HAL
open science

Generalized Benford's Law as a Lie Detector

Nicolas Gauvrit, Jean-Charles Houillon, Jean-Paul Delahaye

To cite this version:

Nicolas Gauvrit, Jean-Charles Houillon, Jean-Paul Delahaye. Generalized Benford's Law as a Lie Detector. Advances in Cognitive Psychology, 2017, 13 (2), pp.121 - 127. 10.5709/acp-0212-x . hal-01745684

HAL Id: hal-01745684

https://hal.science/hal-01745684v1

Submitted on 28 Mar 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Generalized Benford's Law as a Lie Detector

Nicolas Gauvrit [1], Jean-Charles Houillon [2], and Jean-Paul Delahaye [3]

[1] Laboratoire Cognition Humaine et Artificielle, Ecole Pratique des Hautes Etudes, Les Patios Saint-Jacques, Paris, France

[2] Institut d'Enseignement à Distance, University Paris-Saint-Denis, Saint-Denis, France

[3] Laboratoire d'Informatique Fondamentale de Lille, University of Lille I, Centre National de la Recherche Scientifique, Cité Scientifique, Villeneuve d'Ascq, France

ABSTRACT

The first significant (leftmost nonzero) digit of seemingly random numbers often appears to conform to a logarithmic distribution, with more 1s than 2s, more 2s than 3s, and so forth, a phenomenon known as *Benford's law*. When humans try to produce random numbers, they often fail to conform to this distribution. This feature grounds the so-called *Benford analysis*, aiming at detecting fabricated data. A generalized Benford's law (GBL), extending the classical Benford's law, has been defined recently. In two studies, we provide some empirical support for the generalized Benford analysis, broadening the classical Benford analysis. We also conclude that familiarity with the numerical domain involved as well as cognitive effort only have a mild effect on the method's accuracy and can hardly explain the positive results provided here.

KEYWORDS

generalized Benford's law,
fraud detection,
Benford analysis

INTRODUCTION

During the late 19th century, an intriguing phenomenon was discovered by Newcomb (1881): The first significant digit (leftmost nonzero digit) of seemingly random numbers often fails to follow a flat distribution with an equal proportion of 1s, 2s, … , 9s, as one would expect, but instead follows a decreasing distribution, with more 1s than 2s, more 2s than 3s, and so forth. The same phenomenon was later rediscovered and detailed by Benford (1938). According to what is now referred to as *Benford's law* or *Newcomb-Benford law* (NBL), the distribution of the first significant digit X of a "random" number follows a logarithmic law given by $P(X = d) = \mathrm{Log}(1+1/d)$, where Log stands for the base 10 logarithm and d stands for a digit (in the range of 1–9; see Table 1).

TABLE 1.
Proportion of 1s, 2s,…, 9s as First Significant Digit in a Series Conforming to NBL

Digit	1	2	3	4	5	6	7	8	9
Prop (%)	30.1	17.6	12.5	9.69	7.92	6.69	5.80	5.12	4.58

Note. NBL = Newcomb-Benford law.

Many real-world datasets approximately conform to NBL (Hill, 1998; Nigrini, 2012). For instance, the distance between earth and known stars (Alexopoulos & Leontsinis, 2014) or exoplanets (Aron, 2013), crime statistics (Hickman & Rice, 2010), the number of daily-recorded religious activities (Mir, 2014), earthquake depths (Sambridge, Tkalcic, & Arroucau, 2011), interventional radiology Dose-Area Product data (Cournane, Sheehy, & Cooke, 2014), financial variables (Clippe & Ausloos, 2012), and internet traffic data (Arshadi & Jahangir, 2014), were found to conform to NBL. In psychology, NBL was found relevant in the study of gambling behaviors (Chou, Kong, Teo, Wang, & Zheng, 2009), brain activity recordings (Kreuzer et al., 2014), language (Dehaene & Mehler, 1992; Delahaye & Gauvrit, 2013), or perception (Beeli, Esslen, & Jäncke, 2007).

Corresponding author: Nicolas Gauvrit, Laboratoire Cognition Humaine et Artificielle, Ecole Pratique des Hautes Etudes, Les Patios Saint-Jacques, 4-14 rue Ferrus, 75014 Paris, France. Tel: (+33) 6 32 97 85 56.
E-mail: ngauvrit@me.com

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

DOI • 10.5709/acp-0212-x

Although NBL is ubiquitous, not all random variables or datasets conform to it. Scott and Fasli (2001) studied 230 sets of data and found that among them, less than 13% conformed precisely to NBL. Diekmann and Jann (2010), Bonache, Maurice, and Moris (2010), or Lolbert (2008) have also warned against overconfidence in NBL. NBL is not, they argue, a universal law but a property that appears in certain specific (albeit numerous) contexts.

THE SENSITIVITY AND SPECIFICITY OF BENFORD ANALYSIS

Human pseudorandom productions are in many ways different from true randomness (Nickerson, 2002). For instance, participants' productions show an excess of alternations (Vandierendonck, 2000) or are overly uniform (Falk & Konold, 1997). As a consequence, fabricated data might fit NBL to a lesser extent than genuine data (Banks & Hill, 1974). Haferkorn (2013) compared algorithm-based and human-based trade orders and concluded that algorithm-based orders approximated NBL better than human-based orders. Hales, Chakravorty, and Sridharan (2009) showed that NBL is efficient in detecting fraudulent data in an industrial supply-chain context.

These results support the so-called *Benford analysis*, which uses a measure of discrepancy from NBL to detect fraudulent or erroneous data (Bolton & Hand, 2002; Kumar, 2013; Nigrini, 2012). It has been used to audit industrial and financial data (Rauch, Göttsche, Brähler, & Kronfeld, 2014; Rauch, Göttsche, & El Mouaaouy, 2013), to gauge the scientific publication process (de Vries & Murk, 2013), to separate natural from computer-generated images (Tong, Yang, & Xie, 2013), or to detect hidden messages in images' .jpeg files (Andriotis, Oikonomou, & Tryfonas, 2013).

As a rule, the Benford analysis focuses on the distribution of the first digit and compares it to the normative logarithmic distribution. However, a more conservative version of Benford's law states that numerical values or a variable X should conform to the following property: $Frac(\mathrm{Log}(X))$ should follow a uniform distribution in the range of [0,1[. Here, $Frac(x)$ stands for $x—Floor(x)$, $Floor(x)$ being the largest integer inferior or equal to x. The logarithmic distribution of the first digit is a mathematical consequence of this version (Raimi, 1976).

Hsü (1948), Kubovy (1977), and Hill (1988) provided direct experimental evidence that human-produced data conform poorly to NBL. However, in their experiments, participants were instructed to produce numbers with a given number (four or six) of digits. Specifying such a constraint could well induce participants to attempt to generate uniformly chosen numbers or to use a digit-by-digit strategy (repeatedly picking a random digit). Researchers who study the situations in which NBL appears often conclude that one important empirical condition is that the numerical data cover several orders of magnitude (e.g., Fewster, 2009; for a more detailed mathematical account, see Valadier, 2012). Consequently, any set of numbers that are bound to lie in the thousands scale (four digits) or in the hundreds of thousands scale (six digits) will probably not conform to NBL, whether produced by humans or not. Participants' failure to produce data conforming to NBL could just be a consequence of the instructions they were given.

Furthermore, these studies were decontextualized. Participants were asked to give either the "first number that came to mind" or a number chosen "at random", without being told what the numbers were supposed to represent. A "random number with four digits" usually implicitly refers to the uniform distribution (Gauvrit & Morsanyi, 2014)—and the uniform distribution does not conform to NBL. Therefore, the lack of context could prime a non-Benford response, even if participants are, in fact, able to produce series conforming to NBL.

For these reasons, the idea that fabricated numerical data will usually not follow NBL has been questioned. Using a more contextualized design, Diekmann (2007) asked social science students to create plausible regression coefficients, with four-digit precision. Note that, contrary to the case of a four-digit integer, which is bound to fall between 1,000 and 9,999, covering only one order of magnitude, here, the coefficients could run between .0001 and 1, covering four orders of magnitude. Diekmann found that, in this case, the first digit does approximately conform to NBL and concluded that researchers should not only consider the first digit as relevant to detecting fraud but should also look beyond the first digit, toward the conservative version of NBL. Using correlation coefficients makes the task meaningful, and this may explain why Diekmann's participants are not bad at producing plausible *r*s, whereas other samples suggested that human participants would be unable to mimic NBL.

Another study went even further in formulating meaningful tasks by using the type of data known to exhibit a Benford distribution. Burns (2009) asked participants to guess real-world values, such as the US gross national debt or the peak summer electricity consumption in Melbourne. He found that although participants' first digit responses did not perfectly follow a logarithmic law, they conformed to the logarithmic distribution better than to the uniform distribution. Burns concluded that participants are not too bad at producing a distribution that conforms to NBL as soon as the task involves the type of real-world data that do follow NBL.

One limitation of Burns' (2009) study is that it only works at a population level. We cannot know from his data if a particular individual would succeed in producing a pseudorandom series conforming to NBL, since each participant produced a single value. Nevertheless, his and Diekmann's studies certainly suggest that using Benford's law to detect fraud is questionable in general since humans may be able to produce data confirming to NBL, in which case a Benford test will yield many undetected frauds, lacking *sensitivity*. As mentioned above, not all random variables or real-world datasets conform to NBL (and when they do, it is generally only in an approximate manner). Because many real-world datasets do not conform to NBL, a Benford test used to detect fraud not only may have low sensitivity but may also have low *specificity*.

GENERALIZED BENFORD'S LAW

Several researchers (e.g., Leemis, Schmeiser, & Evans, 2000) have studied conditions under which a distribution seems more likely to satisfy NBL. Fewster (2009) provided an intuitive explanation of why and when the law applies and concluded that any dataset that smoothly spans several orders of magnitude tends to conform to NBL. Data limited to one or two orders of magnitude would generally not conform to the law.

To pursue the question of why many data conform to NBL further, the conservative version of the NBL may be a better starting point than the mere first-digit analysis. Recall that in the conservative version, a random variable X conforms to NBL if $Frac(\mathrm{Log}(X)) \sim \mathrm{U}([0,1[)$. In an attempt to show that the roots of NBL ubiquity should not be looked for in the specific properties of the logarithm, Gauvrit and Delahaye (2008, 2009) defined a generalized Benford's law (GBL) associated with any function f as follows: A random variable X conforms to a GBL associated with function f if $Frac(f(X)) \sim \mathrm{U}([0,1[)$. The classical NBL thus appears as a special case of GBL, associated with function Log.

Testing several mathematical and real-world datasets, Gauvrit and Delahaye (2011) found that several of them fit GBL better than NBL. Of 12 datasets they studied, six conformed to the classical NBL, while 10 conformed to a GBL with function $f(x) = \pi \times x^2$, and nine with square-root function. On the other hand, none conformed to GBL with function Log ∘ Log. These findings suggest that a GBL associated with the relevant function—depending on the context—might yield more specific or sensitive tests for detecting fraudulent or erroneous data.

We addressed this question in two studies. In both studies, each participant produced a whole series of values, allowing analyzing the resulting distribution at an individual level. In Study 1, we examined three versions of GBL in four different situations in order to compare the sensitivity and specificity of different types of GBL analyses. Study 2 explored the potential effects of variations of familiarity with the material and of cognitive effort on the productions.

STUDY 1

In Study 1, we examined human pseudorandom productions in four different realistic settings, such as those where a classical Benford's law has been initially observed, and compared these responses with true sample values.

Participants

A sample of 169 adults (63 women) took part in this experiment. Participants were recruited via social networks and e-mails. Ages ranged from 13 to 73 years ($M_{Age} = 40.9$, $SD = 11.6$).

Method

Participants were randomly assigned to one of four groups: cities (n = 41), numbers (n = 44), stars (n = 36), or tuberculosis (n = 48). In each group, participants were informed that a series of 30 numbers had been selected at random from a dataset, and were instructed to produce what they thought would be a credible outcome—that is, to supply 30 plausible numbers. In the *cities* group, the list was the set of populations of the 5,000 most populated US cities. *Numbers* corresponded to the dataset of numerical constants published by Simon Plouffe and described in the experiment as an extensive encyclopedia of mathematical constants. Participants assigned to the *stars* group were told that the dataset was the list of distances from earth to all known visible stars expressed in light-years. Last, the *tuberculosis* group dealt with the set of known country-wise incidences of tuberculosis as measured in 2012.

Real samplings of 30 numbers from the corresponding databases were also performed. The set of populations of the biggest US cities came from an online dataset (http://factfinder2.census.gov/). *Numbers* were randomly selected from the Simon Plouffe database of numerical values (http://www.plouffe.fr). The distances to the stars were read from the HYG2.0 dataset (http://www.astronexus.com/hyg) and multiplied by 3.262 to render them in light-years instead of parsecs. Lastly, the *tuberculosis* dataset was downloaded from the World Health Organization's (WHO) website (http://www.who.int/tb/country/data/download/en/).

Measures

For each set, X of 30 values' (either fabricated or real samples) observed distributions of the fractional parts of $f(X)$, with $f(X) = \mathrm{Log}(X)$, $f(X) = \pi \times X^2$ and $f(X) = \sqrt{(X)}$, were computed. The two last functions were selected on the basis of previous studies indicating that they led to satisfying fits with several numerical datasets (Gauvrit & Delahaye, 2009). The deviation from GBL was measured by the Kolmogorov-Smirnov statistic D—that is, the maximum difference between the cumulative distribution function of $Frac(f(X))$ and the cumulative distribution function of $U([0,1[)$. D thus serves as a proximal measure of conformity to GBL. This statistic is a classical measure of distance between distributions that grounds the classical Kolmogorov-Smirnov test.

Results

As expected, fabricated data were usually less consistent with GBL than real data (see Figure 1, Table 2). There is only one exception to this feature: GBL associated with square function in the case of *numbers*. Depending on the context, different computational variants of GBL seemed more appropriate for segregating true values from fabricated ones. In the case of the largest US cities, for instance, human and real data did not significantly differ in terms of conformity to NBL or GBL with the square function, but they differed when calculated with a square-root function.

To analyze the specificity and sensitivity of a fraud detection tool based on GBL, we drew receiver operating characteristic (ROC) curves (see Figure 2) and computed the areas under the curves (AUCs). As shown in Table 3, different sets of data resulted in different patterns. For the *cities* condition, classical NBL was barely efficient, whereas square root yielded better results. With the Plouffe database, all GBLs were relevant, although the one associated with function $\pi \times X^2$ appeared to be the best one.

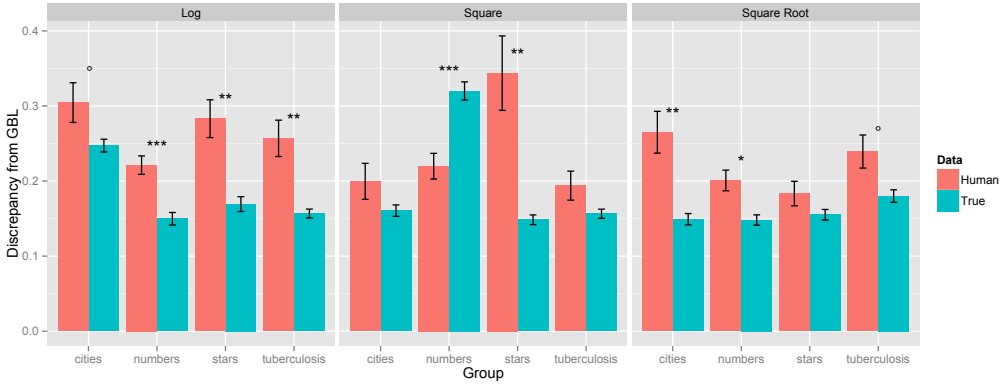

FIGURE 1.

Mean discrepancy ± *SEM* from GBL as measured by *D*, for three functions: Log (corresponding to NBL), $f(x) = \pi \times x^2$ ("Square") and Square root. GBL = generalized Benford's law. °p <.05. *p < .01. **p < .001. ***p < .0001.

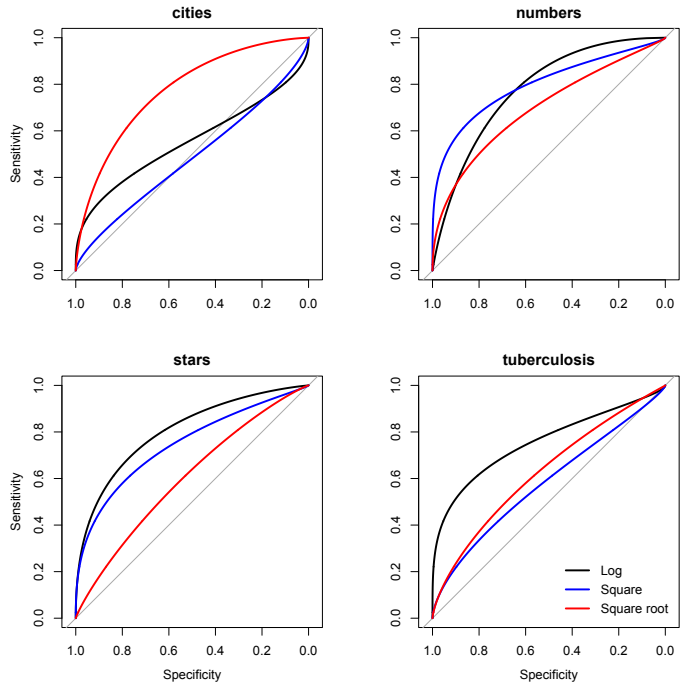

FIGURE 2.

Smoothed receiver operating characteristic (ROC) curves.

TABLE 2.

Results of Two-Sample *T*-Tests (*t*) Comparing the Conformity of Fabricated and Real Data to GBL, With Log, Square and Square Root Function

	Log	Square	Square Root
Cities	2.07 °	1.55	4.01 **
Numbers	4.80 ***	– 4.77 ***	3.41 *
Stars	4.20 **	3.90 **	1.59
Tuberculosis	4.02 **	1.85	2.52

Note. GBL = generalized Benford's law. °p < .05. *p < .01. **p < .001. ***p <. 0001.

TABLE 3.

AUCs With 95% Confidence Intervals (DeLong). In Each Row, the Largest AUC is Bolded

	Log	Square	Square Root
Cities	.55 (.42–.68)	.48 (.36–.61)	**.78** (.68–.88)
Numbers	.78 (.69–.88)	**.81** (.71–.91)	.70 (.59–.81)
Stars	**.81** (.71–.91)	.75 (.64–.86)	.59 (.46–.73)
Tuberculosis	**.75** (.65–.85)	.58 (.47–.70)	.62 (.51–.73)

Note. AUC = Area under the curve.

Discussion

Overall, NBL appeared to be a better means than the other tested variants of GBL for distinguishing between fabricated and real data. However, NBL is not always the best measure, as the *cities* condition showed. Even when NBL was an efficient measure, such as in the *number* condition, some other GBL may have been even better or at least as good as NBL, for example, GBL associated with square, for which the AUC was greater than the NBL AUC. Depending on the type of data one tests, different types of GBL could thus be adviced, either replacing or complementing classical Benford analysis.

A further argument in favor of the GBL analysis is that, with the growing popularity of the Benford analysis, potential swindlers might become aware of the necessity of conforming to the NBL. Alternative methods complementing the classical analysis (for another such method, see Nigrini & Miller, 2009) could thus prove useful, especially in view of the fact that it would be particularly difficult to fabricate data conforming to a whole set of variants of GBL.

STUDY 2

One possible reason why Diekmann (2007) found that humans were able to produce accurate values (r) is that the notion was familiar to the participants (students in the social sciences), a feature that may have had an impact on the outcome. If this is true, our positive results in Study 1 might have been the result of too low a familiarity with the material at hand. In Study 2, we investigated the possible effects of familiarity, as well as that of cognitive effort, on the accuracy of the Benford analysis.

Participants

A sample of 124 first-year psychology students (103 women) from a distant learning program volunteered in the experiment. Ages ranged from 22 to 55 years (M_{Age} = 38.27; SD = 9.08). Participants were recruited by e-mail and voluntarily accepted to participate. We chose distant learning students as participants because, contrary to ordinary students, they have various backgrounds and previous working experiences in a diversity of fields, warranting greater variation in the familiarity with the material.

Method

The experiment was performed online using a Google Form (https://www.google.com/forms/about/). We used country population data, as this was believed to grant a somewhat larger variation in familiarity. Participants were asked to produce series of only 20 data points to lower the risk of tiredness.

Participants were randomly assigned to one of two groups (no time pressure or time pressure condition). Each group included 62 participants. They were informed that they would have to supply a list of 20 values that could be the numbers of inhabitants of 20 randomly selected countries in the world. They were asked to try to guess what these populations could be. They were told that a true sampling would be performed for comparison with their answers. In the no time pressure condition, the instruction was to be "as accurate as possible, taking as much time as needed." In the time pressure condition, the instruction was to be "as fast and accurate as possible." The former condition is known to conduce to superior cognitive effort than the latter (e.g., Maule & Edland, 1997).

Self-reported data on level of expertise in the field of country populations were also collected using a 6-point Likert scale, from 0 = *absolutely naïve about country populations* to 5 = *expert in country populations*. This measure serves as an indication of the participants' familiarity with the material.

Sixty-two true samples of 20 country populations were selected from real-world data (found at http://data.okfn.org/data/core/population#data) to be compared with participants' productions.

Measures

For each set X of 20 values (either fabricated or real samples), observed distributions of $Frac(\mathrm{Log}(X))$ were computed. The discrepancy from uniformity was assessed in each case using the Kolmogorov-Smirnov statistic D.

Results

Reported expertise was rather low (M = 1.3; SD = .97; range of 0–4). To assess the effect of expertise, we performed an analysis of variance (ANOVA), with the dependent variable D and the independent variable Level of Expertise (6). No significant effect was found, $F(4, 119) = 0.64$, $p = .63$. The same procedure yielded nonsignificant statistics for GBL associated with the square-root function, $F(4, 119) = 1.29$, $p = .28$, and with GBL associated with function $\pi \times X^2$, $F(4, 119) = 0.20$, $p = .94$. Correlation analysis showed nonsignificant links between level of expertise and D, for classical NBL, $r(122) = -.11$, $p = .20$, GBL associated with a square-root function, $r(122) = -.14$, $p = .10$, and also with function πX^2, $r(122) = -.06$, $p = .53$.

To assess the effect of cognitive effort, we performed an ANOVA with a dependent variable D and an independent variable Group (2). There was a significant influence, $F(2, 183) = 7.33$, $p < .001$, but a Tukey HSD test showed that the two groups did not differ significantly from one another ($p = .46$), although they both differed from real data values.The same procedure was used to assess the effect of cognitive effort related to other variants of GBL. No effect was found for GBL associated with the square-root function, $F(2, 183) = 1.53$, $p = .22$, or with function πX^2, $F(2, 183) = 0.01$, $p = .99$.

Discussion

Familiarity does not appear to influence the quality of participants' responses in terms of GBL. The experimental procedure used to assess the possible effect of time pressure or cognitive effort yielded no significant results either. These results thus suggested that producing fraudulent data that would remain undetected under the Benford analysis is not necessarily a matter of familiarity or cognitive effort.

It is, however, fair to note two limitations. First, time pressure, although widely used to increase cognitive effort, probably does not result in large differences, especially if a task requires little effort in general. Concerning familiarity with the material, no participants declared

themselves experts (score of 5 on the scale) about country populations, most of the sample lying in the range of 0–3, so that we would not have detected a specific effect only appearing with true experts.

CONCLUDING DISCUSSION

We performed the first investigation of the generalized Benford analysis, an equivalent of the classical Benford analysis, but based on the broader GBL. Results from Study 1 rendered mild support for the generalized Benford analysis, including the classical Benford analysis. They also draw attention to the fact that different types of data yielded different outcomes, suggesting that the best way of detecting fraud using GBL associated with some function *f* would be obtained either by finding the function *f* that best matches the particular data at hand or by combining different analyses. Although the classical Benford analysis was validated in our studies, it occasionally failed at detecting human-produced data as efficiently as other generalized Benford analysis.

The present positive results could have been the result of our sample characteristic, in which participants, contrary to real swindlers, might have put little effort into the task since the stakes were low. Plus, the participants were not highly familiar with the material at hand. To rule out the possibility that our results resulted from such features and GBL would be inapplicable in real situations, Study 2 aimed at demonstrating that cognitive effort and familiarity with the material have little effect on the participants' responses. The data supported this view, although further studies (including higher levels of cognitive pressure and true experts) would be recommended.

With Benford analysis having become more common in fraud detection, new complementary analyses are needed (Nigrini & Miller, 2009). The GBL analysis potentially provides a whole set of such fraud detection methods, which means making it more difficult, even for informed swindlers intentionally conforming to NBL, to remain undetected.

AUTHOR NOTE

Jean-Paul Delahaye initiated the project. Nicolas Gauvrit, Jean-Charles Houillon, and Jean-Paul Delahaye designed and performed the experiments. Nicolas Gauvrit analyzed the data. Nicolas Gauvrit and Jean-Paul Delahaye wrote the paper.

FOOTNOTES

[1] After correction for restriction of range under the hypothesis of a binomial distribution of familiarity, the correlation remains low and almost unchanged ($r = -.14$)

REFERENCES

Alexopoulos, T., & Leontsinis, S. (2014). Benford's Law and the universe. *arXiv:1401.5794*. doi: 10.1007/s12036-014-9303-z

Andriotis, P., Oikonomou, G., & Tryfonas, T. (2013). JPEG steganography detection with Benford's law. *Digital Investigation, 9*, 246-257. doi: 10.1016/j.diin.2013.01.005

Aron, J. (2013). Crime-fighting maths law confirms planetary riches. *New Scientist, 220*, 12. doi: 10.1016/S0262-4079(13)62775-X

Arshadi, L., & Jahangir, A. H. (2014). Benford's law behavior of internet traffic. *Journal of Network and Computer Applications, 40*, 194-205. doi: 10.1016/j.jnca.2013.09.007

Banks, W. P., & Hill, D. K. (1974). The apparent magnitude of number scaled by random production. *Journal of Experimental Psychology, 102*, 353-376. doi: 10.1037/h0035850

Beeli, G., Esslen, M., & Jäncke, L. (2007). Frequency correlates in grapheme-color synaesthesia. *Psychological Science, 18*, 788-792. doi: 10.1037/h0035850 www

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

Bolton, R. J., & Hand, D. J. (2002). Statistical fraud detection: A review. *Statistical Science, 17*, 235-249. doi: 10.1214/ss/1042727940

Bonache, A., Maurice, J., & Moris, K. (2010). Détection de fraudes et loi de Benford: Quelques risques associés [Detection of fraud and Benford's law: The associated risks]. *Revue Française de Comptabilité, 431*, 24-27.

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.

Chou, M. C., Kong, Q., Teo, C. P., Wang, Z., & Zheng, H. (2009). Benford's law and number selection in fixed-odds numbers game. *Journal of Gambling Studies, 25*, 503-521. doi: 10.1007/s10899-009-9145-9 www

Clippe, P., & Ausloos, M. (2012). Benford's law and Theil transform of financial data. *Physica A: Statistical Mechanics and its Applications, 391*, 6556-6567. doi: 10.1016/j.physa.2012.07.063

Cournane, S., Sheehy, N., & Cooke, J. (2014). The novel application of Benford's second order analysis for monitoring radiation output in interventional radiology. *Physica Medica, 30*, 413-418. doi: 10.1016/j.ejmp.2013.11.004 www

Dehaene, S., & Mehler, J. (1992). Cross-linguistic regularities in the frequency of number words. *Cognition, 43*, 1-29. doi: 10.1016/0010-0277(92)90030-L www

Delahaye, J. P., & Gauvrit, N. (2013). *Culturomics: Le numérique et la culture* [Culturomics: The numerical and the culture]. Paris, France: Odile Jacob.

De Vries, P., & Murk, A. J. (2013). Compliance of LC50 and NOEC data with Benford's Law: An indication of reliability? *Ecotoxicology and Environmental Safety, 98*, 171-178. doi: 10.1016/j.ecoenv.2013.09.002 www

Diekmann, A. (2007). Not the first digit! Using Benford's law to detect fraudulent scientific data. *Journal of Applied Statistics, 34*, 321-329. doi: 10.1080/02664760601004940

Diekmann, A., & Jann, B. (2010). Benford's law and fraud detection: Facts and legends. *German Economic Review, 11*, 397-401. doi: 10.1111/j.1468-0475.2010.00510.x

Falk, R., & Konold, C. (1997). Making sense of randomness: Implicit encoding as a basis for judgment. *Psychological Review, 104*, 301-318. doi: 10.1037/0033-295X.104.2.301

Fewster, R. M. (2009). A simple explanation of Benford's law. *The American Statistician, 63*, 27-32. doi: 10.1198/tast.2009.0005

Gauvrit, N., & Delahaye, J.-P. (2008). Pourquoi la loi de Benford n'est pas mystérieuse [Why Benford's law is not mysterious]. *Mathématiques et Sciences Humaines, 182*, 7-16.

Gauvrit, N., & Delahaye, J.-P. (2009). Loi de Benford généralisée [Generalized Benford's law]. *Mathématiques et Sciences Humaines, 186*, 5-15. doi: 10.4000/msh.11034

Gauvrit, N., & Delahaye, J.-P. (2011). Scatter and regularity implies Benford's law… and more. In H. Zenil (Ed.), *Randomness through computation* (pp. 53-69). London, England: World Scientific.

Gauvrit, N., & Morsanyi, K. (2014). The equiprobability bias from a mathematical and psychological perspective. *Advances in Cognitive Psychology, 10*, 119-130. doi: 10.5709/acp-0163-9 www

Haferkorn, M. (2013). Humans vs. algorithms – who follows Newcomb-Benford's law better with their order volume? In F. A. Rabhi & P. Gomber (Eds.), *Enterprise Applications and Services in the Finance Industry* (pp. 61-70). Berlin, Germany: Springer.

Hales, D. N., Chakravorty, S. S., & Sridharan, V. (2009). Testing Benford's law for improving supply chain decision-making: A field experiment. *International Journal of Production Economics, 122*, 606-618. doi: 10.1016/j.ijpe.2009.06.017

Hickman, M. J., & Rice, S. K. (2010). Digital analysis of crime statistics: Does crime conform to Benford's law? *Journal of Quantitative Criminology, 26*, 333-349. doi: 10.1007/s10940-010-9094-6

Hill, T. P. (1988). Random-number guessing and the first digit phenomenon. *Psychological Reports, 6*, 967-971. doi: 10.2466/pr0.1988.62.3.967

Hill, T. P. (1998). The first digit phenomenon: A century-old observation about an unexpected pattern in many numerical tables applies to the stock market, census statistics and accounting data. *American Scientist, 86*, 358-363.

Hsü, E. H. (1948). An experimental study on "mental numbers" and a new application. *Journal of General Psychology, 38*, 57-67. doi: 10.1080/00221309.1948.9711768 www

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. doi: 10.1213/ANE.0000000000000015 www

Kubovy, M. (1977). Response availability and the apparent spontaneity of numerical choices. *Journal of Experimental Psychology: Human Perception and Performance, 2*, 359-364. doi: 10.1037/0096-1523.3.2.359

Kumar, R. (2013). Data mining for detecting carelessness or mala fide intention. *International Journal of Computer Applications, 74*, 8-11.

Leemis, L. M., Schmeiser, B. W., & Evans, D. L. (2000). Survival distributions satisfying Benford's law. *The American Statistician, 54*, 236-241. doi: 10.1080/00031305.2000.10474554

Lolbert, T. (2008). On the non-existence of a general Benford's law. *Mathematical Social Sciences, 55*, 103-106. doi: 10.1016/j.mathsocsci.2007.09.001

Maule, A. J., & Edland, A. C. (1997). The effects of time pressure on human judgment and decision making. In R. Ranyard, R. Crozier, & O. Svenson (Eds.), *Decision making: Cognitive models and explanations* (pp. 189-204). London, UK: Routledge.

Mir, T. A. (2014). The Benford law behavior of the religious activity data. *Physica A: Statistical Mechanics and its Applications, 408*, 1-9. doi: 10.1016/j.physa.2014.03.074

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

Nickerson, R. S. (2002). The production and perception of randomness. *Psychological Review, 109*, 330-357. doi: 10.1037/0033-295X.109.2.330 www

Nigrini, M. (2012). *Benford's Law: Applications for forensic accounting, auditing, and fraud detection* (Vol. 586). Hoboken, NJ: John Wiley & Sons.

Nigrini, M. J., & Miller, S. J. (2009). Data diagnostics using second-order tests of Benford's law. *Auditing: A Journal of Practice & Theory, 28*, 305-324. doi: 10.2308/aud.2009.28.2.305

Raimi, R. A. (1976). The first digit problem. *American Mathematical Monthly, 83*, 521-538. doi: 10.2307/2319349

Rauch, B., Göttsche, M., Brähler, G., & Kronfeld, T. (2014). Deficit versus social statistics: Empirical evidence for the effectiveness of Benford's law. *Applied Economics Letters, 21*, 147-151. doi: 10.1080/13504851.2013.844319

Rauch, B., Göttsche, M., & El Mouaaouy, F. (2013). LIBOR manipulation – Empirical analysis of financial market benchmarks using Benford's law. *Social Science Research Network*. doi: 10.2139/ssrn.2363895

Sambridge, M., Tkalcic, H., & Arroucau, P. (2011). Benford's law of first digits: From mathematical curiosity to change detector. *Asia Pacific Mathematics Newsletter, 1*, 1-6.

Scott, P. D., & Fasli, M. (2001). Benford's law: An empirical investigation and a novel explanation. *CSM Technical Report 349*. Retrieved from Semantic Scholar website: https://pdfs.semanticscholar.org/a358/49266811a69322bf307c035cce5058848b74.pdf

Tong, S. F., Yang, Y. H., & Xie, Y. J. (2013). Distinguishing computer graphics from natural images based on statistical characteristics. *Applied Mechanics and Materials, 380*, 1306-1309. doi: 10.4028/www.scientific.net/AMM.380-384.1306

Valadier, M. (2012). The Benford phenomenon for random variables. Discussion of Feller's way. *arXiv:1203.2518*.

Vandierendonck, A. (2000). Bias and processing capacity in the generation of random time intervals. *Cognitive Science Quarterly, 1*, 205-233.

RECEIVED 12.12.2016 | ACCEPTED 02.03.2017