**Literature Review**

Before testing could begin, suitable background research about random number generators was required. The highlighted materials outlined in this section served as the main sources of information regarding areas such as pseudorandom algorithms, background on the creation of digital random number generation, the use of white noise for true random generation, the use of pseudorandom numbers in cryptography and empirical tests that can be used to evaluate randomness.

Donald Knuth’s 1998 book *The Art of Computer Programming Volume 2: Semi-numerical Algorithms* (**Knuth, 1998**) served as a useful introduction to the theory behind computerised pseudorandom number generators as well as detailed explanations of empirical tests capable of estimating the degree of randomness a sequence displays. These include but are not limited to Birthday Spacings, Serial Correlation and Poker tests. Information regarding these tests proved invaluable throughout the analysis portion of this investigation. The book itself is highly regarded among computer scientists and although it was released twenty-five years ago much of the fundamental theory remains valid with the book itself being a revision of Knuth’s original works released over twenty years before.

Stephen Park and Keith Miller’s 1988 paper *Random Number Generators: Good Ones are Hard to Find* (**Park, Miller, 1988**) was another valuable introduction to random number generators. This paper focused on algorithms used by computer scientists at the time, many of which being highly predictable and rigid in the structure of their ‘random’ outputs such as RANDU (**Wikipedia (X), 2023**) and the practical implementation of a new generator, later coined the Lehmer generator (**Wikipedia (X), 2023**), which aimed to provide a ‘minimal standard’ for a reliably random sequence of numbers with comparatively simple code written in Pascal. Although like Knuth’s book this paper might seem dated, the Lehmer generator remains relevant today. After a review from computer scientists in 1993 (**Marsagilia, Sullivan, 1993**), the Lehmer generator from Park and Miller’s original paper was updated to feature a new base multiplier value, which served to reduce reproducibility and predictability. Both versions of the Lehmer generator are provided as standard in modern languages such as C++11’s minstd\_rand functions (**Wikipedia (X), 2023**).

A pair of similar papers, *Random Bit Sequence Generation from Image Data* by Yas Abbas Alsultanny in 2008 (**Alsultanny, 2008**) and *Random Number Generated from White Noise of Webcam* by Jer-Min Tsai, I-Te Chen and Jengnan Tzeng in 2009 (**Tsai, Chen, Tzeng, 2009**) both focused on the concept of collecting white noise from a webcam image and using it to generate true random number sequences. These papers proved valuable to the project as the outline provided for each of their practical demonstrations on white noise collection, processing and use in generation. Although these papers made use of video data, whereas the white noise collected in this project is audio based, seeing the steps taken in both experiments gave an indication as to how similar steps could be taken regarding an alternate data source.

Likewise, I-Te Chen’s 2013 paper *Random Numbers Generated from Audio and Video Sources* (**Chen, 2013**) also aims to collect and examine random number sequences derived from webcam white noise. The combination of webcam inputs leads to an audio-visual random number generator that produces random sequences from image data as well as taking advantage of audio’s influence (**Chen, 2013**). This paper, while focusing on many of the same steps as the ones outlined previously, proved useful as it gave a demonstration of audio white noise being used to generate random number sequences.

Another area of interest is the future of pseudorandom algorithms. While many of the sources cited in this review focus on the history of random number generators, *Evaluation of splitable pseudo-random generators* released in 2015 by Hans Georg Schaathun (**Schaathun, 2015**) looks at future developments surrounding parallel computing and the creation of parallel number generators. Although the research to be undertaken in this project will not focus on parallel algorithms, the generators outlined could still be adapted to a non-parallel function for testing purposes. The paper uses the serial empirical test described by Knuth (**1998**) which provided a valuable insight into how the test could be used on this project’s data.

*A new approach to analyze the independence of statistical tests of randomness* by Elena Almaraz Luengo, et. al (**Luengo, et. al, 2022**) was a valuable asset when exploring test suite batteries for random numbers such as DIEHARD and DIEHARDER. The paper aimed to outline the main concepts of a variety of suites, including the tests included within them, evaluating the limits of each both in terms of analyzable sequence length and accuracy. Released in 2022, this research presents pros and cons for the most up to date test batteries currently in use, with a focus less on the specifics of each test and instead a wider focus on the performance of the batteries as a whole.

As outlined previously, the successful generation of non-replicable random number sequences is useful for more than just scientific simulations. Cryptography and encryption can both rely heavily on pseudo and true random generators and the paper *Cryptographically secure pseudo-random number generator IP-core based on SHA2 algorithm* by Luca Baldanzi, et. al (**Baldanzi, et. al, 2020**) aims to address this. Due to the sensitive nature of cyber security, the paper features useful details on the need for unpredictability and how random numbers provide secure applications to help combat attacks. This information proved valuable when looking to address potential impacts of this research outside of a purely academic context.

Frederick James and Lorenzo Moneta’s 2020 work *Review of High-Quality Random Number Generators* (**James, Moneta, 2020**) helped to address the need for research such as this, and how testing to find high quality generators benefits mathematicians and computer scientists. As well as providing additional background on pseudorandom number generators, it also provides examples of high-quality algorithms including RANLUX and MIXMAX which show the prominence and need for reliable pseudorandom generators.