ENCODING DIFFICULTY: A PSYCHOLOGICAL BASIS FOR 'MISPERCEPTIONS' OF RANDOMNESS[1]

Clifford Konold
Scientific Reasoning Research Institute
Hasbrouck Laboratory
University of Massachusetts, Amherst

Ruma Falk
Department of Psychology
and School of Education
The Hebrew University of Jerusalem

Abstract

Subjects' ratings of the apparent randomness of ten binary sequences were compared to the time required to memorize those same sequences. Memorization time proved a better predictor of the subjective randomness ratings than measures of the "objective" randomness of the sequences. This result is interpreted as supporting the hypothesis that randomness judgments are mediated by subjective assessments of encoding difficulty. Such assessments are seen as compatible with the information theorists' interpretation of randomness as complexity.

Take a look at the two sequences below. Which sequence, [1] or [2], appears to be the most random?

OXOXOXOXOOOXXXXOXOXOO [1]

OXOXXXXXOXOOOOXOOOXXO [2]

Many will take objection to this question, and understandably so. A recent article by Ayton, Hunt, and Wright (1989) along with a set of published responses in the same journal (Vol. 4, 1991), include a range of arguments for those interested in exploring the debate about the meaning, theoretical status, and psychological investigations, of 'randomness'. We cannot address those issues here.

---

[1] To be presented at the Sixteenth International Conference for the Psychology of Mathematics Education. Durham, New Hampshire (August, 1992). This research was supported, in part, by grant MDR-8954626 from the National Science Foundation to Clifford Konold and by the Sturman Center for Human Development, The Hebrew University, Jerusalem. Opinions expressed are those of the authors and not necessarily those of the sponsoring agencies.

Mathematically, [1] and [2] have the same probability ($.5^{21}$) as any other ordered sequence of the same length of being randomly produced by, for example, flipping a fair coin. On this basis, they could be judged equally random. However, if we consider various attributes of sequences, more of the possible sequences are *like* [2] than they are like [1]. In this sense, [2] might be considered more characteristic of a random process than [1].

One such attribute is the probability of alternation between the two symbols. For every finite binary sequence, we can determine the relative frequencies of the two symbols and the conditional probability of change (or continuity) after a given character in the sequence. Given a sequence length of n, there are n-1 opportunities for a change in symbols. (All but the first character in a sequence can differ from a preceding character). The probability of alternation in a particular sequence, denoted P(A), is obtained by dividing the number of actual changes of symbol-type by n-1. The values of P(A) for [1] and [2] above are 0.7 and 0.5, respectively. When the probabilities of the two symbols are equal, the value of P(A) in large, random samples will tend toward 0.5. This result follows from the principle of independence — regardless of what has already occurred in the sequence, the probability that the next character differs from the previous one is 0.5. Sequences with values of P(A) other than 0.5 occur with less frequency. Additionally, deviations from that modal value are equally probable in the two directions. Thus, sequences with P(A) = 0.7, which contain more alternations than expected, have the same probability of occurring as sequences with P(A) = 0.3, in which there are fewer alternations (longer runs) than expected.

Sequence [2] is considered more random than [1] also from the perspective of information theory. Randomness, in this account, is defined as a measure of *complexity* (Chaitin, 1975; Fine, 1973, chap. 5). Despite the sophisticated computations used in information theory, the notion of randomness as complexity is straightforward: a random sequence is one that cannot be significantly shortened via some coding scheme. This notion can be illustrated with even a simplistic coding convention. For example, the

X O X O X O X O X O X O X O X O X O X O X [3]

perfectly alternating series above can be coded as 10XO 1X (10 repetitions of XO followed by 1 X). By forming the ratio of the number of characters in the *code* (where *10* is considered as one character) to the number in the *sequence*, we can express the complexity (or compressibility) of this sequence as 5/21 = 0.24. Using the same convention, [1] would be coded as 4OX 3O 4X 2OX 2O, for a complexity measure of

12/21 = 0.57. There are 18 characters in the code for [2], only slightly fewer than the 21 original characters; its complexity measure of 18/21 = 0.86 is much nearer the maximum value of 1. Using this coding scheme, [2] would be considered more random than [1], because compared with [1], it cannot be substantially compressed.

Despite mathematical reasons for considering [2] more random than [1], research has shown that most subjects hold to just the opposite. In selecting random sequences, people prefer sequences that include more alternations than typically occur (Falk, 1975, 1981; Wagenaar, 1972). The well-known gambler's fallacy, according to which tails is considered more probable than heads after a run of successive heads, may also be based on the belief that symbols in a random sequence should frequently alternate.

Kahneman and Tversky (1972) have explained these results by suggesting that people rely on error-prone "heuristics." In their account, the judgment that [1] is more random than [2] is based on an incorrect expectation that even small random samples will resemble their parent population (Tversky & Kahneman, 1971). [2] is judged less random because it contains longer runs (e.g., XXXXX) which do not capture or represent the equal distribution of symbols in the population. Random sequences, because they are random, must also avoid obvious patterns. The perfectly alternating [3] is accordingly judged to be less random than [1]. For a sequence to be considered maximally random, it must strike a balance between avoiding simple alternating patterns and maintaining a near equal number of symbol-types in any of its segments.

In the account summarized above, human judgments of randomness are based on the notion of *similarity*. Features of a sample are compared to the corresponding features of a population, and the more similar a sample is to a population, the more likely it is to have come from that population. Our research was designed to investigate an alternative hypothesis — that peoples' perceptions of randomness are based on assessments of *complexity*.

People might assess the complexity of a sequence by gauging how difficult that sequence would be to encode. We frequently are given information that must be copied or memorized. If that information can be reorganized into meaningful "chunks" (cf. Miller, 1965), it can be more efficiently memorized or copied. Chunking is obviously a way of compressing data. Therefore, assessments of "chunkability" are also judgments about difficulty of encoding. We are suggesting that people might make use of this type of assessment in judging the randomness of a sequence.

In the study reported here, we used the time required to memorize a sequence as a measure of encoding difficulty. We compared these times with results from prior research in which subjects rated the perceived randomness of the same sequences. If randomness judgments are rooted in assessments of complexity, we would expect that those sequences which were hardest to memorize would be perceived as most random. Such results would provide evidence of a psychological basis for people's "misperceptions" of randomness — that from the standpoint of human perception, sequences of P(A) = 0.6 *are* more complex, or difficult to encode, than sequences of P(A) = 0.5, and for this reason they are judged as more random. Furthermore, if people base their randomness judgments on the difficulty of encoding, the complexity definition of randomness might prove to be an intuitively compelling introduction to the concept.

Method

Randomness Ratings

Data concerning apparent randomness were obtained in prior research by Falk (1975, 1981). Subjects were shown a set of 10 sequences, which included [1], [2] and [3]. These sequences were of length 21, and comprised two symbols whose frequencies differed by 1. The P(A)s of these sequences ranged from 0.1, 0.2, 0.3 . . . to 1.0. Subjects rated each sequence on a scale that ranged from 1 (not at all random) to 20 (perfectly random). Ratings were obtained from 219 subjects.

Memorization Task

Ten different subjects were individually presented with the same sequences as were used in the rating task. The sequences were presented as shown in Figure 1 on a Macintosh computer.

X X X X X X O O O X X O O O O O O O X X X

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21

Figure 1. Screen display of target sequence P(A) = 0.2

Subjects were instructed to study each sequence until they could reproduce it from memory. When a subject was ready to attempt recall, he or she hit the "return" key. This caused the target sequence to be masked. The subject could then enter a "response" sequence in a field provided on the screen, as shown in Figure 2.

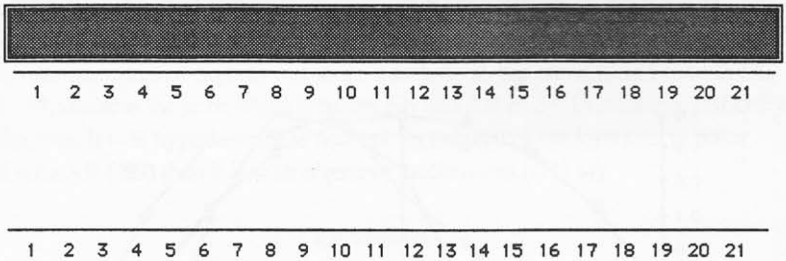

Figure 2. Screen display showing masked target sequence (above) and field for entering response sequence (below).

After entering a response sequence, the subject again hit the return key. If the response sequence was correct, both the target and response sequences were displayed together. The next target sequence could then be displayed by clicking on a "next" button. If the response sequence was incorrect, it disappeared, and the target sequence was again displayed. Subjects continued until they were able to enter the correct sequence.

Subjects were told the computer was recording the total time the target sequence was displayed. They were also told that time spent entering a response sequence was not being recorded and were shown how to use the delete key, which permitted editing a response sequence up to the time the enter key was depressed. They were instructed that the objective was to memorize the sequence as "efficiently" as possible, trying to minimize total viewing time.

The order of presentation of the ten sequences was randomly determined for each subject by the program. These ten experimental sequences were preceded by four practice sequences. The practice sequences had P(A)s of 0.2, 0.9, 0.5, and 0.3 and were always presented in that order. The subjects were not informed that these were practice sequences.

Results

Randomness Ratings

The randomness ratings (denoted AR, for "apparent randomness") for each sequence were averaged over the 219 subjects, and then linearly transformed to range from 0 to 1. Figure 3 shows these averages plotted as a function of P(A). This function peaks at P(A) = 0.6 and is negatively skewed.

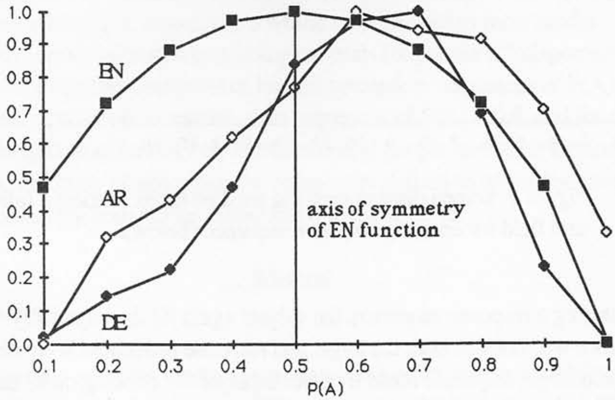

Figure 3. Plot of EN, AR, and DE as functions of P(A).

For comparison purposes, Figure 3 includes values obtained from an "objective" measure of randomness based on the "second-order entropy" (EN) of the sequences (see Attneave, 1959, pp. 19-21). This function peaks at 0.5, and is symmetric[1] around P(A) = 0.5. As reported in the introduction, these data indicate that subjects select as most random, sequences that include more than the expected number of alternations. Indeed, these subjects tended to rate sequences with P(A)s of 0.6, 0.7, and 0.8 as more random than the objectively most random sequence of 0.5.

Memorization Task

The times required to memorize each sequence were first standardized for each subject. For each P(A), we computed the mean of the standard scores over the ten subjects, and then linearly-transformed these to range from 0 to 1. This value, which is our measure of encoding difficulty (DE), is also plotted in Figure 3. The function of DE peaks at P(A) = 0.7.

---

[1] Since in the family of sequences we used there is a sequence with p(A) = 1, but not one with P(A) = 0, the function in Figure 3 is not entirely symmetric.

If encoding difficulty mediates judgments of randomness, than we should expect measures of encoding difficulty to be better predictors of the subjective randomness ratings than are measures of objective randomness. Indeed, the correlation between DE and AR is .89, whereas the correlation between EN and AR is .54. In addition, difficulty of encoding, which was hypothesized to account for subjective randomness, is better correlated with AR (.89) than it is with objective randomness (.71) .

Conclusion

The data presented here offer some support to the hypothesis that judgments of randomness are mediated by subjective assessments of complexity, an assessment that may be accomplished by judging how difficult the sequence would be to encode. The results of the memorization task are preliminary in that they involve only ten subjects, and these were not the same subjects who provided the randomness ratings. We are currently conducting a larger study in which subjects first rate the randomness of various sequences, and then either memorize or copy those same sequences. The copying task allows subjects to enter a sequence in "chunks," copying only what they can easily remember , thus reducing demands on short-term memory.

Though preliminary, our findings do suggest that human judgments of randomness are based in part on the formally sound criteria of complexity. Such a finding could have important implications for instruction. For example, introductions of randomness as a blind process of selection, or as statistical independence, may be difficult to comprehend because students lack prior intuitions into which these ideas can be integrated. Our results suggest that an interpretation of randomness as complexity may have more intuitive appeal to students, and therefore may provide the basis on which an initial understanding of randomness can be constructed.

References

Attneave, F. (1959). *Applications of information theory to psychology: A summary of basic concepts, methods, and results*. New York: Holt, Rinehart & Winston.

Ayton, P., Hunt, A. J., & Wright, G. (1989). Psychological conceptions of randomness. *Journal of Behavioral Decision Making*, 2, 221-238.

Chaitin, G. J. (1975). Randomness and mathematical proof. *Scientific American, 232*, 47-52.

Falk, R. (1975). *Perception of randomness.* Unpublished doctoral dissertation, Hebrew University, Jerusalem (Hebrew with English abstract).

Falk, R. (1981). The perception of randomness. In *Proceedings of the Fifth International Conference for the Psychology of Mathematics Education* (pp. 222-229). Grenoble, France.

Fine, T. L. (1973). *Theories of probability: An examination of foundations.* New York: Academic Press.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology, 3,* 430-454.

Miller, G. A. (1956). The magical number seven, plus-or-minus two, or some limits on our capacity for processing information. *Psychological Review, 63,* 81-97.

Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Review, 76,* 105-110.

Wagenaar, W. A. (1972). Generation of random sequences by human subjects: A critical survey of literature. *Psychological Bulletin, 77,* 65-72.