Peer-reviewed version available at *Fluctuation and Noise Letters* **2020**, *19*, 2050034; doi:10.1142/S0219477520500340

Accepted for publication in Fluctuation and Noise Letters (February 14, 2020).

AN INFORMATION THEORETIC APPROACH TO ORIGINALITY AND BIAS IN SCIENCE

L.B. Kish, C. Singh
Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843-3128, USA

T. Erdelyi
Department of Mathematics, Texas A&M University, College Station, TX 77843-

Abstract. We introduce an information theoretic framework for a quantitative measure of originality to model the impact of various classes of biases, errors and error corrections on scientific research. Some of the open problems are also outlined.

Keywords: Originality; information entropy; bias; political correctness; errors; error correction.

1. Introduction

While there is a consensus that an independent scientific research is superior to a biased one, a deeper mathematical understanding of this fact is yet to be found. Perhaps the most interesting related questions belong the open problem of free will, on which there is a vast literature, particularly in quantum physics. However, only the Chaitin-type, complexity-based approaches to free will have relevance to the problem we are addressing here (e.g. [1-3]).

Chaitin [1,2], similarly to Renyi [4], concludes that deterministic logic operations preserve information , therefore the information entropy of results deduced or calculated from axioms do not contain more than the axioms themselves. Thus the extra information entropy in the publications of such results stems from the uncertain choice of the way of calculations/deductions only. Similar considerations hold for free will: if it exists, it is inherently related to truly random value generation [1,3]. These approaches characterize the amount of new information by the Chaitin-Kolgomorov complexity of data which is however impossible to measure in most of the cases [1].

Our approach is to characterize the originality of a research is different and it does *not* challenge the validity of the efforts mentioned above and preserves their main conclusions. We do not use complexity theory, yet we utilize information theory [5]. The potential advantage of our approach is that the upper limit of new information offered by

Peer-reviewed version available at *Fluctuation and Noise Letters* **2020**, *19*, 2050034; doi:10.1142/S0219477520500340

a given scientific research is its originality, and the impact of biases can be treated at the phenomenological level in modeling.

2. Communication channel approach to scientific research

2.1 A former study

In [7] we proposed an information theoretic approach to visualize the impact of bias on the information content of published research articles. It was inspired by Shannon's picture of the information channel that begins with the Transmitter (Sender) and ends at the Receiver of the message, see Figure 1. The information during its propagation is spoiled by random noise and non-ideal transfer characteristics [5].

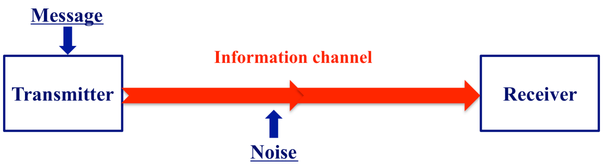

Figure 1. Shannon's scheme of information channel [5].

The information entropy (uncertainty) of the message before it arrives at the Receiver is given by the Shannon formula. In the case of binary data acquisition, for a message of N bit length, Shannon's information entropy is:

$$S_1 = \sum_{j=1}^{N} \sum_{m=0}^{1} p_{j,m} \log_2 \left(\frac{1}{p_{j,m}} \right) \qquad (1)$$

where $p_{j,m}$ stands for the probability of the j-th message bit being in the bit value m $\left(m \in \{0,1\}\right)$. The Shannon formula provides maximum entropy when the probabilities of the different bit values are 0.5, representing the greatest "surprise factor" of the bits upon reception.

For the sake of simplicity, in this paper we assume that the open question to be decided by the research is a single-bit, yes/no question (1 or 0 bit values), such as in today's heated debates about the question "Do human activities cause global warming?". For example, with bit values 1 and 0, we can represent the *yes* and *no* answers, respectively.

Our initial study [7] used a simplified information channel applied to scientific research and publications, see Figure 2. The system begins with the Object under investigation by the research and ends with the Reader of the published paper. In that initial study [7] the

doi:10.20944/preprints202002.0117.v1

focus was on a simple bias of the probabilities that has limited value, as we will show that in Section 2.2.

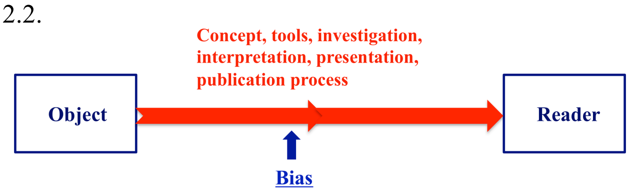

Figure 2. Simplified information channel for research and publication [7].

We represented the *originality* of the of the single-bit (yes/no) question with the information entropy of the answer to it:

$$S_1 = \sum_{m=0}^{1} p_m \log_2\left(\frac{1}{p_m}\right) = p_1 \log_2\left(\frac{1}{p_1}\right) + p_0 \log_2\left(\frac{1}{p_0}\right) = \\ = p_1 \log_2\left(\frac{1}{p_1}\right) + \left(1-p_1\right)\log_2\left(\frac{1}{1-p_1}\right) \tag{2}$$

Suppose the question to be answered is very original in the sense that scientists do not have a clue about the yes/no answer. Thus this open question has maximum entropy (0.5 probability of the bit values). Any bias during the research and publication of the answer would reduce the information entropy and the information gain upon reading the result, see Figure 3.

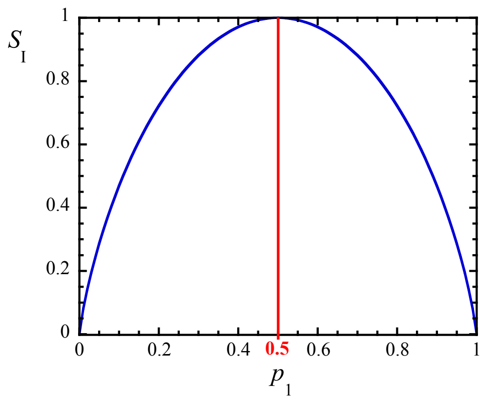

Figure 3. Information entropy versus the probabilities of bit values.

The bias can originate from various sources and the human and the society levels [7].

Peer-reviewed version available at *Fluctuation and Noise Letters* **2020**, *19*, 2050034; doi:10.1142/S0219477520500340

2.2 Problems with the former study

The former approach [7], outlined in 2.1, though valid, has several weaknesses that are limiting its general applicability:

(i) The question may not be fully original in the sense that available scientific knowledge may *correctly or incorrectly* predict the answer with probability different from 0.5. Then the maximum of the curve in Figure 3 is shifted to the left or right.

(ii) Even though it looks like the issue described in point (i) above can be included in the research-bias, different implications of various, available scientific knowledge may correctly or incorrectly predict the answer, thus these biases may amplify or compensate each other.

(iii) Similarly, experimental artifacts can also cause bias, systematic or random. This aspect points toward the issue of errors, that can be deterministic or random, which are basic elements of Shannon's work [5]. These biases and errors may also amplify or compensate each other.

(iv) The problems with compensating of biases and/or errors outlined in points (ii) and (iii) are twofold:

- They would falsely imply a more original research than it actually is.

- Additionally, an explanation based on errors compensating out each other is an incorrect understanding that inhibits finding the scientific truth and it can even nullify the value of the research.

(v) Finally, the research problem may fully by original, yet its answer may have an *inherent/natural bias*, which means that the correct yes/no answer to the question may be statistical with probabilities different from 0.5. For example, at p_1=0.6 and p_0=0.4, Equation 2 would predict a reduced originality of the question even if it is fully original so scientists do not have a clue of answering it, and the research is free of biases and errors.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 February 2020 doi:10.20944/preprints202002.0117.v1

Peer-reviewed version available at Fluctuation and Noise Letters 2020, 19, 2050034; doi:10.1142/S0219477520500340

3. The improved model: biases, errors and error corrections

3.1 The generalized approach with multiple biases and errors

The generalized approach is more in line with Shannon approach than the earlier version because it includes (systematic and random) errors, too (see Figure 4).

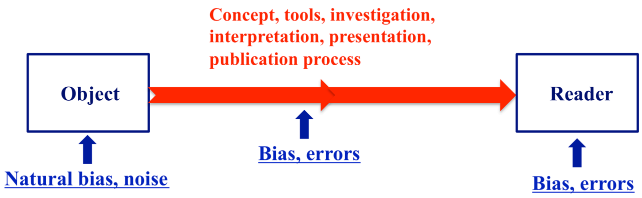

Figure 4. Generalized information channel representation of scientific research. An inherent/natural bias may be present in a statistical fashion and the various biases and errors may amplify or compensate each other.

The pessimistic estimation of the resultant information entropy S reduced by all the biases and errors is a generalization of Equation 2. In the generalized model (see Equation 3), the absolute values of the individual bias and error parameters are used to avoid false positive results due to situations where these factors compensate each other (see point (iv) in section 2.2 above).

$$S = (0.5 + x)\log_2\left(\frac{1}{0.5 + x}\right) + (0.5 - x)\log_2\left(\frac{1}{0.5 - x}\right), \tag{3}$$

where the quantity x is the pessimistic, integrated bias due to the natural, human and instrumental situations, including systematic and random errors:

$$x = |\Delta p_1| + b + \varepsilon \equiv |\Delta p_1| + \sum_i |b_i| + \sum_j |\varepsilon_j|, \tag{4}$$

b_i is the i-th type of research bias, ε_j is the j-th type of research error and the inherent/natural bias is given as $\Delta p_1 = p_1 - 0.5$. Here p_1 is the probability that the bit value is 1, that is, that answer is yes, assuming an ideal, unbiased, error-free assessment. In Equations 3 and 4 we assume that the p, b and ε quantities are uncorrelated.

Note, Equations 3 and 4 however do not contain the valid error corrections done by an ideal scientist.

Peer-reviewed version available at *Fluctuation and Noise Letters* **2020**, *19*, 2050034; doi:10.1142/S0219477520500340

3.2 The generalized approach with multiple biases, errors and error corrections

Even though discarding suspicious data are widely used, it is obvious that there is a danger to throw out original information. Thus, a more correct approach would be different from today's practices by:

Not only reducing the above given ε values by the error correction but also: publishing the suspicious data with the reasons for the doubt about them, and the detailed the error correction protocol.

With the error correction, the pessimistic (worst-case scenario) prediction is:

$$x = |\Delta p_1| + \sum_i |b_i| + \sum_j |\varepsilon_j| + \sum_k |\varepsilon_{ck}| \equiv |\Delta p_1| + E \ , \tag{5}$$

where ε_{ck} is the k-th error correction contribution when the original data and the error correction details are *not* provided in the publication, which is a typical policy nowadays. The sum $E = \sum_i |b_i| + \sum_j |\varepsilon_j| + \sum_k |\varepsilon_{ck}|$ is the "absolute bias and error" that are experimental setup and human related.

3.3 Measure of the quality of research at non-zero natural bias

To characterize the quality of the executed research, we introduce the information entropy reduction factor γ (see Equation 6) that describes the further reduction of the original information entropy of the question (with the inherent/natural bias) by the extra component E representing the scientific biases, errors and unspecified error corrections. The quantity γ is the resultant information entropy S reduced by all the biases, errors and unspecified error corrections (Equations 3, 5) normalized by the reduced entropy solely due to the inherent/natural bias S_{n} :

$$\gamma = \frac{S}{S_{\mathrm{n}}} = \frac{(0.5+x)\log_2\left(\dfrac{1}{0.5+x}\right) + (0.5-x)\log_2\left(\dfrac{1}{0.5-x}\right)}{(0.5+|\Delta p_1|)\log_2\left(\dfrac{1}{0.5+|\Delta p_1|}\right) + (0.5-|\Delta p_1|)\log_2\left(\dfrac{1}{0.5-|\Delta p_1|}\right)} \ , \tag{6}$$

where S_{n} is the information entropy reduced by only the natural bias, that is, when the research is bias and error free due to idealized scientists, instrumentations and conditions.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 February 2020 doi:10.20944/preprints202002.0117.v1

Peer-reviewed version available at *Fluctuation and Noise Letters* **2020**, *19*, 2050034; doi:10.1142/S0219477520500340

The quantity x is the pessimistic, integrated bias due to the natural, human and instrumental situations.

In other words, the factor γ separates the impacts of natural bias from the impact of scientific bias, errors and unspecified error corrections thus it characterizes the research performance: How much is the reduction of the information that the imperfect investigation can extract from the naturally biased question.

3.4 The generalized approach with multiple biases and errors

Figures 5-7 illustrate the behavior of the new measures introduced above for the quality and originality of biased research with errors and error corrections.

Figure 5 illustrates the behavior of S and γ versus the absolute bias and error E at different inherent/natural bias values.

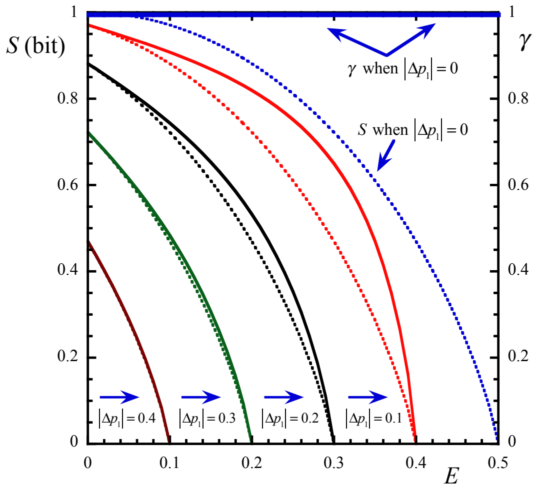

Figure 5. Impact of the inherent/natural bias $|\Delta p_1|$ and scientific bias and errors E on the information entropy. Dashed lines: the information entropy S of the 1-bit question versus E. Solid lines: the information entropy reduction factor γ characterizing the excess reduction of the information entropy due to the scientific bias and errors. The $|\Delta p_1|$ parameters are shown by its values where the arrows points to the relevant curves.

Figure 6 shows that the highest impact of errors and scientific bias E on S is when $x = |\Delta p_1| + E$ is approaching the value of 0.5, which is also obvious from Equations 2 and 5. Figure 7 implies similar conclusion.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 February 2020 doi:10.20944/preprints202002.0117.v1

Peer-reviewed version available at *Fluctuation and Noise Letters* **2020**, *19*, 2050034; doi:10.1142/S0219477520500340

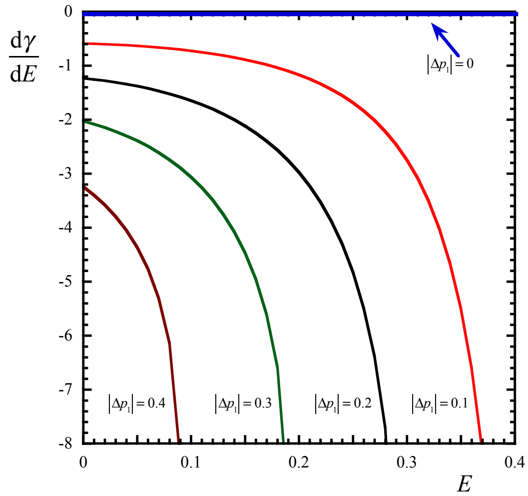

Figure 6. Derivative of the information entropy reduction factor γ versus the absolute bias and error *E*.

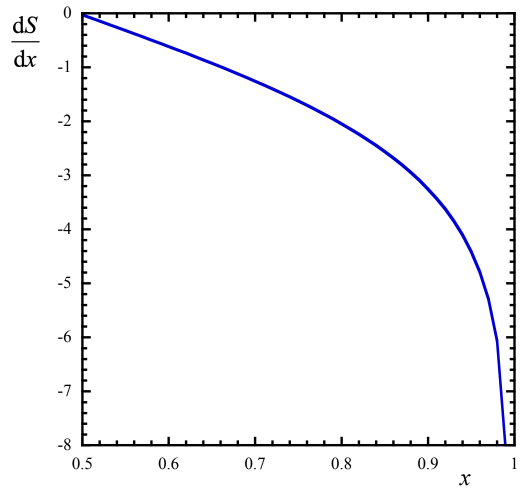

Figure 7. Derivative of the resultant information entropy versus the all biases and errors.

Peer-reviewed version available at *Fluctuation and Noise Letters* **2020**, *19*, 2050034; doi:10.1142/S0219477520500340

4. Some of the open questions

There are many new questions that are triggered by our information theoretic approach and are potential objects of future research. Below we list a few items.

a) *Expert-bias*: This is a rich set of biases. Two examples are below.

We have already quantitatively addressed one potential problem originating from expertise: the issue of error corrections when the inappropriate-looking data of an observation are discarded or corrected without deeper investigation, without publishing the suspicious data and/or the error correction protocol.

Another expertise related matter that can often be seen, for example in scientific debates, is the *justification-by-limitation-bias*. Due to personal bias or society pressures, the scientist is limiting the study to a selected subset of the information space by arguing that significant or decisive effects can only be expected there. A hypothetical example is as follows: Measuring the (seemingly) sinking level of the sea in the vicinity of an island and prematurely concluding that the global sea level is sinking, even though, the island is actually rising and the global sea level is steady.

b) *Evolution of science*: Lakatos [8] points out that mathematics grows by conjectures, proof attempts and critiques of these efforts. This is naturally true also for other fields of science. The process can partially be addressed by our scheme with the essential difference that the error correction is done by other scientists.

c) *Evolution dynamics of bias, errors and error corrections due to society-pressures*: Political and financial pressures in hiring, funding and publication processes can act as a positive feedback to increase bias. A systematic positive feedback can potentially evolve into singularities such as large-scale breakdown in power grid systems (blackouts) [9].

d) *Reader bias and errors*: Influences of personality, society (e.g. media) and background knowledge can yield a heavy bias also on the Reader's interpretation of the publication, where all the above-mentioned examples can be relevant.

5. Conclusion

We have introduced an information-theoretic framework to characterize originality in scientific research. We addressed the impact of various biases, errors and error corrections.

It is important to emphasize that our measures are pessimistic; they show the worst-case scenario. For example, in many cases, discarding spurious data, without publishing them

Peer-reviewed version available at *Fluctuation and Noise Letters* **2020**, *19*, 2050034; doi:10.1142/S0219477520500340

and without describing the detailed error correction protocol, is completely fine because of reasons that are valid with high probability. Yet, such a policy reduces the information entropy of the research and, in a worst-case scenario, it may yield misleading conclusions or even hinder new discovery.

Further research goals are the exploration of potential practical applications of these ideas. For example how could this approach be used by organizations to promote more original research?

6. Acknowledgements

LBK appreciates discussions (2015-16) with Ethan Allen. This research was sponsored by the T3 initiative of Texas A&M University, 2019-20.

References

[1] D.S. Robertson, "Algorithmic Information Theory, Free Will, and the Turing Test", *Complexity* **4** (1999) 25-34.
[2] G. Chaitin, Information, Randomness & Incompleteness, Papers on Algorithmic Information Theory; World Scientific: Singapore, 1990.
[3] I. Stewart, "The Ultimate in Undecidability", *Nature* 332 (1988) 115–116.
[4] A. Renyi, Diary on Information Theory. Wiley, New York, 1987.
[5] C.E. Shannon, W. Weaver, The Mathematical Theory of Communication, University of Chicago Press, Urbana, IL, 1964.
[6] L.B. Kish, D.K. Ferry, "Information entropy and thermal entropy: apples and oranges", *J. Comp. Electr.* **17** (2018) 43–50; https://arxiv.org/abs/1706.01459
[7] L.B. Kish, "Mathematical model of the impact of political correctness on science", feature (conference closing) talk at the 8th international conference on Unsolved Problems of Noise, July 9-12, 2018, Gdansk; unpublished.
[8] I. Lakatos, "Proofs and Refutations: The Logic of Mathematical Discovery"; Cambridge Philosophy Classics: Cambridge University Press 1976.
[9] C. Singh, P. Jirutitijaroen, J. Mitra, "Electric Power Grid Reliability Evaluation: Models and Methods", IEEE Press, Wiley & Sons, New Jersey 2019.