

# The Snow White randomizer: Simple, economical, and the fairest one of all

HERBERT FRIEDMAN

*College of William and Mary, Williamsburg, Virginia 23185*

A convenient and practical method is presented for the generation of unbiased random sequences. The method is based on combining discrete outputs and does not require either precisely specified and unbiased input or complex procedures and equipment.

The problem of producing unbiased random sequences is a major one in parapsychological research and a minor if persistent one in other areas. Machine- or computer-generated random numbers tend to suffer from a variety of problems, particularly sequential dependencies or biased outputs arising from nonrandom states within the device. Prepared lists of "random" numbers (Rand Corporation, 1955) are necessarily subject to the introduction of bias by the selection of an entry point within the list. In many situations involving a limited number of trials or requiring experimental conditions to be presented a predetermined number of times, a restricted random series (Friedman, 1966) has many advantages relative to a purely random series.

In certain research applications, the safest procedure is to generate a random number (specifying a treatment or the correct choice) at the time of the test trial. A variety of devices have been used to this end, ranging from a simple hand-tossed number cylinder (Hamaker, 1949) to triggers controlled by radioactive decay (Schmidt, 1970). An alternative, more convenient, and effective approach is outlined here.

Simple calculation indicates that a small number of biased elements combined under appropriate rules can approximate an unbiased output. For example, two coins that are each likely to land heads 60% of the time ( $p_h = .60$ ) will produce one head and one tail (ht or th) in the proportion of  $2(.6 \times .4)$ , or 48% of the time. Defining State A as one head out of two tosses and State B as zero or two heads leads to the proportion  $A/A + B$  (i.e.,  $p_A$ ) always being closer to 50% than the original  $p_h$  value of each coin (if  $p_h$  or  $p_t \neq 0$ ). If State A in a binary situation is defined as an odd number of Outcome h in the set of N elements, the probability of State A rapidly approaches 50% as the set size increases, even if all the elements are highly biased. Schmidt (1970), in a discussion of a random-number generator, suggested that output bias can be greatly reduced by adding several consecutively generated numbers to produce a "secondary number" (using Modulo 10). The basic formula given by Schmidt is identical to that derived from adding elements in the

binary coin model. For a set of N equally biased elements, the proportion of State A is given by:

$$p_A = \frac{1 + (2 p_h - 1)^N}{2}, \quad (1)$$

while the error or net bias (nb) could be calculated directly,

$$nb = 2^{N-1} (p_h - .5)^N, \quad (2)$$

where  $p_A = .5 + nb$ .

For the situation in which the individual coins are not equally biased,

$$nb = 2^{N-1} (p_{h1} - .5) \times (p_{h2} - .5) \dots (p_{hN} - .5). \quad (3)$$

Therefore, in a set of N coins, if any one (unspecified) coin is perfectly fair ( $p_h = .50$ ), then, by Formula 3,  $nb = 0$  and  $p_A = .50$ , regardless of the degree of bias of the other coins. State A may be viewed as representing every other term in the binomial expansion for  $(H + T)^N$ , that is, those containing an odd number of heads, so that a change in state for any individual element will change State A to State B. It follows that if all the coins but one are glued down (in any given pattern) while the last coin is fair and free to vary, the result would necessarily be  $p_A = .50$ . Therefore, knowledge or control over all but one of the elements in a set would be ineffective in predicting or affecting the series generated if that last element were fair. This last point should be useful in the assignment of blind and double blind research conditions when the subject might be aware of some of the experimental procedures.

If none of the coins is perfectly fair but two or more are free to vary, then  $p_A$  will always be fairer than the fairest element. Formula 3 shows that the net bias decreases rapidly as the number of elements increases, even if all elements are highly (but not totally) biased. For example, with 10 elements each with  $p_h = .80$ , the final  $p_A$  would be .5030 and hardly detectable as

Table 1  
Values of pA, nb, and N(CR2) for Five Identical  
Elements at Several Levels of Bias

ph	pA	nb	N(CR2)
.50	.500000	.000000	∞
.55	.500005	.000005	$4 \times 10^4$
.60	.50016	.00016	>39,000,000
2/3	.50206	.00206	>236,000
.70	.50512	.00512	>38,000
.80	.53888	.03888	662
.90	.66384	.16384	37

not being perfectly fair (an outcome readily confirmed by empirical testing). Randomness is difficult to measure (cf. Chaitin, 1975), but a good working definition of bias for a randomized series employed in research is the number of test trials required to produce a significant outcome, with a critical ratio (CR) of 2, on the basis of that bias alone (Schmidt, 1970):

$$N(\text{CR}2) = 1/(\text{nb})^2. \quad (4)$$

For the above example of 10 coins with  $\text{ph} = .80$ ,  $N(\text{CR}2) > 100,000$ , which is well beyond the number of trials generally used in behavioral studies.

Taking five elements as a practical set size, the probability of State A (pA) for several values of ph, assuming all five elements in the set are equally biased, is shown in Table 1 with the associated values of nb and N(CR2). The outcome is virtually fair over a wide range of bias for the individual elements and, for ph in the range 1/3 to 2/3, five elements (coins) clearly would suffice for acceptable values of pA and N(CR2). In practice, most coins or other available elements (e.g., odd/even outcomes for generated numbers, last digit of telephone book entries, dice, etc.) can be assumed to have ph values suitably close to .50. By expressing the digits 0-9 in binary terms, a group of four unbiased outcomes (A or B) produced as described above will allow the generation of all digits.

When computers or machines are used to produce random sequences, adding several successive digits (Schmidt, 1970) so sharply reduces bias that any of a variety of "random-number generators" would provide acceptable input. If the specification of the items in a series is carried out at irregular or random times, a digital watch is well suited as a substitute for coins or dice. The right-hand digit itself is fair (being in each position 10% of the time), and satisfactory protection against inadvertent prior knowledge of the outcome could be

obtained by adding the hours to the minutes (e.g., 9:28 = 37, 11:05 = 16, etc.) or by taking the momentary readout for seconds (: :01 to : :60). Odd/even may be used for binary decisions, Modulo 6 to replace a die, and Modulo 10 for decimal outcomes. Similarly, the digits of page numbers of books opened at random can be added together and used as above.

In any computer- or machine-generated sequence, the possibility of sequential biasing factors (Davis & Akers, 1974; Gardner, 1977) cannot be ruled out. The use of discrete elements (such as coins) with outcomes combined as proposed here prevents both interactions among the elements and effects persisting from the prior state of a machine. Each coin in the set, even if biased, is clearly independent of the other coins and of its own earlier behavior.

### THE SNOW WHITE RANDOMIZER

Random sequences may very conveniently be produced by repeatedly shaking a cup containing five coins specifying that one, three, or five heads represent State A and an even number State B. Advantages of this procedure are: (1) simplicity with no computer or other complex apparatus or process required; (2) no cues are present prior to the outcome; (3) the state or target for a trial can easily be generated at the time of the response, excellent for gaming strategies and precognition studies; (4) the outcome is unaffected by experimenter expectations or bias; (5) an effectively fair outcome is assured; (6) the total equipment cost can be very reasonable ( $\$ \cong .05$ ).

### REFERENCES

- CHAITIN, G. J. Randomness and mathematical proof. *Scientific American*, 1975, **232**, 47-52.
- DAVIS, J. W., & AKERS, C. Randomization and test for randomness. *Journal of Parapsychology*, 1974, **38**, 393-407.
- FRIEDMAN, H. A restricted random series of numbers. *Psychonomic Science*, 1966, **6**, 311.
- GARDNER, M. ESP at random. *New York Review of Books*, 1977, **24**, No. 12, 37.
- HAMAKER, H. C. A simple technique for producing random sampling numbers. *Proceedings Koninklijke Nederlandsche Akademie van Wetenschappen*, 1949, **52**, 145-150.
- RAND CORPORATION. *A million random digits*. Glencoe, Ill: The Free Press, 1955.
- SCHMIDT, H. Quantum-mechanical random number generator. *Journal of Applied Physics*, 1970, **41**, 262-268.

(Received for publication June 9, 1978.)