Complexity versus Probability: A Comparative Analysis of Simplicity Theory and Bayesian Inference in Cognition and AI

Part I: The Algorithmic-Cognitive Framework: Simplicity Theory

Simplicity Theory (ST), developed primarily by Jean-Louis Dessalles, offers a distinctive and powerful framework for understanding a specific yet fundamental aspect of human cognition: the detection of relevance and the experience of surprise.[1] It posits that the human mind is not merely a passive recipient of information but an active pattern-seeker, exquisitely tuned to identify situations that appear "abnormally simple".[1] This theory departs significantly from traditional probabilistic accounts by grounding its core mechanisms in Algorithmic Information Theory (AIT), proposing that what makes an event interesting is not its low probability per se, but a quantifiable drop in its complexity.[2] By operationalizing abstract concepts from AIT, ST aims to provide a formal, predictive, and descriptive model of cognitive phenomena ranging from conversational narrative to creative insight.

1.1 Foundations in Algorithmic Information Theory (AIT)

The theoretical bedrock of Simplicity Theory is Algorithmic Information Theory, a field that provides a formal, objective measure of complexity.[4] Understanding this foundation is essential to grasping the novel claims and mechanisms of ST.

Core Concept: Kolmogorov Complexity (KC)

At the heart of AIT lies the concept of Kolmogorov Complexity (KC), also known as algorithmic complexity.[4] The KC of an object, such as a string of text or a sequence of numbers, is defined as the length of the shortest possible computer program, written in a fixed universal programming language, that can produce that object as its output and then halt.[4] This provides a measure of the computational resources needed to specify the object, effectively quantifying its intrinsic complexity or incompressibility.[4]

For example, a highly patterned string like "abababababababab" can be generated by a very short program (e.g., "print 'ab' 8 times"), and thus has a low Kolmogorov Complexity. In contrast, a string that appears random, such as "arfbxgkwdlqpvcz," contains no discernible regularities that can be exploited for compression. The shortest program to generate it is essentially the program itself ("print 'arfbxgkwdlqpvcz'"), meaning its KC is approximately the length of the string itself.[6] In this view, simplicity is synonymous with compressibility.[6] This measure is theoretically robust; while the absolute complexity value depends on the choice of programming language, the difference in complexity between two objects is largely independent of the language, differing only by a constant factor.[8]

The Incomputability Problem

A profound and well-known challenge within AIT is that Kolmogorov Complexity is, in the general case, an uncomputable function.[4] Akin to Turing's halting problem, there exists no general algorithm that can take an arbitrary string as input and output the length of its shortest possible program.[4] This is because proving that a given program is the

shortest would require testing all shorter programs to see if they also produce the same output, and there is no universal way to know if these programs will ever halt. This theoretical obstacle would seem to render KC unusable as a practical foundation for any model of cognition, as the human mind must operate within finite time and with finite resources.

ST's Solution: Observer-Dependent, Resource-Bounded Complexity

Simplicity Theory offers a pragmatic and cognitively plausible resolution to the incomputability problem. It posits that the human mind does not and cannot compute the absolute, universal Kolmogorov Complexity of an event.[10] Instead, it computes a resource-bounded and observer-dependent version of complexity.[2] This crucial move transforms complexity from an abstract, objective mathematical property into a subjective, psychological one.

According to ST, the complexity of a situation is always relative to the descriptive and computational resources available to a specific observer at a specific moment.[2] The "shortest description" is not the shortest in any universal language, but the shortest one the observer can personally find using their current knowledge, memory, and pattern-matching abilities. This observer-dependency elegantly resolves many paradoxes. For instance, the lottery draw "12-22-27-37-38-42" has a high complexity for a general observer who sees no pattern. However, for the person who played that exact combination, its description complexity is minimal—it is simply "the numbers I chose".[2] This makes the complexity, and thus the entire framework of ST, contingent on the individual's cognitive state, grounding the theory in psychological reality rather than abstract mathematics.

1.2 The Core Mechanism: Unexpectedness as Complexity Drop

The central conceptual innovation of Simplicity Theory lies not merely in its application of Kolmogorov Complexity, but in its crucial bifurcation of this concept into two distinct, psychologically meaningful measures: generation complexity and description complexity. This division transforms a monolithic mathematical property into a dynamic, relational model of an observer confronting a world. The discrepancy between these two measures is the engine that drives the theory.

Formal Definition

ST's central claim is formalized in its core equation, which defines unexpectedness, U, for a given situation, s, as the difference between its expected complexity and its observed complexity [2]:

$$U(s)=C_{gen}(s)-C_{desc}(s)$$

Generation Complexity (C_gen)

Generation complexity, Cgen(s), represents the observer's model of the causal complexity of the event s.[2] It is the length of the minimal description of all the parameters that must be specified in the "world" for the situation to be generated or to come into existence.[10] In essence,

Cgen is the agent's theory of the world's causal structure—their best guess at the "true" complexity of the process that produced the event. For a fair lottery drawing six numbers from 49, the generation complexity for any specific sequence is high and roughly equal, as it requires the specification of six independent random draws, amounting to a complexity of approximately $6 \times \log 2(49)$ bits.[3]

Description Complexity (C_desc)

Description complexity, Cdesc(s), is the length of the shortest description of the situation s that the observer can mentally construct using their available cognitive resources.[2] This term represents the observer's internal cognitive ability to find a pattern or a compact representation of the observed data. For a "random" lottery sequence like "12-22-27-37-38-42," the observer has no simple rule to describe it, so its

Cdesc is high, close to its Cgen. However, for the sequence "22-23-24-25-26-27," the observer can easily compress it with a simple rule like "the arithmetic sequence starting at 22," making its Cdesc very low.[2]

Unexpectedness (U)

The key output of the theory, unexpectedness (U), is therefore not a measure of absolute simplicity or improbability, but a measure of the *discrepancy* between the agent's theory of the world and the agent's perception of the data. A situation is deemed interesting, relevant, or surprising when it is significantly simpler to describe than the observer's causal model of the world would lead them to expect (i.e., when $C_{desc} \ll C_{gen}$).[1] This "complexity drop" or "randomness deficiency" acts as a powerful cognitive signal.[2] The human mind, in ST's view, is a machine exquisitely tuned to detect this difference.[3] A large positive

U is a cognitive alarm bell that signals, "Pay attention, your model of reality might be incomplete or wrong, because this event is too simple to have been generated by the process you assumed."

1.3 From Unexpectedness to Subjective Experience

Simplicity Theory leverages the concept of unexpectedness to build a quantitative and predictive model of a wide range of subjective human experiences, moving from abstract complexity to concrete cognitive phenomena.

Modeling Interest and Relevance

The theory posits that unexpectedness (U) is one of the two primary ingredients of what humans find interesting (the other being emotional intensity).[10] Events with a high

U value are those that capture our attention, are deemed memorable, and are considered worthy of being reported in conversational narratives or news media.[3] This provides a formal explanation for the "So what?" test of narrative relevance; a story is rejected as uninteresting if it describes an ordinary event where

$C_{gen} \approx C_{desc}$ and thus $U \approx 0$.[12]

This framework explains why we find coincidences so compelling. Meeting a specific friend by chance in a distant foreign city is a highly unexpected event. The generation complexity is enormous (the product of all the independent choices that led both individuals to that exact spot at that exact time). However, the description complexity is much lower, as the event can be compressed by noting the shared identity of the two people and the single location.[12] In contrast, meeting a stranger in that same spot is not interesting, because although the generation complexity is equally high, the description complexity is also high (as there is no simple relation to compress), resulting in a low unexpectedness value.

Defining Subjective Probability

One of the most radical proposals of Simplicity Theory is its redefinition of *ex-post* subjective probability, p, for a unique event:

$p=2^{-U}$

This formulation has several profound implications.[1] First, it is designed to apply to unique, individual situations, a domain where classical probability theory struggles because the objective probability of any fully instantiated, non-repeatable event is effectively zero.[2] Second, it allows for the assessment of subjective probability without requiring knowledge of a complete, well-defined space of alternative events, a major departure from the axiomatic requirements of standard probability.[2] Third, it explains the powerful human intuition that a highly patterned lottery draw like "1, 2, 3, 4, 5, 6" is subjectively less probable than a "random" one, despite both having the same objective probability in a fair draw.[2] The patterned draw has a large complexity drop (

$U>0$), yielding a low subjective probability, while the random draw has no complexity drop ($U \approx 0$), yielding a subjective probability near 1.

A notable feature of this definition is that these subjective probabilities are not required to sum to 1 across different outcomes.[10] Two different, unremarkable lottery draws can both have a subjective probability close to 1. ST argues this is not a flaw, as this constraint only applies to

ex-ante objective probabilities, not *ex-post* assessments of plausibility for individual events.[10]

Broad Applications

The explanatory reach of ST extends to a wide array of cognitive domains. It has been applied to model creativity as a search for solutions that provide an unexpected drop in problem complexity.[11] It has been used to formalize emotional intensity, where the magnitude of an emotion is a function of both unexpectedness and the personal impact of the situation.[11] Furthermore, ST has been operationalized in artificial intelligence and natural language processing, for instance, to solve word analogies by finding the transformation of minimal complexity or to mine for "intuitive" relational explanations in knowledge bases.[1]

1.4 ST as a Descriptive Model of Cognition

It is crucial to position Simplicity Theory within the broader landscape of cognitive modeling. Its primary contribution and philosophical stance is that of a *descriptive* theory of cognition.

Focus on "How" vs. "Should"

Unlike normative theories that prescribe how a perfectly rational agent *should* think, ST's primary goal is to characterize and explain the actual, observable regularities of human thought.[15] It seeks to model the cognitive mechanisms underlying the human faculty for detecting relevance, generating surprise, and structuring narratives.[11] Its aim is to answer the question "How do human minds find things interesting?" rather than "What is the most rational way to process information?".

An Alternative to Probabilistic Accounts

ST is often presented as a direct challenge to the sufficiency of purely probabilistic models for explaining human interest.[11] Proponents argue that frameworks based on Shannon's information theory or Bayesian inference fail to capture the special cognitive status of events that are simple but improbable (like the patterned lottery draw) or deterministic yet surprising (like a predictable lunar eclipse, which is surprising to someone unaware of celestial mechanics).[11] In these cases, the objective probability may be known, but the cognitive experience of surprise is driven by the perceived complexity drop. This suggests that a complexity-based metric may be more fundamental to human intuition than a purely probabilistic one in these specific contexts.

Falsifiable Predictions

Despite its descriptive nature, a key strength of ST is that it is not merely a "just-so story" or a post-hoc rationalization. By defining unexpectedness quantitatively, the theory makes concrete, falsifiable predictions about which events people will find interesting, memorable, and worth communicating.[3] For example, it predicts that the interestingness of a coincidence increases with the complexity of the shared feature and decreases with the simplicity of the context (e.g., meeting a friend is more surprising in a large, distant city than in one's small hometown).[3] This commitment to quantitative prediction lends the theory scientific rigor and distinguishes it from purely qualitative accounts of cognition.

Part II: The Probabilistic-Inferential Framework: Bayesian Model Inference

In stark contrast to the algorithmic-cognitive approach of Simplicity Theory, Bayesian inference provides a comprehensive framework for reasoning and learning under uncertainty, rooted entirely in the axioms of probability theory.[18] While often introduced as a simple rule for updating probabilities, its true power lies in its application as a complete system for model comparison and selection. This system gives rise to an emergent principle of parsimony, known as the Bayesian Occam's Razor, which serves as the primary point of comparison with the complexity-based principles of Simplicity Theory.

2.1 Foundations in Probability Theory

The Bayesian framework is built upon a specific interpretation of probability and a single, powerful theorem that serves as its core inferential engine.

The Bayesian Interpretation of Probability

The Bayesian paradigm interprets probability not as a long-run frequency of an event in repeated trials (the frequentist view), but as a subjective "degree of belief" or confidence that an individual holds about the truth of a proposition.[20] A probability of 1 represents absolute certainty that a proposition is true, a probability of 0 represents absolute certainty that it is false, and values in between represent graded levels of confidence.[20] The central tenet of Bayesianism is that it is rational to update these degrees of belief as new evidence becomes available.[19] This subjective interpretation is fundamental, as it allows probabilities to be assigned to hypotheses (e.g., "the probability that this coin is biased") and not just to the outcomes of random events.

The Core Machinery: Bayes' Theorem for Model Inference

The mechanism for rational belief updating is Bayes' theorem. When applied to the problem of comparing different models or hypotheses in light of data, the theorem is expressed as follows [19]:

$P(H \mid D)=P(D)P(D \mid H)\times P(H)$

Each component of this equation plays a distinct and crucial role in the inferential process:

- **Prior Probability, $P(H)$:** The prior represents the initial degree of belief in a hypothesis, H, *before* any data, D, has been observed.[18] This term allows for the principled incorporation of existing knowledge, assumptions, or experience into the model. For example, in a clinical trial, one might assign a prior belief that a new drug is unlikely to have a large effect.[22] The choice of prior is a critical and sometimes contentious aspect of Bayesian modeling.
- **Likelihood, $P(D \mid H)$:** The likelihood function is the bridge that connects an abstract hypothesis to concrete data. It quantifies the probability of observing the specific data D *if* the hypothesis H were true.[22] For example, if the hypothesis is "this coin is fair" (

H), the likelihood of observing 7 heads in 10 tosses (D) can be calculated directly from the binomial distribution.

- **Posterior Probability, $P(H|D)$:** The posterior is the ultimate output of the inference. It represents the updated, revised degree of belief in the hypothesis H *after* the evidence D has been taken into account.[18] It is a rational synthesis of the prior belief and the information contained in the data. This posterior can then serve as the prior for a subsequent round of belief updating when new data arrives.[20]
- **Evidence (or Marginal Likelihood), $P(D)$:** This term, often treated as a mere normalizing constant in simple parameter estimation problems, is the absolute key to Bayesian model selection.[18] The evidence represents the total probability of the observed data, calculated by averaging the likelihood over all possible hypotheses (or all possible parameter values within a model), weighted by their respective prior probabilities.[24] Formally, for a model with parameters θ:
$P(D)=\int P(D|\theta,H)P(\theta|H)d\theta$
This integral evaluates the *average* predictive performance of the model H across its entire parameter space, not just its performance at the single best-fitting parameter value. As will be shown, this averaging property is the source of the Bayesian Occam's Razor.

2.2 The Bayesian Occam's Razor: An Emergent Principle of Parsimony

One of the most elegant and powerful features of Bayesian model selection is that it provides an automatic, quantitative preference for simpler models without needing to add an explicit penalty term for complexity. This "Bayesian Occam's Razor" is not an ad-hoc addition but an emergent property of the model evidence, P(D).[25] The penalty is not for a model's

syntactic complexity (e.g., the length of its formula), but for its *semantic flexibility*—its capacity to explain too many different outcomes. A model is punished not for being long-winded, but for being non-committal.

The Problem of Overfitting

A fundamental challenge in all of statistical modeling is overfitting.[28] A more complex model (e.g., a high-degree polynomial for curve fitting) can always achieve a better fit to a given set of data points than a simpler model. However, this superior fit often comes at the cost of capturing random noise in the data rather than the true underlying signal. As a result, the overfitted model will make poor predictions for new, unseen data.[28] The principle of parsimony, or Occam's Razor, is the heuristic that we should prefer the simplest explanation that adequately fits the data.[28]

How the Evidence P(D) Automatically Penalizes Complexity

The Bayesian framework provides a formal, mathematical justification for this heuristic. The mechanism operates through the model evidence, P(D). A complex model, by virtue of having more parameters or parameters with wider ranges, is inherently more flexible. It is capable of generating or "predicting" a much wider variety of possible datasets.[25] However, any probabilistic model is governed by what has been called the "Law of Conservation of Belief": the total probability it can assign across all possible datasets must sum to one.[28]

This constraint forces a complex model to "spread its predictive probability more thinly" over the vast space of data it could potentially explain.[25] In contrast, a simpler model, being less flexible, makes more specific, concentrated predictions, focusing its probability mass on a much smaller range of possible datasets.[25]

The "Occam Factor"

As a result of this dynamic, when a specific dataset D is observed that both a simple model (H1) and a complex model (H2) can explain, the simple model will almost always have assigned a higher average probability to that dataset. The model evidence, $P(D|H1)$, will be greater than $P(D|H2)$. This is because the complex model "wasted" much of its prior belief on predicting other datasets that did not end up occurring.[28]

This effect can be formalized by approximating the evidence as the product of the best-fit likelihood and an "Occam factor" [25]:

$$P(D|H) \approx P(D|w_{MP},H) \times \frac{\sigma_{w|D}}{\sigma_w}$$

Here, $P(D|w_{MP},H)$ is the likelihood at the best-fitting (maximum posterior) parameters. The second term is the Occam factor, where σ_w is the prior uncertainty over the parameters and $\sigma_{w|D}$ is the posterior uncertainty after seeing the data.[25] A complex model starts with a large prior uncertainty (

σ_w) and thus is penalized more heavily. The evidence naturally favors models that are not only accurate (high best-fit likelihood) but also predictive and parsimonious (a large Occam factor). The ratio of the evidence values for two competing models, $P(D|H_1)/P(D|H_2)$, which is known as the Bayes Factor, directly quantifies the extent to which the data support the simpler model over the more complex one, thus implementing an "automatic Occam's Razor".[27]

2.3 Bayesianism as a Normative and Descriptive Framework

The Bayesian framework holds a unique dual status in cognitive science, serving as both a benchmark for ideal rationality and a descriptive model of actual mental processes.

The Normative Standard

At its core, Bayesian inference is a *normative* theory.[15] It provides a set of mathematically coherent and provably optimal rules for how a rational agent

should update their beliefs and make decisions in the face of uncertainty. It is a theory of ideal reasoning, grounded in the axioms of probability theory as laid down by figures like Andrey Kolmogorov.[19] Philosophers and statisticians appeal to this normative status to argue that it is the correct way to conduct scientific inference.

The Descriptive Turn: The "Bayesian Brain" Hypothesis

Beginning in the late 20th century, a paradigm shift occurred in cognitive science, where this normative framework was increasingly adopted as a *descriptive* model of the mind.[31] The "Bayesian Brain" hypothesis posits that the brain's fundamental computations are inherently Bayesian.[32] This perspective suggests that neural systems, from low-level perception and motor control to high-level reasoning and language, are constantly performing probabilistic inference to contend with the sparse, noisy, and ambiguous nature of sensory input.[31] For example, the perceptual system is thought to execute an unconscious Bayesian inference, combining prior knowledge about the world with incoming sensory data to form a stable and reliable perception of the environment.[32]

The Is-Ought Duality

This dual status is one of the most powerful and debated aspects of Bayesian cognitive science. The framework is used simultaneously to define "optimal" performance on a cognitive task and to describe the "actual" performance of human subjects.[34] When human behavior aligns with the Bayesian model, it is often interpreted as evidence that the mind is an "optimal" information processor. When behavior deviates, these discrepancies can be analyzed to understand the specific heuristics or biases at play.[34] This creates a powerful but sometimes conceptually blurry link between the "is" of human cognition and the "ought" of rational inference, making Bayesianism a uniquely versatile tool for studying the mind.

Part III: A Comparative Synthesis: Algorithmic Complexity vs. Bayesian Probability

Having established the distinct foundations and mechanisms of Simplicity Theory and Bayesian inference, a direct comparative analysis reveals their deep-seated differences, surprising points of convergence, and complementary explanatory strengths. This synthesis moves beyond a surface-level description to a critical evaluation of their relationship, using the Minimum Description Length (MDL) principle as a crucial conceptual bridge.

3.1 Foundational Axioms: A Tale of Two Theories

The most fundamental divergence between the two frameworks lies in their axiomatic origins and the primary objects of their analysis. This initial comparison is best summarized in a structured format before delving into the detailed arguments.

Table 1: Foundational Comparison of Simplicity Theory and Bayesian Inference

Feature	Simplicity Theory (ST)	Bayesian Model Inference
Theoretical Foundation	Algorithmic Information Theory (AIT) [1]	Probability Theory [18]
Core Metric	Unexpectedness (U): A complexity drop [2]	Posterior Probability ($P(H
Object of Analysis	Unique, individual events (s) [10]	Distributions over hypotheses/parameters (H) [24]
Treatment of Simplicity	Direct calculation: U=Cgen−Cdesc [3]	Emergent property: Penalty via model evidence P(D) [25]
Primary Stance	Descriptive: How minds find things interesting [11]	Normative: How beliefs *should* be updated [16]
Key Challenge	Incomputability of KC; relies on observer-bounded approximations [4]	Requires well-defined, measurable space of events; choice of prior [13]

ST's Critique of Probability

Simplicity Theory's departure from probability theory is not incidental; it is a core part of its motivation. Proponents of ST argue that the axioms of standard probability theory are fundamentally ill-suited for modeling key aspects of human cognition.[13] A primary objection is the requirement of a pre-defined, measurable space of events. This axiom assumes that all possible outcomes can be enumerated in a well-behaved set, which is often not representative of how humans flexibly and dynamically construct concepts and event spaces on the fly.[13] Furthermore, ST claims that probability theory, on its own, cannot adequately explain the cognitive experience of surprise in certain situations. For example, a deterministic event like a predictable lunar eclipse can still be highly surprising to an observer who lacks the correct causal model. Similarly, a highly patterned lottery draw like "1,1,1,1,1,1" is perceived as more surprising and less subjectively probable than a random-looking one, even when their objective probabilities are identical.[13] ST argues that in these cases, the cognitive signal is driven by a complexity drop, not by probability.

Bayesian Critique of AIT/ST

Conversely, from a Bayesian perspective, the foundation of ST appears less rigorous. The primary critique centers on ST's reliance on Kolmogorov Complexity, a theoretically incomputable measure.[4] While ST's solution of using an "observer-dependent, resource-bounded" complexity is cognitively plausible, it introduces a degree of subjectivity that can seem ad-hoc from the standpoint of a formal theory.[10] The choice of what constitutes an observer's "available descriptions" and "computational operations" is not axiomatized and could be difficult to specify and constrain a priori. This contrasts sharply with the rigorous, axiomatic foundation of modern probability theory, which was, ironically, co-developed by Andrey Kolmogorov himself.[19] To a Bayesian, the subjective choice of a prior distribution, while challenging, is arguably more transparent and formally constrained than the subjective choice of a descriptive language in ST.

The "Unique Event" vs. "Distribution" Problem

Perhaps the most profound and clarifying distinction between the two frameworks is their primary explanatory target. Simplicity Theory is fundamentally a theory of *single-instance salience*. It is designed to explain why a specific, unique, non-repeatable event (s) captures human attention and is deemed interesting or memorable.[10] Its core mechanism,

$U=C_{gen}-C_{desc}$, is a post-hoc analysis of the properties of a single, already-observed event.

In stark contrast, Bayesian inference is a theory of *model-based generalization from data streams*. It is designed to update our beliefs about the underlying generative processes (H) that produce data over time.[19] Its core mechanism, Bayes' theorem, is an engine for learning about distributions and making predictions about future events. ST explains the "Aha!" moment; Bayesianism explains the process of learning. This difference in their primary object of analysis—a unique event versus a distribution over hypotheses—is not a minor technicality but a deep conceptual divide that dictates their respective domains of applicability.

3.2 Re-examining Occam's Razor: Two Paths to Simplicity

The apparent chasm between Simplicity Theory's direct, subtractive calculation of complexity and Bayesianism's emergent, probabilistic penalty for complexity is elegantly bridged by the Minimum Description Length (MDL) principle. This reveals that both frameworks are ultimately striving for the same goal—parsimonious explanation—but are approaching it through different, though formally related, mathematical languages.

The Minimum Description Length (MDL) Principle as a Bridge

The MDL principle provides a formal, information-theoretic statement of Occam's Razor: the best hypothesis to explain a set of data is the one that permits the greatest compression of the data.[5] This is typically formulated as a "two-part code," where the total description length,

$L(D)$, for a dataset D using a model or hypothesis H, is the sum of the length of the description of the model itself, $L(H)$, and the length of the description of the data encoded *with the help of* the model, $L(D|H)$ [9]:

$$L(D)=L(H)+L(D|H)$$

The goal of MDL-based inference is to find the hypothesis H that minimizes this total codelength. This naturally trades off model complexity (a more complex model requires a longer description $L(H)$) with goodness-of-fit (a model that fits well allows for a shorter description of the data's deviations, $L(D|H)$).[36]

Connecting MDL to Bayesianism

The pivotal connection between MDL and Bayesian inference comes from Shannon's source coding theorem, which establishes a fundamental relationship between probability and optimal codelength: the most efficient code for an event x with probability $P(x)$ has a length of approximately $-\log_2 P(x)$ bits.[8]

Applying this equivalence to the Bayesian objective of maximizing the posterior probability, $P(H|D)$, reveals a deep connection. Maximizing $P(H|D)$ is equivalent to maximizing its logarithm, $\log P(H|D)$. From Bayes' rule, $P(H|D) \propto P(H)P(D|H)$. Therefore, maximizing the posterior is equivalent to maximizing $\log(P(H)P(D|H)) = \log P(H) + \log P(D|H)$. Minimizing the *negative* of this quantity, $-\log P(H) - \log P(D|H)$, is mathematically analogous to minimizing the MDL two-part codelength $L(H)+L(D|H)$.[37]

This demonstrates that Bayesian model selection can be interpreted as a probabilistic method for finding the most compact description of the data.[38] The Bayesian Occam's Razor is the probabilistic shadow of this complexity-based principle. This reframes the debate: it is not simply "complexity vs. probability," but a more nuanced comparison of different formalizations of the same deep principle of parsimony.

Connecting MDL to ST

While ST's formula $U = C_{gen} - C_{desc}$ is not identical to the MDL two-part code, both are born from the same AIT philosophy that equates learning and understanding with compression. MDL provides the crucial insight that the Bayesian preference for simplicity and the complexity-based preference of ST are not alien concepts. Under certain theoretical conditions, such as the use of a "universal prior" where the prior probability of a hypothesis is determined by its Kolmogorov complexity ($P(H) \approx 2^{-K(H)}$), Bayesian inference and MDL become formally equivalent.[39] This establishes a formal link, showing that both frameworks can be seen as different approaches to the same fundamental problem of finding simple, regular patterns in a complex world.

3.3 Is Bayes' Rule a Special Case of Simplicity Theory?

A provocative hypothesis, advanced by Dessalles and Sileno, suggests that the relationship is not one of parallel approaches but of hierarchy: that Bayes' rule can be viewed as a specific instance of a more general inferential template provided by ST's unexpectedness formula.[13] This claim merits careful and critical examination.

The Dessalles-Sileno Conjecture

The conjecture starts by taking the logarithmic form of Bayes' rule and attempting to map its components onto complexity-based terms from ST. The goal is to show that the probabilistic relationships in Bayesian inference can be reconstructed using the calculus of complexity differences. The proposed mapping is summarized in the table below.

Table 2: An Analytical Mapping of Bayesian Terms to Simplicity Theory Concepts

Bayesian Term	ST Concept / Mapping	Mathematical Formulation (Bayes)	Mathematical Formulation (ST)	Notes / Critique
Posterior $P(M	O)$	**Plausibility** 2–U	$\frac{P(O	M)P(M)}{P(O)}$
Likelihood $P(O	M)$	**Likelihood Complexity** 2–CL	$\int P(O	\theta,M)P(\theta

Prior P(M)	**Prior Complexity** 2–CM	P(M)	CM=Cw(c) (Complexity to generate the cause c)	This maps the prior probability of a model to the causal complexity of a specific cause, which is a significant conceptual leap.[17]
Evidence P(O)	**Environmental Complexity** 2–CE	$\int P(O	M)P(M)dM$	CE=Cw(s) (Total causal complexity of the situation)

Analyzing the Proposed Mapping

A rigorous analysis of this mapping reveals several significant conceptual difficulties that challenge the claim of direct subsumption.

First, there is a fundamental mismatch in the objects being analyzed. The Bayesian posterior, P(M∣O), is a measure of belief in a *model* (M) given an observation (O). ST's core metric, U(s), is a measure of the unexpectedness of a *situation* (s) itself.[13] The mapping requires equating the situation

s with the observation O, but then must find a counterpart for the model M within ST's framework, which is not naturally present.

Second, to make the mapping work, the proponents must introduce new, derived complexity measures that are not part of the original, parsimonious formulation of ST. For example, a "likelihood complexity" (CL) is defined as the complexity to generate an effect from a cause.[17] This suggests that the mapping is not a natural equivalence but a constructed analogy, where the concepts of ST must be retrofitted to resemble the terms in Bayes' rule.

Third, the mapping of the Bayesian evidence, P(O), is particularly problematic. In Bayesian inference, the evidence is the average likelihood over the entire space of models. The proposed ST equivalent is Cw(s), the generation complexity of the single observed situation.[13] This conflates a property averaged over a distribution of hypotheses with a property of a single instance, overlooking the core mechanism of the Bayesian Occam's Razor, which operates precisely through this averaging.

Verdict on the Conjecture

While the formal claim that Bayes' rule is a "specific instance" of ST's unexpectedness formula appears strained and not rigorously proven, the conjecture should not be dismissed outright. Its true value lies not in establishing formal subsumption, but in serving as a powerful and insightful critique of the axiomatic limitations of standard probability theory from a cognitive perspective. It compellingly argues that a complexity-based framework like ST may be more fundamental or general for describing certain types of human inference—especially those involving the detection of relevance in unique events, the experience of coincidence, and the generation of subjective surprise—that fall outside the strict requirements of a measurable event space and probabilistic calculus.

3.4 Explanatory Power and Limitations

Ultimately, the value of any scientific theory lies in its explanatory power. A comparison of ST and Bayesianism reveals that they are not so much direct competitors as they are specialists with different, though sometimes overlapping, domains of expertise.

Domains Where ST Excels

Simplicity Theory provides uniquely intuitive and formal explanations for a class of cognitive phenomena that are often difficult to capture with standard Bayesian models. These include:

- **The "Aha!" Moment and Coincidence:** ST's $U = C_{gen} - C_{desc}$ elegantly models the sudden insight that comes from finding a simple pattern in what seemed to be complex data, as well as the uncanny feeling of a meaningful coincidence.[1]

- **Narrative and Relevance:** The theory provides a formal basis for what makes a story interesting and worth telling, explaining why we communicate about the unusual and not the mundane.[3]
- **Interest in Deterministic Events:** ST can explain why a fully deterministic and predictable event, like a solar eclipse, can still be surprising and interesting to an observer who lacks the correct causal model (i.e., whose Cgen is high).[13]
- **Relevance without Big Data:** ST proposes a mechanism for a system to identify memorable or anomalous events from a data stream without relying on large datasets for statistical training, a key feature of human cognition.[1]

Domains Where Bayesianism Excels

Conversely, the Bayesian framework possesses unparalleled power in domains where ST has limited or no explanatory reach. These are the cornerstones of modern machine learning and rational decision theory:

- **Incremental Learning from Data:** The iterative nature of updating a posterior distribution makes Bayesianism the ideal framework for modeling robust, incremental learning over time as new evidence accumulates.[19]
- **Hierarchical Modeling and Knowledge Integration:** Bayesian methods allow for the construction of complex hierarchical models that can integrate knowledge and uncertainty at multiple levels of abstraction, a key feature of sophisticated reasoning.[14]
- **Principled Quantification of Uncertainty:** The framework provides a complete and coherent system for representing and manipulating uncertainty, not just about data but about model parameters themselves.[20]
- **Optimal Prediction and Decision-Making:** Bayesian decision theory combines posterior probabilities with utility functions to provide a normative standard for making optimal choices under ambiguity.[32]

Part IV: Integration and Future Directions

The comparative analysis reveals that Simplicity Theory and Bayesian inference, rather than being mutually exclusive opponents, can be viewed as holding a complex and potentially synergistic relationship. Moving beyond a simple versus-style comparison to a more integrated perspective illuminates how these frameworks can coexist, inform each other, and jointly contribute to a more complete science of mind and engineering of intelligence.

4.1 A Unified View?: Heuristics, Approximations, and Cognitive Reality

The most productive way to reconcile the two frameworks within cognitive science is to frame ST's mechanism as a descriptive cognitive *heuristic* that efficiently *approximates* a Bayesian *norm*. This reframes their relationship from one of opposition to one of functional, hierarchical integration.

Simplicity as a Heuristic for Bayesian Inference

Full Bayesian inference, with its need to define priors and integrate over complex parameter spaces, is computationally expensive and often intractable for humans to perform in real-time, especially given limited information and cognitive resources.[22] Cognitive science has long recognized that humans rely on a wide range of "fast and frugal" heuristics to make decisions under uncertainty.[42] The principle of parsimony, or Occam's Razor, is one of the most powerful and pervasive of these heuristics, with ample research demonstrating a strong and robust human preference for simpler explanations.[6]

The Bayesian Occam's Razor provides the *normative justification* for this cognitive preference. It demonstrates, through the mathematics of model evidence, that this bias towards simplicity is not arbitrary or merely aesthetic; simpler hypotheses are, in a formal sense, often more probable.[25]

Evidence from Cognitive Science

This theoretical link is supported by empirical evidence. Studies have shown that when faced with uncertain explanations, people use an explanation's simplicity as a direct cue to infer its prior probability and its likelihood of being true.[41] This effect is particularly strong when the formal Bayesian calculation is difficult or when the necessary probabilistic information is missing, suggesting that simplicity serves as a cognitive shortcut to approximate the output of a more complex rational calculation.[41]

Reconciling the Frameworks

This perspective offers a powerful reconciliation. Simplicity Theory can be understood not as an "opponent" to Bayesianism in the cognitive realm, but as a descriptive, algorithmic-level model of the *process* the brain might use to implement a "good enough," resource-bounded version of Bayesian (or more broadly, rational) inference. ST provides a precise, mechanistic account of *how* the brain might implement the simplicity heuristic: by being hard-wired to detect a "complexity drop" via the U = C_{gen} - C_{desc} computation. This mechanism is fast, efficient, and does not require the explicit calculation of probabilities or the integration over belief distributions. In the language of dual-process theory, ST could be seen as describing the fast, intuitive System 1, which generates proposals and flags events based on their salience, while formal Bayesianism describes the slow, deliberative System 2 or the normative ideal that the entire cognitive system strives toward.[34]

4.2 Implications for Artificial Intelligence

The choice between these two frameworks has profound implications for the goals and methods of artificial intelligence research. The question "Which theory is better for AI?" depends entirely on what kind of intelligence one aims to build.

AI based on ST

Adopting Simplicity Theory as a guiding principle would lead to the development of AI systems with a more human-like sense of relevance and interest. The goal would be to create machines that can identify "memorable," "interesting," or "anomalous" events from a massive, unstructured stream of data without extensive prior training or labeled examples.[1] This has direct applications in:

- **Data Summarization:** Automatically generating concise summaries of large datasets or long periods of activity by highlighting only the most unexpected events.
- **Anomaly Detection:** Identifying potential system failures or security threats by flagging situations that are abnormally simple compared to their expected causal complexity.
- **Human-Computer Interaction:** Creating more natural and engaging conversational agents that can understand what a human user would find interesting to discuss, avoiding the reporting of mundane facts and focusing on narratively relevant information.[1]

  The CompLog system, a proof-of-concept that uses ST's inferential mechanisms for reasoning, provides a concrete example of this research direction.1

AI based on Bayesianism

The Bayesian approach is already a well-established and dominant paradigm in a vast portion of modern machine learning and AI. It is the foundation for systems that excel at:

- **Robust Prediction and Classification:** Building models that can make accurate predictions under uncertainty, from medical diagnosis to financial forecasting.
- **Learning from Data:** Developing algorithms that can learn the parameters of complex models from data in a principled and robust manner.
- **Probabilistic Reasoning:** Constructing expert systems, such as Bayesian networks, that can reason about cause and effect in complex domains where knowledge is incomplete.[14]

A Hybrid Future

The most promising future direction for building general and flexible AI likely lies in hybrid systems that combine the strengths of both approaches. One can envision an architecture where powerful Bayesian methods form the backbone of the system, responsible for robust learning, probabilistic reasoning, and optimal decision-making. Layered on top of or integrated with this backbone would be ST-like mechanisms designed to guide the focus of attention. These complexity-based modules could act as a relevance filter, identifying the most salient data points from a high-bandwidth input stream to be fed into the more computationally expensive Bayesian machinery. Such a hybrid system could learn robustly like a machine while attending to the world with the relevance-detecting acuity of a human.

4.3 Conclusion: Complementary Frameworks for a Science of Mind

This comprehensive analysis of Simplicity Theory and Bayesian model inference leads to a nuanced conclusion that rejects a simple adversarial view in favor of a complementary partnership.

Summary of Findings

The two frameworks originate from fundamentally different theoretical traditions—Algorithmic Information Theory and Probability Theory, respectively. This foundational difference leads them to have distinct primary explanatory targets: ST focuses on explaining the cognitive salience of unique, individual events, while Bayesianism provides a general framework for updating beliefs about the generative models that produce data streams. Consequently, their approaches to the principle of parsimony, or Occam's Razor, are distinct. ST employs a direct, subtractive calculation of a "complexity drop" ($U=C_{gen}-C_{desc}$), a post-hoc analysis of a single event's properties. In contrast, the Bayesian Occam's Razor is an emergent property of model evidence, arising from an average-case evaluation of a model's predictive performance across its entire parameter space.

Rejecting a Zero-Sum View

The analysis demonstrates that a zero-sum view, in which one theory must be "right" and the other "wrong," is unproductive and misleading. The question "Which theory is better?" is ill-posed without first specifying the context and the precise explanatory goal. ST provides a superior explanation for the subjective experience of coincidence, while Bayesianism provides a superior framework for optimal learning from statistical data.

A Complementary Partnership

The final verdict is that Simplicity Theory and Bayesian inference offer a powerful and complementary partnership for a comprehensive science of mind. Simplicity Theory provides a unique, formal, and predictive descriptive theory for a specific and crucial set of cognitive faculties: the detection of relevance, the generation of interest, and the structuring of narrative. These are the mechanisms that guide our attention and communication. Bayesian inference, in turn, provides a broad and powerful normative and descriptive framework for the general problem of learning, reasoning, and decision-making under the pervasive uncertainty that characterizes our world. A complete understanding of cognition requires both the specific, algorithmic lens of Simplicity Theory to explain what we find interesting, and the general, probabilistic framework of Bayesianism to explain how we learn from what we find. Together, they offer a richer and more complete picture of the computational architecture of the human mind.

Works cited

1. Jean-Louis Dessalles - Home page, accessed July 17, 2025, https://perso.telecom-paristech.fr/jld/papiers/simplicity.html
2. Simplicity theory - Wikipedia, accessed July 17, 2025, https://en.wikipedia.org/wiki/Simplicity_theory
3. (PDF) Algorithmic Simplicity and Relevance - ResearchGate, accessed July 17, 2025, https://www.researchgate.net/publication/230639494_Algorithmic_Simplicity_and_Relevance
4. Kolmogorov complexity - Wikipedia, accessed July 17, 2025, https://en.wikipedia.org/wiki/Kolmogorov_complexity
5. Kolmogorov Complexity & Minimum Description Length - Statistics How To, accessed July 17, 2025,

https://www.statisticshowto.com/minimum-description-length-kolmogorov-complexity/
6. The simplicity principle in perception and cognition - PMC - PubMed Central, accessed July 17, 2025, https://pmc.ncbi.nlm.nih.gov/articles/PMC5125387/
7. The Simplicity Principle in Human Concept Learning - Rutgers Center for Cognitive Science, accessed July 17, 2025, https://ruccs.rutgers.edu/images/personal-jacob-feldman/papers/feldman_CDPS.pdf
8. The simplicity principle in perception and cognition, accessed July 17, 2025, http://ruccs.rutgers.edu/images/personal-jacob-feldman/papers/feldman_simplicity_2016.pdf
9. Minimum description length - Wikipedia, accessed July 17, 2025, https://en.wikipedia.org/wiki/Minimum_description_length
10. Simplicity Theory, accessed July 17, 2025, https://simplicitytheory.telecom-paris.fr/
11. Simplicity theory: teaching relevances to artificial intelligences - I'MTech, accessed July 17, 2025, https://imtech.imt.fr/en/2016/10/28/simplicity-theory-teaching-relevance-ai/
12. Conversational topic connectedness predicted by Simplicity Theory - eScholarship.org, accessed July 17, 2025, https://escholarship.org/uc/item/8kt3d66b
13. UvA-DARE (Digital Academic Repository), accessed July 17, 2025, https://pure.uva.nl/ws/files/134178934/978_3_031_12429_7_8.pdf
14. Unexpectedness and Bayes' Rule | Request PDF - ResearchGate, accessed July 17, 2025, https://www.researchgate.net/publication/367584737_Unexpectedness_and_Bayes'_Rule
15. Normative vs. descriptive models - (Intermediate Microeconomic Theory) - Vocab, Definition, Explanations | Fiveable, accessed July 17, 2025, https://library.fiveable.me/key-terms/intermediate-microeconomic-theory/normative-vs-descriptive-models
16. Descriptive Decision Theory (Stanford Encyclopedia of Philosophy), accessed July 17, 2025, https://plato.stanford.edu/entries/decision-theory-descriptive/
17. (PDF) Unexpectedness and Bayes' Rule - ResearchGate, accessed July 17, 2025, https://www.researchgate.net/publication/364382008_Unexpectedness_and_Bayes'_Rule
18. Bayesian inference | Introduction with explained examples - StatLect, accessed July 17, 2025, https://www.statlect.com/fundamentals-of-statistics/Bayesian-inference
19. Bayesian inference - Wikipedia, accessed July 17, 2025, https://en.wikipedia.org/wiki/Bayesian_inference
20. Bayesian Statistics: A Beginner's Guide | QuantStart, accessed July 17, 2025, https://www.quantstart.com/articles/Bayesian-Statistics-A-Beginners-Guide/
21. The Basics of the Bayesian Approach: An Introductory Tutorial - The Change Lab, accessed July 17, 2025, https://thechangelab.stanford.edu/tutorials/bayesian-methods/the-basics-of-the

-bayesian-approach-an-introductory-tutorial/
22. 10 Steps to Master Bayesian Inference: A Practical Guide - Number Analytics, accessed July 17, 2025, https://www.numberanalytics.com/blog/10-steps-to-master-bayesian-inference-practical-guide
23. Bayesian Inference - Introduction to Machine Learning - Wolfram, accessed July 17, 2025, https://www.wolfram.com/language/introduction-machine-learning/bayesian-inference/
24. Chapter 12 Bayesian Inference - Statistics & Data Science, accessed July 17, 2025, https://www.stat.cmu.edu/~larry/=sml/Bayes.pdf
25. Model Comparison and Occam's Razor, accessed July 17, 2025, https://www.inference.org.uk/mackay/itprnn/ps/343.355.pdf
26. Bayesian Occam's Razor Is a Razor of the People, accessed July 17, 2025, https://cognition.princeton.edu/document/125
27. SHARPENING OCKHAM'S RAZOR ON A BAYESIAN STROP by William H. Je erys University of Texas at Austin and James O. Berger Purdue Uni, accessed July 17, 2025, https://quasar.as.utexas.edu/papers/ockham.pdf
28. Occam's Razor - Probabilistic Models of Cognition, accessed July 17, 2025, http://probmods.org/chapters/occams-razor.html
29. Simplicity in the Philosophy of Science, accessed July 17, 2025, https://iep.utm.edu/simplici/
30. A note on the evidence and Bayesian Occam's razor, accessed July 17, 2025, https://mlg.eng.cam.ac.uk/zoubin/papers/05occam/occam.pdf
31. Bayesian Cognitive Science, Monopoly, and Neglected Frameworks - Open Access LMU, accessed July 17, 2025, https://epub.ub.uni-muenchen.de/41932/2/BayesCogSci_final.pdf
32. A Realist Perspective on Bayesian Cognitive Science Michael Rescorla Abstract - Philosophy - UCLA, accessed July 17, 2025, https://philosophy.ucla.edu/wp-content/uploads/2016/08/Realist-Perspective-.pdf
33. Bayesian models of cognitive development Elizabeth Bonawitz & Tomer Ullman - Projects at Harvard, accessed July 17, 2025, https://projects.iq.harvard.edu/sites/projects.iq.harvard.edu/files/bbb_chapter20.pdf
34. Discrepancies between normative and descriptive models of decision making and the understanding/acceptance principle - PubMed, accessed July 17, 2025, https://pubmed.ncbi.nlm.nih.gov/10328857/
35. Behavioral Decision Making in Normative and Descriptive Views: A Critical Review of Literature - MDPI, accessed July 17, 2025, https://www.mdpi.com/1911-8074/14/10/490
36. A Tutorial Introduction to the Minimum Description Length Principle - CWI, accessed July 17, 2025, https://homepages.cwi.nl/~pdg/ftp/mdlintro.pdf
37. Minimum Message Length and Kolmogorov Complexity - CiteSeerX, accessed July 17, 2025, https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=4ef08582dc5c

11b3286b463307dd6ee739d2f8c6
38. MDL, Bayesian Inference and the Geometry of the Space of ..., accessed July 17, 2025, https://www.sas.upenn.edu/~vbalasub/public-html/Inference_files/MDLChapter.pdf
39. (PDF) Minimum Description Length Induction, Bayesianism, and Kolmogorov Complexity, accessed July 17, 2025, https://www.researchgate.net/publication/301882515_Minimum_Description_Length_Induction_Bayesianism_and_Kolmogorov_Complexity
40. Minimum Description Length Induction, Bayesianism, and ... - arXiv, accessed July 17, 2025, https://arxiv.org/pdf/cs/9901014
41. Simplicity as a Cue to Probability: Multiple Roles for Simplicity in ..., accessed July 17, 2025, https://cognition.princeton.edu/sites/g/files/toruqf3386/files/cognitive_science_-_2022_-_vrantsidis_-_simplicity_as_a_cue_to_probability_multiple_roles_for_simplicity_in_evaluating.pdf
42. Decision-Making Models | Cognitive Psychology Class Notes - Fiveable, accessed July 17, 2025, https://library.fiveable.me/cognitive-psychology/unit-11/decision-making-models/study-guide/ZfFUqJu7OgvQsuNu
43. Bayes and the Simplicity Principle in Perception - Rutgers Center for Cognitive Science, accessed July 17, 2025, https://ruccs.rutgers.edu/images/personal-jacob-feldman/papers/feldman_psychreview2009.pdf