Journal of Scientific Exploration, Vol. 24, No. 3, pp. 437–475, 2010 0892-3310/10

RESEARCH

Soal's Target Digits: Statistical Links Back to the Source He Reported After All

RODERICK GARTON

School of Psychology, Faculty of Science and Engineering, University of Tasmania, Australia
rgarton@utas.edu.au

Abstract—Given abiding questions in the establishment of fraud in the Soal and Goldney (1943) study of "precognitive telepathy," retrieval was attempted of a sample of its target series from their reported source, viz., final digits of 7-figure logarithms. Distinct from earlier efforts, but consistent with Soal's statements, the length of retrievals was not assumed, and it was hypothesized that retrievals should most frequently occur within the first 20 pages of the published source. Testing 30 published series largely marked as fraudulent, their retrieval was indicated in comparison to chance-control sources, and the early-entry hypothesis also was supported. These findings were maintained when exhaustively and exclusively searching for the longest possible retrievals, and the earliest of entries per retrieval. Additionally, Benford's Law for the distribution of leading digits offered theoretical expectations that were matched by each chance-control source, but surmounted by Soal's reported source, precisely in the range indicated by his method. Alternative logarithmic sources could not reproduce these effects. While reserving implications for the population of target series, it could appear that Soal derived the target series as he originally reported. Clarification and elaboration of extant fraud scenarios are offered by this interpretation.

Keywords: Professional Ethics—Peer Evaluation—Precognition

The terminal critique by Markwick (1978) of the Soal and Goldney study (1943) of "precognitive telepathy" offered strong evidence that the target series had been manipulated to score spurious hits. Markwick's results could even suggest particular digits that had been manipulated. Yet questions remained as to the extent of fraud, and the manner in which fraud was perpetrated. This was because the identification of manipulated digits in Markwick's study was dependent on identifying target series that were reused from one run into another, while there were only limited indications of such reuse. What has prevented an advance on this issue, and what more can be done?

The target series were random samples of the digits 1 to 5 that, on most of the runs, were reportedly drawn from a published source of random numbers and used to indicate which of five target alternatives was to be guessed by the participant, Shackleton. Earlier studies had suggested that Soal could have stacked some target series with additional 1s, which he converted into other digits at some point during each run of the experiment so as to match the guesses. This was based on an allegation by Albert, an agent in a couple of the sittings, that she had seen Soal altering 1s into 4s and 5s, plus some statistical evidence supporting this possibility (see Markwick, 1985, for a review). However, apart from questions concerning the reliability of Albert's testimony, the original records, having been lost, could not be examined for any such alterations, so that there was little opportunity to identify particular runs, let alone trials, on which these manipulations might have occurred. This meant that the extent of fraud remained inexplicable, and the very practice remained dubious.

A potential solution introduced by Medhurst (1971) involved identifying the source of the target series, and then noting how the series, as eventually used, differed, if at all, from the source. However, all efforts to identify the target series from their reported source failed. Then, Markwick (1978) discovered that Soal occasionally reused some target series from one run into another. Reuse itself was not suspicious—it could be accounted for on the basis of convenience or accident, rather than being a necessary part of any manipulation. However, instances of reuse now permitted that the originally used series could be treated as the source, and deviations from exact duplication should be able to identify any spurious hits in the copy. Markwick could, indeed, identify such deviations in the form of one or more "extra digits" apparently being haphazardly inserted among some of the duplicated series. These "extra digits" were found to be disproportionately associated with hits, so suggesting that they were the very digits that were manipulated.

With Soal having died in 1975, the publication of these findings in 1978 was followed by immediate withdrawal of almost all support by his former colleagues. Still, limitations of the fraud scenario remained. It was unclear how Soal could have practically perpetrated fraud, particularly when he had no access to the target sheets during the experiment, and when independent observers were responsible for scoring the runs. Additionally, reuse was found on only a very small proportion of runs, such that the evidence suggestive of fraud was very limited. Markwick's final result implicated reused series within only 13 runs (plus two within the later Soal–Bateman study) of the 529 runs in the Soal–Goldney study (2.5%), these 13 runs being confined to 7 of the 40 sittings (17.5%) in which 78 runs were administered, 64 of which were administered under conditions predicted to yield above-chance scores. Markwick's (1985) later qualification of her findings stipulated that fraud was "virtually conclusive" for some runs within two of the 40 sittings, and the remainder of her evidence implicated, with the status of "suggestive only," runs within another six sittings.

How Were the Targets Sourced?

Toward extension and clarification of the fraud or any other model of these results, we need to closely review how the target series were sourced. Soal reportedly derived the target digits, for the most part, from a published source of 7-figure logarithms, viz., *Chambers' Tables* (Soal & Bateman, 1954, Soal & Goldney, 1943). Attempting to retrieve the target series from this source has depended on several assumptions concerning how the target series were compiled and eventually sourced for use within a run. There were published statements on the issue, but they are not as specific as this objective requires. What the Soal–Goldney report tells us is that:

> S.G.S. [i.e. Soal], before coming to the sitting, fills in the A [target] divisions on all the sheets to be used by (EA) [i.e. experimenter with the agent] with a random sequence of the digits 1, 2, 3, 4, 5. In general S.G.S. prepares these lists from the last digits of the seven-figure logarithms of numbers selected at intervals of 100 from Chambers' Tables. (See *Proc*. xlvi, 156.) In some cases, however, Tippets' [*sic*] random numbers were used. These lists are compiled by S.G.S. at his lodgings in Cambridge with no-one present but himself, and they are kept under lock and key until the day of the sitting. (Soal & Goldney, 1943:38–39)

While quite descriptive, this statement already bears some ambiguity: What are we to make of the qualification "in general"? Does this qualify the identity of the published source, and/or the method of using it? Ambiguity is exhausted neither by the reference to *Tippett's tables*, nor by the citation of Soal's (1940) *Fresh light* report, of which the relevant section is as follows.

> I had at my disposal 1200 [ESP "Zener"] cards, there being 240 of each symbol. I first associated with each of the symbols +, O, Star, Rectangle, Wave the respective numbers 1, 2, 3, 4, 5. I then provided myself with Chambers's *Seven-figure Mathematical Tables*, and read from them the *last* digits of the logarithms of the following numbers:
>
> 10078, 10178, 10278, … 99978.
>
> The numbers chosen were thus taken at intervals of 100, so as to ensure that the last digits in the logarithms should be independent. If the digit happened to be one of the numbers 1 to 5 the digit was entered on the list, or more exactly the corresponding symbol was written. If the digit happened to be 0 or 6, 7, 8, 9, it was not entered. From this sequence I thus obtained a random series of about 450 cards. The process was then repeated with, say, the following numbers:
>
> 10043, 10143, 10243, … 99943,
>
> and so on until a list of 1000 cards had been compiled. The actual cards were then chosen one by one according to the above list from the 1200 cards in my possession. (Soal, 1940:156)

These statements should assure us of the source's identity, and how the source was accessed. However, they do not tell us how the digits were compiled for the Soal–Goldney study. This is because, in the earlier study described in the above quote, the targets were arranged as 25-card *decks*, which permitted that the digits be immediately encoded into the ESP symbols that formed the deck. But for the Soal–Goldney study, while each run also was constituted of 25 trials, no decks were used. Instead, the digits were shown, one at a time, to the agent, directing her to select, during the run itself, one of the five randomly ordered targets (usually animal pictures or names) for that trial.

Retrieving the Targets from Their Source

In the first report of an effort to retrieve the targets from their source, Medhurst (1971) offered no statement concerning how he assumed the digits to have been compiled and eventually sourced for use in a run of the experiment. Yet we can surmise something of his assumptions from how he proceeded, as follows.

There were 372 runs among a total of 529 for which the targets could have been compiled, as above, from random numbers; naturally excluded were runs using "counters" as a real-time source of target digits, and three runs in which Soal tested the effect of target randomness. Given limited computer resources, Medhurst applied the economical approach of seeking segments of six target series, each of six digits in length, from the sitting (no. 16) to which Albert's allegation pertained, and in which Albert served as agent.

Why series of exactly six digits in length? This followed Medhurst's estimation of the number of series of the digits 1 to 5 that could theoretically be retrieved from *Chambers' Tables* at different lengths. Too low a criterion (less than six) would yield too many chance retrievals, making too difficult a manual check of the matched series with published tables, which Medhurst felt obliged to do. Still, too high a criterion (greater than six) was unreasonable as Medhurst considered that "there is always the possibility that Dr Soal was interrupted during his compiling of particular runs and returned to the Chambers Tables [*sic*] at a different point," such that "it would be a pity to employ long sequences" (Medhurst, 1971:50). Searching for series of six digits in length was therefore most resourceful and reasonable. Later, Medhurst expanded his sample to include sittings other than those implicated in the Albert allegation: He selected six more series, two each from Sittings 6 and 31, and another two from series prepared by Wassermann for Sitting 34—each, again, of six digits in length.

Note that Medhurst (1971), as quoted above, described the "compiling of particular runs" in the process of using *Chambers' Tables*. This suggests that he assumed that Soal compiled discrete 25-digit runs directly from the *Tables*. This key point will shortly be amplified.

As for his results, Medhurst was indeed able to successfully locate at least one entry-point in the final digits of 7-figure logarithms that, by proceeding at intervals of 100, produced each one of the 12 six-digit series that he tested. One entry-point produced an 8-digit series, and two produced a 7-digit series. Retrievals were not possible, however, for series of a greater length. This was ascertained by obtaining the next three digits of each retrieved 6-digit series from the logarithms, by 100-step intervals, and comparing these with the three that followed the series used in the experiments. No matches of these 9-digit series were possible. Immediately after presenting this result, Medhurst offered the following as his conclusion:

> In each case it is apparent that these sequences were not, in spite of Dr Soal's assertion, derived from the Chambers Tables [*sic*] in the way specified, and there appear to be enough of them to make it clear that the complete sequences for that session was [*sic*] not so derived. (Medhurst, 1971:51)

What were the features of his results that compelled this conclusion? Medhurst clearly did not offer it on the basis that *none* of the tested target series could be retrieved; some, at least, were likely to be chance matches. But why did he insist on matches of *nine* digits in length? This is nowhere explained in the report; it only appears in the Results section. In the absence of a rationale, we could expect that, if a 9-digit match were found, a 10-digit match would be insisted upon, and so on. Perhaps Medhurst even expected to find the entire 25-digit series for each run, as appears to be the reasoning behind his phrases "enough of them" and "complete sequences for that session." Medhurst did not reveal this (or any other) criterion to his readers, and we might well have expected something else, considering his surmise that Soal was unlikely to have produced "long sequences." While this allowance was necessary to justify his method, it disappeared at the point of conclusion, and no statement of particular criteria by which to judge an adequately sized match was offered in its place.[1] Why this was the case will shortly be offered (specifically, because none of the kind is possible); but for now, we can best turn to efforts to extend these searches, and ask if they surmounted these problems.

Two reports describe attempted replication and extension of Medhurst's effort. Scott and Haskell (1974:44) reported, in a sentence, that they "extended" Medhurst's search "by applying his method to samples taken from every sitting in which Soal reported having used prepared random numbers, without finding a single identifiable sequence." As in Medhurst's report, no criterion as to what constituted an "identifiable sequence" was specified. Markwick (1978:251) also briefly reported replication attempts. These were described as exploratory searches: "to try out some ideas which had occurred to me for extending and modifying Dr. Medhurst's search technique." Testing the 12

9-digit series reported by Medhurst, five approaches for their retrieval were attempted, e.g., searching the series in reverse, and using 6-figure logarithms. No quantitative statement of results was offered apart from the admission that "None of these efforts met with any success" (Markwick, 1978:251). Markwick (1978) returned to the question of retrieving the series after identifying what appeared to be "duplicated sequences" of target digits from some runs into others, as described above. Markwick reasoned that those portions that were fully "duplicated," without any interruption by suspicious "extra digits," were likely to be "manipulation-free," and hence they should stand a better chance of being retrieved from *Chambers' Tables*. As for method and results, it was reported that "I selected a number of suitable sub-sequences and carried out a further computer search, but this again drew a blank—as did a search on sequences taken from the Stewart data" (Markwick, 1978:254). Additionally, Markwick reported "a preliminary computer search through the 41600 digits comprising *Tippett's tables*, in four directions," but that this "also failed," again, by no specified criterion.[2]

These past three efforts to retrieve the Soal–Goldney target series share similar limitations. Absent were explications of assumptions framing the tests, and the criteria of success, apart from such subjective and formless criteria as "enough long matches" and "identifiable sequences." Instead of comparison of observations with expectations (the foundation of statistical argument), only standalone statistics (e.g., "drew a blank") were reported. In the absence of a statement as to what constituted a "success," past efforts appear to have been based on the assumption that the target series were compiled (without error) directly from their source into discrete 25-digit runs; such that long, 25-digit series should be retrieved in toto and de novo for each run. However, this assumption has no historical basis or logical necessity, and against it there is much contrary information. Does this information yet inform a more objective and reproducible approach?

Revisiting the Assumptions

Pratt (1971) pointed out that all of Soal's published statements, as well as the logic of convenience, suggested that, when compiling the target digits for the Soal–Goldney study, Soal initially created a large pool of digits that he then haphazardly entered, as needed, in order to retrieve a 25-digit series—rather than creating so many discrete sets of 25 digits at the outset directly from the *Tables*. This was procedurally conditioned given that Soal had no reason, in this study, to compile 25-card decks. However, as we have seen, the description of procedure for the Soal–Goldney study relied upon citing Soal's earlier study, such that we do not have an explicit statement on this point for the Soal–Goldney study. Yet if we can generalize from the study with Stewart—

which followed the Soal–Goldney study, and in which the method of 100-step intervals through *Chambers' Tables* was again reportedly used for target construction—we have such an explicit statement: "A large stock of the digits from 1 to 5 were [*sic*] prepared . . ., and S.G.S. generally drew upon this stock of numbers to meet the requirements of the tests" (Soal & Pratt, 1951:194). Pratt (1971) quoted a similar statement from Soal and Bateman (1954) that referred to taking the 25-digit series "from a large pool," which, incidentally, Medhurst (1971:48) also quoted, while evidently drawing no implications from it for his method. Markwick (1978) also referred to Soal's use of a "pool" of target digits, but it is unclear how, if at all, this impacted on her hypotheses and methods. The construction of a digit pool was also indicated by Soal's (1971:202) much later statement on the point, albeit when under pressure to conform memory with Medhurst's conclusion, when he was aged 82; Soal here described himself compiling "a very long list of the digits 1 to 5."

Such logical and evidential considerations indicate that Soal compiled a pool of digits. We must next ask: How did Soal enter and use this pool? Did he mark it where every new entry-point commenced? There would seem to be no reason to keep such a record; the pool could be cut up into 25-digit series later on, and then there would be no telling if any series, as eventually used in a particular run, had been produced from one or more passes through the *Tables*. In this way, it is quite likely that any 25-digit series, as finally used, could represent digits sourced from more than one entry-point into *Chambers' Tables*.

Further on the question as to how Soal used the pool of digits derived from *Chambers' Tables*, did he record the entry-points he used and avoid reusing them as he added to this pool? If not, some fortuitous duplications of series would occur. But this does not even depend on entry-point reuse. Take Soal's example entry-point of 10,078, and the average length of Markwick's (1978) "duplicated sequences," which can be calculated from her Table 4 as about 18. The 25 digits that can be read off from *Chambers' Tables*, by the 100-step method, from this entry-point, can be matched to the length of 18 digits from seven additional entry-points, e.g., 10,178, 10,778, 11,178. Accordingly, it is quite likely if not inevitable that Soal (and others assigned to the task) would enter the table at a point that could overlap rows and columns previously searched, and so obtain the same sequences, building up a pool of random digits that contained "duplicated sequences." There is no statement by Soal or others that they sought to avoid duplications fortuitously happening in this way, and there is, indeed, no reason to expect that they should have done so. Accordingly, we should expect that, when attempting to retrieve the target series from *Chambers' Tables*, any one series can be returned more than once.

When accessing this pool, did Soal mark off series once he used them? Did he attempt to avoid reusing series, or parts thereof? This is unlikely, for there is no theoretical necessity to do so; and Soal would have had to reuse series if the pool he started using, when the study commenced, in January 1941, did not amount to the 9,300 (25 × 372) digits that would eventually be required by April 1943, when the study ended. Accordingly, it could come as no surprise to find that Soal reused digits. Markwick's (1978) findings suggest that such reuse could be performed directly from the record of previous sittings, rather than the pool, for—as an alternative way of accounting for some of the "hits" on "extra digits" —Soal appears to have haphazardly eliminated digits that had previously been subject to hits, and inserted other digits along the way, or at the end, making up for those eliminated. Otherwise, as Markwick showed, he occasionally duplicated, with reversals and so on, one or more sheets, or parts thereof, at a particular time. This practice would serve to further dissociate the series from the *Tables*. Given, in these ways, initial fortuitous reuse from the pool, and secondary planned reuse from series already selected from the pool, the series from all runs should all the more not be expected to be retrievable, as complete nor unique 25-digit series, from *Chambers' Tables*.

Discontinuity by interruption, as raised by Medhurst, must also be considered, and not only interruptions of the "knock on the door" type. Informative on this point is Rosenthal's (1987) review of 15 psychological studies that reported the incidence of simple numerical recording errors (mainly of basic summing and copying; including a study of telepathy). The average error rate was 0.71% (weighted for number of recordings), while it achieved a maximum of 4.17% for a study with only 96 observations; with little more than 1,000 recordings, error rates ranged from 1.59% to 3.17%. Now one pass through *Chambers' Tables* at 100-step intervals provides about 450 digits in the 1–5 range. The study of Rosenthal's with *N* recordings nearest to this value (360) produced an error rate of 2.5%. As 450 digits equates to 18 25-trial segments, and this error rate predicts about 11 errors within the 450, it is quite likely that, with a random distribution of errors over the working space, and non-overlapping segments, we should find almost two-thirds of the 25-trial runs to be discontinuous with their source. This is a conservative estimate of the error rate, given that working through *Chambers' Tables*, while leaping over 100 values, requires a constant exercise of spatial attention, orienting one into the correct row and column within a page, while the ordinal position of the correct row can change in relation to the blocks in which rows are organized. Short of being surprised by such extensive errors in using mathematical tables, it should be noted that errors are known to "abound" in these tables themselves (e.g., Uhler, 1938), and that final digits are known to be particularly prone to unreliable recording and derivation (Preece, 1981).

Errors were, in fact, clearly indicated in the work of others assigned to the task of compiling the target series. These were produced for two sittings by

Wassermann (Sittings 27 and 34), and those for another sitting were produced by Blascheck (Sitting 28). Wassermann (1975) (a physicist) offered that it would be no wonder to him if the digits could not be retrieved from their source as he had found that reliably reading and recording digits was, for him, habitually impossible. Also, those prepared by Blascheck were noted by Medhurst (1971:52) to be "grossly non-random" and "not compiled in any even plausible way." Soal had also noted that Blascheck evidently omitted pairs of digits showing repetition (doublets) (Markwick, 1978:272). Also, the later Stewart series were compiled by the same method, but with errors (Soal & Pratt, 1951). That Soal himself was prone to making errors of this type is indicated by the errors found by appointed checkers and others of the Soal–Goldney run records (Markwick, 1976, Pratt, 1974:99–103, Scott & Haskell, 1974:71–72, 1975:226, Soal & Goldney, 1943:87). As an estimate of the extent of such errors, it was reported by Scott and Haskell (1975) that 22 errors were found by the checkers appointed by Soal, and that this figure covered 18 sittings. With the average number of runs per sitting being 13.225, this would cover about 238 runs, so indicating an error within almost 1 in 10 runs. This is not an inordinate error rate for these tasks; Martin and Stribic (1940) found, in their studies, almost 6% of 1,000 25-trial runs to be affected by recording error. On top of errors in transcribing digits from the source to the pool, these findings also oblige us to add a source of error in copying the target digits from the pool to the target sheets.

Another indication of Soal's propensity for recording errors is provided by a privately circulated publication for which he was himself responsible (Soal, 1966). This contains many handwritten corrections of the digits within it; a copy held by the present author even bears a handwritten front-cover note by Soal that, in its digits, is incorrect: "Table 1 page 4 is corrected," he advised, whereas the table—the only one in the volume—appears on page 10.

Medhurst's (1971) assumption that "it is a straightforward procedure to cast one's eye down a particular column, reading off final digits" (p. 49) was insensitive to these facts and fallibilities.

Implications for Retrieving the Soal–Goldney Target Series

For all these reasons, again, it should not be expected that complete 25-digit series, as used in any particular run, could be retrieved, in toto and de novo, from a single entry-point in *Chambers' Tables*. We should also not expect retrieval of series of any particular length, and the probability of retrieval quickly shrinks as length increases. Only matches of considerably shorter lengths than 25 should be expected. The manner in which Soal constructed the pool of digits, and how its digits were transferred to the lists finally used, then, limit the manner in which Medhurst's objective can be pursued, and the implications we might take from his results, and so of those who replicated his efforts. What more objective criteria could be applied to the question as to whether the target series were derived from 7-figure logarithms?

Chance-expectation values for the number of matches, for any length of series retrieved, can be theoretically or empirically determined. Deriving these values must make some account of the fact that the final digits of 7-figure logarithms are likely to fail basic tests of randomness. As Medhurst, for one, noted, randomness—in the senses of sequential independence and equiprobability of alternatives—was "not necessarily the case with these tables" (Medhurst, 1971:50), and that "the last digits form a far from random sequence" (p. 53); this was, indeed, the reason that Soal selected final digits at intervals of 100 (Soal, 1940). However, comparing the number of matches yielded by the logarithms with a theoretical value, or even one derived from a random pool of digits, could favor the logarithms should the target series fortuitously share its properties of non-randomness. Markwick (1978) reported failure to retrieve the target series from 7-figure logarithms on the basis of shared non-randomness, which suggests that the issue is not a problem, and that theoretical values or randomly generated digits could offer valid bases of comparison. However, given the unfalsifiability of certain hypotheses in this field, it is just as well to keep the possibility of shared non-randomness in mind.

This suggests the necessity of an empirical approach where we attempt to retrieve the target series from *N* "control" lists of digits constituted by the same specifications as the final digits of 7-figure logarithms. From these, we can obtain *N* retrieval counts to compare with the number retrieved from *Chambers' Tables*. The simplest control can involve so many randomly reordered lists of the final digits of 7-figure logarithms themselves. This partly takes care of the non-randomness issue as the frequency distribution of the digits 1 to 5 in these lists would be identical to that in the original; we only leave to chance their sequencing.

If we are also concerned to ensure that our random samples share the sequential properties of the 7-figure logarithms, we can, alternatively, hold order constant while transforming the values 1–5 within the 7-figure source through all their possible permutations. For example, permuting the digits 1 to 5 in the order of 12354 renders the series 044957 as 055947. Continuing thus with all 119 permutations alternative to 12345, we obtain 119 lists that share with the true list of final digits its digit sequences as well as frequency distribution.

However, while this approach should assure us of the independence of results from simple effects of non-randomness, it is likely to overestimate the number of matches with non-permuted control digits. For instance, if a target series is comprised of an extended sequence of the digits 1, 2, and 3, and the permutation only involves the digits 4 and 5, then the permuted final digits will naturally yield a match on the same basis as would the "true," non-permuted digits. Another limitation is the small number of permutation samples: 119 in addition to the "true" series. In order to more generally assess the role of any non-randomness in the lists, it will, thirdly, be informative to obtain the expectation values by searching lists that are purely constructed on a random basis, while being composed of the same elements, to the same size, as the "true" series.

Taking expectation values from these randomly shuffled, digit-permuted, and randomly generated lists should meet any concern for the randomness of the control data.[3] Hereafter, these digit sources are respectively referred to as the *shuffled*, *permuted*, and *random* sources.

How do we search these potential sources of the Soal–Goldney target series? Assuming extraction from a pool of digits, the target series must be tested at all points at which they can be cut; for any digit could represent the entry-point from which it and its following digits were derived. Then, it must be expected that identical series across runs will share the same entry-point; any one series could be found at more than one entry-point; and a series could be matched at an entry-point already covered by a match of the same series. The fact that there is no a priori basis for deciding at which digit within a series a search for a match should commence is accommodated by the proposed approach of empirical control, as any overlapping or multiple retrievals should be equally likely whatever the source of their retrieval.

Hypotheses

Empirical expectation. That Soal did *not* source his digits from a table of 7-figure logarithms in the manner he described defines the null hypothesis: Retrieving the series from *shuffled*, *permuted*, and *random* lists should be no more or less successful than deriving them from the list of final digits offered by *Chambers' Tables*. The number of retrievals from the chance-control sources that are greater than or equal to the number obtained from the "true" source (i.e. *Chambers' Tables*) defines the probability value by which to assay the "true" count. Additionally, where the counts are normally distributed, the number of matches and their variance over these lists offer empirical values by which to test the hypothesis. In this case, if the number retrieved from the 7-figure logarithms is significantly greater than the number retrieved from the chance-control sources, we might reasonably reject the null hypothesis in favor of the alternative hypothesis that the target series were sourced from a table of 7-figure logarithms.

Entry-point (antilogarithmic) intervals. How should an entry-point be chosen if performing Soal's task? Medhurst (1971:49) offered that "Of course, in accordance with Dr Soal's published procedure the starting point in the tables can be arbitrary." However, there is nothing factually "of course" about the entry-point being arbitrary. Logarithms in *Chambers' Tables* for numbers below 1,000 are necessarily in a different format from those proceeding from 1,000, and the logarithms for four-figure numbers (1,000 to 9,999) are read from within the tables from 10,000 that follow (Pryde, 1930). In the above-quoted description of the use of *Chambers' Tables*, Soal (1940) gave 10,078 and 10,043 as examples of entry-points. This suggests that he (1) ignored the table of logarithms for 1–999, (2) only used the tables that offered the most consistent format throughout the publication, and (3) intended to make a continuous pass through *Chambers' Tables* from the earliest of its most useful entry-points. Medhurst, by the way, assumed as much, for his searches only commenced from the logarithm of 10,000.

These observations suggest a further confirmation of Soal's procedure would be the preponderance of entry-points into the *Tables* in, say, the 10,000 to 19,999 range, relative to the chance match within this range, and to the eight equally sized ranges further into the *Tables*. Essentially, verifying that the target series are the *final* digits of *logarithms* involves identifying an effect of the *leading* digits of their *antilogarithms*. The null hypothesis is that the entry-points for any series retrieved from the 7-figure logarithms is no more likely to have a value below 20,000 than those retrieved from the chance-control sources, nor to differ in number of retrievals from entry-points greater than or equal to 20,000. While match probability slightly decreases as entry-points approach 100,000, this should be reproduced by the control sources. Finding a preponderance of entry-points in the 10,000–19,999 range—with 1 as the most likely leading digit—would offer support for the source of the series as reported by Soal and Goldney (1943).

Hence we have the following two hypotheses:

Hypothesis 1, that there will be a greater number of matches from the 7-figure logarithms than from the chance-control sources; and

Hypothesis 2, there will be a greater number of matches from the 7-figure logarithms for the range of entry-points (i.e. antilogarithms) from 10,000 to 19,999, relative to those obtained in this range among the chance-control sources, and relative to other entry-point intervals.

The latter prediction has an interesting association with Benford's Law for the proportion of leading digits among so-called "anomalous numbers" (see Hill, 1995, and Raimi, 1976, for reviews). Benford (1938) observed that in many naturally occurring series—from numbers in the pages of *Readers' Digest* to measures of black body radiation—there was a logarithmic decline in the frequency of leading digits in synchrony with their ordinal position. The Law is quite formally specific in describing this distribution: The proportion of figures commencing with the digit 1 is ≈30.10%; that commencing with the digit 2 is ≈17.61%; and so on, following the equation for leading digit *i*:

$$P(i) = \log_{10}\left(\frac{i+1}{i}\right) \tag{1}$$

This distribution has been found to particularly hold for figures describing a relatively unlimited range, covering several orders of magnitude, with no obvious internal relationships (Fewster, 2009, Smith, 2006). Accordingly, Benford (1938) reported that the leading digits of street addresses published in *American Men in Science* were in excellent agreement with the Law, but this outcome would not have been so likely for, say, the ages in years, or IQs, of these persons.

Quite pertinently to the present enquiry, the original observation of this distribution was made by Newcomb (1881)—the mathematician–astronomer and first president of the American Society for Psychical Research—with respect to the usage of a table of logarithms: "how much faster the first pages wear out than the last ones" (Newcomb, 1881:39), he noted, and thereupon he described the law now attributed to Benford. It can yet be expected that, for a process as described by Soal, where there is a deliberately repeated, if not exclusive, use of the earliest pages of a table of logarithms, the proportion of entry-points commencing with 1 should *surmount* that expected by Benford's Law. After all, the Law only describes the naturally occurring bias in the distribution of leading digits. Perhaps, however, we cannot expect a retrieval operation to satisfy this prediction unless we know which particular series were sourced from their particular antilogarithms; for in attempting to retrieve all of the series as originally sourced, we will no doubt include many spurious matches, which could even lead to the observed proportion falling short of that predicted by Benford's Law. Hence, at this stage, we can best rely on empirical expectation values. It will be useful, however, to keep this theoretical prediction in mind when interpreting results and planning further search operations.

EXPERIMENT 1. FIXED RETRIEVAL LENGTHS

Method

Original Target Series

The previous studies on this issue have all commenced with the records of the target series as held by the Society for Psychical Research. These are not known to be digitally available. In any case, the most accurate record of the series was retained by Pratt (1974); Markwick's (1978) findings were, indeed, checked with this record before publication. However, these records, too, have not been published, and Pratt's archives are uncatalogued (Matlock, 1987). This situation necessitated an economical reliance on target series that have been published. Complete 25-digit series for 18 runs have been published in the papers of Scott and Haskell (1974) and Markwick (1978),[4] while Medhurst (1971) reproduced two 12-digit series, and 10 9-digit series. In total, a sample of 30 target series, from perhaps as many runs, from eight sittings, comprising a total of 564 digits, could be obtained from published sources. While limited in representativeness of the 372 runs for which random numbers were reportedly prepared, this sample represented more than twice as many series as tested by Medhurst. Sampling limitations are addressed in the Discussion, but it should be noted at the outset that the sample was largely constituted of series previously implicated in claims of fraud.

With respect to the Scott and Haskell (1973, 1974) critique that suggested a non-uniform distribution of the target digits, it should be preliminarily noted that the distribution of digits within this sample of target series was quite uniform, well reproducing the distribution within the table of 7-figure logarithms themselves (outside of the digits 0 and 6–9); see Table 1. With respect to Markwick's (1978) finding that some series were copied but with reversal of order, the order of final use on the score sheets was here of interest, such that both the reversed and nonreversed series were represented in the sample.

TABLE 1

Mean and Proportional Frequencies per Digits 1–5 in Sampled Target Series and Logarithms

Digit	Mean per Target Series	SD per Target Series	% in Target Series	% in 7-Figure Logarithms
1	3.80	1.79	20.21	19.93
2	4.00	1.46	20.57	19.70
3	3.70	1.74	19.68	20.10
4	4.76	2.20	21.10	20.12
5	3.71	2.11	18.44	20.15

Computation of Chambers' Tables

A file bearing the final digits of 7-figure logarithms was compiled by taking the natural logarithm of each number between 1 and 100,000, inclusively, dividing it by the logarithm of 10 (yielding its common logarithm), and rounding the result to 7 decimal places. This conforms to the method by which

Chambers' Tables is constructed (Pryde, 1930),[5] and used by Medhurst (1971) to computationally reproduce them. The standard Perl functions *log* and *sprintf* were used to perform the computation. All of Medhurst's matches were replicated by searching with this source, at his published entry-points. Retrieving novel series generated by Soal's method—as performed by Medhurst to validate his file (Pratt, 1971)—was also successful in this case, for series of 25 digits in length. These results suggested that the file was effectively equivalent to those final digits published in *Chambers' Tables*.

Computation of Chance-Control Series

As raised above, we should want for chance control (1) a random organization of *Chambers' Tables*; (2) permutation of the digits 1 to 5 of the final digits of the *Tables*; and (3) a randomly generated set of digits fulfilling the specifications of the *Tables*.

Fulfilling the first of these ends, 1,000 randomly ordered lists of the final digits of 7-figure logarithms were compiled by effecting a Fisher–Yates shuffle of the digits with the Mersenne Twister algorithm[6] as the basis of randomization. This algorithm was re-seeded prior to generating each list by a 32-bit integer randomly generated by the software PCQNG; which is described by its purveyors as a truly random event generator.[7]

Programmatic construction of the *permuted* source involved looping through the final digits of the 7-figure logarithms for 1 to 100,000, once for each permutation of the digits 12345 (e.g., for 12354, 12534). On each loop, and for each digit in the range of 1 to 5, the program exchanged the value for that corresponding to the digit in its position in the current permutation. So, for example, for the permutation 12354, the digit 4 in the list of final digits became 5, and the digit 5 became 4. This procedure yielded 119 files based on the final digits of 7-figure logarithms, one for each permutation of the series 12345.

For a purely *random* sample, PCQNG was used to construct 240 files each consisting of 100,000 digits in the range of 0 to 9.

A few hours were required for the random shuffling of 1,000 samples of the final digits, and several days were required for generation of the PCQNG data. Given these and other aspects of method, with confidence it can be stated that replication of the here-described method should reproduce the values here reported, within at least two or three decimal places.

Search Logic

Matches of the Soal–Goldney series were sought by the same search routine for all sources, including taking digits at 100-step intervals from their ten-thousandth digit. Searches progressed through each digit from each series, counting up the length by which its subsequent digits appeared in the "true" and chance-control sources. If a match failed because it was not a digit in the range 1–5, the next 100th digit in the source was tested for identity. If a match failed because it was not identical to the tested series, while being in the range 1–5, the test of that digit was aborted, and the next digit in the series (otherwise, the first digit of the next series) was tested. In this way, every digit of the target series was tested for identity of itself and its subsequent digits with every digit of the 7-figure logarithmic and three chance-control sources.

Results

Number of Matches Compared to Empirical Expectation

Hypothesis 1 stated that there would be a greater number of matches from the 7-figure logarithms than from the chance-control sources. Probability values (p_e) were obtained by summing the counts in the chance-control sources that were greater than or equal to the "true" count, and dividing by the sample size.[8] For brevity, particular series that were matched will be referred to as retrievals; given the probabilistic basis of their identification, they should naturally be always understood as *ostensible* retrievals.

10-digit series. The largest number of continuous digits that could be retrieved was 10. There were three matches of this length, two from run 25-1a, and another from run 24-2b. Naturally, no meaningful test of hypotheses could be offered by such a small-sized distribution. Given that there was no clear basis to pre-empt the length of series to be matched, further searches were conducted for matches of nine digits in length. This length is equal to that ultimately tested by Medhurst (1971).

9-digit series. Of the 30 sample series, one continuous segment of nine digits in length from eight series was retrieved from the 7-figure logarithmic source. Counting all the ways in which these series could be retrieved from this source, there were 16 matches. These counts are listed in the "*N* series matched" and "*N* matches" columns, respectively, in Table 2, together with the corresponding mean retrievals from the chance-control sources. It can be noted from Table 2 that about twice as many matches were yielded from the 7-figure logarithms than from the chance-control sources; and these counts were at least two standard deviations beyond each chance-control mean count. These results, singly and in combination, represent independent confirmations of Hypothesis 1. However, these results depend upon a tally of only 16 values from the "true" source. A distribution this small can well be suspected of offering spurious confirmations. Therefore, before progressing to tests of Hypothesis 2, it was considered useful to assay the number of retrievals by testing for 8-digit series. This should offer sufficient data to overcome spurious confirmations, and, indeed, a more optimal number of observations for testing Hypothesis 2.

8-digit series. The number of 8-digit series matched with the 7-figure logarithms was 19, 63% of the sample. When accounting for all possible segmentations of these 19 series, 57 retrievals were obtained, compared to about 40 from each chance-control source (see Table 2). As indicated by the values of p_e, the "true" value of 57 retrievals of eight digits in length offered, in comparison to the values obtained by the chance-control sources, repeated confirmation of Hypothesis 1 that the target series were more likely to be retrieved from a table of 7-figure logarithms than could be expected on the basis of various specifications of chance.

TABLE 2

Number of Matches from 7-Figure Logarithms; Averages (SDs) per Control Sources

Source	8-Digit Series			9-Digit Series		
	N Series Matched	*N* Matches	p_e	*N* Series Matched	*N* Matches	p_e
7-figure logarithms	19.00	57.00		8.00	16.00	
Shuffled source (*N* = 1000)	17.18 (2.31)	39.91 (7.91)	.017	5.15 (2.02)	7.26 (3.32)	.016
Permuted source (*N* = 119)	17.28 (2.26)	41.41 (9.08)	.059	5.44 (1.92)	7.76 (3.49)	.034
Random source (*N* = 240)	17.30 (2.22)	39.93 (7.65)	.021	5.09 (1.90)	6.94 (3.01)	.017

Entry-Point Analysis

Hypothesis 2 stated that there would be a greater number of matches from the 7-figure logarithms for the range of entry-points from 10,000 to 19,999, relative to chance-control sources and other entry-point intervals. Counts were taken of how many segments of the target series could be retrieved from an entry-point within the intervals of 10,000 to 19,999; 20,000 to 29,999; and so on, until 90,000 to 99,999; for each source of digits.

9-digit series. A tentative stab at the hypothesis was firstly offered by looking only at the retrievals of target segments to the length of nine digits. At this length, the mean number of matches per interval from the 7-figure logarithmic source was 1.78 (*SD* = 1.72). The count for the number of matches from the key interval of 10,000–19,999 within the 7-figure logarithmic source was 5. This was the maximum value over the entry-point intervals; almost 2

*SD*s beyond the mean count, and constituting 31% of all entry-points from this source. From the *shuffled* source, we obtain p_e = .009. From the *permuted* source, we obtain p_e = .0252. From the *random* source, we obtain p_e = .0125. With appreciation of the small numbers involved, these analyses singly and collectively offer tentative confirmation of Hypothesis 2.

8-digit series. This outcome was maintained for retrievals of 8-digit series. The mean number of matches per interval for the 7-figure logarithmic source was 6.33 (*SD* = 4.36); the 10,000–19,999 range bore the maximum of 14 (25%); almost two standard deviations beyond the mean. How often did the chance-control sources offer at least as great a number of retrievals within this range? From the *shuffled*, *permuted*, and *random* sources, respectively, the "true" count deviated from that obtained by chance with p_e = .003, zero, and .00417, indicating that the observed "true" count (of 14) was again reliably higher than any count obtainable by way of the chance-control sources.

Looking over the entire range of nine entry-point intervals, from 10,000 to 99,999, the goodness-of-fit of the counts from the "true" source was very low in relation to each chance-control source. Percentages of shared variance between the "true" and chance-control counts only ranged from 1% to 5%, while, among themselves, the chance-control sources shared from 44% to 86% of their variances. There was, in fact, little to tell between the outcomes from the chance-control sources. This can readily be seen in Figure 1, where counts of the matches per source and entry-point intervals for 8-digit series are presented. The number of matches per source within the 1–999 entry-point range is also presented—amounting to the preliminary first *four* pages of *Chambers' Tables*. As part of the rationale for Hypothesis 2, it was reasoned that Soal would ignore these pages; and, indeed, as Figure 1 shows, not a single count was obtained from the target series for the 7-figure logarithms in the corresponding range. The sudden beyond-chance surmount in the subsequent range—comprising the next 20 pages of *Chambers' Tables,* starting from 10,000—is also clearly observable.

Again, the statistical results singly and collectively compelled rejection of the null hypothesis, and acceptance of Hypothesis 2, the hypothesis of early entry into the *Tables*. Accordingly, we can conclude that the target series were quite *unlikely* to have been obtained by random selection from tables constructed in the manner of the last digits of 7-figure logarithms, but *quite likely* to have been derived from within the first 20 pages of a source such as *Chambers' Tables*.

It can further be noted from Figure 1 that the rationale for Hypothesis 2 was also represented in the next most frequent entry-point interval being the second interval, corresponding to the next 20 pages of *Chambers' Tables*. This data pattern was also seen with the 9-digit series, when the second-highest count (4)

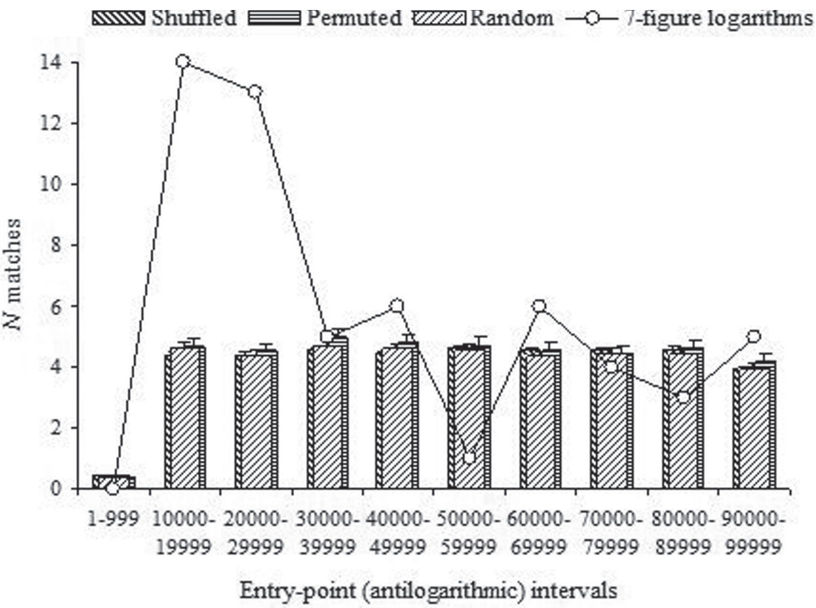

Figure 1. Retrieval counts (mean + SE) per entry-point interval and digit sources for target series of 8 digits in length.

fell within the 20,000–29,999 bin, and the remaining bins bore counts ranging only from 0 to 2.

A more exacting post hoc hypothesis. This result encouraged some extra confidence in the rationale for the hypothesis of early entry, viz., by suggesting the further hypothesis that the 14 matches of 8-digit series should, for the most part, have been obtainable within an even more restricted range of the digits 10,000 to 14,999—i.e. within what would amount to the first *10 rather than 20* pages of *Chambers' Tables*—if, indeed, the result reflected the physical conditions of obtaining digits from a published table of 7-figure logarithms. The count of retrievals with entry-points within range from 10,000 to 14,999 was 10; only 4 retrievals occurred in the subsequent range of 15,000 to 19,999. Nothing of this character appeared in the chance-control sources: e.g., the *shuffled* source gave an even split of 2.19 and 2.20 as mean counts within these intervals; and, more generally, within the 10,000–14,999 range, only miniscule chance retrievals were obtained, the means ranging from 2.19 to 2.48, with *SDs* ranging from 1.76 to 2.11. Compared to the "true" value of 10 counts, it is clear, without statistical testing, that this extended hypothesis of early entry into *Chambers' Tables* was reasonably confirmed.

Discussion

In comparison to empirical expectation, the number of retrievals of the Soal–Goldney target series from a source such as *Chambers' Tables* was observed, in Experiment 1, to be exceedingly more likely than that obtainable on the basis of chance. Not only were significantly more retrievals of the published target series possible from the source and by the method originally reported by Soal and Goldney (1943), but the particular series that were retrieved bore the character of having been constructed by a process identical to that described in their report, and as suggested by reviewing the procedure and conditions. That is, we can quite graphically and statistically appreciate that a published table of 7-figure logarithms was sourced by skipping its first four pages, preferentially taking entry-points within its first 20 pages for antilogarithms greater than 10,000, then the next 20 pages, and thereon making a continuous sweep through the publication.

As the results were reliable over differently constructed chance-control sources, we can be confident that the results were independent of any question of the underlying randomness of the source. That is, it made very little difference whether the chance-control sources were composed of digits that were identical in their sequencing and frequency distribution with a table of the final digits of 7-figure logarithms, or merely shared their frequency distribution, or were, indeed, fully generated as independent ("truly random") events.

One 9-digit match was obtained with a series sourced from Medhurst's (1971) study, i.e. within Sitting 16. That this was not identified by Medhurst as a match exposes the limitation of his method: He did not permit matches from all possible segmentations of a series, but only those he himself delimited on some arbitrary basis.

An arguable limitation to conclusions is that the gross number of series retrieved (see the "*N* series" column in Table 2) did not significantly deviate from the numbers retrieved by chance. However, this level of analysis does not take into account the number of possible segmentations of each series. There is always the possibility—especially for the longer, 25-digit series—that some segments of the series should be more likely to be retrieved than others, given the assumptions reviewed in an earlier section ("Revisiting the Assumptions"). The number of times any segment of the series could be retrieved therefore represents the most appropriate level of analysis—not the number of series themselves that can be retrieved. On this basis, there was no question that the 7-figure logarithmic source yielded significantly more retrievals than expected by chance. Naturally, some retrievals were merely chance matchings, and we must include overlapping and repeated segments of the series in the counts. This was theoretically necessary as we have no a priori basis for segmenting the series. Yet the present method ensured that such overlaps and repeats were just as likely to occur from the chance-control sources as they were from the key 7-figure logarithmic source. We can therefore be confident that the results, based on the number of segments of each series that was retrieved, reliably lead us to support of the alternative hypotheses.

This reliable confirmation of the hypothesis of early entry into *Chambers' Tables* suggested that an even more stringent test would involve requiring that the match of any segment of a series, at the *largest* possible length, that occurred at the *earliest* point into the *Tables,* should be taken *exclusively* of any other possible matches. Any remaining digits in the series should be treated in the same way, so that the search strategy is both *exclusive* and *exhaustive*. This would have the particular advantage of permitting the chance-control sources to return retrievals at greater lengths than found for the "true" source. This reiterative, exclusive, and exhaustive search strategy formed the basis of a second experiment.

EXPERIMENT 2. EXHAUSTIVE, EXCLUSIVE, AND EARLIEST RETRIEVALS

On the basis of the results of Experiment 1, it was feasible to attempt to retrieve the Soal–Goldney target series by an exclusive and exhaustive search. Also, it would be desirable to not have a fixed segment length for all series, but to vary the length, starting from the size of the series itself, and reiteratively searching for retrieval, by an ever-decreasing range, until a maximal series length. Effectively, we abandon Hypothesis 1, concerning the gross number of matches, while Hypothesis 2—the hypothesis of early entry—remains relevant. Replication of the post hoc confirmation of the hypothesis of early entry for the first 5,000—rather than 10,000—of entry-points into (or 10 rather than 20 pages of) *Chambers' Tables* was also hypothesized.

Moreover, it was considered that this procedure should force the chance-control sources to represent Benford's Law, while the corresponding antilogarithms with a leading digit of 1 by way of the "true" source should *surmount* Benford's Law. This would naturally arise when the search strategy preferentially sampled entry-points into the table with the leading digit 1, with the leading digits 2, 3, and up to 9 being successively ever less likely. Perhaps a limitation was that the range of antilogarithms covered only one order of magnitude: After leaving the first range with leading digits of 1 (from 10,000 to 19,999), we do not return to them by sampling for logarithms beyond 99,999. Still, it could be reasonably hypothesized that Benford's Law would predict the distribution of figures so contrived, and that the search strategy reasonably fulfilled what is known of the Law's conditions (Fewster 2009, Hill, 1995, Raimi, 1976). In this way, hypothesis testing could involve theoretical as well as empirical expectation.

Method

The sample of target series, and the chance-control files, were identical to those used in Experiment 1. The search logic, however, was modified so that, initially, only the one maximally sized retrieval was accepted for each target series, and then only by the earliest of entry-points retrieving this length. Whatever portion was not thus matched was then searched for independent retrieval at another entry-point. Segments of less than five digits in length were not searched; this criterion being imposed prior to any searches being conducted, in the interest of economizing on search time, and considering that the approach would be sufficiently informative without these additional searches.

An exceptionally long amount of time was required to search for retrievals on these bases, and, accordingly, the chance-control sources were identical to those used in Experiment 1, excepting that the first 240 of the 1,000 files of the *shuffled* source (and so as many as constituted the *random* source) were tested. Testing these samples itself required 12 days of continuous computerized searching to complete.

Results

The retrieval counts by the method of exhaustive search for maximal-length segments at the earliest antilogarithms within each source are presented in Table 3. Naturally, as expected, the gross number of matches was identical over each source of digits. However, there was a divergence in derivation frequency via the "true" 7-figure logarithmic source versus chance expectation, per each chance-control source, for the number of retrievals within the first 5,000 digits. Specifically, 18 retrievals were obtainable by way of the "true" source, compared to an average of about 11.67 from each chance-control source (with *SD*s ranging from 2.93 to 3.38). Values of p_e are listed in Table 3, and these can be compared with those for the neighboring region of the next 5,000 digits. There was a total conformance within this neighboring range to empirical expectation, while the small values of p_e for the primary range were strongly consistent with the hypothesis that the target series were derived by early entry into a table of 7-figure logarithms.

Figure 2 permits assaying this result in the context of counts from all entry-point intervals. The observably smooth decline in counts per chance-control sources well reflects the bias toward early intervals incorporated in the search criteria. In fact, the number expected within each range on the basis of a logarithmic number line—given in Figure 2's dotted line—was tightly matched by the chance control sources. These represent a near-perfect representation of Benford's Law. This can be better appreciated by considering the 10,000-digit intervals. For example, for the *shuffled* source, there was an average of 19.98

TABLE 3
Retrieval Counts for Longest Non-Overlapping Segments per Series

Source	*N* Matches	% Matches 10,000–14,999	p_e	% Matches 15,000–19,999	p_e
7-figure logarithms	64.00	28.13		12.50	
Shuffled source ($N = 240$)	65.39	17.83	.063	12.72	.596
Permuted source ($N = 119$)	65.51	17.88	.042	13.13	.613
Random source ($N = 240$)	65.27	17.81	.017	12.24	.521

retrievals from the first 10,000-digit interval up to 20,000; comprising exactly 30.56% of the average total (over the 240 files) of 65.39 retrievals, which is precisely the proportion of antilogarithms with the leading digit of 1 predicted by Benford's Law. With the average count in the next 10,000-digit interval being 11.10, we have 16.98%, where Benford's Law predicts 17.61%; a match-to-law with a trivial sampling error. The corresponding proportions were 31.02% and 17.44% for the *permuted* source, and 30.05% and 17.24% for the *random* source. Quite unambiguously, chance matching of the logarithmic final digits conformed to law for the antilogarithmic leading digits.

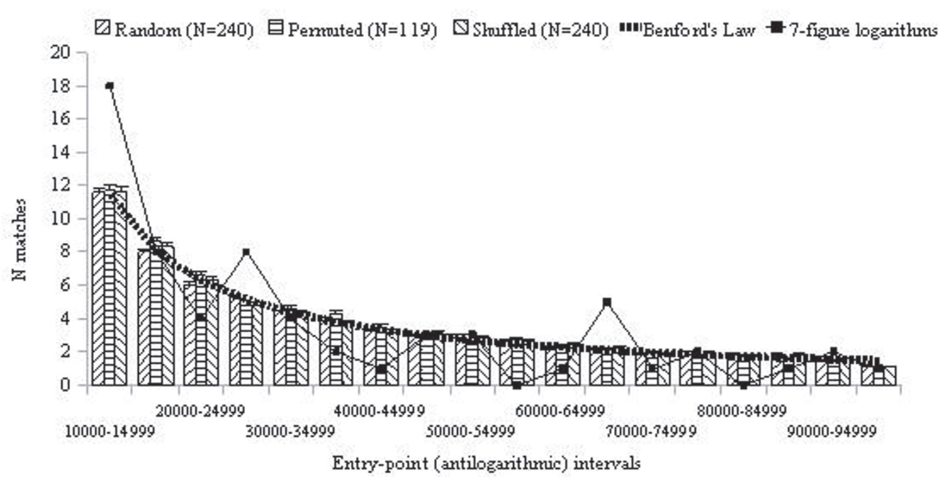

Figure 2. Retrieval counts (mean + SE) per entry-point interval and digit sources between 10,000 and 100,000, with exhaustive search of earliest and longest segments, and per Benford's Law for leading digit *i* (Equation 1).

Still with reference to Figure 2, search of the Soal–Goldney target series from the “true” source, as reported by Soal, yielded a very sharp deviation at the outset from chance-wise and lawful expectation; but an always chance-wise and lawful decline thereafter. The Soal-wise count of antilogarithms with 1 as their leading digit comprised 41% of all retrievals—clearly surmounting the 30.10% predicted by Benford’s Law for a chance operation, and reproduced by each of the chance-control series. With 18.75% in the next 10,000-digit interval, we have a sudden return to Law. The minor peaks observable in Figure 2 at intervals starting at 25,999 and 65,999 were not significant; e.g., $p_e = .163$ and .0750, respectively, in comparison to the number of *random* retrievals at these intervals.

This deviation can be further identified and assessed in comparison with binomial-theoretical expectation;[9] the associated *p*-value indicated as p_b. By testing the deviation of the observed from expected proportional frequencies, we find that the expected first-digit frequency (of .301) for the digit 1 was significantly surmounted by way of the 7-figure logarithms ($1p_b = 0.0473$), but not by way of random digits ($1p_b = 0.446$).

This close representation of Benford’s Law by the chance-control series extended to the second digits, i.e. when splitting the antilogarithms commencing with 1 down the middle. To obtain the proportion of *pairs* of leading and subsequent digits *ij* in accordance with Benford’s Law, the *general significant-digit law* (Hill, 1995, Eq. 4) can be rewritten as:

$$P(ij) = \log_{10}\left(1 + \frac{1}{ij}\right) \tag{2}$$

In order to obtain the proportion expected for a range of consecutive digit pairs such as 10, 11, 12, etc., we sum the proportions obtained from Equation 2 for each pair included in the range. This gives for leading digits in the range of 10 to 14 the expectation of 17.61%. The subsequent and equally sized range of leading digits from 15 to 19 is expected to comprise 12.49%. Referring to the subtotals in Table 3 shows, yet again, a *very* close fit of each of the chance-control sources with these theoretically expected values. The Soal–Goldney target series, however, *surmounted* the expectations of Benford’s Law precisely *and only* in the earliest range up to 14,999 when retrieved from the 7-figure logarithms. That is, the effect observed for all entry-points commencing with 1 was restricted to the range of 10 to 14 as leading digits, which equates to the antilogarithms in the first *10*—rather than *20*—pages of *Chambers’ Tables*. By the binomial-theoretical, we obtain, quite significantly, $1p_b = .0252$ for the first 10–14 range, but, quite negligibly, $1p_b = .556$ for the subsequent 15–19 range, indicating the predicted surmount.

Manifestly insignificant were the values of $1p_b$ of .262 and .176 for the same ranges, respectively, by way of the random digits.

Could the Soal–Goldney series—as retrieved from *Chambers' Tables*—do even better in relation to Benford's Law—by the disproportion in this range being accounted for by even the very earliest pages of the *Tables*? To answer this question, the data presented in Table 4 were sought; for brevity, only the *random* source of the chance-control data are given in addition to the proportions expected by Benford's Law; while Figure 3 presents the same data but with respect to all chance-control sources. The percentages expected by Benford's Law can be seen to have been very closely reproduced by the *random* control series (by summing, for each of its 240 files, those entry-points that yielded the target series that commenced with 10, 11, 12, etc., and dividing by *N* retrievals from the full 100,000 digits over the 240 files). Only one value among these (for entry-points commencing with 12) significantly deviated from Benford's Law, but it was of a minor negative deficit that was quickly overcome by neighboring values, and hence of chance character.

TABLE 4
Theoretical and Observed Proportional Frequencies for Antilogarithms 10,000 to 19,999

Leading Digits	% Expected by Benford's Law	Observed for Soal–Goldney Series from Random Digits		Observed for Soal–Goldney Series from 7-Figure Logarithms	
		%	$1p_b$	%	$1p_b$
10	4.14	4.18	.4163	9.38	.0489
11	3.78	3.96	.1161	6.25	.2227
12	3.48	3.23	.0474	4.69	.3850
13	3.22	3.34	.2023	6.25	.1509
14	3.00	3.10	.2376	1.56	.4248
Subtotal	17.61	17.81	.2627	28.13	.0252
15	2.80	2.63	.0985	3.13	.5387
16	2.63	2.58	.3489	1.56	.4951
17	2.48	2.40	.2648	0.00	.2001
18	2.35	2.38	.3992	4.69	.1904
19	2.23	2.25	.4203	3.13	.4186
Subtotal	12.49	12.24	.1760	12.50	.5559
Grand sum	30.10	30.05	.4462	40.63	.0473

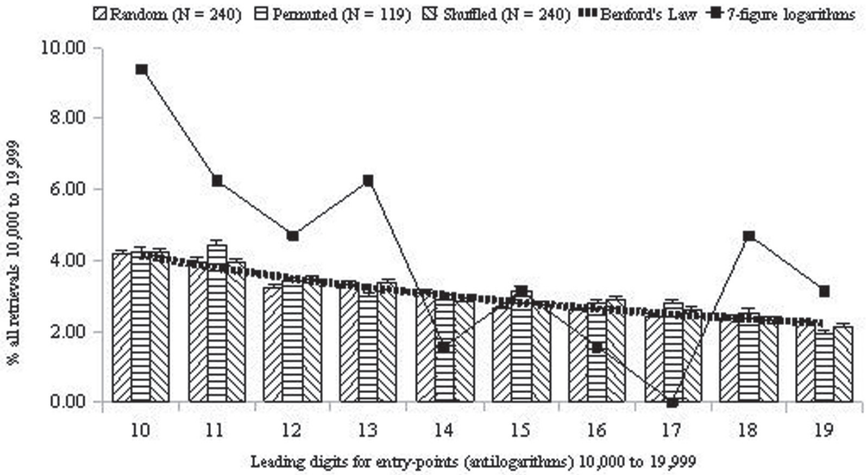

Figure 3. Theoretically and empirically expected, and observed, retrievals for entry-points between 10,000 and 19,999, with exhaustive search of earliest and longest segments (mean + SE), and per Benford's Law for leading digit pairs *ij* (Equation 2).

When, however, we examine the percentages for the target series retrieved from Soal's 7-figure logarithmic source, we see that, yes, indeed, the effect thus far observed was essentially contributed by the *very earliest* of entry-points into the *Tables*. That is, not only was the effect of early entry restricted to antilogarithms commencing with digits in the range of 10 to 14, but *within this range* the greatest contribution was given by those antilogarithms commencing with 10 (i.e., 10,000 to 10,999). This equates to preferential derivation of the target series from within the first *two* pages of *Chambers' Tables*. The observed proportion in this range was more than twice that theoretically and empirically expected. No proportion other than that for these very first two-page entry-points differed from expectation by such a scale.

A further characterization of the results is offered by Nigrini's (1996) *distortion factor* (*DF*), a proportional measure based on the deviation of the observed from the expected mean for a Benford-conforming series, after scaling the data within the range of 10 to 100. With respect to our hypothesis, we should expect *negative DF* values, indicating that the values in the data tended to be smaller than expected for series conforming to Benford's Law. From Nigrini's Equation 6, we obtain, for the antilogarithms yielded by searching through *Chambers' Tables*, $DF = -0.107$, indicating that, in the predicted direction, there was a very substantial 11% excess of lower-valued leading digits. From Nigrini's Equation A9, we obtain the variance of *DF* expected for the relevant

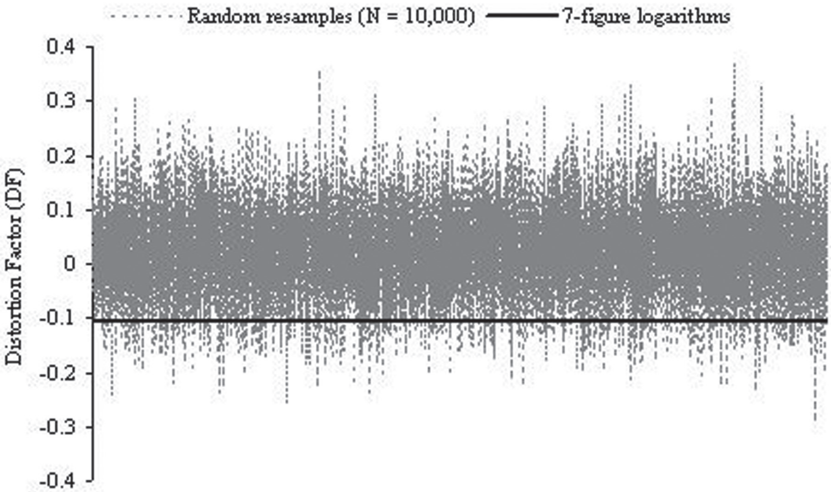

Figure 4. Contra-Benford Distortion Factors for antilogarithms yielding the Soal–Goldney target series by 7-figure logarithms, and randomly resampled random digits.

sample size, and, assuming normality, therewith construct a Z value. Here we find that the *DF* for the antilogarithms based on *Chambers' Tables* tended toward the conventional level of significance: $Z = -1.343$, $1p = 0.0897$; reflecting a strong effect being weighed in terms of a relatively small sample size of 64. For the random control data, its more than 15,000 observations should have been enough to offer suggestiveness for even a weak effect; we obtain $Z = 0.558$, $1p = 0.577$ (*ns*).

For a distribution-free approach that is sensitive to the difference in sample sizes, we can take the *DF*s for the set of antilogarithms yielded by each of the 240 random files, and assess the proportion of the 240 *DF*s that are less than or equal to the one obtained by way of *Chambers' Tables*. This yields $p_e = .0625$. More formidably, we can take, say, 10,000 samples of $N = 64$ from all 15,664 entry-points offered by way of the 240 *random* source files. Sampling with replacement from the yield of all these files at once—using the Mersenne Twister algorithm re-seeded on a random basis, after every 1 to 20 samples by PCQNG—yielded a comparable value of $p_e = 0.0526$ (see Figure 4). In conclusion, even when not looking specifically at the antilogarithms commencing with 1, or the range 10–14, or only 10, we find that the distribution of antilogarithms offered by *Chambers' Tables* reliably tended to surmount Benford's Law, specifically within what amounts to the earlier rather than the later of the *Tables'* pages.

Discussion

The results of Experiment 2 immediately recall the quote from Soal's (1940) *Fresh light* report detailing, with examples, the procedure he followed in deriving the target series. Soal gave examples of his entry-points as antilogarithms commencing with the digit pair 10—and this is precisely what we find substantiated by attempting to retrieve the allegedly fraudulent target series by retracing the method he reported for their construction. The deviations we find all bear the mark of a human hand drawing digits from *Chambers' Tables* by preferential entry into its earliest pages—clearly beyond the natural bias for entries in this range.

More generally, the results showed that the effect observed in Experiment 1 indicative of the Soal–Goldney target series having been sourced by early entry within *Chambers' Tables*—as originally reported—was robust even under the conditions that (i) only the earliest of entries from all sources should serve as the basis of chance estimation and observed count; and (ii) the longest possible retrievals from each target series were exhaustively obtained. Under these conditions, a positive disproportion of matches was expected for the retrieval counts true to Soal's method, versus those expected on the bases of chance and Benford's Law. With already stringent search criteria and sampling that might well have favored the chance-control sources, this hypothesis was confirmed at ever more concise resolutions of the relevant antilogarithms.

Summarily, there was no reason to doubt that the target series were sourced other than how they were originally reported to have been, while positive evidence was acquired that they were indeed thus sourced.

EXPERIMENT 3: ALTERNATIVE LOGARITHMIC SOURCES

Upon concluding his study, Medhurst (1971) confronted Soal with the conclusion that the target series could not have been compiled as Soal had originally reported. Trusting the validity of Medhurst's finding, Soal—the octogenarian—was distressed by this information, and he offered in explanation that, perhaps, he had formed the erroneous recollection that the publication he used in 1941 to prepare the target series was the same that he had used in his earlier studies; perhaps he had even sampled the digits at intervals other than 100 (Soal, 1971).

If a table of logarithms other than *Chambers' Tables* had, then, been used, it might offer a larger number of retrievals than that obtainable from a source of 7-figure logarithms. Alternatively, if the 7-figure logarithms produce the highest retrieval counts, we could be only more confident in the accuracy of Soal's reporting, and all the more assure ourselves that the number retrieved from the 7-figure logarithms has not been a mere fluke based on some fortuitous distributional properties of the digits 1–5 in the 7-figure logarithms. Markwick

(1978), indeed, offered a test based on a table of 6-figure logarithms, and one allowing for discrepancies from intervals of 100. Readers were informed that these efforts "met with no success."

What other forms of logarithmic tables might be tested? From surveys and catalogues of mathematical tables (especially Comrie, 1948, Fletcher, Miller, Rosenhead, & Comrie, 1962, Henderson, 1926), it can be learned that, by 1941, a wide array of tables was available, differing by logarithmic precision and antilogarithmic range. The 4-figure tables appear to have been the most popular, followed by 6-figure tables; but also prominent were 10-, 12- and 20-place logarithms for integers in the range of 1 or 10,000 to 100,000, plus several series offering 5- to 8-place logarithms in a similar range. Additionally, a search of library reserves for such materials revealed (most conspicuously) a 7-place source for the numbers 20,000 to 200,000 (Sang, 1915), and a 16-place source of *natural* logarithms for the numbers 1 to 100,000, in two volumes (and the decimal numbers from 0.0001 to 10.0000 in a further two) (Lowan, 1941).

There is no positive indication that Soal resourced any of these publications, and his suspicion that he was in error in describing *Chambers' Tables* as his source was only made—as we may thus far be obliged to conclude—under dubious compulsions to do so. However, as a test of alternative logarithmic sources could provide a comparative assay of the robustness of the idiosyncrasy thus far observed for the 7-figure source, searches for the target series were conducted on these sources, with the same hypotheses as applied to Experiment 1. Given the arbitrariness of the possible interval alternative to 100 through the tables, this factor cannot here be informatively pursued.

Method

The method was identical to that used in Experiment 1, including the same sample of target series and the fixed-length search logic, but only attempting retrievals of *eight* digits in length. Perl functions were again used to produce the files of final digits for logarithms taken at precisions of 4, 5, 6, 8, 10, 12, 16, and 20 decimal places, always assuming the same method of construction as used for *Chambers' Tables*.[10] These files were searched in the range of 10,000 to 99,999. Additionally, 7-figure logarithms in the range of 20,000 to 199,999, were searched.

For the lower range, no new chance-control sources were necessary; expectation and variance could be reliably obtained from the *random* source, which offered a normal distribution of retrieval counts. In order to provide for empirical expectations up to 200,000 digits, the economical approach was taken of combining and shuffling the digits from each one of the 240 *random* source files with a file randomly selected without replacement from the *shuffled* source, thus producing 240 files of 200,000 digits 0–9.

Results

Retrieval counts are presented in Table 5. For ease of comparison, the first row re-presents the counts obtained via the "true" source of 7-figure logarithms from Experiment 1. It will be readily noted that the alternative precisions produced retrieval counts that were generally within the range of chance expectation. The proportional frequencies of counts within the *random* chance-control source that were at least as great as that observed per source are represented in the p_e column of the table. Assuming normality and using the mean and variance of the *random* source as the basis of assaying deviation essentially yielded the same indications of significance. For the chance-control files of the count from the table of 7-figure logarithms up to 200,000, we obtain $Z = -0.859$, $1p = .805$. However, there were some deviations among the alternative sources with respect to the gross number of retrievals that will merit discussion.

As for the proportion of counts that fell in the critical range of entry-points (<20,000), the eight alternative sources generally produced what could be expected of a uniform distribution of counts over the nine possible intervals (about 11%). As listed in Table 5, the p_e values for deviation of the counts from expectation indicated that, according to the criterion of early entry, there was absolutely no practical use of logarithmic tables apart from the table of 7-figure logarithms. For the two-volume set of 16-figure natural logarithms, an indication of artifactuality would have been obtained if, at the start of the second volume (from 50,000), some disproportion in matching were obtained; however, quite unlike the result for the first volume, there was not a single retrieval among the first 5,000 antilogarithms of this second volume. For the source of 200,000 entries, we can only start with the second interval of 20,000 to 29,999; and here, as already indicated in the results for Experiment 1, there was some modest deviation—secondary in significance to the first interval—of the "true" count of 13 from that predicted by the chance control, in this case 4.833. All the counts at further intervals into the tables, up to 199,999, were quite in conformity with chance; the highest remaining count being only 7, for the very last interval.

Discussion

A search for the sample of the Soal–Goldney target series from logarithmic sources other than those reported by Soal and Goldney (1943) failed to reproduce the results of Experiment 1 obtained with the "true" reported source. This was with respect to the number of matches obtainable by any segmentation of the target series, and the particular number of matches that fell in the critical range that amounted to the first 20 pages of *Chambers' Tables*.

TABLE 5
Matches of Target Series with Alternative Logarithmic Sources

Source	*N* Series Matched	*N* Matches	p_e	% Matches within 1st 9,999	p_e
Range 10,000 to 99,999					
7-figure logarithms	19	57	.021	24.56	.004
4- figure logarithms	6	66	.000	13.64	.079
5- figure logarithms	10	35	.754	5.71	.892
6- figure logarithms	15	38	.583	7.89	.796
8- figure logarithms	16	46	.229	17.39	.108
10-figure logarithms	18	40	.496	15.00	.350
12- figure logarithms	20	51	.088	7.84	.629
16- figure logarithms (natural)	18	32	.867	15.63	.496
20- figure logarithms	20	52	.075	13.46	.213
Range 20,000 to 199,999					
7-figure logarithms	23	70	.817	6.04	.021

An idiosyncratic result was obtained for the deviation of the gross number of matches via the 4-figure logarithmic source. This was the only case that showed a deviation greater than that obtainable by way of the "true" source. However, this result was based on retrievals from a very small number of target series; 65% of its 66 retrievals coming from repeated retrievals of a small range of digits within run 24-1a, almost all from the entry-point interval of 60,000 to 64,999. Performing a search with this series using the method of Experiment 2 revealed that its maximal retrieval length was 9, and involved matches of only three target series. This result for the 4-figure logarithms must therefore be received as an aberration, based on some fortuitous correspondence of a small range of digits. However, in any subsequent studies, it would appear to be useful to assess derivation from the 4-figure as well as 7-figure logarithmic tables.

Additionally, it can be noted that the 20-figure logarithms gave some marginally appreciable deviation from expectation of the gross number of matches by p_e; there were only five more matches (about 9% of the total) by way of the 7-figure logarithms in comparison to both alternative sources. Whatever might be the correct interpretation of this marginal result (including chance), its occurrence in the presence of the very clear confirmation of Hypothesis 2 renders the latter only more remarkable: Even when the alternative logarithmic sources give retrievals on par with, or even in excess of, the "true" source, they cannot match it for indication of early entry within their published tables.

General Discussion

After examining the question as to the source of the Soal–Goldney target series, it appeared inadvisable, on historically evidential and logical grounds, to assume capacity to retrieve complete runs of 25 digits by retracing the manner in which they were reportedly generated (Soal & Goldney, 1943). Not even any long series of a particular length could be assumed to be retrievable, contrary to the unstated but apparent assumptions of Medhurst (1971) and those who replicated his efforts (Markwick, 1978, Scott & Haskell, 1974). This critique accorded with comments by Pratt (1971) on Medhurst's research; and, when so informed, searches were conducted that indicated that the null hypothesis of non-derivation from a published method of using 7-figure logarithms was most unlikely, relative to chance-control sources based on shuffling or permuting the final digits of the logarithmic source, and randomly sampling its range of digits. Also, the proportion of entry-points in the range 10,000–19,999, among these matches, was consistently greater for the 7-figure logarithmic than the chance-control sources. These entry-points from the logarithmic source represented those within the first 20 pages of *Chambers' Tables* (starting from 10,000) that accorded with the finer details of Soal's description of his method. By effectively reducing the size of the bins within which the antilogarithms were tested—from 10,000 to 5,000 and then 1,000—this pattern was reliably observed, without any increase in sample size, to be restricted to what amounted to the first *10,* and then the first *two,* pages of *Chambers' Tables.* This sign of early entry into the *Tables* was clearly indicated as non-artifactual by its lack of reproduction by alternative logarithmic sources as well as each chance-control source. It was also robust against the allowance for chance-control retrievals to be entirely based on the earliest of entry-points, and in relation to Benford's Law concerning the distribution of leading digits. Given that it was reasoned that Soal would have started his searches, in the main, within these first 20 pages, and the extant evidence indicated as much, these results suggest that the target series in the sample were generally obtained in the manner originally reported by Soal and Goldney (1943).

Relation to Earlier Efforts at Target Retrieval

The conclusion suggested by these results is at odds with those offered by Medhurst (1971), Scott and Haskell (1974), and Markwick (1978) for their searches of the target series. This discrepancy is simply effected given that the earlier reports lacked explicit statement of their search assumptions, and the criteria by which the likelihood of derivation from *Chambers' Tables* could be adduced. Earlier results were compared with neither theoretical nor empirical expectation; readers were only offered something of a standalone statistic: "enough of them," "no identifiable match," and "drew a blank." In contrast, the present findings have been based on judging retrieval against reliable and replicable empirical and—where appropriate—theoretical values of what should be expected; and the approach has been consistent with all the documented facts, and has relied on no novel assumptions regarding how the digits were sourced, compiled, and eventually used. Given these differences between the studies, we can expect no comparability of their results.

Generalizability

The present results have been based on 30 target series, from perhaps as many runs, and eight sittings, from a possible 372 runs over 40 sittings. Are we yet permitted to draw conclusions about the population of target series from these results?

Simply in terms of sample size, and the publication status of previous assays of the target series, this appears, by precedent, to be permissible. Table 6 presents the counts and proportions relevant to this comparison. In comparison to Medhurst's (1971) study, which involved only 12 segments from 6 to perhaps 12 runs, the sample is extremely ample; and, indeed, encompasses and goes beyond his sample. Also, the scale of the present sample is similar to that on which claims of data manipulation have been based. Specifically, Hansel's (1959) statistics were based on 51 runs within 6 sittings, restricted to those involving a "rapid-rate" of target assignment, but then not even all runs under this condition. While Scott and Haskell (1973, 1974) tested all of the relevant target series, their eventual claim pertained to the 14 runs of Sitting 16, selected on the basis of Albert's allegation, and then the 12 runs of Sitting 8, and the first 6 of the 20 runs of Sitting 17, selected on the basis of a search for a data pattern identical to that observed in Sitting 16; see Pratt (1974:104) concerning this post hoc basis of their results, and Scott and Haskell (1975:222) for a rebuttal. Markwick's (1978) final statistical result—as best as can be figured from her tables—was based on only 13 runs administered within 7 sittings, the 41 "extra digits" encompassed within these runs selected from an original yield of 93. Several subjective criteria were employed in this cull, e.g., of what were described as "weak," "ambiguous," and "apparently discrepant" "extra digits," as well as "unmatched" digits. Then, the statistical result offered on this remainder amounted to a conservative confirmation of the null. In contrast, the present results are based on an *objectively limited* sample of 30 target series from 8 sittings that often equates with or exceeds these former studies in terms of sample size, while relying on neither subjective nor post hoc processes in its constitution, and being based on conventional statistical arguments. However, the entire target series were not available for testing the hypotheses; the sample was drawn from what of the population has been published, rather than randomly selected, such that it might be reasonably argued that conclusions should be constrained to the sample itself but could be well predicted for the population.

TABLE 6
Sample Sizes in This and Previous Studies of the Soal-Goldney Target Series

Study	Sittings		Runs		Trials	
	N	% 40	*N*	% 529	*N*	% 12,650
Hansel (1959)	6	15.00	51	9.64	1,139	9.00
Medhurst (1971)	4	10.00	12	2.27	114	0.90
Scott & Haskell (1974)	2.3	5.75	32	6.05	768	6.07
Markwick (1978)	7	17.50	13	2.48	266	2.10
Present study	8	20.00	30	5.67	564	4.46

Note: For Medhurst (1971), and those series of the present study derived from his study, run and trial numbers are the maximum possible, including all tested digits, given that runs were not identified in the report. For Scott and Haskell (1974), *N* trials is given by the number of +1 trials in their 32 runs. For Markwick (1978), *N* trials consists of the length of the "interrupted duplicated sequences" listed in her Table 7, less those involving the Stewart study. Percentages of trials are taken with respect to the number of possible +1 hits (or +2 for rapid-rate) conducted under telepathy, clairvoyance, and other conditions, less missed trials (Soal & Goldney, 1943:95–97); or for on-trial hits if there was a suggestion to score on-trial (Soal & Goldney, 1943:49,55).

Implications for the Fraud Scenarios

What implications of these results are there for the fraud scenarios? Most clearly, Medhurst's (1971) conclusion that the target series could not have been derived as originally reported must be queried: When respecting the conditions of their production, particularly as advised by Pratt (1971), the target series bear the marks of having been produced from a table of 7-figure logarithms, as originally reported. Scott and Haskell repeatedly referred to Medhurst's *null* conclusion as corroborative but *necessary* evidence in support of their investigation of Albert's allegation; they defined its evidentiality as equal to the loss of the original records, stating (somewhat ambiguously) that "These had to happen on our hypothesis and the fact that they did happen provides a little further confirmation" (Scott & Haskell, 1975:222). This necessary "evidence" can, in the light of the present results, be judged to be of quite arguable merit, if not nonexistent—even if we only consider the results for the sub-sample tested by Medhurst and upon which Scott and Haskell relied.

However, while the present results confirm Medhurst's original hypothesis, they do not serve his objective to vindicate Soal; and it would be wrong to interpret these results as somehow implying Soal's innocence. Our methods only involved, and the results confirmed, that "retrieval" should be possible for any small segments of target series. This was a necessary assumption given such procedures as haphazardly compiling the series into, and drawing them from, a pool of digits; copying errors; reusing prior series with various transformations, including omission of prior hits and reversals—*and/or* such manipulations as stacking the series with 1s, and altering the 1s into other digits. In this way, it involves no paradox to hold that the target series were sourced as originally reported, but also manipulated.

Furthermore, we need not compel the fraud scenario to predict that the points at which the "retrieved" series were segmented should, by and large, be those points at which Markwick found "extra digits" to occur.[11] Consider, for instance, the manner in which Experiment 2 offered "retrieval" of run 25-1a. This involved, firstly, a 10-digit match, surrounded by two 6-digit matches, accounting for 22 of its 25 digits, as follows (the different segments separated by dashes).

(5)143125(32)-2543251314-232154

Those digits that did *not* match (shown above in parentheses) were two from the start of the series, and two before the 10-digit match commenced. Those digits that Markwick identified as manipulated (underlined in the above) were *included* in the "retrieved" 10-digit segment. This cannot be used to suggest, however, that the indication of these digits as manipulated is somehow "wrong." We have, after all, only obtained this "retrieval" by a statistical process, and can make no claim that any particular case represents the "real" source of the digits from *Chambers' Tables*. Secondly, we cannot be sure that in any case of reuse the first-used series is the original and the later-used series is its copy—such that any "extra digits" may, in fact, be the result of omissions from one or the other series, rather than insertions into the later-used series. Also, the scenario of manipulation suggested by Markwick's study does not limit manipulation to those digits that appear to be altered in the process of duplicating an already-used series; that limitation is only in the nature of her evidence, not in its implications. The above example could, after all, have been drawn from the original pool of digits, and manipulated at that point. The unmatched digits, then, could point to manipulated digits that were not apparent by Markwick's method.

The original objective, as stated in the beginning, was, indeed, to identify the source of the target series so that Markwick's findings of data manipulation could be extended beyond cases of reused target series. Some ideas for future research, consistent with the interest in extending if not confirming Markwick's findings, can be offered. The present research suggests that we can pursue this objective by simulating the pool on the basis of its reported source, i.e. *Chambers' Tables*. However, we need to suppose that Soal made use of "extra digits" and other alterations not only when reusing target series *from one run to another*, but also when transcribing the digits *from the pool onto the target sheets*—and perhaps even when copying digits from the *Tables* into the pool. Then, on the basis of Markwick's and the present findings, we could hypothesize that those digits within a series that fall outside the starts and ends of any particular "retrievals" are more likely than not to occasion hits; and that they could mostly have been additional 1s that were altered into 4s and 5s. But this cannot be predicted for all series; the manipulations would not, for instance, offer any advantage when Soal had no sure access to the target sheets. So we could restrict our hypothesis to those runs when Soal acted as the agent's experimenter, and/or was involved in the final scoring. When, as in Sittings 23, 24, and 25, sheets with "extra digits" appear to have been used in runs when Soal was responsible for the guess sheets, we could suppose that this simply occurred because Soal had prepared many sheets for manipulation, and he only needed to reduce the proportion of 1s *before* the sitting to that predicted by chance, and then manipulated (only) the guesses during the sitting. In this case, the hits should *not* tend to fall at the junction of any "retrievals." Clearly, to support the necessary assumptions and various factors implied in these predictions, a larger sample of target series than has been presently available is required.

There are a couple confirmations of previously raised points in the context of the fraud scenario that can be mentioned. Firstly, the present results inform against the particular contra-fraud model of fortuitous reuse. This is because it is far less likely that duplications should fortuitously arise if Soal commenced his searches within only some few early pages of the *Tables* rather than freely entered them all over the volume. This interpretation is not clear-cut; as previously noted, for Soal's 10,078 example of an entry-point, we can find several long duplications commencing with antilogarithms less than 11,000. Still, it might be tentatively concluded that support for the hypothesis of early entry is incompatible with a model of fortuitous reuse. Additionally, the footnoted statement by Soal and Bateman (1954) that specified the dates of the sittings for which *Tippett's tables* were used must be inaccurate. The statistical details suggest that *Tippett's tables* were used for *all* the runs within these sittings. As the deviations from chance in the present study were somewhat dependent on including runs from these sittings, accepting the present results implies that the targets for these runs were more likely to have been generated from a table of 7-figure logarithms than a source such as *Tippett's*.

In summary, the present results suggest that (1) the fraud scenarios cannot be clearly rationalized on the basis that the target series cannot be retrieved from the reported source, but that (2) extending the search for evidence of data manipulation by identifying the source of the target series might well return to the originally reported source.

Notes

[1] It must be noted that Medhurst was chronically ill by the time he composed this report (Barrington, 1971, Goldney, 1974); it was published two months after his decease. What was published appears to be an unrevised manuscript, given, as others subsequently noted (Scott, 1971; *Editor's note* on p. 203 of the same volume), that it contains several statistical and linguistic errors and ambiguities, including a crucial statistical test that Medhurst offered regarding the target digits, whereas it is clear that he confused these with the response digits. The presently described limitations—suggesting at least a hurriedness in Medhurst's conclusion—must be put to the same account.

[2] Soal and Goldney (1943) did not specify the runs for which *Tippett's tables* were used. Yet Soal and Bateman (1954:137n) later specified these as Sittings 24, 25, and 26. Still, the method of entering *Tippett's tables* was not described, nor was it stated that they were used for *all* 42 runs of these sittings. This might, however, be assumed to have been the case, as, following this statement, the result for "480 (+1) trials" on these sittings was given. This number corresponds to the 20 telepathy runs, at "normal-rate," with the usual agent (Elliott), among the 42 runs that were administered within these 3 sittings.

[3] An alternative and somewhat more economical approach would be to generate random target series, or reorder and permute the original targets, and test them for retrieval from the 7-figure logarithms in comparison to the original targets. This approach should yield the same results as using the true target series against randomly constructed search lists. Initial indications were, indeed, that the approaches yielded identical results. The present approach was, however, adopted, as it was considered to offer greater face validity to apply randomization to the source digits rather than to "tamper" with the original target series.

[4] A report by Pratt (1951) reproduced targets and responses for 2 25-digit runs from Sitting 32. However, these were represented by the letters A to E, which, the text explained, substituted for the target initials (E, G, L, P, Z). As the digits 1 to 5 were randomly assigned to the targets upon every second run, it is not clear how they correspond to letters A to E. Being reliant on the record of digits, the present study could not include these series in the sample.

[5] This copy of the *Tables* in fact extends to 100,009; but in deference to other editions that were available at the time, the present study sought for matches only up to the logarithm for 100,000.

[6] This was as implemented by the Perl module Math::Random::MT, available from http://www.cpan.org

[7] Specifications of this software are available from http://comscire.com/Home

[8] Some readers might prefer to calculate p_e with the addition of 1 to the numerator, and perhaps also with the addition of 1 to the denominator. Multiplying the stated probabilities by the sample size, as given in Table 2, permits such calculations. The results obtainable thereby represent no meaningful restriction on the results here reported. Additionally, standard normal deviates and associated probabilities were calculated. The chance-control counts were not always normally distributed, as indicated by the (extremely conservative) Kolmogorov–Smirnov test, although the values for skewness and kurtosis were always quite small, and there was typically little difference between the mean and median counts. These values are available from the author by request.

[9] For example, the *random* source yielded 654 among 15,664 entry-points that commenced with the digits “10”. This compares with 648.375 entry-points according to Benford’s Law (i.e. when entering “10” into Equation 2, we obtain .04576, which we then multiply by 15,664). The binomial distribution is referred to in relation to 15,664 trials, 654 “hits,” and a theoretical probability of .04576.

[10] With respect to conventional limits in computational resources, a different set of functions—from the Perl module Math::BigFloat—was required to generate and store the decimal strings for the 16- and 20-figure logarithms than those with shorter decimal strings. Testing this module by using it to also produce the final digits of base-10 7-figure logarithms, and comparing them with those produced by the method otherwise used (which simply relied on Perl’s *log* and *sprintf* functions) revealed 110 discrepancies, presumably attributable to different rounding conventions. A sample of about 20 of these discrepancies was checked for identity with *Chambers’ Tables* (Pryde, 1930), and the standard (*log-sprintf*) method was found to give final digits *always* agreeing with the *Tables*. Accordingly, results for the 16- and 20-figure logarithmic sources should be treated with particular reserve.

[11] This point is raised in consideration of an interpretation canvassed by an anonymous reviewer of this article.

References

Barrington, M. R. (1971). Dr. R. G. Medhurst 1920–1971 [Obituary]. *Journal of the Society for Psychical Research*, *46,* 124–126.

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

Comrie, L. J. (1948). *Chambers’s Six-Figure Mathematical Tables*. London: Chambers.

Fewster, R. M. (2009). A simple explanation of Benford’s Law. *American Statistician, 63,* 26–32.

Fletcher, A., Miller, J. C. P., Rosenhead, L., & Comrie, L. J. (1962). *An Index of Mathematical Tables* (2nd ed.). Oxford: Blackwell Scientific.

Goldney, K. M. (1974). The Soal–Goldney experiments with Basil Shackleton (BS): A personal account. *Proceedings of the Society for Psychical Research, 56,* 73–84.

Hansel, C. E. M. (1959). Experimental evidence for extra-sensory perception. *Nature, 184,* 1515–1516.

Henderson, J. (1926). *Bibliotheca Tabularum Mathematicarum. Part I: Logarithmic Tables*. London: Cambridge University Press.

Hill, T. P. (1995). A statistical derivation of the significant-digit law. *Statistical Science, 10,* 354–363.

Lowan, A. N. (Ed.) (1941). *Table of Natural Logarithms*. New York: Federal Works Agency.
Markwick, B. (1976). [Letter]. *Journal of the Society for Psychical Research, 48,* 287–288.
Markwick, B. (1978). The Soal–Goldney experiments with Basil Shackleton: New evidence of data manipulation. *Proceedings of the Society for Psychical Research, 56,* 250–277.
Markwick, B. (1985). The establishment of data manipulation in the Soal–Shackleton experiments. In: *A Skeptic's Handbook of Parapsychology* edited by P. Kurtz. Buffalo, NY: Prometheus. pp. 287–312.
Martin, D. R., & Stribic, F. P. (1940). Studies in extrasensory perception: III. A review of all University of Colorado experiments. *Journal of Parapsychology, 4,* 159–248.
Matlock, J. G. (1987). Archives and psychical research. *Journal of the American Society for Psychical Research, 81,* 233–255.
Medhurst, R. G. (1971). The origin of the "prepared random numbers" used in the Shackleton experiments. *Journal of the Society for Psychical Research, 46,* 39–55.
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. (1996). A taxpayer compliance application of Benford's Law. *Journal of the American Taxation Association, 18,* 72–91.
Pratt, J. G. (1951). The reinforcement effect in ESP displacement. *Journal of Parapsychology, 15,* 103–117.
Pratt, J. G. (1971). [Letter]. *Journal of the Society for Psychical Research, 46,* 199–202.
Pratt, J. G. (1974). Fresh light on the Scott and Haskell case against Soal. *Proceedings of the Society for Psychical Research, 56,* 97–111.
Preece, D. A. (1981). Distributions of final digits in data. *Statistician, 30,* 31–60.
Pryde, J. (Ed.). (1930). *Chambers's Seven-Figure Mathematical Tables*. London: Chambers.
Raimi, R. A. (1976). The first digit problem. *American Mathematical Monthly, 83,* 521–538.
Rosenthal, R. (1987). How often are our numbers wrong? *American Psychologist, 33,* 1005–1008.
Sang, E. (1915). *A New Table of Seven-Place Logarithms of All Numbers from 20000 to 200000*. London: Charles and Edwin Layton.
Scott, C. (1971). [Letter]. *Journal of the Society for Psychical Research, 46*, 252–253.
Scott, C., & Haskell, P. (1973). 'Normal' explanation of the Soal–Goldney experiments in extrasensory perception. *Nature, 245,* 52–54.
Scott, C., & Haskell, P. (1974). Fresh light on the Shackleton experiments? *Proceedings of the Society for Psychical Research, 56,* 43–72.
Scott, C., & Haskell, P. (1975). Fraud in the Shackleton experiments: A reply to critics. *Journal of the Society for Psychical Research, 48,* 220–226.
Smith, S. W. (2006). *The scientist and engineer's guide to digital signal processing* (revised edition). San Diego, CA: California Technical.
Soal, S. G. (1940). Fresh light on card-guessing—Some new effects. *Proceedings of the Society for Psychical Research, 46,* 152–198.
Soal, S. G. (1966). *ESP—A Scientific Evaluation by C. E. M. Hansel: A Reply*. London: Powlsons.
Soal, S. G. (1971). A note on the evaluation of card-guessing experiments [Letter]. *Journal of the Society for Psychical Research, 46,* 202–203.
Soal, S. G., & Bateman, F. (1954). *Modern Experiments in Telepathy*. London, UK: Faber.
Soal, S. G., & Goldney, K. M. (1943). Experiments in precognitive telepathy. *Proceedings of the Society for Psychical Research, 47,* 21–150.
Soal, S. G., & Pratt, J. G. (1951). ESP performance and target sequence. *Journal of Parapsychology, 15,* 192–215.
Uhler, H. S. (1938). Log π and other basic constants. *Proceedings of the National Academy of Sciences, 24,* 23–30.
Wassermann, G. D. (1975). The soul of Soal [Comment]. *Journal of the Society for Psychical Research, 48,* 89–91.