Granular Interaction Thinking Theory in Open Science: A Novel Approach for Enhancing the Plausibility of Social Sciences

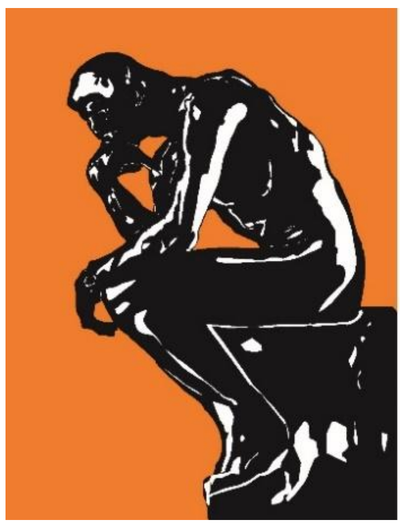
Minh-Hoang Nguyen [1,*], Viet-Phuong La [1,2], Quan-Hoang Vuong [1,3]

[1] Centre for Interdisciplinary Social Research, Phenikaa University, Hanoi, Vietnam

[2] A.I. for Social Data Lab (AISDL), Hanoi, Vietnam

[3] Professor, University College, Korea University. 145 Anam-ro, Seongbuk-Gu, Seoul 02841, South Korea

* Correspondence: hoang.nguyenminh@phenikaa-uni.edu.vn



March 09, 2025

[*Original working draft v1 / Un-peer-reviewed*]

"No utility compares to that of escaping an illicit dream, alive and better off, still intact!"

— "The Philosophy of Awakening"; Wild Wise Weird (2024)

Abstract

The reproducibility crisis in social sciences has revealed significant weaknesses in conventional research practices, including selective publication, questionable statistical methods, and opaque peer review processes. This paper introduces Granular Interaction Thinking Theory (GITT) as a novel framework for understanding the plausibility of scientific findings, conceptualizing knowledge validation as a structured entropy-reduction process. Within this framework, open science practices—such as open data, open review, and open dialogue—initially increase informational entropy by exposing inconsistencies. However, through iterative refinement, they ultimately enhance the robustness and plausibility of scientific knowledge. To systematically assess a study's plausibility, we propose the Scientific Plausibility Index (SPI)—an entropy-based metric that integrates data and method transparency, replication success rates, and community engagement. Additionally, leveraging artificial intelligence (AI) and natural language processing (NLP), a dynamic plausibility-tracking system could be developed to detect unreliable claims early and accelerate scientific self-correction. This shift—from closed, one-time assessments to continuous, community-driven evaluation models—ensures that scientific knowledge remains rigorously tested and updated. We advocate for shifting from closed, one-time assessments to continuous, community-driven evaluation models, ensuring that scientific knowledge remains rigorously tested and updated. Ultimately, aligning incentives with entropy-reducing mechanisms—such as rewarding replication efforts and fostering open discourse—can cultivate a research culture that prioritizes robustness, transparency, and cumulative knowledge-building.

Keywords: Granular Interaction Thinking Theory (GITT); informational entropy; knowledge plausibility; scientific credibility; entropy-reduction process

1. Introduction

In recent years, confidence in social science findings has been eroded by widespread reproducibility failures. Numerous high-profile results have failed to replicate under scrutiny, raising concerns about the credibility of published research (Camerer et al., 2018; Open Science Collaboration, 2015). A 2016 survey of 1,576 scientists published in Nature found that over 70% had been unable to reproduce another researcher's results, and more than half had failed to replicate their own experiments (Baker, 2016a).

Several conventional research practices have contributed to this crisis. Questionable research practices, such as *p*-hacking (i.e., manipulating statistical analyses to achieve significance), HARKing (i.e., hypothesizing after results are known), and the selective publication of positive results have long gone unchecked in many fields (Baker, 2016a; Vuong et al., 2019). This has led to an abundance of statistically significant but fragile findings that fail to withstand replication. Meanwhile, the traditional peer review system, often operating behind closed doors, has struggled to filter out irreproducible claims. With limited transparency, errors and biases remain hidden, allowing spurious conclusions to propagate as accepted knowledge.

The net effect is a body of research with high entropy, where findings appear disorderly and unpredictable upon replication attempts (Camerer et al., 2018; Open Science Collaboration, 2015).

The severity of this issue has been demonstrated through large-scale replication projects. For example, an ambitious effort to replicate 21 major social-behavioral experiments published in *Nature* and *Science* between 2010 and 2015 succeeded in only 13 cases (62%), with replication effect sizes averaging 50% of the originals (Camerer et al., 2018). These outcomes highlight a credibility gap: if two studies of the same phenomenon produce widely divergent results, which—if either—reflects social reality? This epistemic uncertainty reflects a high-entropy system, where disorder dominates, and findings lack plausibility and reproducibility.

Solving the plausibility crisis in social science requires more than procedural reforms—it demands a theoretical framework that explains how reliable knowledge emerges from the research process. Here, we introduce Granular Interaction Thinking Theory (GITT) as a lens for understanding and systematically improving scientific plausibility (Vuong & Nguyen, 2024b, 2024c). Originally developed from quantum mechanics (Hertog, 2023; Rovelli, 2018), Shannon's Information Theory (Shannon, 1948), and Mindsponge Theory (Vuong, 2023), GITT explores how value creation in complex systems arises through dynamic, multi-state interactions among numerous information units. Previously, GITT has been applied to analyze the advantages and limitations of rejection mechanisms within the knowledge production system (Vuong & Nguyen, 2024b).

From the GITT perspective, the publishing, peer-review, and rejection systems function as mechanisms to reduce entropy in the scientific knowledge system (Vuong et al., 2024; Vuong & Nguyen, 2024b). Despite the rise of alternative dissemination channels (e.g., preprint repositories), scientific journals remain the dominant mode of knowledge validation. This reflects a community consensus that structured peer review adds value by filtering and refining information before publication. However, the current system is not infallible—editors and reviewers operate under cognitive constraints, subjectivity biases, and time limitations, allowing some unreliable findings to pass through. Once published, such results gain credibility by default, increasing their likelihood of being cited, stored, and reused in subsequent research. When these flawed findings are later debunked, the broader trust in published knowledge erodes, heightening uncertainty and amplifying epistemic entropy (Moodie, 2019).

Vuong (2017) previously argued that open science practices—including open data, open review, and open dialogue—can enhance the plausibility of social science findings. Grounded in GITT, we expand on this idea, proposing that while open science initially increases entropy, it facilitates the granular interactions necessary to eventually reduce uncertainty and improve plausibility robustly. This theoretical framework leads to an important forward-looking question:

- Can we develop entropy-based metrics to assess the plausibility of scientific findings?

If entropy serves as an indicator for uncertainty, then key factors such as data accessibility, reproducibility rates, and the degree of community scrutiny might be quantified into an Scientific Plausibility Index that indicates a study's credibility. By articulating this possibility, we set the stage for an interdisciplinary approach to diagnosing and mitigating the plausibility crisis in social science.

2. Open dialogue, open data, and open review through the lens of GITT

Granular Interaction Thinking Theory (GITT) posits that macro-level order emerges from structured micro-level interactions (Vuong & Nguyen, 2024a, 2024c). Drawing from quantum mechanics and information theory, GITT suggests that stable systems—whether markets, cognitive processes, or bodies of scientific knowledge—are built upon granular interactions at their smallest scales. In physics, macroscopic reality is shaped by interactions among quantum particles; Wheeler (2002)'s concept of "it from bit" encapsulates the idea that order arises from fundamental units of information.

GITT extends this principle to human cognition and scientific knowledge production, proposing that psychological and epistemic processes function analogously to quantum systems with three key characteristics (Rovelli, 2018; Vuong & Nguyen, 2024b):

- Granularity: Information, including human cognition and knowledge systems, is finite and discrete.
- Relationality: Cognitive and epistemic processes arise through interactions between existing knowledge and newly acquired information.
- Indeterminacy: Future knowledge states are probabilistic rather than deterministic, making scientific knowledge an inherently evolving entity.

Given the finite capacity of any individual to process information, knowledge production is a dynamic, multi-state process requiring contributions from multiple individuals. Insights from earlier stages of knowledge production (State 1) serve as resources for subsequent developments (State 2). This iterative process involves interactions among new observations, theoretical formulations, and accumulated knowledge.

Because scientific findings are not unequivocally determined by past research, maximizing the probability that valid knowledge transitions from State 1 to State 2 is crucial for maintaining an effective knowledge system. Editors and reviewers play a key role in this process, assessing a study's plausibility—the extent to which a claim, hypothesis, or finding appears credible based on available information (e.g., data, theory, methods, logic). When a study passes peer review, it is effectively deemed to have achieved sufficient plausibility for integration into the next stage of knowledge production.

Despite the importance of editorial and peer-review processes, their accuracy is constrained by inherent limitations. One primary concern is subjectivity—the notion that peer review is entirely objective is a common misconception. Editors and reviewers, as humans, are inevitably influenced by personal biases, disciplinary perspectives, and cognitive constraints

(Smith, 2006). Consequently, they may be more inclined to accept research aligned with their knowledge pool and worldviews while rejecting contradictory findings (Vuong, 2023; Vuong et al., 2025). Beyond subjectivity biases, capacity constraints also pose a challenge. Editors and reviewers operate under time, energy, and resource limitations, often lacking the tools or methods necessary to fully assess the plausibility of a study beyond the contents of the manuscript. The "publish or perish" culture exacerbates these issues by motivating unethical practices such as peer-review manipulation and fraudulent reviewer accounts (Kulkarni, 2016), further complicating the credibility evaluation process.

According to GITT, determining the plausibility of a piece of knowledge requires insights generated from its interactions with existing knowledge, including prior data, established theories, and methodological critiques. The conventional publishing system represents a low-density filtering process, which may allow unreliable claims to slip through due to limited interactions with knowledge of few editors and reviewers. In such a scenario, the responsibility of assessing plausibility should extend beyond editors and reviewers to include broader community engagement. A larger number of informed participants in the evaluation process will increase the likelihood of accurately assessing a study's credibility.

In particular, a new published finding in social science—such as an experimental result or a survey correlation—initially has certain plausibility as it has passed through the peer review process. Nevertheless, such plausibility is subject to change through a series of post-publication interactions, including peers' evaluation, evidence using new data, independent replication attempts, robustness checks, theoretical scrutiny, practical applications, etc. Through these interactions, the scholarly community collectively determines the level of trust that should be placed in the finding.

This dynamic plausibility assessment process underscores the importance of open, structured interactions in knowledge production. Open data, one of the fundamental pillars of open science, plays a key role in this process, as it helps make the information on which the study is grounded accessible for scrutiny (MacMillan, 2014; Ramachandran et al., 2021). When datasets, analysis code, and materials are publicly accessible, any researcher can serve as a potential reviewer and replicator, significantly increasing the micro-level interactions that a finding undergoes. Rather than relying solely on authors', editors', and reviewers' credibility alone, peers can independently verify results, recalculate statistics, test alternative models, identify anomalies, and attempt reproductions, thereby narrowing the space of possible truths. Empirical evidence indicates that studies with publicly shared datasets tend to exhibit greater credibility, while those where data remain inaccessible are more prone to statistical inconsistencies and weaker evidential support (Wicherts et al., 2011).

Similarly, open peer review also plays a crucial role in facilitating post-publication plausibility assessment (Ross-Hellauer, 2017). Traditional closed peer review limits evaluation to a small set of binary decisions (accept/reject), offering little support for further refinement through post-publication discourse. By contrast, publishing review reports alongside manuscripts can help foster broader community engagement and creates a more iterative and dynamic knowledge validation process, where later peers' evaluation can take into account the former evaluations of editors and reviewers (Ross-Hellauer & Görögh, 2019). Open community dialogue further facilitates this process by enabling continuous scholarly engagement through forums, conferences, social media, and post-publication peer review (Nosek et al., 2015; Vuong, 2017). Rather than treating publication as the final stage, open dialogue frames it as the beginning of an interactive, post-publication validation process. The resulting scholarly discourse generates meaningful interactions that enhance insights, reduce uncertainty, and improve the plausibility of research findings.

However, it should be noted that this process is not linear—greater transparency initially increases entropy in the knowledge system by possibly enabling a surge of conflicting information into scientific discourse. Such entropy (i.e., uncertainty or disorder) can be reflected through the formula of Shannon (1948) as follows:

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

The entropy (missing information or uncertainty) of a random variable X with possible outcomes $\{x_1, x_2, \dots, x_n\}$ and corresponding probabilities $\{P(x_1), P(x_2), \dots, P(x_n)\}$ is represented by $H(X)$. $P(x_i)$ is the probability of the outcome x_i. Each probability $P(x_i)$ represents how likely each outcome x_i is to occur. Applied to scientific knowledge, entropy provides a quantitative metaphor for plausibility.

- High entropy ($H(X)$ is large) ⟶ Low plausibility: A study's outcome is highly uncertain, meaning no existing evidence can definitely validate or reject the knowledge, and future research could confirm, refute, or significantly alter its conclusions.
- Low entropy ($H(X)$ is small) ⟶ High plausibility: A study is well-validated, with new data, evidence, and explanations likely to align with existing findings.

The level of informational entropy of a study depends on the its credibility status (e.g., 'highly credible', 'moderately credible, 'neutral', and 'not credible'), x_i, and $P(x_i)$ probability that x_i happen. The credibility statuses and their probability to happen are determined by subsequent peers' evaluation, new data, independent replication attempts, robustness checks, theoretical scrutiny, and methodological evaluations supporting or contradicting the study X. The iterative plausibility assessment process aligns with a Bayesian perspective, where prior beliefs about a claim are continuously updated as new evidence emerges (Vuong et al., 2022).

Specifically, after study X is published, its claim is considered to have high plausibility (i.e., low-entropy state), as its results are deemed credible by the editors and reviewers, and there are no other information indicating otherwise. However, when the post-publication plausibility assessment process continue, three main scenarios will happen:

- Scenario 1: If a study's findings are consistently replicated and supported by subsequent evidence, peers' evaluation, the probability of its validity approaches 1, and entropy trends toward zero (greater certainty).
- Scenario 2: If negative evaluations, evidence using new data contradicting the study's results, failed replications, and theoretical or methodological critiques arise, the informational entropy $H(X)$ of study will rise rapidly as new credibility statuses (e.g., 'moderately credible', 'not credible') of the study's claim starts to emerge, reflecting the growing uncertainty surrounding the claim.
- Scenario 3: If overwhelming contradictory evidence emerges, the scientific community may totally or partially reject the claim, or in some cases retract the study if it is involved in unethical practices. In such cases, the probability of 'not credible' of the study X will approximate 1, ultimately reducing entropy after an initial increase.

When the third scenario happens, it mirrors a physical system in a metastable state, where certain claims appear valid but later collapse when rigorous verification exposes flaws. The Fredrickson-Losada positivity ratio theory exemplifies this metastable knowledge state well. Fredrickson and Losada's (2005) positivity ratio theory, which proposed a precise 2.9 tipping-point ratio of positive to negative emotions for human flourishing (Oransky, 2013). Despite widespread citations, the model rested on fundamentally flawed mathematics. The claim remained unchecked for years, illustrating a temporary low-entropy state. However, until Brown et al. (2013) deconstructed the model, showing "no theoretical or empirical justification for the use of differential equations drawn from fluid dynamics, a subfield of physics, to describe changes in human emotions over time." While the paper of Brown et al. (2013) initially increased the entropy (uncertainty) of the positivity ratio by showing that it is unfounded, the 'credible status' of the ratio shifted rapidly to the 'not credible status' as the modelling element in the Fredrickson and Losada's (2005) paper was partially withdrawn (Oransky, 2013). Nevertheless, the increasing disorder within the knowledge system induced by the withdraw of Fredrickson and Losada's (2005) ratio is tremendous, as studies based on Fredrickson and Losada (2005) would fall under scrutiny and attributed with 'not credible' status, requiring more validation in the future to effectively reduce the uncertainty.

In this sense, the rising disorder induced by open science practices, like open data, open review, and open dialogue, is necessary, as they help accelerate the plausibility assessment process. As independent researchers engage in validation efforts, replication studies, and reanalysis, unreliable claims are identified and corrected (including retractions where necessary), guiding scientific knowledge toward a lower-entropy state characterized by greater reliability and coherence. The faster the scientific achieve such a robust low-entropy state, the less the negative impacts caused by the collapse of metastable knowledge state.

In general, from the GITT perspective, the plausibility of scientific findings emerges as a dynamic property shaped by continuous community interactions over time. Open science practices, e.g., open dialogue, open data, and open dialogue, function as a long-term entropy-reduction mechanism for the knowledge's plausibility status by accelerating and structuring information interactions. While increased transparency initially introduces greater variability and uncertainty, open science provides the necessary conditions for filtering out errors and inconsistencies quickly, reducing the damages inflicted by unreliable knowledge, and ultimately fostering a more stable, credible, and verifiable body of knowledge.

3. Rethinking Peer Review as a Granular Interaction Process

The traditional peer review system requires reform through the lens of Granular Interaction Thinking Theory (GITT). Currently, peer review functions as a one-time gatekeeping event—studies are evaluated by (typically) two or three anonymous reviewers, and if they pass this limited scrutiny, they enter the literature with an implicit seal of validity. However, this model has several shortcomings: minimal interactions, lack of transparency, and weak incentives for rigorous critique.

First, the limited number of reviewers means that each manuscript undergoes only a few interactions before publication. Significant flaws can go undetected simply because the probability of available reviewers catching every issue is low—especially in interdisciplinary research, where gaps in expertise may lead to overlooked errors. The case of the Fredrickson & Losada positivity-ratio paper is a prime example. None of the reviewers had sufficient background in nonlinear dynamics to scrutinize Losada's equations, allowing a deeply flawed mathematical model to pass peer review, be published in a prestigious journal, and navigate hundreds of subsequent studies Brown et al. (2013). GITT theory suggests that increasing both the number and diversity of review interactions would enhance amount and diversity of knowledge (e.g., existing knowledge, established theories, methodological critiques, etc.) that the study has to interact with before being published, eventually increasing the the quality of published research.

In practice, this calls for a shift toward open peer review. Journals and publishing platforms could expand the review process by soliciting broader community feedback, for example, by sharing submissions as preprints and inviting public commentary. Some journals are already experimenting with iterative, community-driven review models—such as *eLife*'s "publish, then review" approach (Eisen et al., 2020) and *Peer Community In*'s public evaluation of preprints (https://peercommunityin.org/). These models reframe peer review not as a one-time judgment but as an ongoing process of refinement. Moreover, open peer review also addresses the system's opacity. Under the traditional model, readers have no visibility into how rigorously a paper was vetted; a weakly reviewed study and a thoroughly scrutinized one appear identical in print. By making reviewer feedback and discussions publicly accessible, open peer review introduces an additional layer of accountability (Ross-Hellauer & Görögh, 2019; Ross-Hellauer, 2017). Reviewers may be more diligent, knowing their critiques are part of the scientific record, and authors can no longer conceal inconvenient criticisms from readers.

Open peer review also enables post-publication readers to contribute, fostering a dynamic scholarly dialogue. If an issue is raised but not fully resolved during the initial review, later commentators can build upon it. Platforms like *PubPeer* have demonstrated the value of post-publication discourse, yet traditional journals have often resisted or ignored such input. Bridging this gap—such as by formally integrating PubPeer comments into journal websites—would move the system toward a “living” peer review model, where published papers remain open to ongoing scrutiny and refinement interactions (Oransky, 2024).

Another critical reform is incentivizing peer review and replication as core scholarly activities. The current academic reward structure prioritizes novel publications over the careful examination of existing work (Horta & Jung, 2024). This imbalance has led to an oversupply of submitted papers—many of questionable quality—and a shortage of qualified, motivated reviewers. Addressing this requires systemic changes, such as recognizing peer review contributions in academic evaluations, allocating dedicated funding for replication studies, and creating journals or recognition badges for “verified results.” If reviewers and replicators are acknowledged as making contributions as vital as those of authors, the volume and quality of review interactions will increase. Each published finding could then be accompanied by independent confirmatory reports or formal commentaries, significantly enhancing confidence in the research.

Incentivization aligns the entropy-reduction process with researchers’ self-interest, ensuring that essential interactions—rigorous reviews and replication attempts—occur. Instead of viewing peer review as a single gate, it should be restructured into a multi-node network, where papers accumulate credibility through continuous engagement with reviewers, discussants, and replicators. The role of journals and institutions should shift toward facilitating and documenting this process. Some steps are straightforward: adopting open review policies, integrating post-publication commentary, and mandating open data and methods as a prerequisite for review. Other steps require cultural shifts, such as encouraging researchers to publish critical commentaries without fear of gatekeeping. Platforms like *PubPeer* and *F1000Research* are making progress in this regard by normalizing public critique.

Ultimately, reframing peer evaluation as an iterative, community-wide dialogue—rather than a closed-door process controlled by a select few—will significantly and robustly reduce the entropy of the scientific knowledge base. By identifying errors earlier and converging more rapidly on reliable findings, this approach ensures a more rigorous and self-correcting research ecosystem.

4. The Potential of Developing Entropy-Based Metrics for Research Creadibility

A key implication of this framework is the potential to develop quantitative metrics for assessing the “entropy”, or plausibility, of a research finding or body of work. In other words, we can ask: How much uncertainty surrounds this result, and to what extent has the scientific community imposed certainty through validation? Traditional metrics, such as citation counts or journal impact factors, do not directly capture a study’s plausibility or credibility. An entropy-based approach, however, would focus on indicators of transparency, reproducibility, and verification.

To this end, we propose a Scientific Plausibility Index (SPI), which incorporates at least three key components into a continuous, objective measure: 1) Data Openness, 2) External Validation, and 3) Community Engagement.

Data Openness – Measured by whether data and code are publicly accessible, the completeness of documentation, and the extent to which materials are reusable (Vuong et al., 2024). Studies with fully open data and code would score high, indicating robust low entropy, as others can readily verify the work. Meanwhile, those with proprietary or undisclosed data would score low, indicating high entropy, as critical information is missing. Empirical evidence supports the importance of this factor. Wicherts et al. (2011) found that papers with readily shared data tended to have fewer statistical errors. Recognizing this, funders and journals have started prioritizing data transparency—the National Institutes of Health (NIH), for example, now mandates data management plans, and some journals award badges for open data (Baker, 2016b; National Institutes of Health, 2023).

External Validation – This measures the extent to which a finding has been independently reproduced or verified. Key indicators include the number and outcomes of direct replications, whether the finding has been tested in preregistered multi-team studies, and whether it has been confirmed using different datasets or methods. Resources such as the Psychology Reproducibility Project and systematic reviews could provide empirical inputs. A high score in this category indicates multiple successful replications, indicating robust low entropy—greater plausibility, whereas a low score reflects either a lack of replication attempts or predominantly failed replications, indicating high entropy—greater uncertainty or potential falsehood.

Community Engagement – This component measures the level of scrutiny and discussion a work has received—essentially, the extent of "granular interaction" it has undergone. Key metrics may include the number of citations classified as confirmatory or disputational (e.g., using tools like scite.ai that categorize citations based on supporting or contradicting evidence), mentions in post-publication review forums, inclusion in policy or practice as a form of external validation, and overall visibility within the field. A paper that is widely mentioned and analyzed—particularly tend to undergo more entropy-reducing iterations than one that is published and largely ignored. However, while high engagement does not inherently confirm validity (as even debunked ideas can attract widespread discussion), low engagement leaves uncertainty high. By incentivizing researchers to invite scrutiny and encouraging journals to facilitate open commentary, this component promotes a more transparent and self-correcting scientific process (Vuong, 2020). Ultimately, such practices would contribute to a higher SPI score, reinforcing the plausibility of published findings.

Calculating such an index in practice would require aggregating data from multiple sources, including open data repositories, replication databases, and citation indices. However, in the era of digital scholarship, this is becoming increasingly feasible. Defense Advanced Research

Projects Agency (n.d.)'s SCORE program (Systematizing Confidence in Open Research and Evidence) has already explored automated approaches to predict replicability, demonstrating the potential of machine learning and knowledge graphs in assessing research reliability. The SCORE team developed algorithms that generated a "confidence score" for social science papers by integrating various factors: sample size, p-values, prior literature, and network relationships (e.g., the research lab behind the study and its connections to established theories) (Cohen, 2022). Their findings showed that combining micro-level features (statistical and textual details of a paper) with macro-level features (its position within the broader scientific network) produced the most accurate predictions (Cohen, 2022). This aligns with GITT, which emphasizes both localized information processing and emergent system-wide patterns in evaluating credibility. A formal Scientific Plausibility Index (SPI) could build on such work, providing a quantitative credibility measure that journals, funders, and readers could use—similar to how impact factors, despite their limitations, serve as a proxy for prestige.

Adopting entropy-based metrics could also reshape funding and policy priorities. Research programs demonstrating low entropy—such as those committed to open science practices and rigorous validation—could be prioritized for funding. Similarly, hiring and promotion criteria could incorporate SPI-type metrics as evidence of research quality, reducing the current overemphasis on sheer publication volume. By sending a clear message that a low-entropy research portfolio—one rich in transparency, scrutiny, and confirmed discoveries—is highly valued, such a system would create a positive feedback loop, reinforcing best practices and fostering a more reliable, self-correcting scientific culture.

5. Future Directions: Toward a Dynamic Scientific Plausibility Index

Looking ahead, we envision expanding these ideas into a dynamic Scientific Plausibility Index—functioning much like a stock ticker for scientific claims. As new evidence emerges—whether a replication succeeds, a major critique is published, or key data are released—the plausibility "value" of a finding would dynamically adjust (Vuong et al., 2025). Implementing this vision could leverage advances in artificial intelligence (AI) and natural language processing (NLP) to monitor the literature in real time. AI systems could be trained to scan new papers and preprints, assess their impact on a particular claim, and update an entropy-based credibility score accordingly. This process would resemble weather forecasting, where models continuously integrate new data to refine predictions—except here, the forecast pertains to the likelihood of a claim being valid.

Some early efforts in this direction already exist. The DARPA SCORE project has produced machine-predicted confidence scores for social science claims, while platforms like Curate Science aggregate replication results across multiple studies (https://curatescience.org/). Building on these, one could imagine a centralized dashboard tracking credibility scores for key claims in the social sciences. Such a system might display:

Claim X – Credibility: 78% (up from 75% last month)

Latest evidence: [+] Successful replication (Wang et al., 2025); [-] Data inconsistencies (Smith, 2025)

This tool would not only provide researchers and policymakers with real-time credibility assessments but also highlight high-entropy claims—those with unresolved uncertainties or limited transparency—helping to prioritize further testing and data sharing.

Beyond technological advances, cultural and institutional changes are needed to fully realize an entropy-reduction framework. A major challenge is the low prestige of replication studies and null results. In the current system, a young researcher who attempts to replicate a famous result—and finds it does not hold—may risk harming their career. To change this, the scientific community need to 1) elevate the status of replication research, possibly through dedicated high-impact journals focused on replications, and 2) normalize the publication of null findings and failed replications, ensuring they are treated as valuable contributions rather than inconvenient setbacks.

Some disciplines are already making progress. The Registered Reports initiative commits journals to publishing studies based on methodological rigor rather than outcome significance, reducing bias and encouraging more balanced scientific records (Soderberg et al., 2021). If widely adopted in the social sciences, this practice could help build a more orderly body of evidence—one that does not simply consist of a stack of significant findings while null results remain hidden.

Finally, education and training play a crucial role in shaping a scientific culture that prioritizes rigor, transparency, and cumulative knowledge-building. Introducing informational entropy and open science principles in research methodology courses could help emerging social scientists internalize these values early in their careers. If students are taught to view research not as a one-off result but as part of a larger knowledge network that must be continuously refined and pruned, they may develop a stronger inclination toward collaboration, replication, and data sharing. Simple demonstrations—such as showing how uncertainty in results decreases as sample size or replications increase—can make the concept of entropy reduction tangible. Over time, this could foster a mindset where scientists instinctively ask: “How can I reduce the entropy of this finding?” In other words, how can a study be designed and its results disseminated in a way that maximizes certainty, verifiability, and long-term reliability?

The future of credible social science will likely be driven by both technological innovations and cultural shifts in scientific practice. Developing metrics and tools that track and communicate entropy levels in research findings can help identify weak points early, directing resources toward clarifying uncertain claims rather than perpetuating them. One practical policy shift could be requiring funding proposals to include an “entropy impact statement”—alongside the usual pratical and theoretical implication section—explicitly addressing how the research will enhance clarity and reduce uncertainty (or, if poorly designed, risk adding confusion). Aligning incentives in this way reinforces the core goal of a self-correcting, cumulative science that justifies public trust (Vuong, 2018).

After all, the true aim of open science is not just openness for its own sake, but improved reliability—ensuring that the knowledge we produce forms a coherent, reproducible understanding of social phenomena, rather than a fragmented collection of unverified claims.

References

Baker, M. (2016a). 1,500 scientists lift the lid on reproducibility. *Nature*, *533*, 452-454. https://doi.org/10.1038/533452a

Baker, M. (2016b). Digital badges motivate scientists to share data. *Nature*, *10*, 19907. https://doi.org/10.1038/nature.2016.19907

Brown, N. J., Sokal, A. D., & Friedman, H. L. (2013). The complex dynamics of wishful thinking: the critical positivity ratio. *American Psychologist*, *68*(9), 801–813. https://doi.org/10.1037/a0032850

Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T.-H., Huber, J., Johannesson, M., . . . Pfeiffer, T. (2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nature Human Behaviour*, *2*(9), 637-644. https://doi.org/10.1038/s41562-018-0399-z

Cohen, J. (2022). The reproducibility crisis in science – These researchers have a fix. https://viterbischool.usc.edu/news/2022/11/the-reproducibility-crisis-in-science-these-researchers-have-a-fix/

Defense Advanced Research Projects Agency. (n.d.). *SCORE: Systematizing Confidence in Open Research and Evidence*. https://www.darpa.mil/research/programs/systematizing-confidence-in-open-research-and-evidence

Eisen, M., Akhmanova, A., Behrens, T., Harper, D., Weigel, D., & Zaidi, M. (2020). Peer Review: Implementing a" publish, then review. *eLife*, *9*, e64910. https://doi.org/10.7554/eLife.64910

Hertog, T. (2023). *On the origin of time: Stephen Hawking's final theory*. Random House. https://www.google.com/books/edition/On_the_Origin_of_Time/lIBTEAAAQBAJ

Horta, H., & Jung, J. (2024). The crisis of peer review: Part of the evolution of science. *Higher Education Quarterly*, *78*(4), e12511. https://doi.org/10.1111/hequ.12511

Kulkarni, S. (2016). What causes peer review scams and how can they be prevented? *Learned Publishing*, *29*(3), 211-213. https://doi.org/10.1002/leap.1031

MacMillan, D. (2014). Data sharing and discovery: What librarians need to know. *The Journal of Academic Librarianship*, *40*(5), 541-549. https://doi.org/10.1016/j.acalib.2014.06.011

Moodie, G. (2019). Fudged research results erode people's trust in experts. *Nature Index*. https://www.nature.com/nature-index/news/fudged-research-results-erode-people-trust-experts-research-scientists

National Institutes of Health. (2023). *Final NIH Policy for Data Management and Sharing*. National Institutes of Health Retrieved from https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-013.html

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., . . . Christensen, G. (2015). Promoting an open research culture. *Science*, *348*(6242), 1422-1425. https://doi.org/10.1126/science.aab2374

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716. https://doi.org/10.1126/science.aac4716

Oransky, I. (2013). Fredrickson-Losada "positivity ratio" paper partially withdrawn. *Retraction Watch*. https://retractionwatch.com/2013/09/19/fredrickson-losada-positivity-ratio-paper-partially-withdrawn/

Oransky, I. (2024). 'The PubPeer conundrum:' One view of how universities can grapple with a 'waterfall of data integrity concerns'. *Retraction Watch*. https://retractionwatch.com/2024/09/04/the-pubpeer-conundrum-how-universities-can-grapple-with-challenges-to-research/

Ramachandran, R., Bugbee, K., & Murphy, K. (2021). From open data to open science. *Earth and Space Science*, *8*(5), e2020EA001562. https://doi.org/10.1029/2020EA001562

Ross-Hellauer, T., & Görögh, E. (2019). Guidelines for open peer review implementation. *Research Integrity and Peer Review*, *4*(1), 4. https://doi.org/10.1186/s41073-019-0063-9

Ross-Hellauer, T. J. F. (2017). What is open peer review? A systematic review. *F1000Research*, *6*, 588. https://doi.org/10.12688/f1000research.11369.2

Rovelli, C. (2018). *Reality is not what it seems: The journey to quantum gravity*. Penguin. https://www.google.com/books/edition/Reality_Is_Not_What_It_Seems/fsQiDAAAQBAJ

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*(3), 379-423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Smith, R. (2006). Peer review: a flawed process at the heart of science and journals. *Journal of the Royal Society of Medicine*, *99*(4), 178-182. https://doi.org/10.1177/014107680609900414

Soderberg, C. K., Errington, T. M., Schiavone, S. R., Bottesini, J., Thorn, F. S., Vazire, S., . . . Nosek, B. A. (2021). Initial evidence of research quality of registered reports compared with the standard publishing model. *Nature Human Behaviour*, *5*(8), 990-997. https://doi.org/10.1038/s41562-021-01142-4

Vuong, Q.-H. (2017). Open data, open review and open dialogue in making social sciences plausible. *Scientific Data Updates*. https://doi.org/10.48550/arXiv.1712.04801

Vuong, Q.-H. (2018). The (ir)rational consideration of the cost of science in transition economies. *Nature Human Behaviour*, *2*, 5. https://doi.org/10.1038/s41562-017-0281-4

Vuong, Q.-H. (2020). Reform retractions to make them more transparent. *Nature*, *582*(7811), 149. https://doi.org/10.1038/d41586-020-01694-x

Vuong, Q.-H. (2023). *Mindsponge theory*. Walter de Gruyter GmbH. https://www.amazon.com/dp/B0C3WHZ2B3/

Vuong, Q.-H. (2024). *Wild Wise Weird*. AISDL. https://www.amazon.com/dp/B0BG2NNHY6

Vuong, Q.-H., Ho, T. M., & La, P. V. (2019). 'Stargazing'and p-hacking behaviours in social sciences: some insights from a developing country. *European Science Editing*, *45*(2), 54-55. https://europeanscienceediting.org.uk/articles/stargazing-and-p-hacking-behaviours-in-social-sciences-some-insights-from-a-developing-country/

Vuong, Q.-H., La, V.-P., & Nguyen, M.-H. (2024). Critical remarks on current practices of data article publishing: issues, challenges, and recommendations. *Data Science and Informetrics*, In Press.

Vuong, Q.-H., La, V.-P., & Nguyen, M.-H. (2025). Informational entropy-based value formation: A new paradigm for a deeper understanding of value. https://philarchive.org/rec/VUOIEV

Vuong, Q.-H., & Nguyen, M.-H. (2024a). *Better economics for the Earth: A lesson from quantum and information theories*. AISDL. https://www.amazon.com/dp/B0D98L5K44/

Vuong, Q.-H., & Nguyen, M.-H. (2024b). Exploring the role of rejection in scholarly knowledge production: Insights from granular interaction thinking and information theory. *Learned Publishing*, e1636. https://doi.org/10.1002/leap.1636

Vuong, Q.-H., & Nguyen, M.-H. (2024c). Further on informational quanta, interactions, and entropy under the granular view of value formation. *The VMOST Journal of Social Sciences and Humanities*. https://doi.org/10.2139/ssrn.4922461

Vuong, Q.-H., Nguyen, M.-H., & La, V.-P. (2022). *The mindsponge and BMF analytics for innovative thinking in social sciences and humanities*. Walter de Gruyter GmbH. https://www.amazon.com/dp/B0C4ZK3M74/

Wheeler, J. A. (2002). Information, physics, quantum: The search for links. In A. Hey (Ed.), *Feynman and computation* (pp. 309-336).

Wicherts, J. M., Bakker, M., & Molenaar, D. (2011). Willingness to share research data is related to the strength of the evidence and the quality of reporting of statistical results. *PLOS ONE*, *6*(11), e26828. https://doi.org/10.1371/journal.pone.0026828