Why is there Math in my Archaeology?

John Justeson and the Limits of Archaeological Interpretation

James Scott Cardinal Jennifer Ann Loughmiller-Cardinal

Fifty years ago, what was arguably the most important paper ever written for modern work in quantitative archaeology was published in American Antiquity. Unfortunately for its author, and generations of archaeologists, few took notice of it at the time. With few citations, more than half of which have occurred in just the last few years, its elegance and mathematical precision went largely unappreciated – even by the growing cohorts of computational and quantitative archaeologists whose work would have greatly benefited from it. In this paper, we demonstrate that John Justeson's 1973 article "Limitations of Archaeological Inference" was not only accurate and precise in its implications, but also very much still at the forefront of archaeological thought… even if the field at large doesn't yet realize it.

A Gentle Introduction to Information Theory

What is now known as "Information Theory" began with a paper written by Claude Shannon, titled "A Mathematical Theory of Communication" (1948), resulting from his work in cryptography at Bell Labs. At the heart of Shannon's theory was the idea that *information* is fundamentally tied to the reduction of *uncertainty.* Shannon approached information not in terms of meaning, but as a measure of the *reduction of uncertainty* within a system of communication.

The influence of telecommunication and cryptography on Shannon's theories are obvious, but the underlying concepts quickly found new applications and implications in other fields of study. By linking information to uncertainty and statistical probabilities, Shannon's abstracted and highly generalized model of information and communication could be adapted to studying all manner of systems. It would not be too long after the theories described in Shannon's technical paper were expanded and republished in a book a year later as "The Mathematical Theory of Communication" (Shannon and Weaver 1949) that they would begin to appear in disciplines ranging form physics to physiology – and, of course, archaeology.

Information, Entropy, and Surprisal

Shannon proposed a particular relationship between information and uncertainty in terms of statistical probabilities. He derived a quantitative measure of that uncertainty derived from the concept of *entropy* used to describe disorder in the thermodynamics of physical systems. Shannon, however, repurposed entropy to refer to the average uncertainty contained in a system given by the equation:

$$H(X) = -\sum_{i=1}^{n} p(x_i)\ \log_2\ p(x_i)$$

What this equation is describing is the total entropy H of some system X that contains n discrete attributes or elements $(x_1, x_2, ... x_n)$. The entropy is equal to the negative sum, over all n features, of each element's probability of occurrence $p(x_i)$ times the $\log_2$[1] of that probability.

The higher the entropy of a system, indicated by a larger value for H, the more uncertainty or randomness there is to the elements of X. Somewhat counterintuitively, the more uncertain or random a system the more information it conveys. Remember that Shannon defines information as the reduction of uncertainty. The greater the uncertainty (i.e., high entropy), the more potential information the system is capable of producing because there is greater uncertainty to reduce.

To see how, we need to understand what Shannon defined as *surprisal.* Surprisal, also known as self-information, is a measure of how surprising or unexpected a specific event is based on its probability. In essence, surprisal measures the information content of a specific outcome – i.e., rare events carry more information than common ones because they are less expected. Low probability events, those that occur infrequently, are highly surprising. Conversely, high probability events are not.

Consider it this way – if an event is nearly certain to occur, you would *already* be expecting it to happen when it does. Its occurrence tells you nothing that you did not already know. It is only when something happens that we did *not* expect (i.e., we are surprised) that it is providing *new* information. Therefore, surprisal (denoted as $I(x)$) is the potential *information* contained in a single event based on its probability $p(x)$:

$$I(x) = -\log_2 p(x)$$

Surprisal is zero for events that are certain (i.e., the probability $p(x) = 1$), and grows larger as the probability of the event decreases (Figure {#figure:surprisal_example}). Exceedingly rare events, by contrast, would be very surprising to witness and approaching "infinitely" surprising as the probability of the event goes to zero (i.e., $lim_{p(x)\to 0}\ I(x) = \infty$).

[1] $\log_2$ refers to the base-2 logarithm.

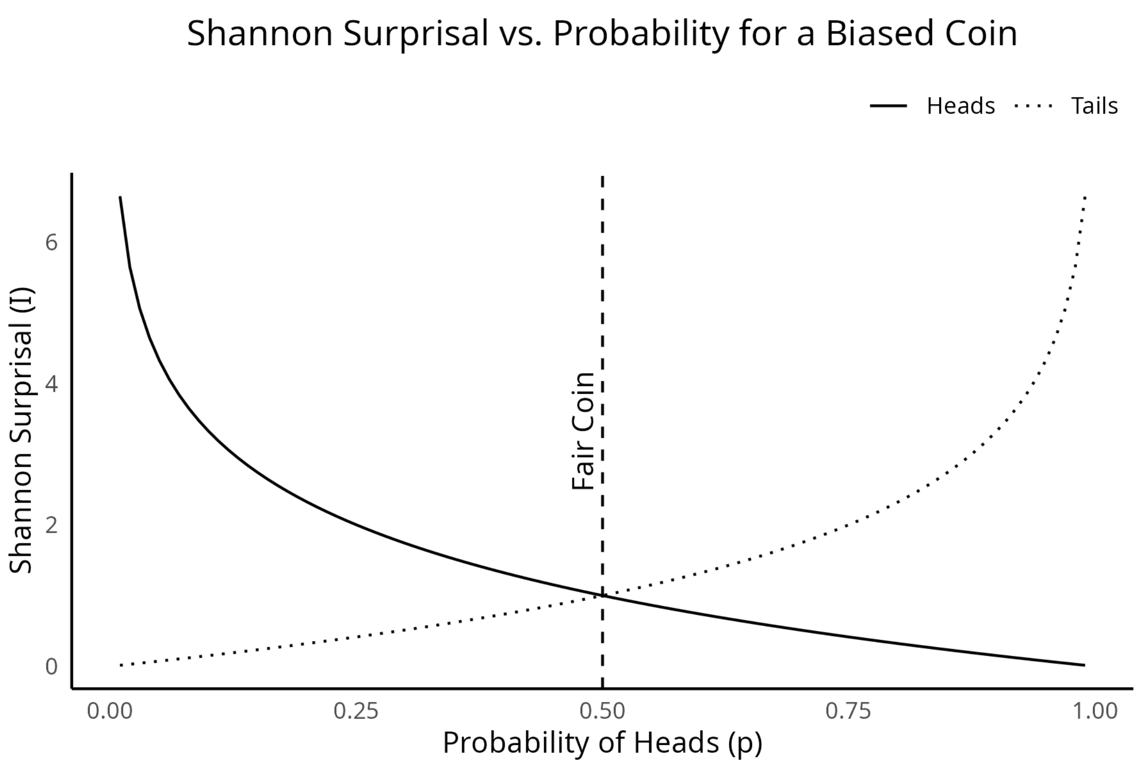

Figure 1: The surprisal $I(x)$ of a coin flip x (i.e., lands "heads" or "tails") as the probability $p(x)$ of landing "heads" ranges from 0 to 1 for a "biased" coin. A "fair" coin would land on heads or tails with equal chances or $p(x) = 0.5$.

Entropy represents the *average* surprisal over all possible outcomes from a probability distribution. It quantifies the overall uncertainty or unpredictability of a system or source of information. The higher the entropy, the more information the system is capable of producing, since there is greater uncertainty about which outcome will occur.

Entropy is highest when all outcomes are equally likely, and decreases as we gain more information to anticipate whether or not that event is likely to occur (Figure {#figure:entropy_example}). Information is therefore the reduction of that uncertainty or entropy when a new event is observed. We have learned more about the underlying probabilities for future events.

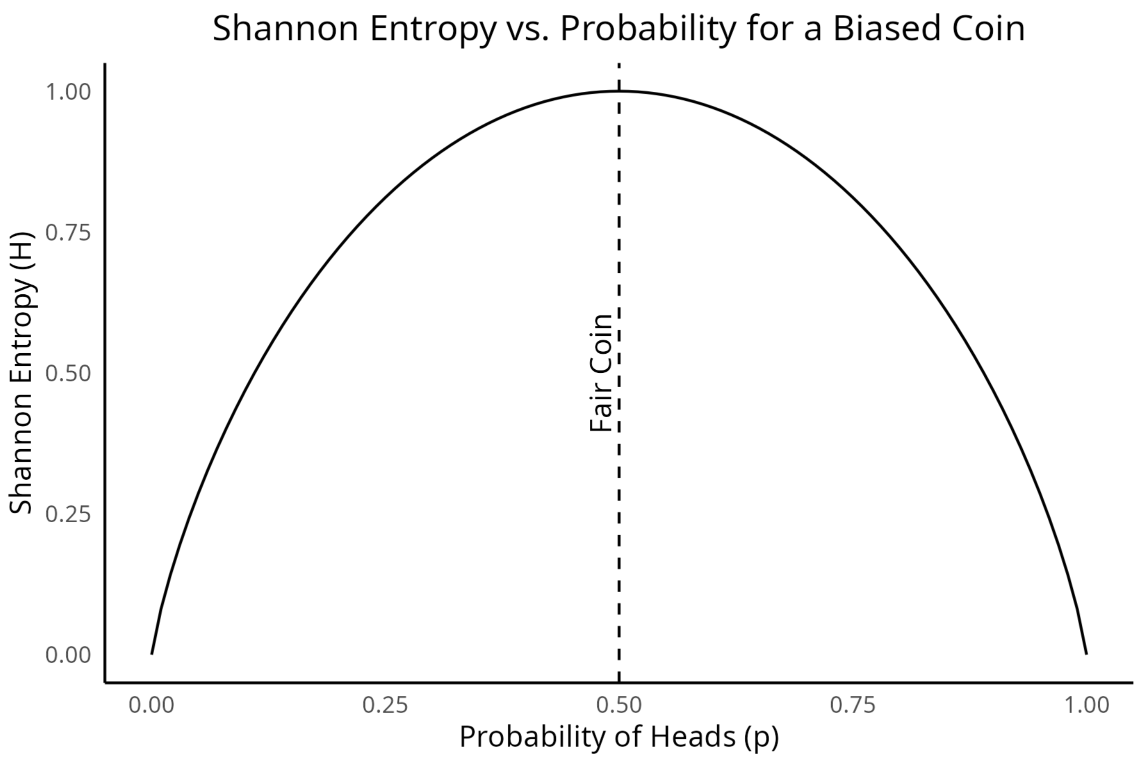

Figure 2: The overall system entropy $H(X)$ for biased coin flips as the probability $p(x)$ of landing "heads" ranges from 0 to 1. A "fair" coin with $p(x) = 0.5$ is the system with the most uncertainty, since either outcome ("heads" or "tails") is equally possible.

For the first time, scientists had a way to *quantify* information. Shannon had defined information in a way that made it possible to measure and *analyze* it mathematically, based solely on its statistical structure and independently of its content or meaning.

Information theory has evolved over the last few decades into a highly diverse discipline in its own right, with broad applications. Shannon, however, developed the theory towards one particular application – communication. Specifically, he was looking for a way to understand how information could be efficiently and reliably transmitted across communication systems, especially in the presence of noise or interference.

Channel, Signal, and Noise

Under Shannon's model of communication, the relationships between information, channel, signal, noise, and channel capacity define the core aspects for transmitting data effectively. Information represents the content or message that needs to be conveyed, which can vary in complexity. In this context, entropy is a measure of the inherent complexity of the information a message might contain – i.e., higher entropy indicates greater variability in *potential* messages.

A communication *channel* is the medium or system through which information is transmitted. Channels connect sender to receiver, and are characterized by their capacity to handle information. This *channel capacity* (C) represents that maximum rate at which information that may reliably be transmitted across that a channel. It places an upper bound on how much information can be sent through such a channel of communication with an arbitrarily low rate of errors, given by:

$$C = \max_{p(x)} I(X \mid Y)$$

where $I(X \mid Y)$ is the *mutual information*[2] between the transmitted variable X and the received variable Y. It measures the amount of information *shared* between X and Y or, put another way, how much of what was transmitted by the sender is retained and correctly *understood* by the receiver. The capacity C for the channel, then, is where maximum amount of information can be correctly transmitted with the fewest number of errors or misunderstandings by the receiver.

Information is transmitted across a communication channel as *signals*, which are encoded representations of the information. A signal is is defined as the physical embodiment of information that is transmitted across a communication channel from a sender to a receiver. Shannon treats signals as mathematical entities that *encode* data using a sequence of symbols, typically binary (0s and 1s), representing the discrete or continuous probability states of information. This encoding enables information to be manipulated, stored, and transmitted efficiently, with the ultimate goal of achieving maximum fidelity in the presence of noise or interference. Shannon's model abstracts signals into probabilistic terms, allowing for quantification of the information.

Channels, however, are not perfect. They can introduce disturbances known as *noise*, which interferes with the signal and can alter the received message, creating a challenge in accurate data transmission. The more noise present, the harder it is to reliably convey information. Noise is essentially random disturbances or fluctuations in the transmission of information along a channel that interfere with the signal. Noise can distort or obscure messages, increasing the probability of errors in decoding them.

[2] Remember, $I(x) = -\log_2 p(x)$ is the *surprisal* value of event x that represents the information conveyed by that event. The *mutual information* between two events can be though of as the information conveyed when *both* events occur simultaneously.

Since channel capacity is the maximum rate at which information can be transmitted over a channel without errors, excess noise degrades capacity by introducing errors. So, channel capacity depends on both the *bandwidth* of the channel (i.e., the allowable range of possible signal frequencies) and the *signal-to-noise ratio* (often simply called "SNR"). Shannon's theory shows that for a channel to transmit information efficiently, the signal must be strong enough to overcome noise, but *not* so strong that it leads to unnecessary redundancy in the message encodings.

This balance maximizes the channel's capacity, allowing the most efficient transfer of information while minimizing error. This gives us another way[3] to find a channel's capacity, given by:

$$C = B \log_2 \left(1 + \frac{S}{N}\right)$$

where B is the bandwidth of the channel, S is the power of the signal, and N is the noise. The signal-to-noise ratio (SNR) $\frac{S}{N}$ represents, a measure of how strong the signal is relative to the noise. As noise (N) gets larger relative to signal (S), the SNR starts dropping closer towards zero. Since $\log_2(1+0) = 0$ that means that, no matter how large its ideal bandwidth, the channel's capacity C goes towards zero as well.

A Brief History of Quantitative Analysis and Information Theory in Archaeology

The integration of quantitative methods into archaeology during the 20th century profoundly transformed the methods by which archaeologists tried to understand the past. By the 1950s, methodological innovations in quantifying archaeological analysis, in works such as Brainerd (1951) or Heizer and Cook (1956), highlighted the value of statistical frameworks in chronology-building and site analysis, establishing a precedent for rigorous quantitative inquiry in archaeology. This push intensified during the 1960s with the advent of the "New Archaeology" championed by figures such as Lewis Binford.

New Archaeology, later termed Processual Archaeology, advocated for an explicitly scientific approach grounded in hypothesis testing, formal quantification, and systems theory (Kendall 1969; Binford 1981; Raab and Goodyear 1984). David Clarke's texts *Models in Archaeology* (1972) and *Analytical Archaeology* (1978) formalized these aspirations by advocating for complex statistical models and systems theory to provide explanatory frameworks in the study of past human behavior. By situating quantitative methods at the heart of analysis, Processual Archaeology sought to go beyond mere description to causal understanding, particularly through middle-range theories that sought correlations between artifacts and behavioral processes (Binford 1981; Schiffer 1983).

[3]This way to calculate channel capacity is more common (and often much more practical) in telecommunication applications, such as those Shannon was studying, in which such things as "channel bandwidth" and "signal power" can be directly measured or otherwise experimentally ascertained.

By the 1970s, quantitative archaeology had begun to consider some of the conceptual elements of Shannon's information theory, such as its introduction in Doran's (1970) early applications of systems theory and simulation. Justeson (1973) notes the work of Fred Gorman (1970) as possibly the first *formal* mathematical application of Shannon's theory to the quantitative analysis of archaeological collections. Justeson's own offering (1973) explored the potential for information-theoretic concepts as a tool for addressing inferential challenges in archaeology. He demonstrated how Shannon's entropy measures could be applied to quantifying the integrity of archaeological *signals* by considering the formation of the archaeological record itself in terms of *channel* and *capacity*. The rapid developments in computer applications further fueled this expansion, facilitating the adoption of statistical methods across archaeological contexts (Kintigh and Ammerman 1982; Kintigh 1984; Richards 1998; Djindjian 2015).

The slow adoption and application of methods increasingly influenced by Shannon's information theory (albeit not often explicitly so) allowed archaeologists to assess patterns in artifact distribution and site organization with new mathematical precision, later inspiring applications in cultural transmission and inter-group interaction studies (Dickens and Fraser 1984). Through the 1980s, however, critiques of quantitative and rigidly "science-like" approaches to cultural phenomena began to emerge, primarily from post-processual theorists who argued for subjective interpretations and a focus on agency and meaning rather than structural functionalism (Klejn 1977).

Despite such critiques, quantitative methods, including information-theoretic approaches, continued to evolve and become an indispensible feature of archaeological methods. In recent years, advances in entropy and information measures emerging from developments in machine learning and data science have been increasingly applied to artifact analysis, as seen in works such as Paige and Perreault (2022) or Río, López-Hernández, and Chaparro Velázquez (2024), employing entropy to examine variability in stone tool production strategies. These newer studies align with a trend towards recognizing the flexibility of quantitative tools to address a broader array of archaeological questions, marking a shift away from the purely deterministic frameworks of early Processualism towards a more nuanced integration between methods and theories (Nolan 2020; Gheorghiade et al. 2023).

From the 1950s to today, quantitative analysis in archaeology has progressed somewhat independently from larger theoretical movements and critiques. Rather than the overarching epistemological ambitions of the early "New" or Processual archaeologists for a scientific objectivism, quantitative methods have instead become part of the standard toolkit of contemporary archaeological practice.

The Evolution of Information Theory in Archaeology

The integration of *formal* information theory into archaeological research has had a much slower evolution. In the late 1960s and early 1970s, inspired by Shannon's ideas, archaeologists began to experiment with these concepts to analyze the transmission of cultural traits, the integrity of artifacts, and the uncertainty inherent in archaeological data. More explicit and formal applications Shannon's model were much slower to appear. The mathematical and computational complexity of such models largely exceeded the capabilities at the time, and there remained substantial debate regarding their limitations in addressing the complexities of human behavior and cultural evolution.

Although not the first, Michael Schiffer's work (1972) is likely the best known of the early applications of a theory of information in archaeology. He tried to formalize the disruption of information flow caused by post-depositional processes, termed as "systemic and archaeological contexts." Schiffer emphasized how the information contained within artifacts could degrade over time due to various environmental and cultural factors, introducing "noise" into the archaeological record. This idea aligned with Shannon's theory of communication, where noise distorts messages as they pass through a channel. Schiffer's subsequent work (1983) on formation processes expanded on this, demonstrating how entropy, a measure of disorder, influences the amount of reliable information that survives in archaeological contexts.

Schiffer's book "Formation Processes of the Archaeological Record" (1987) is still among the most influential applications of information concepts to archaeology, even though Schiffer did not explicitly use Shannon's framework. Schiffer introduced the idea that archaeological sites are the result of two key processes – i.e, cultural formation processes involving human behaviors that create and modify archaeological deposits, and natural formation processes through non-human agents such as erosion or animal activity that affect the archaeological record.

The notion of information loss in these processes echoed Shannon's concepts of entropy and noise. Schiffer's emphasis on understanding how archaeological data are transformed before and after deposition mirrors the concerns of information theory regarding how signals are distorted through transmission. By conceptualizing the archaeological record as a series of transformations from its original state, Schiffer advanced a model that paralleled Shannon's information processing, where each formation process acts as a filter, introducing "noise" and altering the original "message."

John Justeson (1973), however, explicitly applied Shannon's concepts directly to archaeological inference by demonstrating and quantifying the theoretical limits of reconstructing past human behavior through fragmentary data. He focused on how entropy could quantify uncertainty and signal degradation, although he cautioned that oversimplification sometimes results when human complexity is reduced to mathematical models. Justeson's objective was to try and formalize the analysis of the inherent *limitations* of such interpretations. He derived a complete mathematical formulation to assess whether any given assemblage of archaeological features contained *sufficient* signal to rigorously identify underlying patterns. It directly addressed the tension between abstract quantitative frameworks and the nuances of particular cultural trajectories, a critique that has persisted in the field, but sought out a methodological compromise that could actually *quantify* that inherent uncertainty.

Justeson's work helped begin a dialogue within archaeology about the inherent limitations of inference from incomplete datasets (e.g., Sullivan 1978; Plog 1978; Hayden 1984), contributing to the development of more cautious and methodologically rigorous approaches to interpreting the archaeological record. His use of Shannon's ideas encouraged some archaeologists to critically evaluate the reliability of their data, and the extent to which they could justifiably infer past behaviors or cultural practices. Unfortunately, the sophisticated mathematical and computational understanding required for the article's *quantitative* implications and applications seems to have relegated it to relative obscurity. [4]

It would not be until the 1980s that other scholars substantially applied formal information theory to model cultural interactions. Dickens and Fraser (1984; notably citing Justeson 1973) used Shannon's idea of channel capacity to study the flow of cultural information in the Middle Woodland Period, seeking to quantify how much cultural interaction could be detected within the archaeological record. Similarly, Renfrew (1983) explored the idea of culture as a communication system, where information flows between individuals and groups. He applied Shannon's concept of information transmission to study how cultural signals travel and degrade over time, though he acknowledged the complexity of non-linear dynamics in human societies, which challenge the assumptions of equilibrium-based models.

More recently, the use of information theory in archaeology had broadened, particularly in studies of cultural transmission. Crema, Kandler, and Shennan (2016) advanced Shannon's ideas by applying equilibrium and non-equilibrium models to study cultural transmission from frequency data. They used these models to reveal how cultural traits spread and stabilize within populations, providing quantitative insights into processes that are often difficult to observe directly in the archaeological record. Similarly, Carrignon, Alexander Bentley, and O'Brien (2023) used information theory to estimate transmission rates, applying Shannon's communication model to measure the uncertainty associated with the diffusion of cultural traits.

Gheorghiade et al. (2023) expanded Shannon's concept of entropy into a framework they called "Entropology" that posits entropy measures to better understand archaeological data. They critique the traditional applications of information theory for focusing too narrowly on entropy without accounting for the broader complexity and uncertainty of archaeological contexts. This critique echoes the central debate surrounding the use of information theory in archaeology – i.e., while it offers valuable tools for formalizing the study of cultural transmission and data integrity, it does not capture the intricate and chaotic nature of human historical exigencies.

Another major critique of these applications, such as that by Raab and Goodyear (1984), concerns the oversimplification of human behaviors when abstract models like those derived from Shannon's theories are applied. They argue that middle-range theory, which often uses these models, fails to capture the full complexity of human action. Zubrow (1972) similarly critiqued the difficulty of accounting for environmental and social variables when applying information-theoretic frameworks. Despite this, some scholars such as Nolan (2020) have proposed to refine these models. Nolan assessed entropy, noise, and channel capacity to evaluate the significance (in the technical and regulatory cultural resources sense of the term)

[4] It is worth noting that John had published this article while still a graduate student at Stanford University, before completing his Masters. The article has seen a recent and substantial resurgence of attention, garnering more citations within the last ten years than it had in the previous four decades.

of archaeological data, particularly focusing on how much information about past societies could be accurately recovered from the fragmented and noisy record.

The use of Shannon's information theory in archaeology has evolved from early models of data degradation and cultural transmission to more sophisticated frameworks that incorporate the entropy and uncertainty in teh archaeological record. Scholars like Schiffer (1972), Justeson (1973), and Renfrew (1983) laid the foundation, while modern researchers like Nolan (2020), Crema, Kandler, and Shennan (2016), and Gheorghiade et al. (2023) have expanded these concepts to address the challenges posed by incomplete and noisy archaeological records. However, the ongoing debate highlights the tension between the precision offered by information theory and the complex realities of human history, questioning the extent to which these mathematical models can truly capture the richness of the past.

Ironically, the ensuing debates largely failed to recall that one of them already provided a roadmap for determining *exactly* that extent quite early on in the venture.

The Limitations of Archaeological Inference

In Justeson's 1973 article, he introduces Shannon's theory of communication as a means to formalize the analysis of inherent inferential limitations in archaeological interpretation. In the introduction, he firmly situates the paper within what was, at the time, growing theoretical tensions between conflicting goals within archaeology. Some were advocating for a "new" archaeology focused on "predictive behavioral science" while the "traditional" archaeology's aim was the reconstruction of "social and cultural histories" (Justeson 1973, 131). Justeson viewed the distinction as merely opposing "poles on a continuum of research commitments" and instead posed a slightly different question – is there a way to determine whether or not we were actually capable of doing *either*, given a particular archaeological source?

The article is presented in two parts. The first ("A Theoretical Framework") introduces the relevant aspects of Shannon's theory of communication and posits the analogous relationships between those processes and the nature archaeological data. He makes the argument that it is not sufficient just to describe the archaeological record as an information channel, but that is is specifically a channel of a particular *type* that allow its interpretation. The second part ("Application of Information-Theoretic Measures") illustrates how specific measures for the fidelity or integrity of the archaeological record can be derived from Shannon's model. He presents a set of formal tools through which archaeologists could calculate these specific measures from observations of artifact attributes to determine whether sufficient information existed within an assemblage to be *interpretable*.

Rather than present a predominately conceptual framework, as previous works had done (e.g., Doran 1970; Schiffer 1972; Clarke 1973), the objectives of the article were more ambitious. Justeson aimed to use Shannon's formal models to demonstrate the specific properties and capabilities of the archaeological record to transmit information. Working backwards from the observation that the archaeological record functioned as a communication channel, he showed that such a channel must also possess the formal properties of a certain type of communication network. Furthermore, he showed that the encoding of information carried by that channel would need to take a particular form in order for it to successful convey interpretable information about the past. In essence, "…the archaeologist is in the position of the code-breaker tapping a channel with whose code he is not fully familiar by means of another channel" (Justeson 1973, 134).

"Part I – A Theoretical Framework"

The paper builds from the hypothesis that archaeological interpretation is fundamentally limited by the quantity of information that can be extracted from the archaeological record. Like the other early archaeological invocations of information theory, Justeson described the archaeological record as a degraded and incomplete set of signals from past behavior that are transmitted through the "channel" of the archaeological record. Each artifact or feature would represent a small, noisy fraction of the original cultural system.

Justeson's main departure from the others was in that he applied Shannon's concept of entropy *directly* to the assessment of the degree of uncertainty that might be incorporated into archaeological interpretations simply by the nature of that channel. Remember, Shannon linked reduction of uncertainty to information. Justeson focused on highlighting how *noise* in the archaeological record – due to processes such as taphonomy or excavation biases – interacts with the inherent *entropy* entailed by the archaeological features or processes that *encode* past behaviors. Those interactions innately affect the *capacity* of the archaeological channel to reliably transmit information.

Therefore, the inherent limit of archaeological inference would be the limits (i.e., the "upper bound" in mathematical terms) of the channel's capacity given a certain amount of noise. Past those limits, *de*coding the source signal (i.e. the *behavior*) would become highly susceptible to more ambiguous, unreliable, or even spurious interpretations. To find – and *calculate* – that limit, Justeson needed to specify the nature of the transmission channel and its properties and identify (and prove) the existence of a coherent system of encoding.

Basic Concepts of Information Theory and Their Archaeological Correlates Whereas Shannon described information in terms of the reduction of uncertainty, Justeson notes that information can also be thought of in terms of *contrasts*. That is, information can be seen as a way we are able to distinguish the qualities or attributes of one type of thing from those of another. Information, then, is how we determine categories. In that case, information reduces the uncertainty of correctly assigning a thing or event to a given category.

That information is, in the archaeological case, *encoded* through the expression of material traits that reflect past behaviors as an input message or source *signal*. That message or signal is transmitted across a channel, which has particular characteristics and limitations. These properties allow for the introduction of *noise* that may affect the input signal and alter or obscure the original message.

Figure 3: Schematic representation of information transmission ("Fig. 1" in Justeson 1973, 133).

> The code joining signal to channel is a trivial construct: the input signal is the material assemblage laid down, and their laying down is the "code" that commits them to the channel.
>
> The "code" by which the output signals are recovered is, likewise, the process of recovery. The code joining the human behavior to its material consequents as committed to the earth is a code which is not at all trivial; it is the crucial concern of the archaeologist, since it is through it that cultural description is attempted.
>
> The above description of the total communication process becomes somewhat more complex if we reconsider the input signals as being primarily the inputs and outputs of a communication process that was operative within the original culture. In that case it becomes clear that the archaeologist is in the position of the code-breaker tapping a channel with whose code he is not fully familiar by means of another channel. (Justeson 1973, 134)

Channel Classification, Channel Properties, and Codes ...

> The basic distinction is between memoryless channels and channels with memory. In a memoryless channel, any transmission is unaffected by any other transmission, and the elements of the input signal are unaffected by the other elements of the signal; in a channel with memory, past history does have a bearing upon later transmissions. Another distinction is between finite and infinite input and/or output alphabets. The input alphabet is the set of symbols drawn upon to form the input signal, which the output alphabet is the set of symbols drawn upon for the output signal. The symbols of these alphabets have been mentioned earlier in this report as "elements." (Justeson 1973, 135)

...

> If the empirically measured parameters are not consistent with the relationship between them that is required by the theory for a given material or behavioral system, then the data by which that system is to be interpreted cannot have a consistent susceptibility to decoding; that is, there will be no basis for deriving a coherent archaeological interpretation of the data that will accurately reflect the prehistoric situation. Thus, the question of the existence of a code is one of primary importance for our considerations. (Justeson 1973, 136)

The code is defined mathematically as a system of N ordered pairs consisting each of an input sequence u_i of n alphabetic symbols and a set A_i of output sequences, where the N sets of output sequences do no have any members in common, and where the probability that the sequence received when u_i is transmitted will be among the members of the set A_i will always be greater than or equal to $1 - \lambda$ where λ is greater than 0 and less than or equal to 1. In symbols, it is "a system

$$\left\{(u_1, A_1), \dots, (u_N, A_N)\right\}$$

where the u_i are n-sequences, the A_i are disjoint sets of n-sequences and

$$P\left\{v(u_i) \in A_i\right\} \geqslant 1 - \lambda,\ i = 1, \dots, N$$

… we shall call it a code (n, N, λ)" (Wolfowitz 1961, 51–52). The expression $v(u_i)$ represents the signal received when u_i is sent, while the term *n-sequence* is an input or output signal of length n.

The parameters n, N, and X therefore specify the code for the channel. Without such a code we cannot really speak of information being transmitte dor received, hence there is really no basis for speaking of the existence of a channel. We can find out if there is a code for he archeological channel by finding if values we compute for N, n, and λ are consistent with the requirements of a code for a discrete finite-memory channel. Inparticular, the value of N is related to that of n by the formula $N = 2^{n(C-\epsilon)}$, where C is the *channel capacity* – a measure of the ability of the channel to transmit information – and ϵ is a positive constant. C may be determined, often only with great labor, from the relation

$$C = \max_{\pi} \left\{ \sum_j \left[\sum_i \pi_i w(j|i) \log_2 \sum_i \pi_i w(j|i) - \sum_i \pi_i w(j|i) \log_2 \sum_i \pi_i w(j|i) \right] \right\}$$

where $\pi = (\pi_1, \dots, \pi_k)$ is any probability distribution, $w(j|i)$ is the probability of receiving j if i is sent – j can be null – and k is the number of elements in the input alphabet…

If the empirically measured parameters are not consistent with the relationship between them that is required by the theory for a given material or behavioral system, then the data by which that system is to be interpreted cannot have a consistent susceptibility to decoding; that is, there will be no basis for deriving a coherent archaeological interpretation of the data that will accurately reflect the prehistoric situation. Thus, the question of the existence of a code is one of primary importance for our considerations.

Part II – Application of Information-Theoretic Measures

Extrapolation of the Prehistoric Distribution of Design Elements

The first step in using this technique is to tabulate the values of N_t, the number of attributes occurring exactly t times in the sample, where the values of t range from 1 to T, and T is the maximum number of times any attribute occurs. Then M_t, the number of attributes occurring at least t times, is calculated and tabulated. M_t can be calculated from the formula:

$$M_t = \sum_{i=t}^{T} N_i$$

M_1 is the number of different design elements occurring in the sample.

Finally, L_t is calculated and tabulated. Its formula is

$$L_t = \sum_{i=t}^{T} M_i$$

L_t has no meaning except when $t = 1$, at which point it is the number of occurrences of attributes in the sample.

From the curve obtained, a value at $t = 0$ can be extrapolated to give M_0, the number of attributes occurring at least 0 times in the sample – the total number of attributes in the system. Since $L_0 = M_0 + L_1$, the extrapolation of L_0 provides an alternate check on M_0.

Noise Levels

To calculate the noise factor for each design element [$\psi(r)$, where r is the rank by decreasing frequency of the design element and $p(r)$ is the frequency of the element of rank r] (*sic*), let

$$\begin{aligned}\psi(r) &= P(\text{receiving design element } r \text{ given that } r \text{ was sent}) \\ &= P(\text{receiving } r \mid r \text{ was sent})\end{aligned}$$

But $P(A|B) = P(A \text{ and } B) \div P(B)$[5], so

$$\begin{aligned}P(r \text{ sent and } r \text{ received}) &= P(r \text{ sent} \mid r \text{ received}) \cdot P(r \text{ received}) \\ &= P(r \text{ received} \mid r \text{ sent}) \cdot P(r \text{ sent})\end{aligned}$$

[5]This is known as the *conditional probability* of an event. In other words, what is the probability of event A if we know event B has happened (i.e., $p(A|B)$ or "probability of A *given* B"), which is equal to the probability that A and B occur together ($P(A \text{ and } B)$ or $P(A \cup B)$) divided by the probability the B happens ($P(B)$).

But $P(r \text{ received} \mid r \text{ sent}) = \psi(r)$; $P(r \text{ sent}) = p_E(r)$, the extrapolated frequency for the design element r; $P(r \text{ sent} \mid r \text{ received}) = 1$; and $P(r \text{ received}) = p(r)$. Thus, $\psi(r) = p(r)/p_E(r)$. To find the parameter λ of our code we must find the minimum of the $\psi(r)$ values,

$$\begin{aligned}
\min_r \; \psi(r) &= \min_r P(r \text{ received} \mid r \text{ sent}) \\
&= \min_r 1 - P(r \text{ not received} \mid r \text{ sent}) \\
&= \max_r P(r \text{ not received} \mid r \text{ sent}) \\
&= 1 - \lambda
\end{aligned}$$

... the observed values should never be higher than the extrapolated, and that the mean value should be 1, since

$$\begin{aligned}
\overline{\psi} &= \sum_r p_E(r)\psi(r) \\
&= \sum_r p_E(r) \left[p(r) \div p_E(r)\right] \\
&= \sum_r p(r) = 1
\end{aligned}$$

Existence of a Code

The value of $C - \epsilon$ must be less than 1 to be at all restrictive, since 2^n is the number of different combinations than can be gotten from n elements, so that there will always be at most 2^n code words. Since the value of ϵ must lie between 0 and C, our range of values for $N = 2^{n(C-\epsilon)}$ is not limited any further. There will, therefore, at least if $C > 1$, be a value of ϵ for which the actual value of N is achieved, so that a code does exist in this case.

Information Distortion

System Dynamics

$$\begin{aligned}
H' &= -\sum_{i=1}^{k} p(x_i) \log_2 p(x_i) \\
&= -\sum_{i=1}^{k} \frac{1}{k} \log_2 \frac{1}{k} \\
&= -\log_2 \frac{1}{k} \\
&= \log_2 k
\end{aligned}$$

$$h = H/H' \text{ and } h_E = H_E/H'_E$$

Binary Coding and Its Applications

To set up the binary code for an attribute system, the attributes should first be ranked by frequency from highest to lowest; then the frequencies are divided into 2 groups with equal frequency totals, or with totals as nearly equal as is possible. The first group receives the code digit 0, and is made up of the higher-frequency elements, the second receives the digit 1 and is made up of the lower-frequency elements. Then the process is applied to each of these subgroups, and then continually to the resulting subgroups until all the attributes have been isolated in single-attribute groups. This process is represented schematically by Fig. 6. The binary codings themselves are given in Tables 1 and 2.

Given this data we compute H^*, the information content of the system in terms of binary coding, or, in other words, the maximum information retrievable for a given frequency distribution, by the formula

$$H^* = \sum_{i=1}^{k} p(x_i)b(x_i)$$

where $b(x_i)$ is the number of digits in the binary code for attribute xi, and the other quantities the formula are as in the last section. The ration $h^* = H/H^*$ is then a measure of the coding efficiency.

References Cited

Binford, Lewis R. 1981. "Behavioral Archaeology and the "Pompeii Premise"." *Journal of Anthropological Research* 37 (3): 195–208. https://doi.org/10.1086/jar.37.3.3629723.

Brainerd, George W. 1951. "The Place of Chronological Ordering in Archaeological Analysis." *American Antiquity* 16 (4): 301–13. https://doi.org/10.2307/276979.

Carrignon, Simon, R. Alexander Bentley, and Michael J. O'Brien. 2023. "Estimating Two Key Dimensions of Cultural Transmission from Archaeological Data." *Journal of Anthropological Archaeology* 72 (December): 101545. https://doi.org/10.1016/j.jaa.2023.101545.

Clarke, David L., ed. 1972. *Models in Archaeology.* London: Methuen & Co.

———. 1973. "Archaeology: The Loss of Innocence." *Antiquity* 47 (185): 6–18. https://doi.org/10.1017/s0003598x0003461x.

———. 1978. *Analytical Archaeology.* 2. ed. New York: Columbia University Press.

Crema, Enrico R., Anne Kandler, and Stephen Shennan. 2016. "Revealing Patterns of Cultural Transmission from Frequency Data: Equilibrium and Non-Equilibrium Assumptions." *Scientific Reports* 6 (1). https://doi.org/10.1038/srep39122.

Dickens, Roy S., and Martin D. Fraser. 1984. "An Information-Theoretic Approach to the Analysis of Cultural Interaction in the Middle Woodland Period." *Southeastern Archaeology* 3 (2): 144–52.

Djindjian, François. 2015. "Computers and Mathematics in Archaeology, Anatomy of an Ineluctable Success!" In *CAA2014: 21st Century Archaeology*, edited by F. Giligny, F. Djindjian, L. Costa, P. Moscati, and S. Robert, 1–6. Archaeopress Publishing Ltd. https://doi.org/10.2307/jj.15135883.4.

Doran, James. 1970. "Systems Theory, Computer Simulations and Archaeology." *World Archaeology* 1 (3): 289–98. https://doi.org/10.1080/00438243.1970.9979448.

Gheorghiade, Paula, Vaiva Vasiliauskaite, Aleksandr Diachenko, Henry Price, Tim Evans, and Ray Rivers. 2023. "Entropology: An Information-Theoretic Approach to Understanding Archaeological Data." *Journal of Archaeological Method and Theory* 30 (4): 1109–41. https://doi.org/10.1007/s10816-023-09627-4.

Gorman, Frederick. 1970. "Symbol as Strategy: Toward a Synthetic View of Material Culture." Field Museum of Natural History.

Hayden, Brian. 1984. "Are Emic Types Relevant to Archaeology?" *Ethnohistory* 31 (2): 79. https://doi.org/10.2307/482057.

Heizer, Robert F., and Sherburne F. Cook. 1956. "Some Aspects of the Quantitative Approach in Archaeology." *Southwestern Journal of Anthropology* 12 (3): 229–48. https://doi.org/10.1086/soutjanth.12.3.3629082.

Justeson, John S. 1973. "Limitations of Archaeological Inference: An Information-Theoretic Approach with Applications in Methodology." *American Antiquity* 38 (2): 131–49. https://doi.org/10.2307/279360.

Kendall, David G. 1969. "Some Problems and Methods in Statistical Archaeology." *World Archaeology* 1 (1): 68–76. https://doi.org/10.1080/00438243.1969.9979428.

Kintigh, Keith W. 1984. "Measuring Archaeological Diversity by Comparison with Simulated Assemblages." *American Antiquity* 49 (1): 44–54. https://doi.org/10.2307/280511.

Kintigh, Keith W., and Albert J. Ammerman. 1982. "Heuristic Approaches to Spatial Analysis in Archaeology." *American Antiquity* 47 (1): 31–63. https://doi.org/10.2307/280052.

Klejn, Leo S. 1977. "A Panorama of Theoretical Archaeology." *Current Anthropology* 18 (1): 1–42. https://doi.org/10.1086/201846.

Nolan, Kevin C. 2020. "Bringing Archaeology into the Information Age: Entropy, Noise, Channel Capacity, and Information Potential in Archaeological Significance Assessments." *Quality & Quantity* 54 (4): 1171–96. https://doi.org/10.1007/s11135-020-00980-0.

Paige, Jonathan, and Charles Perreault. 2022. "How Surprising Are Lithic Reduction Strategies? The Information Entropy of the Modes a-i Framework." *Lithic Technology* 48 (3): 237–52. https://doi.org/10.1080/01977261.2022.2113699.

Plog, Stephen. 1978. "Social Interaction and Stylistic Similarity: A Reanalysis." In *Advances in Archaeological Method and Theory*, 143–82. Elsevier. https://doi.org/10.1016/b978-0-12-003101-6.50011-1.

Raab, L. Mark, and Albert C. Goodyear. 1984. "Middle-Range Theory in Archaeology: A Critical Review of Origins and Applications." *American Antiquity* 49 (2): 255–68. https://doi.org/10.2307/280018.

Renfrew, A. Colin. 1983. "Divided We Stand: Aspects of Archaeology and Information."

American Antiquity 48 (1): 3–16. https://doi.org/10.2307/279815.

Richards, Julian D. 1998. "Recent Trends in Computer Applications in Archaeology." *Journal of Archaeological Research* 6 (4): 331–82. https://doi.org/10.1007/bf02446083.

Río, Fernando del, Rodrigo López-Hernández, and Claudia Chaparro Velázquez. 2024. "On Information, Entropy, and Early Stone Tools." *Molecular Physics*, February. https://doi.org/10.1080/00268976.2024.2310644.

Schiffer, Michael B. 1972. "Archaeological Context and Systemic Context." *American Antiquity* 37 (2): 156–65. https://doi.org/10.2307/278203.

———. 1983. "Toward the Identification of Formation Processes." *American Antiquity* 48 (4): 675–706. https://doi.org/10.2307/279771.

———. 1987. *Formation Processes of the Archaeological Record*. University of New Mexico Press.

Shannon, Claude E. 1948. "A Mathematical Theory of Communication." *Bell System Technical Journal* 27 (4): 623–56. https://doi.org/10.1002/j.1538-7305.1948.tb00917.x.

Shannon, Claude E., and Warren Weaver. 1949. *The Mathematical Theory of Communication*. Urbana, Illinois: University of Illinois Press.

Sullivan, Alan P. 1978. "Inference and Evidence in Archaeology: A Discussion of the Conceptual Problems." In *Advances in Archaeological Method and Theory*, edited by Michael B. Schiffer, 1:183–222. Elsevier. https://doi.org/10.1016/b978-0-12-003101-6.50012-3.

Wolfowitz, Jacob. 1961. *Coding Theorems of Information Theory*. Englewood Cliffs, N.J.,: Prentice-Hall.

Zubrow, Ezra B. W. 1972. "Environment, Subsistence, and Society: The Changing Archaeological Perspective." *Annual Review of Anthropology* 1 (1): 179–206. https://doi.org/10.1146/annurev.an.01.100172.001143.