

Lucky numbers: Choice strategies in the Pennsylvania Daily Number game

ANDREA R. HALPERN and SCOTT D. DEVEREAUX
Bucknell University, Lewisburg, Pennsylvania

We examined the amount of money bet during a week of Pennsylvania's Daily Number game. In this game, players receive a predetermined payoff for picking the 3-digit number (000 to 999) drawn on that day. The betting distribution was distinctly nonuniform, and we identified several betting patterns, such as picking triples and avoiding double 9s. In addition, we asked separate groups of subjects to rate selected numbers for randomness, luckiness, and perceived history of winning; to categorize numbers; and to free associate to numbers. We propose that people seem to choose highly patterned, available, and/or "lucky" numbers. People apparently do not bet numbers that reflect the random process of the game (they do not utilize a representativeness heuristic).

In this paper, we explore the betting strategies that people employ in what should objectively be a strategy-free situation. The Pennsylvania Daily Number game is one of several lotteries run by the state. Every day except Sunday, a number from 000 to 999 is drawn at random. An important point to note is that the game is not parimutuel: payoffs do not depend on the total amount of money bet on that number. For instance, winners receive a payoff of 500 to 1 on a straight bet, regardless of how many other people win on that day. This eliminates the necessity for analyzing complex strategies arising from perceived need to avoid popular numbers, etc. Therefore, number choices in this game should reflect only psychological and not game-strategic influences.

A number of hypothetical strategies for choosing numbers exist that would result in nonuniform distributions of preferences. One strategy assumes both that the bettor knows that the winning number is selected at random, and that he or she seeks to imitate that process by in turn selecting a random-looking number. That is, a bettor might use a representativeness heuristic (Kahneman & Tversky, 1972) and select a random-looking number. Although it is difficult to specify exactly what such random-looking numbers would be, we should expect to see an avoidance of numbers with obvious patterns: triples (111), runs (456), or other sequences (246).

A second strategy assumes that people note and remember previous winning numbers (not necessarily consciously). At the time of data collection in this study, the game had been running approximately 7½ years, and thus approximately 2,200 winning numbers had been selected. Use of this knowledge and the representativeness heuristic

could lead to the Gambler's Fallacy: a tendency to avoid choices that have been recently (or frequently, in this case) selected, in the mistaken belief that successive draws are in fact dependent on previous draws.

Use of an availability heuristic might lead to yet a different pattern of choices. Tversky and Kahneman (1973) have shown that people tend to overestimate the probability of items that are easy to recall. If subjects are using availability in playing the Daily Number, they might in fact choose numbers that had appeared frequently in the past, or those with obvious patterns that might come easily to mind (contrary to representativeness).

The first study consisted of a descriptive analysis of a week's worth of bets in the Pennsylvania Daily Number. In the second to fourth studies, some specific hypotheses about what was governing the choice behavior were tested, by comparing data from experimental subjects with those from actual bettors.

STUDY 1

Method

Data generously provided by the Pennsylvania Bureau of State Lotteries consisted of the amount of money bet in the Daily Number on each of the 1,000 possible choices for the period July 31 to August 7, 1984 (excluding Sunday, August 5). Bets can vary from \$.50 to \$5.00, and the Bureau of State Lotteries collects data only on the total amount bet on each number, not on the number of people betting. For most analyses, the amount bet was summed across the 7 days. However, data for each day were available for separate analyses where appropriate.

Results

A graph of the total amount bet for each possible betting choice is shown in Figure 1. The trends most noticeable in this display are: several prominent peaks, a pattern over the entire range of numbers, and a pattern within each group of 100 numbers (deciles). The range of bets was \$1,925 to \$147,383, with a median of \$9,278.¹

An obvious feature on the graph is the clear preference for tripled digits: 777, 333, and 111 are most popular,

We gratefully acknowledge the assistance of the Pennsylvania Bureau of State Lotteries in providing the Daily Number data. We thank Lari Roberts, Barbara Maynard, Susan Milton, and Lisa Ludemann for their help in data collection and analysis. Address correspondence to Andrea R. Halpern, Department of Psychology, Bucknell University, Lewisburg, PA, 17837.

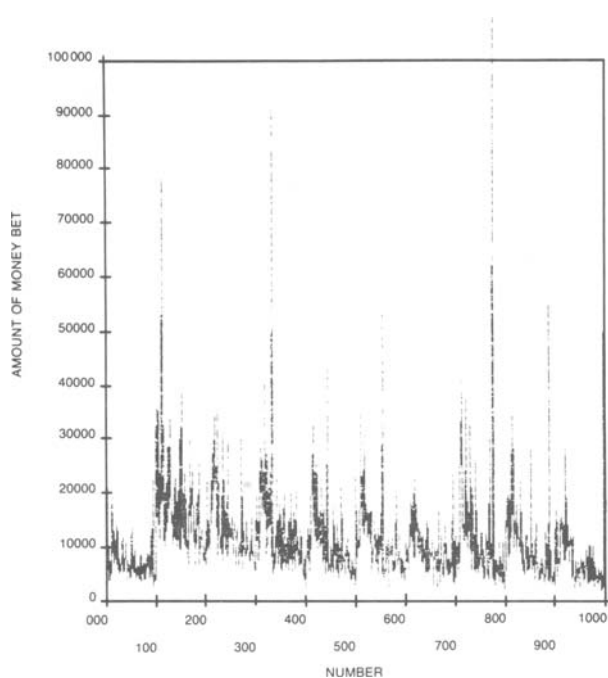


Figure 1. Amount of money bet on each of the 1,000 possible numbers in the Pennsylvania Daily Number game during the week July 31 to August 7, 1984.

followed by 888, 555, 999, and 444. Choice 222 comes in at rank 13, behind 714, 318, 151, 721, and 511. However, the remaining triples are somewhat less popular, with 000 at rank 45, and 666 at rank 152. The lower ranking of 000 is consistent with a general trend toward avoidance of numbers beginning with 0. The relative unpopularity of 666 could be due to the association of that number with Satanism. The poll in Study 4 showed that 31 of 54 respondents spontaneously mentioned this association.

Over the range of numbers, there is a trend toward preference for numbers with lower cardinality, with the exception of those in the 001 to 099 range. But superimposed on that trend is a more microscopic trend towards preferring lower numbers within deciles. Thus, the overall pattern resembles a succession of scallops, descending from the 100s to the 900s, with the largest drop from the 100s to the 400s.

Less popular numbers tend to have the same number in two places (e.g., 994 or 665). Of the 25 numbers least often bet, 21 fit this description. Only 18 of the top 100 numbers bet have doubled digits. Even more striking is the avoidance of double 9s. Of the 27 double-9 choices, 18 are in the 100 least bet numbers, and the 8 least preferred bets are all double 9s.

Although we may generalize that people prefer smaller numbers, we must consider the avoidance of the 000 to 099 range. The most heavily bet number in that range was only rank 72 (093). This pattern may result from a strategy of avoiding extremes, or from the use of an availability

heuristic. In most situations other than the Daily Number, numbers under 100 are referred to without the prefixed 0s. Bettors may simply be unaccustomed to producing numbers like "006." Even the potentially meaningful choice of 007 was only rank 945 here.

Birthdays and anniversaries. One possible explanation for the popularity of bets in the lower range of each decile is that people are betting significant dates. This would also help account for the avoidance of 0xx numbers, which cannot possibly be dates. If this hypothesis is correct, we should see a decrease in bets between, for example, 131 (or January 31) and 132 (not a possible date). This in fact happened for seven of the nine months that can be represented in this system. For those months, the mean change between the numbers of interest was a loss of 226 ranks, and the mean decrease in money bet was \$7,045.

Current date. Another betting strategy would be to choose the current date. The hypothesis was confirmed for each date in the sample. On the average, a number played on the date it represented gained 168 ranks and \$1,872, relative to its mean for the week.

Previous day's winner. If people are using a representativeness heuristic, one prediction is that bettors will avoid playing the previous day's winning number. The same winner 2 days in a row might be considered highly unlikely to someone viewing the draws as nonindependent events.

To test this hypothesis, the amount bet and rank on each day were noted for the previous day's winning number. This was compared against the mean performance of that number. The results showed that the previous day's winner was bet less heavily than average for 5 of the 7 days. For all 7 days, the average decrease was \$831 and 103 ranks. Thus we may weakly confirm that people display their notion of a self-correcting random process by this strategy.

Frequency of winning. As noted in the introduction, people may remember the winning history of numbers, and choose to bet previous winners (if they consider them to be "lucky"), or avoid them (another manifestation of the Gambler's Fallacy). To examine this, we correlated the overall rank of each number in the betting data with its frequency of winning since the game was instituted. These frequencies were not uniform (ranging from 0 to 8, where the expected value was 2.2), allowing a meaningful correlation. The Pearson correlation was $-.16$, which was significant, with df of 998. This may suggest that subjects avoid numbers that have won in the past. However, with the correlation accounting for only 2.6% of the variance, we must be very cautious about this conclusion.

Other lucky numbers. Sevens are traditional lucky numbers, and in addition to 777, quite a few numbers beginning with 7 were popular, including the traditional lucky 711 (rank 20). Numbers containing multiples of 7 (714 and 721) were popular, and indeed all the permutations of these two trios were in the top 90 ranks, which may reflect the popularity of "boxing" a bet: the payoff is 80 to 1 if any permutation of a bet wins.

Other regular patterns examined were ascending and descending runs with intervals of one and two digits (e.g., 123, 135, 987, 753). Ascending runs with intervals of one digit were bet more heavily than average, with all such numbers receiving bets above the median. Bets on the other pattern types were above the median amount only slightly more than half the time.

To summarize, the distribution of bets was markedly nonuniform. Several strategies were identified, some of which were consistent with the use of the availability or the representativeness heuristic. Subjects did not, however, appear to be choosing random-looking numbers. Study 2 was carried out partly to assess this possibility more rigorously.

STUDY 2

In order to determine whether people are choosing random-looking numbers, we must gather judgments as to what numbers in fact look random. In addition, we wanted to know whether people might increase or decrease bets according to a *perceived*, rather than actual, winning history of numbers; and finally, whether a perceived "luckiness" of numbers could affect bets.

Method

Subjects. Sixty-one adults of various backgrounds studying or attending summer programs at Bucknell volunteered to serve as respondents. All were Pennsylvania residents.

Materials. A questionnaire was prepared, on which respondents were asked to perform the following three tasks: (1) Rank a set of 11 numbers according to how random each looked, (2) rank a set of 11 numbers according to how lucky each one looked, and (3) rank a set of 11 numbers according to how often you think each has won in the Daily Number.

Three sets of 11 to-be-rated numbers were chosen by first selecting threesomes of numbers comprising betting ranks 1 (most heavily bet), 2, and 3; another, ranks 99, 100, and 101; and so on, to ranks 998, 999, and 1,000. One set consisted of numbers from rank 1, 99, 199, etc.; the second contained ranks 2, 100, 200, etc.; the third contained ranks 3, 101, 201, etc. Three forms of the questionnaire were constructed by rotating the number sets against the tasks. Within each task, the numbers to be rated were presented in a random order.

Procedure. The respondents were familiarized with the Daily Number game if necessary, and then they filled out one form of the questionnaire, which took about 10 min. Approximately one third of the respondents received each form.

Results and Discussion

Averaging the rankings across the three forms of the questionnaire, randomness was found to be significantly negatively correlated with luckiness ($r = -.79$, $p < .01$). That is, the lucky numbers were also the patterned (nonrandom) numbers. Though not statistically significant, a moderately high correlation was found between luckiness and actual selection ($r = .50$), and a negative relationship was found between randomness and actual selection ($r = -.46$). Moderate positive correlations were also found between perceived previous winning and randomness ($r = .40$), and a negative relationship was found between perceived previous winning and luckiness

($r = -.45$). Perceived previous winning did not correlate with actual selection.

Thus, although random-looking numbers are rated as more likely to have won in the past, people avoid random-looking numbers in actual betting. This is consistent with the tendency to avoid previous winners shown in Study 1. The opposite tendency is seen with lucky numbers: these numbers seem to be lucky and bet heavily despite their perceived losing record.

This study supports the idea that luckiness and randomness are usually incompatible when people are choosing draws for the Daily Number. In addition, a tendency to bet patterned and/or lucky numbers was shown. Because the correlations did not quite reach significance, an alternative method to quantify this idea of random and lucky numbers was explored in Study 3.

STUDY 3

In this study, the subjects were asked to categorize a set of numbers as being lucky, unlucky, or neutral; or in a second task, as random, nonrandom, or neutral. "Lucky" numbers were expected to have low numerical (heavily bet) ranks, and "random" numbers to have high ranks.

Method

Subjects. Sixty adults from the same pool as was used for Study 2 participated; however, none had participated in Study 2.

Materials. A total of 120 numbers was selected for study. These comprised numbers bet heavily (ranks 1 to 40), moderately (ranks 480 to 520), and infrequently (ranks 960 to 1,000). These were divided into two sets of odd- and even-numbered ranks, respectively.

The two number sets were rotated against two tasks. In the first, the respondents were asked to "decide whether each combination of 3 digits appears random, nonrandom, or neutral." The subjects placed an appropriate code letter next to each randomly presented number. The second task called for categorization as lucky, unlucky, or neutral numbers.

Procedure. Each subject performed both tasks, with half the subjects classifying by randomness first. Each task utilized a different number set, so that the subjects saw a different group of 60 numbers for their second task. Both tasks together took about 15 min.

Results and Discussion

For each subject's choices of "random," "nonrandom," and "neutral" numbers, respectively, the mean ranking in the actual betting was calculated. An analysis of variance was computed separately for each task. The between factor was number set, and the within factor was category. Number set had no main effect, nor did it interact with classification in either task, so it will not be considered further. For both tasks, the effect of classification was significant. For the luckiness task, the mean ranking for lucky numbers was 428; for unlucky, 500; and for neutral, 512 [$F(2,116) = 3.89$, $p < .05$]. A Newman-Keuls tests revealed that lucky numbers were significantly more heavily bet than unlucky or neutral numbers, which did not differ between themselves ($p < .05$).

For randomness, the mean ranks were 475 for random numbers, 474 for nonrandom, and 558 for neutral [$F(2,116) = 4.41, p < .05$]. A Newman-Keuls test confirmed that the mean ranks of random and nonrandom numbers were identical, each differing from the mean rank of neutral numbers.

Thus, following our prediction, "lucky" numbers were indeed bet more heavily than other kinds. However, both random and nonrandom numbers were bet more heavily than neutral numbers. One explanation for this might be that many of the least popular numbers had doubled digits, as noted in Study 1, and these may of course be seen as "patterned" as well.

STUDY 4

In Study 1, we suggested that people might be betting highly available numbers. One way numbers might be available is for them to be tied to associations in memory. In Study 4, we pursued this idea further by asking subjects to generate associations to each of the 1,000 possible lottery picks. Numbers eliciting the largest number of associations should also be frequently picked.

Method

Subjects. The participants were 215 Bucknell students.

Materials. Because it was impractical to have each subject consider all 1,000 numbers, four forms were created, each containing a random subset of 250 numbers.

Procedure. The subjects were mostly tested in groups. The instructions were as follows: "Look at the list of numbers and write down any which have special meaning (either cultural or personal). Also note numbers which you consider to be 'good' or 'bad.'" An example of "2001" eliciting "Space Odyssey" was given. Approximately one quarter of the subjects saw each sheet, so that each number was presented to an average of 54 people.

Results and Discussion

We eliminated from consideration responses that were either uninterpretable ("mm in 1 ton" for 760) or the many that were idiosyncratic ("my grandmother's address") or applicable only to this sample of people ("our area code"). Examples of responses that we allowed to stand were "airplane" for 727, "TV show" for 222, and any mention of a number's being lucky, unlucky, aesthetically pleasing, etc.

A total of 98 associations was given for the 25 most-bet numbers (17 of those numbers elicited at least one association), but only 8 associations for the bottom 25 (6 of them elicited associations). This suggests that availability does play a role in the selection of top choices. One interesting note is that 38 of the 98 associations to top numbers included the word "lucky," while only two mentions of "lucky" were found in the bottom 25.

We also tested for a relationship between number of associations (testing only those numbers with at least two associations) and rank in betting over the entire set of numbers, but we found a correlation close to 0. However, it should be noted that the number of associations was very far from being normally distributed over the set, so any correlation should be viewed with caution.

GENERAL DISCUSSION

To summarize, players tended to avoid the endpoints of the range of numbers, the upper end of each decile, numbers with doubled digits, and previous winners. Preferred numbers included tripled digits, dates (especially the current date), trios containing 1, 4, and 7; or 1, 2, and 7, and some ascending runs.

Can we find any support for use of a representativeness heuristic in this game? Clearly, bettors were not in general picking numbers that mimicked the random process of the game. We did find modest support for the operation of a representativeness heuristic over the long term, as previous winning numbers suffered somewhat in later betting. However, this strategy cannot predict many of the most and least preferred numbers. Among the top three preferences, 777 had won more frequently than average in the past; 333 and 555 had won about an average number of times. On the other hand, of the two most frequently winning numbers (eight wins), 667 was bet at rank 810, while 695 was bet at rank 66.

Availability fares better as an explanation for some of the preferences. Players seem to bet numbers that have associations in memory (although they avoid such recently encountered numbers as the previous day's winner). Significant dates or the current date may simply be highly available patterns. If bettors generate selections by mentally scanning numbers from 1xx to 9xx, the former class may again be more available, and are therefore bet more heavily than the latter. And as mentioned above, numbers beginning with 0 may be difficult to produce, although they do reasonably well at generating associations.

However, these hypotheses leave a number of betting patterns unexplained. For instance, runs should be highly available, but only one type was bet more heavily than average. If tripled digits are highly available (once the first digit is picked, the same choice for the second and third digits is an easy path), then doubled digits should be available as well, but instead they are avoided.

We wish to propose, therefore, that some numbers are viewed as being lucky. As just mentioned, the popularity of triples *could* be explained by availability, but we may wonder why number 999 was very available (rank 6), while similar numbers were so weak in memory (998 was rank 986; 997 was rank 999, etc.). Perhaps triples are considered to be lucky numbers. Indeed, nearly all of the triples elicited multiple associations of goodness or luckiness (or unluckiness, in the case of 666). The high proportion of bets on numbers containing 7s and multiples of 7 seems also to appeal to traditional notions of lucky numbers.

We do not mean to suggest that choosing "lucky numbers" is a particularly irrational strategy. It is as rational as any strategy in a game where strategies are irrelevant. In general, people are wont to see patterns or skill in randomness (see Gilovich, Vallone, & Tversky, 1985), and it seems obvious that many lottery players try to use "skill" in this inappropriate context. In fact, in discussing this research with other people, we have been asked whether we were trying to figure out the game's "system," or been asked for advice in picking a winning number. Conversely, we have attempted to talk to players about what numbers they pick and why, but these efforts have been met with suspicion.

Finally, we must note that we have so far been unable to explain the popularity of several highly preferred numbers. Of the most popular 25 numbers, 16 can be accounted for by one of the strategies we have so far described. However, the popularity of 318, 151, 511, 101, 105, 815, 115, and 518 remains more obscure, despite some similarity in the presence of 1s and 5s.

REFERENCES

- GILOVICH, T., VALLONE, R., & TVERSKY, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology*, 17, 295-314.
- KAHNEMAN, D., & TVERSKY, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 3, 430-454.
- TVERSKY, A., & KAHNEMAN, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5, 207-232.

NOTE

1. Similar graphs of each decile are available from the first author.