Advances in Social Sciences Research Journal – Vol. 12, No. 10
Publication Date: October 12, 2025
DOI:10.14738/assrj.1210.19469.
Zhang, M. C. (2025). Strong Institutional Objectivity: Designing Knowledge Systems Beyond Individual Bias and Power Asymmetry. *Advances in Social Sciences Research Journal, 12(10).* 57-69.

SERVICES FOR SCIENCE AND EDUCATION

Strong Institutional Objectivity: Designing Knowledge Systems Beyond Individual Bias and Power Asymmetry

Michael Che Zhang
The Bolles School, Jacksonville, Florida, USA

ABSTRACT

Recent developments in cognitive neuroscience, particularly predictive processing models, demonstrate that human perception operates through active inference rather than passive observation (Clark, 2016; Friston, 2010). This finding, combined with previous research on heuristics and biases, challenges the empirical viability of the traditionally conceived version of individual-level objectivity. (Kahneman, 2011; Hohwy, 2013). Simultaneously, digital systems exhibit measurable tendencies toward homophily reinforcement and selective exposure amplification that worsen the information environment (Pariser, 2011; Bakshy et al., 2015). Therefore, this paper develops a framework of strong institutional objectivity that synthesizes insights from predictive processing, social epistemology, and power analysis to reconceptualize objectivity as an emergent institutional property rather than an individual action.. Through comparative case analysis of scientific reform initiatives and algorithmic mediation systems, I identify specific institutional design features that demonstrably improve error detection and correction rates while addressing structural exclusions that compromise knowledge production. The analysis yields testable propositions about the interaction between procedural controls and participation structures in determining epistemic outcomes, providing both a theoretical synthesis and an empirically grounded model for designing knowledge systems that achieve reliability through institutional mechanisms.

INTRODUCTION

The reliability of human knowledge systems faces empirical challenges from two directions:

1. Human Bias
2. Flawed digital information systems

First, cognitive science has documented inherent biases that lead away from rational inference in human judgment, with predictive processing models suggesting these biases originate in the fundamental architecture of perception itself (Clark, 2016; Kahneman, 2011). Second, digital information technologies that transmit an increasing proportion of knowledge flow exhibit measurable tendencies toward reinforcing rather than correcting these biases (Vosoughi et al., 2018; Sunstein, 2009).

These empirical findings necessitate the reconceptualization of objectivity as an institutional property that emerges from specifically structured social processes, rather than an individual cognitive achievement. This paper develops a framework that integrates insights from cognitive science, social epistemology, and science studies to identify design principles for knowledge-producing institutions that can achieve reliable error correction despite operating through biased individual cognition.

The contribution is threefold: (1) synthesizing disparate literatures on cognitive bias, social epistemology, and institutional design into a unified analytical framework; (2) deriving testable propositions about the interaction between procedural controls and participation structures; (3) identifying specific, measurable indicators for assessing institutional epistemic performance.

LITERATURE REVIEW

Cognitive Architecture and Epistemic Limitations

Predictive Processing Models:

Contemporary neuroscience has converged on predictive processing as a unifying framework for understanding perception and cognition, with profound implications for empiricist epistemology (Clark, 2013, 2016; Friston, 2010; Hohwy, 2013). Under this model, observation is fundamentally inferential rather than passive: the brain maintains hierarchical generative models that actively construct percepts by predicting sensory input and updating these predictions based on prediction errors. This architecture directly challenges the empiricist assumption that careful observation can provide theory-neutral access to empirical reality.

If perception involves active model-based inference rather than passive recording, then observation becomes inherently theory-laden in a neurologically literal sense. What observers report seeing depends not only on external stimuli but also on their prior theoretical commitments, which are embedded in predictive models. Studies using ambiguous stimuli demonstrate this epistemological problem empirically: identical evidence yields radically different observational reports depending on observers' priors (Sterzer et al., 2008). The bistable perception of Necker cubes, binocular rivalry, and the 2015 dress color phenomenon illustrate how prior expectations can overwhelm sensory data in determining conscious experience (Lafer-Sousa et al., 2015).

Furthermore, neuroimaging evidence reveals the extent of this challenge for objective observation. BOLD signals in the visual cortex reflect not just stimulus properties but also prior expectations, with unexpected stimuli generating larger prediction error signals in hierarchically organized cortical areas (Kok et al., 2014). Most problematically for epistemic practice, under conditions of strong priors and degraded sensory input—precisely the conditions that characterize much scientific observation—the brain can generate percepts with minimal or absent corresponding stimuli (Powers et al., 2017). When theoretical commitments are strong and empirical evidence is ambiguous, individual observers are neurologically predisposed to "observe" what their models predict rather than what the data actually contain, undermining the possibility of theory-independent empirical constraint.

Heuristics and Biases Program:

The predictive architecture documented by neuroscience aligns with four decades of behavioral research on judgment and decision-making. The heuristics and biases program initiated by Kahneman and Tversky has identified over 100 systematic deviations from normative rationality, many of which can be understood as consequences of predictive processing (Kahneman, 2011; Gilovich et al., 2002).

Confirmation bias—the tendency to search for, interpret, and recall information that confirms prior beliefs—represents perhaps the most epistemically problematic bias. Meta-analyses reveal effect sizes ranging from d = 0.80 to 1.20 across various domains, with more pronounced effects for emotionally charged or identity-relevant beliefs (Nickerson, 1998). The bias manifests in multiple cognitive processes, including biased information search (Wason, 1960), biased interpretation of ambiguous evidence (Lord et al., 1979), and biased recall that favors confirming instances (Frost et al., 2015).

Additional biases with epistemic consequences include availability heuristic (overweighting easily recalled instances; Tversky & Kahneman, 1973), anchoring effects (insufficient adjustment from initial values; Epley & Gilovich, 2006), and base rate neglect (underweighting prior probabilities; Kahneman & Tversky, 1972). These biases exhibit remarkable consistency across cultures, albeit with variations in magnitude (Henrich et al., 2010).

Social Epistemology and Collective Knowledge Production

Falsificationism and Its Limits:

Karl Popper's falsificationist program represented an early attempt to address confirmation bias through methodological prescription (Popper, 1959). By making attempted refutation rather than confirmation the hallmark of the scientific method, falsification-ism elegantly converts the search for disconfirming evidence from a cognitive weakness into a methodological strength.

However, subsequent work revealed fundamental limitations. The Duhem-Quine thesis posits that hypotheses are never evaluated in isolation but rather in conjunction with auxiliary assumptions, rendering definitive falsification impossible (Duhem, 1906; Quine, 1951). Lakatos's methodology of scientific research programs demonstrated how scientists rationally protect core theories through modifications to auxiliary hypotheses, rather than abandoning theories when faced with anomalies (Lakatos, 1970).

Paradigms and Social Structure:

Thomas Kuhn's historical analysis revealed that scientific communities operate within paradigms—shared frameworks of theoretical commitments, methodological standards, and exemplary problems that enable normal science but create resistance to fundamental change (Kuhn, 1962). Paradigms are maintained through social mechanisms, including graduate training, peer review, and career incentives that reward work within established frameworks.
This social dimension implies that objectivity cannot be achieved solely through method. The same social structures that enable cumulative progress within paradigms can also perpetuate systematic biases and resist correction. Studies of peer review show reliability coefficients of only 0.23-0.34, indicating that acceptance decisions are influenced by factors beyond scientific quality (Bornmann et al., 2010).

Longino's Procedural Account:

Helen Longino provides the most developed account of how objectivity can emerge from social processes despite individual bias (Longino, 1990, 2002). Her model identifies four conditions necessary for a community to achieve objectivity through critical discourse:

First, there must be recognized venues for criticism—such as journals, conferences, and other forums — where critical evaluation occurs publicly. These venues must be genuinely accessible rather than merely formal. Analysis of publication patterns reveals that 87% of authors in top journals are affiliated with 20% of institutions, indicating limited accessibility in practice (Larivière et al., 2013).

Second, communities must show uptake of criticism by modifying theories and methods in response to valid objections. Citation analysis reveals that negative citations comprise only 2.4% of references, suggesting limited engagement with criticism (Catalini et al., 2015).

Third, there must be publicly recognized standards for evaluation. While allowing theoretical pluralism, communities need sufficient overlap in evaluative criteria to enable productive disagreement. Studies of interdisciplinary review show that divergent standards reduce funding success by 15-30% (Bromham et al., 2016).

Fourth, communities require tempered equality of intellectual authority. Extreme hierarchies enable dominant perspectives to ignore criticism from marginalized positions. Bibliometric analyses show that papers with famous last authors receive 5.8 times more citations independent of quality metrics (Sekara et al., 2018).

However, these procedural conditions were developed before the full implications of predictive processing and systematic bias research were understood. Given that individual observers are neurologically predisposed to perceive theory-consistent evidence and systematically ignore disconfirming information, Longino's framework faces a deeper challenge: even when critical venues exist and uptake occurs, the criticism itself may be filtered through the same biased cognitive architecture that generated the original claims. Suppose community members are subject to confirmation bias and motivated reasoning. In that case, criticism may be selectively attended to, interpreted in ways that preserve existing beliefs, or dismissed based on source rather than content. The procedural safeguards, while necessary, may be insufficient to overcome the systematic nature of cognitive bias documented in the heuristics and biases literature.

Power, Knowledge, and Epistemic Exclusion

Foucauldian Analysis:

Michel Foucault's genealogical method reveals that knowledge production always occurs within matrices of power that determine what can be said, who can speak authoritatively, and which questions are legitimate (Foucault, 1977, 1980). Scientific disciplines do not simply discover pre-existing truths but actively constitute their objects of study through specific techniques of observation, measurement, and classification.

The concept of "regimes of truth" captures how different historical periods and institutional contexts establish distinct criteria for what counts as true or false. These regimes operate through mechanisms that include funding priorities (determining which questions are investigated), publication practices (determining which findings become visible), and credentialing systems (determining who is considered an expert).

Empirical studies support Foucault's analysis. Industry funding increases the odds of pro-industry conclusions by a factor of 3.6 (OR = 3.60, 95% CI: 2.63-4.91) even when controlling for study quality (Lundh et al., 2017). Publication bias toward positive results means that 68% of null findings never enter the scientific record (Franco et al., 2014).

Foucault's analysis poses a fundamental challenge to procedural accounts like Longino's. While Longino assumes that properly structured critical discourse can overcome bias through rational evaluation, Foucault suggests that the very categories of 'rational' and 'objective' are themselves products of power relations rather than neutral epistemic standards. From a Foucauldian perspective, Longino's procedural safeguards may simply institutionalize existing power structures under the guise of objectivity—her 'publicly recognized standards' embed dominant perspectives, her 'venues for criticism' remain accessible primarily to credentialed experts, and her 'tempered equality' preserves hierarchies through the rhetoric of meritocracy.

However, this creates a productive tension rather than mere contradiction. Foucault's genealogical method reveals how power shapes knowledge, but offers limited guidance for designing better epistemic institutions. Longino's procedural framework provides concrete mechanisms for organizing critical discourse, but lacks adequate analysis of how power relations can subvert these procedures. A synthesis must acknowledge that procedures operate within power matrices while maintaining that some procedural arrangements are more conducive to error correction than others. The challenge becomes designing institutions that are both procedurally rigorous and reflexively attentive to their own power dynamics.

Standpoint Epistemology:

Standpoint theory, developed primarily within feminist epistemology, argues that social position affects epistemic position—what can be known and how it can be known (Harding, 1991, 2004). The theory makes two key claims: (1) knowledge is socially situated, and (2) marginalized social positions can yield epistemic advantages for understanding certain phenomena.

The strong version of standpoint theory posits that oppressed groups possess privileged access to knowledge about social relations, as they have a lesser investment in ideological justifications for existing arrangements. Empirical support is evident in cases where the inclusion of marginalized perspectives has led to the correction of scientific errors. The exclusion of women from cardiac research provides a paradigmatic example: studies conducted primarily on male subjects missed sex differences in symptom presentation, leading to systematic misdiagnosis in women (Pope et al., 2000).

Epistemic Injustice:

Miranda Fricker's work on epistemic injustice provides a framework for understanding how social power creates systematic distortions in knowledge through credibility dynamics (Fricker, 2007). With her work, she shows that testimonial injustice occurs when prejudice causes deflated credibility attributions. Experimental studies have shown that identical statements receive different credibility ratings depending on the speaker's demographics, with women and minorities facing credibility deficits of 10-25% (Moss-Racusin et al., 2012).

Hermeneutical injustice occurs when marginalized groups lack access to interpretive resources that enable them to articulate their experiences effectively. The development of concepts like "sexual harassment" and "microaggression" illustrates how conceptual resources shape what can be known and communicated about social reality.

Digital Mediation and Algorithmic Epistemology

Filter Bubbles and Echo Chambers:

Digital information systems create new dynamics in the production and circulation of knowledge. Personalization algorithms designed to maximize engagement can create filter bubbles—information environments that systematically exclude challenging perspectives (Pariser, 2011). However, empirical evidence is mixed. While some studies find decreased exposure diversity (Bakshy et al., 2015), others find that digital media increases incidental exposure to diverse sources (Fletcher & Nielsen, 2018).

The effects are contingent upon the network structure and user behavior. Users with homogeneous social networks experience greater ideological isolation, while those with diverse networks may encounter more varied content than through traditional media. Platform design choices significantly influence outcomes: chronological feeds show greater content diversity than algorithmically curated feeds (Levy, 2021).

Misinformation Dynamics:

False information spreads differently than accurate information in digital networks. Analysis of 126,000 stories on Twitter found that false news spreads six times faster than trustworthy news and reaches 20 times more people (Vosoughi et al., 2018). The differential is not explained by bot activity but appears to result from human psychology: false information tends to be more novel and emotionally arousing.

Correction attempts face additional challenges online. The continued influence effect—whereby false information continues to influence judgments even after correction—is amplified in digital environments where corrections may not reach all who saw original false claims (Lewandowsky et al., 2012). Backfire effects, where corrections strengthen false beliefs, occur primarily for politically motivated reasoning (Nyhan & Reifler, 2010).

THEORETICAL SYNTHESIS: STRONG INSTITUTIONAL OBJECTIVITY

Integration Framework

The reviewed literatures converge on several key insights:

1. Individual cognition operates through predictive models that create systematic biases
2. Social processes can either amplify or mitigate individual biases depending on structure
3. Power relations shape what questions get asked and whose answers count
4. Digital systems introduce novel dynamics that existing institutions are not designed to handle

Strong institutional objectivity integrates these insights through four design principles:

- **Cognitive Realism:** Accept predictive, biased cognition as a fixed constraint rather than a defect to eliminate. Design institutions that work with human psychology as it actually operates.

- **Procedural Systematization:** Implement formal mechanisms that make error detection and correction independent of individual virtue. These must target specific, empirically documented failure modes.
- **Power Reflexivity:** Build continuous analysis of power dynamics into institutional operations. Make exclusions visible and correctable.
- **Epistemic Pluralism:** Structure institutions to harness multiple partial perspectives rather than seeking a single, complete view. Diversity becomes instrumental to reliability.

Operational Model

The framework translates into specific institutional features:

- **Distributed Verification:** Multiple independent validation pathways prevent both intentional fraud and unintentional self-deception. The Cancer Reproducibility Project demonstrates this principle: parallel replications across laboratories revealed that only 11% of findings were completely reproduced (Errington et al., 2021).
- **Structured Disagreement:** Adversarial collaboration requires opposing theoretical camps to jointly specify decisive tests. Studies show this reduces disagreement persistence by 68% compared to traditional debate (Mellers et al., 2001).
- **Inclusive Participation:** Beyond demographic diversity, institutions must incorporate epistemically relevant stakeholder perspectives that hold genuine agenda-setting power. The Patient-Centered Outcomes Research Institute model shows 23% improvement in research relevance metrics when patients participate in study design (Frank et al., 2014).
- **Radical Transparency:** Complete openness, from funding to raw data, enables distributed error checking. Studies comparing transparent and opaque research find 49% higher error detection rates in transparent work (Wicherts et al., 2011).

EMPIRICAL CASES

Case A: Registered Reports in Psychology

Intervention Structure:

Registered Reports involve peer review before data collection based on introduction, hypotheses, methods, and analysis plans. Studies that pass review receive in-principle acceptance regardless of the results. As of 2023, 307 journals offer this format, up from 1 in 2013.

Mechanisms and Outcomes:

Registered Reports target specific failure modes:

- P-hacking (eliminated by pre-specification): Studies show researchers degrees of freedom can inflate false favorable rates from 5% to 61% (Simmons et al., 2011)
- Publication bias (reduced by result-independent acceptance): 61% of Registered Reports report null results versus 5-8% of traditional papers
- HARKing (prevented by timestamp): 39% of traditional papers show evidence of hypothesizing after results are known

Meta-analysis of 71 Registered Reports versus 152 traditional replications shows:

- Statistical power: 0.89 vs 0.65

- Effect size inflation: 1.07 vs 2.94
- Replication rate: 86% vs 36% (Scheel et al., 2021)

Power Dynamics:
Despite procedural improvements, Registered Reports show persistent power asymmetries:
- Submission burden disadvantages under-resourced labs
- 78% of published Registered Reports come from WEIRD populations
- Senior author demographics show no improvement over traditional publishing

Adding diversity requirements to Registered Report protocols could address these limitations. Strong institutional objectivity suggests that additional modifications are needed to address both cognitive and power limitations simultaneously. First, **mandatory adversarial review**: each Registered Report could require one reviewer explicitly tasked with identifying weaknesses in the proposed design, who would be compensated equally to supportive reviewers. This targets confirmation bias in peer review while creating career incentives for critical evaluation.

Second, **community stakeholder validation**: for research on human populations, affected communities could review study designs before acceptance, with veto power over research they deem extractive or harmful. This addresses both the WEIRD population bias and the broader issue of who determines legitimate research questions. Studies show that community-reviewed research produces findings that are 34% more likely to be implemented in practice, suggesting both epistemic and ethical benefits (Israel et al., 2012).

These modifications exemplify how procedural reforms must be coupled with power redistribution to achieve genuine epistemic improvement rather than merely technical compliance.

Case B: Algorithmic Content Curation
System Architecture:
Major platforms use engagement-optimization algorithms combining:
- Collaborative filtering (user-user similarity)
- Content-based filtering (item-item similarity)
- Deep learning models predicting engagement probability

Facebook's algorithm makes ~500,000 ranking decisions per user per day (Backstrom, 2016).

Measured Effects:
Empirical studies document:
- Exposure diversity: 13% reduction in cross-cutting content for median user (Bakshy et al., 2015)
- Emotion contagion: 0.07 SD increase in negative posts following negative feed manipulation (Kramer et al., 2014)
- Misinformation velocity: False news travels 6x faster than true news (Vosoughi et al., 2018)

- Polarization: 28% increase in affective polarization among heavy social media users (Levy, 2021)

Importantly, these documented harms stem from specific algorithmic design choices rather than inherent properties of digital systems. Current engagement optimization prioritizes immediate emotional arousal and return visits, metrics that systematically favor false and polarizing content. Alternative optimization targets—such as user-reported satisfaction with information quality, diverse source engagement, or constructive dialogue metrics—would yield different epistemic outcomes, as demonstrated by the intervention experiments below.

Intervention Experiments:
Platform modifications show measurable effects:

- Chronological feeds: 19% increase in content diversity
- Friction for sharing: 25% reduction in misinformation spread (WhatsApp)
- Downranking borderline content: 50% reduction in hate speech engagement
- Context labels: 11% reduction in sharing of labeled false content (Facebook, 2021; Reis et al., 2020)

DERIVED PROPOSITIONS

Based on the theoretical synthesis and empirical cases:

- **P1:** In engagement-optimized systems, exposure to epistemically relevant diversity decreases by 10-30% relative to chronological or editorial baselines, conditional on network homophily > 0.4.
- **P2:** Procedural interventions (e.g., Registered Reports) combined with participation mandates yield 20-40% greater reliability improvements than procedures alone, as measured by replication rates and effect size accuracy.
- **P3:** Institutions implementing ≥4 epistemic hygiene practices show 15-25% improvement in predictive accuracy and replication rates over 3-5 year periods, controlling for field and resource effects.

DESIGN IMPLICATIONS

Institutional Design Principles

Based on the analysis, institutions seeking epistemic reliability should implement design principles that target the specific failure modes identified:

- **Targeting Predictive Processing Biases:**
    1. **Systematic Adversarialism:** Institutionalized devil's advocacy and red-teaming counter confirmation bias by mandating engagement with disconfirming evidence. Implementation: 10% of review effort allocated to adversarial examination.
    2. **Iterative Verification:** Multiple independent replications address the tendency for strong priors to generate theory-consistent "observations." Implementation: Replication requirement for tenure decisions.
- **Addressing Power-Based Exclusions:**
    3. **Stakeholder Inclusion:** Affected populations participate in agenda-setting with veto power over extractive research, countering epistemic injustice and hermeneutical gaps. Implementation: Community advisory boards with budget authority.

4. **Transparent Provenance:** All knowledge claims must include traceable origins (methods, funding, conflicts), making power influences visible for correction. Implementation: Blockchain-based research registries.

- **Mitigating Algorithmic Amplification:**

5. **Incentive Alignment:** Reward structures that value error correction equally with discovery, countering engagement optimization that favors sensational over accurate content. Implementation: Citation metrics that weight corrections positively.

LIMITATIONS

Several limitations bound these findings:

1. **Field Specificity:** Replication dynamics vary across disciplines. Psychology findings may not generalize to physics or history.
2. **Measurement Challenges:** The epistemic quality of information because metrics can be gamed (Goodhart's Law).
3. **Resource Constraints:** Proposed interventions require significant resources, potentially exacerbating global inequalities.
4. **Platform Opacity:** Proprietary algorithms prevent complete evaluation of digital mediation effects.

CONCLUSION

This analysis demonstrates that achieving epistemic reliability under conditions of predictive cognition and digital mediation requires fundamental institutional redesign. Individual-level interventions, such as critical thinking training, while valuable, cannot overcome the systematic biases documented in cognitive science. Similarly, purely procedural reforms that ignore power dynamics fail to address how structural exclusions compromise knowledge production.

"Strong institutional objectivity provides a framework for designing knowledge systems that achieve reliability through the interaction of procedures, incentives, and participation structures. Unlike cognitive training approaches that attempt to eliminate individual bias, this framework works with predictive cognition as a fixed constraint, using institutional design to harness diverse biased perspectives rather than seeking unbiased individuals. Unlike purely procedural reforms that assume the rational uptake of criticism, it anticipates how power dynamics will subvert formal procedures and builds countermechanisms directly into the institutional architecture. The result is knowledge systems capable of systematic error correction even when populated by cognitively biased actors operating within power hierarchies—a form of collective epistemic reliability that emerges from institutional structure rather than individual virtue.

The empirical cases demonstrate that such designs can enhance replication rates, mitigate effect size inflation, and increase the detection of errors. However, implementation requires sustained effort across multiple institutional levels and continuous monitoring to prevent gaming and capture. Future research should experimentally test the derived propositions through randomized institutional interventions, develop more sophisticated measures of epistemic quality, and explore how the framework applies to emerging technologies, such as large language models and automated science. The goal is not perfect objectivity—an impossibility given cognitive constraints—but knowledge systems that fail better: detecting and correcting errors faster while expanding the range of perspectives that contribute to collective understanding.

References

American Heart Association. (2021). *Heart Attack Symptoms in Women*. https://www.heart.org/en/health-topics/heart-attack/warning-signs-of-a-heart-attack/heart-attack-symptoms-in-women

Anderson, E. (1995). Feminist Epistemology and Philosophy of Science. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*.

Anderson, E. (2004). *How Not to Criticize Feminist Epistemology: A Review of Scrutinizing Feminist Epistemology*. http://www-personal.umich.edu/~eandersn/hownotreview.html

Barnes, B. (1982). *T.S. Kuhn and Social Science*. Columbia University Press.

Bijker, W. E., Hughes, T. P., & Pinch, T. (Eds.). (1987). *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology*. MIT Press.

Bishop, M. A., & Trout, J. D. (2005). *Epistemology and the Psychology of Human Judgment*. Oxford University Press.

Borgman, C. L. (2015). *Big Data, Little Data, No Data: Scholarship in the Networked World*. MIT Press.

Boyd, R. N. (1983). On the Current Status of Scientific Realism. *Erkenntnis, 19*(1/3), 45-90.

Budapest Open Access Initiative. (2002). *Read the Budapest Open Access Initiative*. https://www.budapestopenaccessinitiative.org/read

Burnyeat, M. F. (1976). Protagoras and Self-Refutation in Plato's Theaetetus. *The Philosophical Review, 85*(2), 172-195.

Chambers, C. D. (2013). Registered Reports: A new publishing initiative at Cortex. *Cortex, 49*(3), 609–610.

Chambers, C. D., Feredoes, E., Muthukumaraswamy, S. D., & Etchells, P. J. (2014). Instead of 'playing the game' it is time to change the rules: Registered Reports at AIMS Neuroscience and beyond. *AIMS Neuroscience, 1*(1), 4-17.

Clark, A. (2016). *Surfing uncertainty: Prediction, action, and the embodied mind*. Oxford University Press.

Collins, P. H. (1986). Learning from the Outsider Within: The Sociological Significance of Black Feminist Thought. *Social Problems, 33*(6), S14-S32.

Foucault, M. (1977). *Discipline and punish: The birth of the prison* (A. Sheridan, Trans.). Pantheon Books.

Foucault, M. (1980). *Power/Knowledge: Selected Interviews and Other Writings, 1972-1977* (C. Gordon, Ed.). Pantheon Books.

Friedman, M. (1992). *Kant and the Exact Sciences*. Harvard University Press.

Fuller, S. (2003). *Kuhn vs. Popper: The Struggle for the Soul of Science*. Icon Books.

Green, A. (2021). *The Social Contexts of Intellectual Virtue: Knowledge as a Team Achievement*. Oxford University Press.

Grier, M. (2001). *Kant's Doctrine of Transcendental Illusion*. Cambridge University Press.

Harding, S. (1991). *Whose science? Whose knowledge? Thinking from women's lives*. Cornell University Press.

Heesen, R., & Bright, L. K. (2020). Is Peer Review a Good Idea? *The British Journal for the Philosophy of Science, 72*(3), 635-663.

Hohwy, J. (2013). *The predictive mind*. Oxford University Press.

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

Kant, I. (1998). *Critique of pure reason* (P. Guyer & A. W. Wood, Trans.). Cambridge University Press. (Original work published 1781/1787).

Kidd, I. J. (2016). Intellectual Humility, Confidence, and Argumentation. *Topoi, 35*(2), 395-405.

Kourany, J. A. (2010). *Philosophy of Science After Feminism*. Oxford University Press.

Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd ed.). University of Chicago Press.

Lakatos, I. (1970). Falsification and the Methodology of Scientific Research Programmes. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the Growth of Knowledge* (pp. 91-196). Cambridge University Press.

Leary, M. R., et al. (2017). Cognitive and Interpersonal Features of Intellectual Humility. *Personality and Social Psychology Bulletin, 43*(6), 793-813.

Longino, H. E. (1990). *Science as social knowledge: Values and objectivity in scientific inquiry*. Princeton University Press.

Longino, H. E. (2002). *The Fate of Knowledge*. Princeton University Press.

Mellers, B., Hertwig, R., & Kahneman, D. (2001). Do frequency representations eliminate conjunction effects? An exercise in adversarial collaboration. *Psychological Science, 12*(4), 269-275.

Mercier, H., & Sperber, D. (2017). *The Enigma of Reason*. Harvard University Press.

Munafò, M. R., et al. (2017). A manifesto for reproducible science. *Nature Human Behaviour, 1*(1), 0021.

National Academies of Sciences, Engineering, and Medicine. (2017). *Fostering Integrity in Research*. The National Academies Press.

Nielsen, M. (2012). *Reinventing Discovery: The New Era of Networked Science*. Princeton University Press.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2*(2), 175–220.

Nosek, B. A., Spies, J. R., & Motyl, M. (2012). Scientific Utopia: II. Restructuring Incentives and Practices to Promote Truth Over Publishability. *Perspectives on Psychological Science, 7*(6), 615-631.

Nyhan, B., & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior, 32*(2), 303–330.

O'Connor, C., & Weatherall, J. O. (2019). *The Misinformation Age: How False Beliefs Spread*. Yale University Press.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716.

Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. Penguin UK.

Pinnick, C. L., Koertge, N., & Almeder, R. F. (Eds.). (2003). *Scrutinizing Feminist Epistemology: An Examination of Gender in Science*. Rutgers University Press.

Plato. (2004). *Theaetetus* (J. McDowell, Trans.). Clarendon Press.

Popper, K. R. (1959). *The logic of scientific discovery*. Hutchinson.

Popper, K. R. (1963). *Conjectures and Refutations: The Growth of Scientific Knowledge*. Routledge & Kegan Paul.

Popper, K. R. (1972). *Objective Knowledge: An Evolutionary Approach*. Oxford University Press.

Putnam, H. (1975). *Mathematics, Matter and Method: Philosophical Papers, Vol. 1*. Cambridge University Press.

Rabinow, P. (Ed.). (1984). *The Foucault Reader*. Pantheon Books.

Rescher, N. (1997). *Objectivity: The Obligations of Impersonal Reason*. University of Notre Dame Press.

Roberts, R. C., & Wood, W. J. (2007). *Intellectual Virtues: An Essay in Regulative Epistemology*. Oxford University Press.

Rouse, J. (1996). *Engaging Science: How to Understand its Practices Philosophically*. Cornell University Press.

Ryan, S. (2018). Wisdom. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2018 ed.).

Schiebinger, L. (2001). *Has Feminism Changed Science?* Harvard University Press.

Siegel, H. (1987). *Relativism Refuted: A Critique of Contemporary Epistemological Relativism*. D. Reidel Publishing Company.

Siegel, S. (2017). The rationality of perception. In *The Nature of Perception* (pp. 624-649). Oxford University Press.

Solomon, M. (2001). *Social Empiricism*. MIT Press.

Sunstein, C. R. (2009). *Going to extremes: How like minds unite and divide*. AEI Press.

Tetlock, P. E., & Mitchell, G. (2009). Adversarial collaboration: A process for conducting constructive controversial research. In *Social and psychological bases of belief and disbelief* (pp. 199-228).

Thornton, S. (2023). Karl Popper. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Winter 2023 ed.).

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124–1131.

UNESCO. (2021). *UNESCO Recommendation on Open Science*.

Vaccarino, V., Parsons, L., Every, N. R., Barron, H. V., & Krumholz, H. M. (1999). Sex-based differences in early mortality after myocardial infarction. *New England Journal of Medicine, 341*(4), 217-225.

Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology, 12*(3), 129–140.

Wenger, N. K., Speroff, L., & Packard, B. (1993). Cardiovascular health and disease in women. *New England Journal of Medicine, 329*(4), 247-256.

Whitcomb, D., Battaly, H., Baehr, J., & Howard-Snyder, D. (2017). Intellectual Humility: Owning Our Limitations. *Philosophy and Phenomenological Research, 94*(3), 509-539.

Wylie, A. (2003). Why Standpoint Matters. In R. Figueroa & S. G. Harding (Eds.), *Science and Other Cultures: Issues in Philosophies of Science and Technology* (pp. 26-48). Routledge.

Young, K. (2004). The Body as a Template: The Case of the Male and Female Body in Medicine. *Feminism & Psychology, 14*(2), 315-319.