The Possibility Of Uniform Pseudo-Random Number Generation By A Group Of Humans

Samuel Toluwalope Ogunjo
Department of Physics
Federal University of Technology Akure
Akure, Nigeria
stogunjo@futa.edu.ng

Emmanuel Jesuyon Dansu
Department of Mathematical Sciences
Federal University of Technology Akure
Akure, Nigeria
ejdansu@futa.edu.ng

Oluwagbenga Abiodun Fakanye
Dept. of Community Health Sciences
University of Manitoba,
Winnipeg, Canada
fakanyeo@myumanitoba.ca

Ibiyinka Agboola Fuwape
Department of Physics
Federal University of Technology Akure
Akure, Nigeria
iafuwape@futa.edu.ng

2020 International Conference in Mathematics, Computer Engineering and Computer Science (ICMCECS) 978-1-7281-3126-9/20/$31.00 ©2020 IEEE 10.1109/ICMCECS47690.2020.240869

***Abstract*—The ability of humans to generate numbers that are really random has always been a subject of debate. This paper investigated the possibility for a group of humans to serve as random number generators. A total of 2344 students, who were not pre-informed to avoid bias, from different faculties within the Federal University of Technology Akure were asked to choose a random number between 1 and 10. Using various statistical tests, we sought answers to the possibility of predictors like participant's test score, gender, age and school influencing their choice of random numbers. We discovered that the numbers generated are highly random and chaotic despite number 1 being the most selected number across all predictors that was considered. Our study found that gender, test score, age did not significantly influence the choice of number while faculty showed a significant relation.**

Keywords*—*chaos, random numbers, human generators, predictors, stochastic processes

I. Introduction

Reference [29] defined a random sequence as a sequence of numbers, *n*, within some bounded range, where it is not possible to predict n_{k+1} from any combination of preceding values n_i, $i=0,1,...,k$. According to [16], a random number sequence is uniformly distributed over all possible values and each number is independent of the numbers generated before it. A sequence of numbers would be described as random if an observer were unable to determine a formula for exactly predicting each number in the sequence [20]. Generally, random numbers have elements with every possible statistical feature of unpredictability. These elements satisfy given tests of stochasticity that are universal [17].

In as much as man would like to be able to predict events and occurrences around him, randomness and unpredictability are integral parts of his existence. Since the human mind and senses are wired to observe patterns, human beings might be seen as poor random number generators. For instance, it is extremely difficult for the mind to conceive words that are totally unconnected to initially spoken words [12]. In mathematics and other fields of study, random number generation has been a serious subject of discussion as it has great applications in psychology, psychiatry [7], cryptography [21], Geographic Information System [19], gaming [3, 26], visual arts [20], computer simulations[10], and other fields [6].

The importance of random numbers has made its generation a major research focus. The supposition that randomness occurs in nature is the basis for many theories in science and it constitutes the bedrock of quantum mechanics, as such, it can be assumed that natural occurrences like radioactive decay can be used to generate random numbers [8]. The use of computers and software to generate random numbers has been seen in algorithms such as the Mersenne twister [18] and Algorithm AS 183 [31]. Also, chaos based methods have been proposed and implemented for random number generation. The limitations of computation have drawn attention to random number generators in nature as well as human capabilities. It is still contentious that humans can consciously generate random numbers.

Reference [27] investigated the choices of children and adults in the process of generating random digit sequences. Confirming earlier researches, it was discovered that children and adults tend to have biases for big and small numbers respectively. Researchers have identified factors which can influence number generation by humans to include lateral head turn [15], mental state [22], reading direction [9], hand and mouth action [11] and composite body movement [5]. The study in [24] affirms previous findings revealing the effect of active head rotation on the randomization of numbers in adults who were observed to generate smaller numbers during left rather than right rotation. As opposed to adults, random number generation in elementary school children did not significantly differ between active left/right head rotation.

According to [30], the subjectivity of randomness is used in psychology to explain quite a number of research outcomes. Random number generation by individuals has been posited to be a consequence of health conditions. Random number generation ability is observed to be impaired in patients with early signs of early schizophrenia [4]. Subjects have also been found to exhibit number biases in multiples of 1, 10, and 100 [1]. Furthermore, it has been shown that small number bias show a leftward shift in the number line [14].

Statistical testing is used to verify that random number generators (RNGs), which are critical in building cryptographic algorithm parameters, actually produce numbers that are really random. Certain metrics are thus made use of to establish the randomness of cryptographic RNGs [25]. Reference [28] considered the various approaches that had been developed to evaluate the dynamics

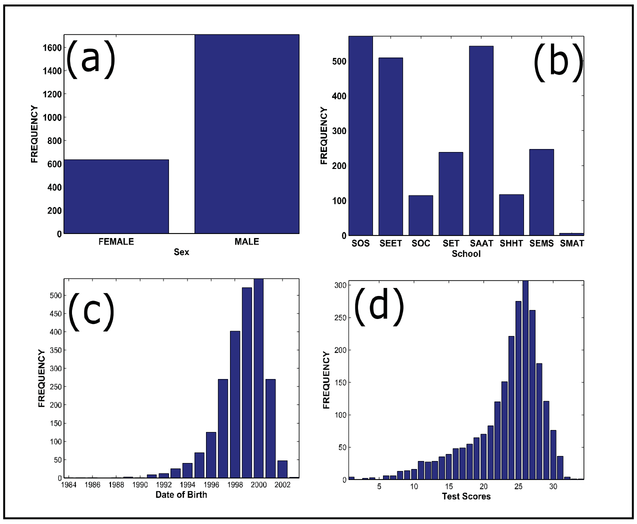

Fig. 1: Distribution of participants based on (a) gender (b) school (c) year of birth; and (d) quiz score.

of random generation of numbers and incorporated the measurement scales into a computer programme known as RgCalc which analyzed the quality of human attempts at such numbers. The programme also provided, for the purpose of comparison, computer-generated pseudorandom sequences.

In this study, we aim to investigate the hypothesis that a group of students with different backgrounds can generate random numbers. Furthermore, we aim to study the influence of age, intellectual capabilities, faculty and gender on both the choice and the randomness of numbers chosen by the students considered.

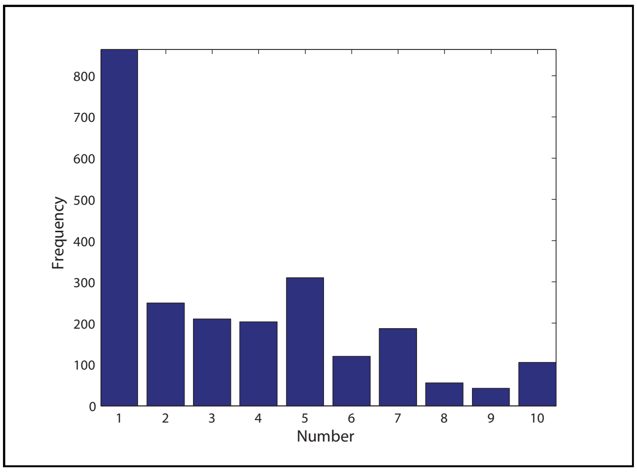

Fig. 2: Descriptive statistics of the chosen number - mode.

In the rest of this paper, we describe the methodology in section 2; the results are discussed in section 3; and we make concluding remarks in section 4.

II. METHODOLOGY

The data for this investigation was obtained from first year undergraduate students of the Federal University of Technology Akure, Ondo State, Nigeria during General Physics I class. A questionnaire, with quiz based on the subject matter taught, was administered on-line and the students were required to fill in any two numbers from 1 to

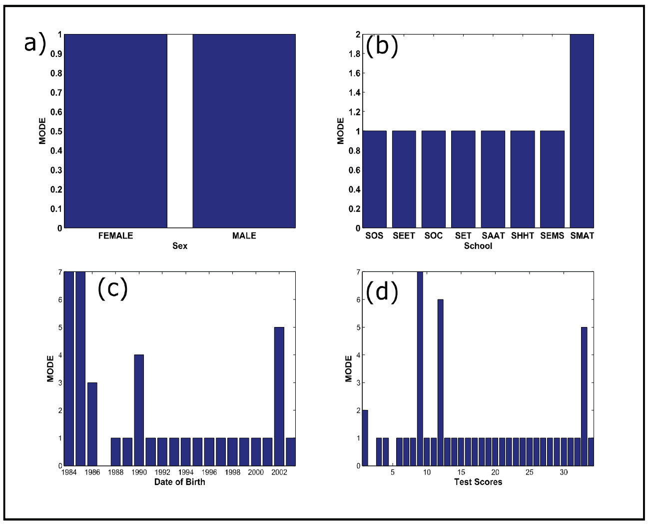

Fig. 3: Mode of chosen numbers based on (a) gender (b) school (c) year of birth; and (d) quiz score.

10 before beginning the quiz. A total of 3200 students partook in the experiment. To avoid bias, the students were not informed about the experiment or the purpose and nature of the experiment before, during and after the responses.

The R programing language and SAS university edition were used in the statistical analyses of the data obtained for the study. The data was filtered to remove incomplete, duplicate and unrealistic entries such as the ones with date of birth set at 2013 and other outliers. A total of 2344 valid entries were used for this experiment. The distribution of the entries based on gender, school, year of birth and test scores are presented in Figures 1(a)--(d) respectively. The schools considered are School of Engineering and Engineering Technology (SEET), School of Sciences (SOS), School of Computing (SOC), School of Environmental Technology (SET), School of Health and Health Technology (SHHT), School of Management Technology (SMAT) and School of Earth and Mineral Sciences (SEMS).

The mode of the data set was used as a descriptive statistics for the analysis. Entropy and recurrence quantification analyses were employed to ascertain if the numbers generated are periodic, stochastic or deterministic. To test for randomness, the run test for randomness was used while One Sample Kolmogorov Smirnov test, Anderson-Darling test and Jarque Bera test were used to test for normality in the data. Also, we used a multinomial multivariate logistic regression to assess the dependency of the choice of number selected by the participants on their unique attributes. This helped to check if the observed attributes influence the choices of random numbers generated, thereby revealing whether the numbers are random or not.

Logistic regression was adopted in the analysis because of its ability to explain dependent variables that are not continuous in nature as they relate to a set of explanatory variables. The response variable measured in this study is assumed to be a categorical variable between 1 and 10. The main focus here is not how the predictors affect the response variable but to answer the research question 'Do the predictors influence the responses?' To this end, we fit several univariate logistic regressions on each response individually considering student's intelligence, gender, age and faculty as the explanatory variables.

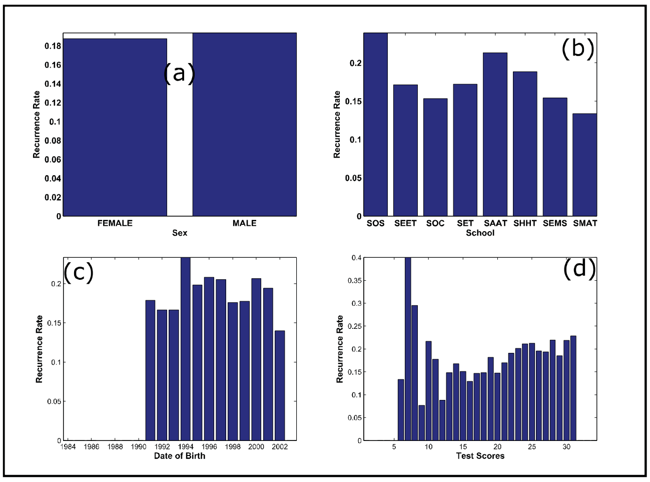

Fig 4: Entropy of chosen numbers based on (a) gender (b) school (c) year of birth; and (d) quiz scores.

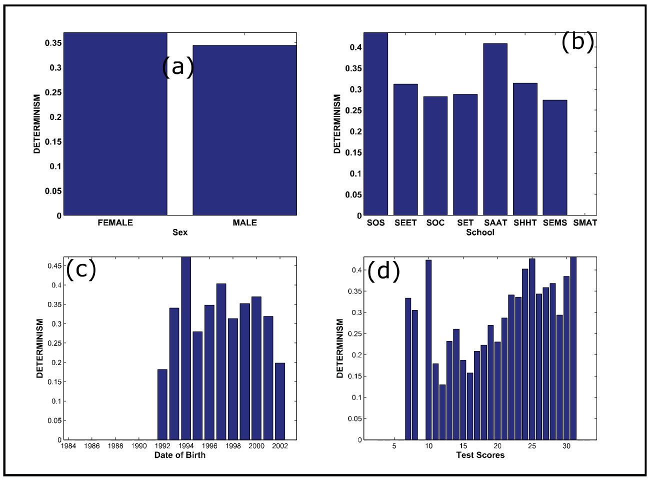

Fig 5: Recurrence rate of chosen numbers based on (a) gender (b) school (c) year of birth; and (d) quiz score.

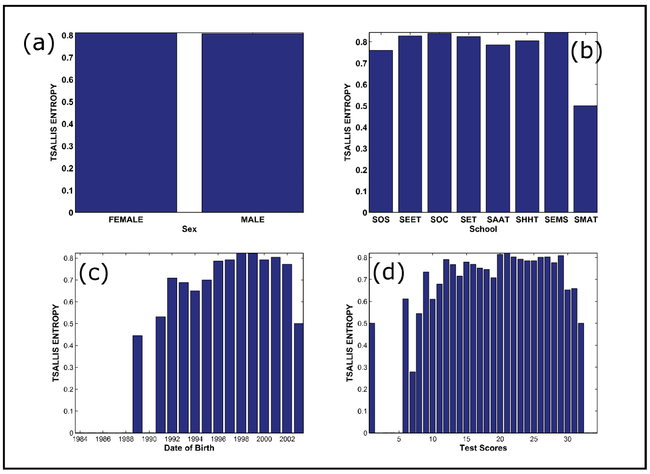

Fig 6: Entropy of chosen numbers based on (a) gender (b) school (c) year of birth; and (d) quiz scores.

III. RESULTS AND DISCUSSION

The mode of the whole data was computed (Figure 2) as well as the mode based on the predictors (Figure 3). From Figure 2, it can be observed that the number 1 has the highest frequency with over 800 entries. This is followed by the numbers 5 and 2. The number with the least entry was observed to be 9. To study the pattern of numbers chosen, the mode of the entries were considered based on gender, date of birth, school and quiz performance. This result seems to deviate from other experiment where 7 is the preferred number among several participants [2, 13]. T]he differences could be attributed to many factors including individual preferences, ethnic and religious biases amongst other factors. Using an open ended option, [23] found that a general population seems to favour chosen number. numbers in the range *7,000–7,999*.

TABLE 1. STATISTICAL TEST FOR NUMBERS GENERATED BASED ON SCORES IN A GIVEN TEST. SIGNIFICANT RESULTS AT 95% CONFIDENCE INTERVAL ARE INDICATED WITH AN *.

Scores	runstest	kstest	adtest	jbtest
1	0	1*		
3	0	0		
4	0	1*		
6	0	1*	0	0
7	0	1*	1*	1*
8	0	1*	1*	0
9	0	1*	0	0
10	0	1*	1*	0
11	0	1*	1*	1*
12	0	1*	0	0
13	0	1*	1*	0
14	0	1*	1*	0
15	0	1*	1*	0
16	0	1*	1*	0
17	0	1*	1*	1*
18	0	1*	1*	0
19	0	1*	1*	1*
20	0	1*	1*	0
21	0	1*	1*	1*
22	0	1*	1*	1*
23	0	1*	1*	1*
24	0	1*	1*	1*
25	0	1*	1*	1*
26	0	1*	1*	1*
27	0	1*	1*	1*
28	0	1*	1*	1*
29	0	1*	1*	1*
30	0	1*	1*	1*
31	0	1*	1*	1*
32	0	1*		
33	0	1*		
34	0	0		

In Figure 3a, number 1 is the favourite of both genders. Students from all Faculties also chose 1 as their favourite number except students from the School of Management Technology who prefer the number 2 (Figure 3b). In Figure 3c, the number 1 is the most common among

TABLE 2. STATISTICAL TEST FOR NUMBERS GENERATED BASED ON YEAR OF BIRTH IN A GIVEN TEST. SIGNIFICANT RESULTS AT 95% CONFIDENCE INTERVAL ARE INDICATED WITH AN *.

Year of birth	runstest	kstest	adtest	jbtest
1984	0	1*		
1985	0*	1*		
1986	0*	1*		
1988	0*	0		
1989	0*	1*		
1990	0*	1*		
1991	0*	1*	1*	0
1992	0	1*	0	0
1993	0*	1*	1*	0
1994	1*	1*	1*	1*
1995	0*	1*	1*	1*
1996	0*	1*	1*	1*
1997	0*	1*	1*	1*
1998	0*	1*	1*	1*
1999	0*	1*	1*	1*
2000	0*	1*	1*	1*
2001	0*	1*	1*	1*
2002	0*	1*	1*	0
2003	0*	0		

All the students born in 1984 and 1985 tend to choose the number 7, while those born in 1986, 1990 and 2002 prefer the numbers 5, 3 and 4 respectively. A similar trend was observed in statistics based on test scores. The students generally prefer the number 1 with the exception of those who scored 1, 9, 12 and 33 who prefer the numbers 2, 7, 6 and 5 respectively.

TABLE 3. STATISTICAL TEST FOR NUMBERS GENERATED BASED ON GENDER IN A GIVEN TEST. SIGNIFICANT RESULTS AT 95% CONFIDENCE INTERVAL ARE INDICATED WITH AN *.

Gender	runstest	kstest	adtest	jbtest
Female	0	1*	1*	1*
Male	0	1*	1*	1*

Results for statistical tests based on different categories are presented in Tables 1–5. According to the results of the runs test, all the scores are in random order, although not significant. Similar results were obtained for gender. Significant random numbers between 1 and 10 were chosen by all the students based on their year of birth except, non-significant random numbers chosen by students born 1984 and 1992, and significantly non-random numbers by those born in 1994. All faculties considered showed that the chosen numbers are random, albeit, non-significant. An exemption was found in the analysis of chosen numbers by students in SAAT where the numbers were not random.

The Kolmogorov Smirnov (KS) and Anderson-Darling (AD) tests were used to determine if the numbers chosen are from a normal distribution.

TABLE 4.STATISTICAL TEST FOR NUMBERS GENERATED BASED ON FACULTY IN A GIVEN TEST. SIGNIFICANT RESULTS AT 95% CONFIDENCE INTERVAL ARE INDICATED WITH AN *.

School	runstest	kstest	adtest	jbtest
SOS	0	1*	1*	1*
SEET	0	1*	1*	1*
SOC	0	1*	1*	1*
SET	0	1*	1*	1*
SAAT	1*	1*	1*	1*
SHHT	0	1*	1*	1*
SEMS	0	1*	1*	1*
SMAT	0	1*	0	0

KS test based on test scores do not come from a standard normal distribution except students who score 3 and 34 while AD test showed that numbers chosen by students who scored 6, 9 and 12 are not from a population with normal distribution. Numbers generated by students born in 1988 and 2003 come from a standard normal distribution using the KS test while the AD test showed that only students born in 1992 generated numbers that follow a normal distribution. Numbers chosen by both gender and faculties were also found not to be normally distributed except those from SMAT.

TABLE 5. LOGISTIC REGRESSION ANALYSIS TABLE

	Test Statistics	Chi-Square	p-Value
Grade	Likelihood ratio	14.4830	0.1062
	score	14.7608	0.0977
	Wald	14.6134	0.1021
Gender	Likelihood ratio	14.9443	0.0925
	score	15.0455	0.0897
	wald	14.8770	0.0944
Gender	Likelihood ratio	9.8452	0.3632
	score	13.0472	0.1605
	wald	7.2173	0.6145
Faculty	Likelihood ratio	124.8393	< .0001
	score	129.973	< .0001
	wald	123.3042	< .0001

To determine if the chosen numbers followed a normal distribution with an unknown mean and variance, the Jaque-Bera test was conducted. Results showed that all the numbers generated based on test score, gender, school, and year of birth are from a normal distribution with unknown mean and variance.

The result presented in Table 5, represents the univariate analysis of each predictor on the choice of number. we deduced from the analyses that all the predictors except faculty has a significant relationship with choice of number at $\alpha=0.05$.

To investigate the complexity in the chosen numbers by students, entropy of the time series was analyzed based on gender, age, school and quiz score. The results obtained are presented in Figure 4. The entropies of the numbers chosen based on gender have identical values of Tsallis entropy of 0.8. This implies that the complexity of numbers chosen by the two gender are identical. The complexity of chosen number based on schools are in the range *0.7–0.9* except the School of Management Technology. This exception might be due to the small data available for the Faculty. Entropy values between *0.6* and *0.8* were obtained for numbers generated by students born in 1992 and 2002. Similar ranges of entropy values were obtained for test scores.

To determine if the numbers generated by the different categories are chaotic, recurrence quantification analysis based on the concept of recurrence plot was conducted. Two parameters: recurrence rate and determinism were considered and the results are presented in Figures 5 and 6. Low values (closer to 0) of recurrence rate and determinism indicates chaos while high values (close to 1) are periodic. Based on gender, there was a very small difference between the recurrence rate and determinism values. Both tend towards a chaotic value for both gender. In Figures 5b and 6b, students from the School of Sciences and School of Management Technology have the highest and lowest values respectively. All values are less than *0.25*, hence, we can infer deterministic tendencies. In Figures 5c, the recurrence rate values lie in the range $0.13 < RR < 0.25$ while the corresponding determinism values (Figure 6c) are in the range $0.15 < DET < 0.45$. This shows varying complexity in number generated by students of different ages. With a few exceptions, the complexity in number generated tend to reduce with increasing scores 5d and 6d. The exceptions could be found in students who scored 7, 8, 10 and 11.

IV. CONCLUSION

In this study, we examined the effects of age, gender, faculty of choice and intelligence using student performance in a course on random number generation by a group of students. The descriptive analysis carried out showed that number one was mostly selected among students across gender, schools, ages and intelligence levels, however, there was no confirmatory analyses support it. Though the result from the logistic regression analysis showed that faculty predicts choice of number but no further backing to prove if it favours selection of number one compared to others.

Considering this population, we can infer that the numbers generated by the different categories show high degrees of randomness and low dimensional chaos. This shows that humans can hardly generate truly random numbers due to certain sociological and psychological factors. Further studies are required to establish the effect of race, ethnicity, tribe, sexual orientation, dominant hand and religion on random number generation.

ACKNOWLEDGMENT

The authors appreciate the Department of Physics, Federal University of Technology, Akure, Ondo State, Nigeria for permission to carry out the research.

REFERENCES

[1] J. C. Baird and E. Noma. "Psychophysical study of numbers," Psychological Research, vol. 37 no. 4, pp. 281–297, 1975.

[2] A. Bellos. "Alex through the looking glass: how life reflects numbers, and numbers reflect life," A&C Black, 2014.

[3] D. A. Berg, R. A. Luciano Jr, and A. Saffari. "Central random number generation for gaming system," 1998. US Patent 5,779,545.

[4] K. K.-S. Chan, C. L.-M. Hui, J. Y.-M. Tang, C. P.-Y. Chiu, S. K.-W. Chan, M. M.-L. Lam, and E. Y.-H. Chen. "Random number generation deficit in early schizophrenia," Perceptual and Motor Skills, vol. 112 no. 1, pp. 91–103, 2011. PMID: 21466082.

[5] X. Cheng, H. Ge, D. Andoni, X. Ding, and Z. Fan. "Composite body movements modulate numerical cognition: evidence from the motion-numerical compatibility effect," Frontiers in Psychology, vol. 6 no. 1692, 2015.

[6] D. DiCarlo. "Random number generation: types and techniques," Liberty University, pp. 1–32, 2012. A Senior Thesis submitted in partial fulfillment of the Requirements for graduation in the Honors Program.

[7] F. J. Evans. "Monitoring attention deployment by random number generation: An index to measure subjective randomness," Bulletin of the Psychonomic Society, vol. 12 no. 1, pp. 35–38, 1978.

[8] M. Figurska, M. Staczyk, and K. Kulesza. "Humans cannot consciously generate random numbers sequences: a polemic study," Medical Hypotheses, vol. 70, pp. 182–185, 2008.

[9] S. M. Gobel, C. A. Maier, and S. Shaki. "Which numbers do you have in mind? Number generation is influenced by reading direction," Cognitive processing, vol. 16 no. 1, pp. 241–244, 2015.

[10] H. Gould, J. Tobochnik, and W. Christian. "An introduction to computer simulation methods," vol. 1. Addison-Wesley New York, 1988.

[11] S. Grade, A. Badets, and M. Pesenti. "Influence of finger and mouth action observation on random number generation: an instance of embodied cognition for abstract concepts," Psychological research, vol. 81 no. 3, pp. 538–548, 2017.

[12] R. T. Kneusel. "Random numbers and computers," Springer, 2018.

[13] P. L. Knowles. "The blue seven is not a phenomenon," Perceptual and Motor Skills, vol. 45 no. 2, pp. 648–650, 1977.

[14] T. Loetscher and P. Brugger. "Exploring number space by random digit generation," Experimental Brain Research, vol. 180 no. 4, pp. 655–665, 2007.

[15] T. Loetscher, U. Schwarz, M. Schubiger and P. Brugger. "Head turns bias the brain's internal random generator," Current Biology, vol. 18 no. 2, pp. R60–R62, 2008.

[16] G. Marsaglia. "Random number generation," In: Encyclopedia of Computer Science, pp. 1499–1503, John Wiley and Sons Ltd., Chichester, UK.

[17] P. Martin-Lof. "The definition of random sequences," vol. 9, Elsevier, 1966.

[18] M. Matsumoto and T. Nishimura. "Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator," ACM Transactions on Modeling and Computer Simulation (TOMACS), vol. 8 no. 1, pp. 3–30, 1998.

[19] K. V. Niel and S. W. Laffan. "Gambling with randomness: the use of pseudo-random number generators in GIS," International Journal of Geographical Information Science, vol. 17 no. 1, pp. 49–68, 2003.

[20] A. M. Noll. "Computers and the visual arts," Design Quarterly, vol. 66/67):64{71, 1966.

[21] C. S. Petrie and J. A. Connelly. "A noise-based IC random number generator for applications in cryptography," IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications, vol. 47 no. 5, pp. 615–621, 2000.

[22] S. Rosenberg, N. Weber, M.-A. Crocq, F. Duval, and J.-P. Macher. "Random number generation by normal, alcoholic and schizophrenic subjects," Psychological Medicine, vol. 20 no. 4, pp. 953–960, 1990.

[23] E. C. Sanford. "On the guessing of numbers," The American Journal of Psychology, vol. 14 no. 3/4, pp. 383–401, 1903.

[24] C. Sosson, C. Georges, M. Guillaume, A.-M. Schuller, and C. Schiltz. "Developmental changes in the effect of active left and right head rotation on random number generation," Frontiers in Psychology, vol. 9 no. 236, pp. 1–14, 2018.

[25] J. Soto. "Statistical testing of random number generators," In: Proceedings of the 22nd National Information Systems Security Conference, 1999.

[26] I. S. Telnaes. "Electronic gaming device utilizing a random number generator for selecting the reel stop positions," 1984. US Patent 4,448,419.

[27] J. N. Towse, T. Loetscher, and P. Brugger. "Not all numbers are equal: preferences and biases among children and adults when generating random sequences," Frontiers in psychology, vol. 5 no. 19, 2014.

[28] J. N. Towse and D. Neil. "Analyzing human random generation behavior: a review of methods used and a computer program for describing performance," Behavior Research Methods, Instruments, & Computers, vol. 30 no. 4, pp. 583–591, 1998.

[29] S. Volchan. "What is a random sequence?" The American Mathematical Monthly, vol. 109 no. 1, pp. 46–63, 2002.

[30] W. Wagenaar. "Generation of random sequences by human subjects: a critical survey of literature," Psychological Bulletin, vol. 77 no. 1, pp. 65–72, 1972.

[31] B. A. Wichmann and I. D. Hill. "Algorithm as 183: an efficient and portable pseudo-random number generator," Journal of the Royal Statistical Society, Series C (Applied Statistics), vol. 31 no. 2, pp. 188–190, 1982.