GLI®

GLI Europe BV
Diakenhuisweg 29-35
2033 AP Haarlem
The Netherlands

Tel +31 (0)88 220 6600
www.gaminglabs.com

Chamber of Commerce
Leiden nr. 28117769
VAT Identification number
NL 8184.73.393.B.01

RvA Registration Number of Accreditation applicable to this Report:	
Testing	L372
Certification	C577

Report Type:	Evaluation and Certification
Report Date:	16 August 2019
Issuing Laboratory:	GLI Europe B.V.
Evaluating Laboratory:	GLI Europe B.V.
Recipient:	Evenbet Gaming Limited Dar il - Hena, Triq Il- Kbira Mosta MST1019 Malta
Tested against Requirements	GLI-19 – Interactive Gaming Systems v2.0
Jurisdiction:	Non Jurisdictional
Manufacturer	Evenbet Gaming Limited Dar il - Hena, Triq Il- Kbira Mosta MST1019 Malta
Submitter	Evenbet Gaming Limited Dar il - Hena, Triq Il- Kbira Mosta MST1019 Malta
Product Name	**Evenbet RNG, version 6.8.1**
Description of the product tested	**libshuffling.so**
Request Date:	As requested per manufacturer's letter received 5 August 2019.
Evaluation Period:	6 August 2019 / 9 August 2019
Internal Reference	RN-120-ESF-19-01
Result	Pass
Internal methods used reference:	WI-MA-006, PC-TC-001

Technical Evaluation authorized by:

Martin Britton
Managing Director

RANDOMNESS REPORT FOR THE EVENBET RNG, VERSION 6.8.1

The intent of this report is to indicate that **Gaming Laboratories International**, **LLC** (GLI) has completed its evaluation of the Evenbet random number generator (RNG), version 6.8.1, provided by Evenbet Gaming Limited.

SECTION I — SCOPE OF TESTING

Evenbet Gaming Limited submitted the required materials to GLI in order to conduct a random number generator analysis on the Evenbet RNG, version 6.8.1. The scope of this analysis was limited to software verification, source code review, and data analysis. The RNG was tested for its ability to randomly produce outcomes for the card (poker type) games.

The Evenbet RNG, version 6.8.1 was evaluated against the RNG-specific requirements of the following technical standards:

- GLI-19 – Interactive Gaming Systems v2.0

The software being certified herein contains a cryptographically strong Random Number Generator (RNG) and as such, obsoletes the necessity of background cycling to maintain unpredictability when in use.

SECTION II — SOFTWARE VERIFICATION

Verify+ by Kobetron™ signatures for the Evenbet RNG, version 6.8.1 are as follows:

File	Version	Type	Signature
libshuffling.so	6.8.1	Kobe4	02UU
		MD5	37D8786C9E911094F535554FD4D924F4
		SHA-1	D3AE03BF095E2CC7B5292403F3A4AA7B1B5E31BA
		Kobe40	2A644H439315770F9F3U192UAPUCUPHFP0695980
		CDCK	9591

Table 1. Digital Signatures

SECTION III — SOURCE CODE REVIEW

Evenbet Gaming Limited submitted appropriate documentation and full source code which pertains to the generation of random numbers. GLI reviewed the source code provided by tracing the path of the RNG application from the initiation of the draw to the selected output of random numbers. GLI inspected the source code, where practicable, in an attempt to find any undisclosed switches or parameters having a possible influence on randomness and fair play. GLI assessed the ability of the RNG to produce all numbers within the desired range.

SECTION IV — DATA ANALYSIS

The game configuration and parameters for the data obtained and tested are listed in Table 2. GLI performed a data format check on each data set listed in order to confirm that the game parameters were correctly represented in the data analyzed.
A set of numbers is said to be drawn without replacement if a number can only be selected once within the same draw.

Data Set	Range	Positions	Replacement	Draws
Poker_52Deck	0-51	52	No	50,000,000
Poker_36Deck	16-51	36	No	50,000,000
Binary for NIST	0-4,294,967,295	1	N/A	31,250,000
Binary for DieHard	0-4,294,967,295	1	N/A	3,000,000

Table 2. Game Parameters

For a summary of the statistical tests applied to each data set, see Appendix A. For a description of the overall test methodology and a description of each test used, see Appendix B.
Overall, the RNG passed the battery of tests for each configuration at the 99% confidence level.

SECTION V — SUMMARY

Overall Evaluation of the Random Number Generator

GLI's conclusion based upon the tests applied to the Evenbet RNG, version 6.8.1 data is that this random number generator has exhibited random behavior and is suitable for the applications as described herein. If a game utilizes a different range or a different number of selections from the included ranges, the RNG should be resubmitted to test that set of parameters.

APPENDIX A: Statistical Test Summary

					Test Names													
Data Set	**Range**	**Positions**	**Replacement**	**Draws**	**Runs**	**Serial Corr.**	**Interplay Corr.**	**Adj. Max-Min**	**Adj. High-Low**	**Adj. Blocks**	**Coupon Coll.**	**Duplicates**	**Overlaps**	**Permutation**	**Tot. Dist.**	**Tot. Dist. by Pos.**	**DieHard**	**NIST**
Poker_52Deck	0-51	52	No	50,000,000	X	X	X	X	X	X	X	X	X	X	X	X		
Poker_36Deck	16-51	36	No	50,000,000	X	X	X	X	X	X	X	X	X	X	X	X		
Binary for NIST	0-4,294,967,295	1	N/A	31,250,000														X
Binary for DieHard	0-4,294,967,295	1	N/A	3,000,000													X	

Table A 1. Tests Applied

APPENDIX B: Test Descriptions

B.1 **Definitions**. The following terms apply to the below test descriptions. Randomness Device or Random Number Generator (RNG) output may be collected multiple numbers at a time. Each set of numbers is called a draw. Each individual number has a particular order within the draw. This is referred to as the number position.

B.2 **Distribution Comparisons**. Many of the tests compare an observed numerical distribution with an expected distribution. Unless otherwise specified, this is done by means of a statistical chi-square goodness-of-fit test. The value chi-square is computed in the standard way. If k is a possible value, o_k is the observed count of that value, and e_k is the expected count:

$$\chi^2 = \sum_k \frac{(o_k - e_k)^2}{e_k}$$

In the case where expected counts are too small for accurate use of the above formula, values are 'binned' together to ensure an appropriate minimum expected count. The resultant value for chi-square is compared against the distribution for the appropriate number of degrees of freedom. Unusually high (distribution mismatch) or unusually low (insufficient randomness) chi-square values can be causes for data failure.

B.3 **Meta-testing**. Evaluation of groups of p-values may include a meta-test for extremity of high or low p-values, a meta-test for frequency of high or low p-values, and a meta-test for uniformity of p-values, as appropriate.

B.4 **Confidence Level**. The statistical tests conducted by GLI are done at a particular confidence level. Common confidence levels used include 95%, 98%, and 99%, depending on jurisdictional requirements, and intended use of the RNG. High confidence level testing has low risk of mistakenly failing a good RNG, but higher risk of passing a bad RNG. Lower confidence level testing has increased power of detecting bad RNGs, while also increasing the risk of false failures of good RNGs. Specifically, the confidence level represents the probability that an ideal source of randomness would pass the testing. If an RNG passes statistical tests at a given confidence level, passage at all higher confidence levels is implied.

B.5 **Tests**. Some tests are only applicable to certain types of data. Some tests may be applied only to a portion of the data. Some tests may require that the data be parsed, binned, or otherwise transformed, as necessitated by data format.

APPENDIX B: Test Descriptions

Adjacency Blocks:

For each draw, the data is first sorted. Then the amount of contiguous blocks of numbers is counted. These statistics are then compared against the expected. For example, if a draw consists of the numbers:

$$1, 5, 4, 2, 6, 9$$

the data would be sorted and separated into blocks. The resulting statistic would be 3.

Adjacency High-Low:

For each draw, the number of local extrema ('highs' and 'lows') in the data is recorded and compared with the expected distribution. These are also referred to as 'turning points'. For example, if a draw consists of the numbers:

$$1, 3, 5, 7, 2, 9$$

there would be one local maximum (7) and one local minimum (2). The resulting statistic would be 2.

Adjacency Max-Min:

For each draw, the difference between the maximum and minimum values is calculated and recorded. This is compared with the expected theoretical distribution. For example, if a draw consists of the numbers:

$$2, 3, 6, 7, 4$$

the resulting statistic would be 5, the difference between the maximum value (7) and the minimum value (2).

Coupon Collector's:

The Coupon Collector's Test is applied positionally. The data is parsed until all possible values have been observed, then the number of values checked is recorded and the count is restarted. This is compared with the expected distribution. For example, if the set of all possible values is {0, 1, 2} and the first position of each draw is:

$$1, 0, 1, 0, 2, 0, 1, 2, \ldots$$

then all values are observed in the first position by the fifth draw. All values are then observed within the next 3 draws, so the first two statistics for the first position would be 5 and 3.

APPENDIX B: Test Descriptions

DieHard:

The DieHard Battery of Tests is a standard assessment of the randomness in raw outcomes generated from an RNG. The collection, designed by George Marsaglia, tests for a variety of patterns in the individual binary bits of RNG output. GLI uses a custom implementation to conduct DieHard testing.

Duplicates:

The Duplicates Test counts the number of times a draw is exactly duplicated in the data. In the case that a particular draw is repeated more than twice, every possible way to generate a duplicate is counted. This is compared against the theoretical distribution to verify that the number of duplicate draws falls within expected bounds. For example, consider the dataset consisting of the following draws of two numbers each.

$$
\begin{array}{l}
a)\ 1,3 \\
b)\ 4,1 \\
c)\ 1,3 \\
d)\ 1,3 \\
e)\ 4,1 \\
f)\ 3,1
\end{array}
$$

The duplicate pairs are (*a*,*c*), (*a, d*), (*c, d*), and (*b, e*), for a total of 4 duplicates. *(f)* is not counted as a duplicate since the draw must match in order as well as values.

Interplay Correlation:

The Interplay Correlation Test measures statistical correlation between different positions of the same draw. For each pair of positions, statistical correlation is calculated as in the Serial Correlation Test. In the case of without replacement data, an adjustment is made to account for the expected resulting negative correlation.

APPENDIX B: Test Descriptions

NIST Test Suite:

The following "bitwise" tests from the NIST Test Suite were applied:

- Frequency (Monobits) Test
- Frequency Test within a Block
- Run Test
- Test for the Longest Run of Ones in a Block
- Binary Matrix Rank Test
- Discrete Fourier Transform (Spectral) Test
- Non-Overlapping Template Matching Test
- Maurer's "Universal Statistical" Test
- Linear Complexity Test
- Serial Test
- Approximate Entropy Test
- Cumulative Sums (Cumsum) Test
- Random Excursions Test
- Random Excursions Variant Test

Overlaps:

The Overlaps Test compares consecutive draws for overlapping values. The number of overlapping values is recorded for each pair of draws. This observed distribution of overlaps is then compared against the expected distribution. For example, if the following draws are observed consecutively,

$$a)\ 1, 4, 5, 6$$
$$b)\ 4, 1, 7, 6$$

the number of overlaps would be 3, representing the values 1, 4, and 6.

APPENDIX B: Test Descriptions

Permutation:

The Permutation Test is a test applicable to data that represents a reordering of numbers. Each draw can be considered as a permutation of the original ordering. Every permutation can be decomposed into disjoint cycles, which represent the possible positions a number would occupy if the same permutation is applied repeatedly. For each draw, three statistics are collected based on the cycle decomposition:

- The number of cycles.
- The size of the smallest cycle.
- The size of the largest cycle.

Each of these statistics generates a distribution of observations which are compared with their respective expected distributions. For example, if the following draw were observed as a reordering of the numbers from 1 to 6,

$$1, 3, 5, 4, 2, 6$$

the cyclic decomposition would be (1)(2 3 5)(4)(6). 1, 4, and 6 remain in their original positions, so they form their own cycles. The values 2, 3, and 5 are shuffled, so they form a single cycle together. The total number of cycles is 4, the smallest cycle has size 1, and the largest cycle has size 3.

Runs:

The Wald-Wolfowitz Runs Test is applied to each position within the draw. A center is established, typically the data median, and the number of ‘runs’ above and below the center are tallied. Values exactly equal to the center are discarded. This is compared to the expected distribution, which depends on the number of values above and below the center. For example, if the numbers drawn at a particular position were:

$$2, 3, 1, 5, 4, 7, 3, 2, 3, 2, 3, 2, 6, 7, 3, 5$$

and the established center were the data median of 3, the data would be parsed for runs above 3 and runs below 3.

$$\underbrace{2, 3, 1}, \overbrace{5, 4, 7, 3}, \underbrace{2, 3, 2, 3, 2}, \overbrace{6, 7, 3, 5}$$

This would be counted as 4 runs.

APPENDIX B: Test Descriptions

Serial Correlation:

The Serial Correlation Test measures statistical correlation between consecutive draws of the same position. For each position, the sample Pearson correlation coefficient is calculated. If *X* represents the first number, and *Y* the number that follows, then the coefficient is:

$$r = \frac{cov(X, Y)}{s_X s_Y}$$

where *s* denotes the sample standard deviation. The coefficients are used to generate a *p*-value for each position.

Total Distribution:

The Total Distribution Test is a simple tally of all observed values throughout the data. This is compared with the expected distribution. Typically, the expected distribution is a uniform distribution. In the case of unequal weighting of values, an appropriate discrete distribution is used.

Total Distribution by Position:

The Total Distribution by Position Test tallies the observed distribution of values for each position within the draw. Each of these distributions is then compared with the expected.