). Reducing the Expectation Gap. *The CPA Journal, 78(6),* 20-26.