

# Against the “Meme Machine”: Normativity, Qualia and “Self”

**Yutong Zhang**

Department of Electrical and  
Computer Engineering, McGill  
University, Montreal, Quebec,  
Canada, H3A 0G4  
yutong.zhang3@mail.mcgill.ca

## Abstract:

This paper critiques the strong “meme theory” advanced by Susan Blackmore, which claims that all mental phenomena like consciousness, agency, and selfhood can be reduced to the replication and competition of memes. While memetics offers valuable insights into cultural evolution, its explanatory power becomes limited when extended to the philosophy of mind. Two main challenges have developed. First, the problem of normativity: humans evaluate ideas by standards of truth, rationality, and ethics, capacities that cannot be fully explained by replication success alone. Drawing on the notion of innate cognitive structures, the paper argues that evaluative reasoning reveals a layer of mental autonomy beyond memetic causation. Second, the problem of consciousness: eliminative interpretations within strong memetics fail to account for the irreducible reality of subjective experience (qualia). Even if memes construct a “selfplex” that produces an illusory sense of self, the persistence of phenomenological experience resists full reduction. The paper concludes that memetics remains insufficient as a complete theory of mind, despite its insights on cultural theory.

**Keywords:** Memetics, Meme Theory, Consciousness, Normativity, Qualia

## 1. Introduction

The concept of the “meme” was first introduced by Richard Dawkins in *The Selfish Gene* (1976) [1], as a cultural replicator analogous to the function of genes. This proposal sparked the field of “memetics,” which is a framework accounting for the cultural transmission and progression in (Darwinian-styled) evolutionary terms [2]. An earlier version of memetic theory often argues that all ideas, religions, cultural practices or artifacts were adopted and transmitted

through the mechanism of variation, competition and inheritance of memes, which shares the essential properties of genes: “high-fidelity, fecundity, and longevity” [1].

Years later, leading scholars like Susan Blackmore advanced what has come to be known as “strong meme theory”, most prominently in “*The Meme Machine*” (1999) [3]. Here, memes are not only seen as cultural replicators but also as autonomous “forces” shaping human identities. Blackmore describes hu-

mans as “meme machines” and argues that our perception of “self” is in fact a “selfplex” (a cluster of memes that creates the illusion of a unified and autonomous subject) [3]. In her view, not only self-identities and agency but also consciousness itself could be reducible, and therefore, what feels like a central experiencing “self” is just a construct generated through memetic processes [4].

This paper aims to critique the strong meme theory on two central limitations. Firstly, about normativity, which is the “capacity for human beings to evaluate beliefs and practices based on standards” (beyond the memetic fitness) [5]. Second, the reality of subjective experience (qualia) [6] cannot be branded as an illusion. Since such an illusion is still something it feels to have, this suggests that qualia cannot be dismissed outright. Overall, these gaps suggests that while memetics offers valuable insights into cultural progression and the shaping of selfhood, but it cannot exhaustively account for the nature of the mind; memes should be seen as parts of the story, not the whole.

## 2. The Problem of Normativity

### 2.1 Normative Elements and Evaluative Capacities

The most pressing challenge to strong meme theory would be the discussion of normativity, namely, the human capacity to evaluate ideas according to standards of truth, rationality, or ethics, rather than mere replication success. Yujian Zheng argues that memetics cannot provide a sufficient account of this evaluative dimension, since memes are thought to replicate regardless of how justified they are [5]. Similarly, Kate Distin emphasizes that cultural transmission “presupposes information-processing rules” [7] that cannot themselves be reduced to memes. In summary, both critics emphasize the cognitive structures that are not simply memeplexes but provide the standards by which memes are assessed. Such structures may include innate capacities for reasoning and judgment that enable humans to judge the validity of arguments (or between ethical and unethical choices). Chomsky’s theory of the “Universal Grammar” [8] demonstrates that language capacities are not reducible to mere operant conditioning or the cultural copying mechanism (of memeplexes), but instead rest on innate structures in the human mind. If even a domain like language requires such innate foundations, then normative reasoning, which involves evaluating truth and ethical standards, should also depend on some cogni-

tive capacities that cannot be purely reduced to memes. If strong memetics reduces all agency to the competition of memes, it obscures this normative layer in “self,” avoiding the question of how humans are able to reject certain memes precisely because they are irrational or unjustified. Take the scientific shift from geocentrism to heliocentrism for example. From the framework of the strong meme theory, in the early modern period, the geocentric worldview dominated as a powerful and prevalent memeplex. It was institutionally embedded and reinforced by religious or political forces. By contrast, the heliocentric model initially had little memetic appeal, and meanwhile, it contradicts popular religious narratives and demands complex mathematical reformulations [9]. But over time, heliocentrism gradually displaced geocentrism, and became the new paradigm. When applying a purely memetic account to explain the shift, one might suggest that the heliocentric memeplex became more “adaptive” to its environment, or it offered predictive accuracy, integrated with emerging mathematical techniques, and eventually aligned with broader thoughts in the Enlightenment movement. However, by reducing the transition entirely to memetic “fitness,” it misses what made the new paradigm attractive in the first place. Its rise was not simply a matter of replication, but of truth-tracking, in that heliocentrism better corresponded to empirical reality, and this “normative” dimension (to judge one theory as superior to another) was crucial in the adoption of the new paradigm.

### 2.2 Challenge from Postmodernist and Memetic Replies

Admittedly, a following attack from the memetics (or parallel arguments from postmodern thinkers) would likely target on the understanding of “truth.” On this view, the pursuit of objective scientific truth may not be some independent faculty of reason, but a historically or culturally formed cluster of memes, which gained dominance because it was adaptive in Western intellectual culture. Foucault, for example, suggests that knowledge is inseparable from power, so the very standards of “truth” would be entangled with institutional and political forces [10]. Similarly, a “strong-memetic theorist” might hold the stance that heliocentrism prevailed not because it was objectively “true” in a pure absolute sense in epistemology, but because the “truth-seeking” memeplex gradually displaced older theological or dogmatic memeplexes, aided by shifts in cultures and technology.

However, this argument doesn’t entirely cancel the “nor-

mative element.” Even if our understanding of truth is shaped by memes, the capacity to judge between competing claims would still be independent of the memetic process. Agents (knowers) still try to distinguish between better or worse reasons and between accuracy and uncertainties, even if their judgements are never free of cultural or societal influence. The point made is not that normativity operates in a vacuum, but rather that evaluative standards work within traditions and contexts in a way similar to what philosophers describe as “reflective equilibrium” [11]. There is no “point from nowhere” outside culture from which truth is judged, but this does not imply that all judgments collapse into memetic succession. Without such evaluative capacities, the distinction between one paradigm replacing another seems like an arbitrary succession. The very idea that “truth-seeking memes” could outcompete others already presupposes a cognitive environment in which evaluation is possible. The goal is not to prove that everyone has an innate “moral truth detector”. Instead, the argument is that humans have an internal, reflective, and non-deterministic process for arriving at ethical judgments that cannot be reduced to a purely external cultural transmission mode. In other words, memetics may describe the vehicles of cultural change, but it cannot dispense with the normative faculties that make some memplexes persuasive as “truths” rather than as mere fashions.

### 2.3 Normativity and Evolutionary Framework

Following the analogy that many made to link genes to memes, genetic succession is rarely arbitrary in biological evolution. This is because it is constrained by an external environment (including climate, predators, or food sources) that genes themselves cannot fully control [12]. By contrast, strong meme theory often suggests or leads to the idea that memes both shape and are shaped by their environment. Religious memplexes would establish institutions that then favor certain beliefs, or political memplexes create structures that sustain themselves. Yet if memes both generate and adapt to their own environment, succession risks collapsing into circularity (memes succeed because they create the very conditions of their success). The analogy with genetics therefore breaks down unless there is an independent constraint on memetic survival. Consequently, it could be suggested that normativity provides the evaluative standards of rational and ethical justification and serves as an external check on memplexes, analogous to how physical environments constrain genes. Without this normative dimension, the

difference between a scientific paradigm shift and mere arbitrary succession of cultural fashions becomes unintelligible.

By describing the self as a memplex, strong meme theory reduces it to a passive vehicle of replication. On the contrary, the presence of normativity indeed shows that the self is where evaluation happens. One might object by saying that standards of truth or ethics are simply the prevailing norms of a culture, but this misses the fact that individuals do not merely inherit standards, but also reject or refine them. The act of judgment is to decide that one claim is more justified than another, so it cannot be reduced to memes competing for attention, whenever it already presupposes evaluative faculties (that allow memes to be filtered in the first place). In this sense, the self cannot be collapsed into a memplex: it includes the capacity to stand back from prevailing standards and apply rules of reasoning or justification.

## 3. The Problem of “Qualia”

### 3.1 “Eliminativism” and “Illusionism” of Qualia

A second major difficulty for strong meme theory concerns qualia [6]: the subjective character of experience, or what it is like to see red, feel pain, or doubt one’s own existence [13]. For thinkers such as Susan Blackmore, consciousness and qualia are not fundamental but rather illusions generated by the “selfplex”. On her account, the sense of an inner subject experiencing vivid mental states is itself a construction produced by memes, and therefore no more real than the self that allegedly inhabits those states [14]. Daniel Dennett takes an even more radical approach by denying the need to posit qualia at all. In his rejection of the “Cartesian theater,” Dennett argues that there is no inner stage where raw experiences are presented to a self; consciousness is nothing beyond the distributed processes of information handling in the brain [15]. Both versions aim to dissolve the problem of experience by denying its independence. However, this paper argues that these accounts fail to eliminate or explain the phenomenon. The illusion of experience is still experienced, and the denial of qualia undermines the very distinction between conscious and unconscious states.

In Blackmore’s version of illusionism, where consciousness is thought to be illusions produced when memes compete for replication, leading to the formation of me-

meplex (selfplex). Therefore, there isn't a "continuous stream of rich and detailed experiences, happening one after the other to a conscious person, this is the illusion" [3]. This enticing argument seems to be aligning with popular attacks coming from the field of philosophy of mind, but with the use of memetics theory. However, it is not difficult to prove that they both face the similar challenges of branding consciousness as illusion while not dismissing it entirely. For a broader sense of subjectivity and agency, the "self" is considered to feel and know of the Qualia, and therefore, would still experience this illusion of "subjective experience." As David Chalmers pointed out, this is still not explanatory of "why it is like this to be a conscious agent. [16]" The problem is not evaded easily, because even an illusion is something that is experienced. The very feeling of being misled, of "mistaking a mirage for water, is itself a phenomenal state" [16]. In fact, Blackmore pointed out the nuance between different illusionist accounts, with herself clearly stating that there might exist an authentic or "real" form of consciousness, as illusions are based on truths. It then raises a critical question in epistemology: how we could ever know of the "real" consciousness. Thus, similar "illusionist" accounts risk similar challenges of explaining why it is like something to undergo the illusion in the first place.

On the other hand, the eliminative account of consciousness seemingly avoids the issue by denying the existence of qualia or conscious experience as a whole, while applying a physical account of the mind using memetic framework. Some (memetic) ideas inherit eliminativist strategies from philosophy of mind by denying that qualia exist in any substantive sense [17]. The self, on this account, is a memeplex whose talk of consciousness is just another replicating construct. Yet this position faces the same problem as eliminativism in general: without qualia, there is no meaningful distinction between conscious and unconscious states, and the very language of 'illusion' collapses. If no one actually experiences anything, then memes cannot be said to create an illusion of consciousness as they merely produce verbal behaviors [18]. This leaves memetics unable to account for why conscious subjects consistently report a sense of experience, a gap that undermines its explanatory power. This is further complicated by Dennett's rejection of the so-called "Cartesian theater" — the idea that there is a central stage in the mind where experiences are presented to an inner observer. Eliminativists deny such a theater in order to dissolve the problem of qualia, but in doing so, they remove the very subject

who could be deceived by an illusion [18]. Memes aren't deceiving a subject but rather activating behavioral chains. Since Dennett famously rejects the notion that experiences appear to an internal observer (avoiding infinite regress of "homunculi"), this move destroys the only possible locus for an illusion to be experienced. If there's no observer at all, even metaphorically, then memes cannot create an illusion because there's no such entity to be deceived. Then, memetics theory must choose to either concede that there's some kind of standpoint or "subject" (and face the "hard problem of consciousness"), or to deny one and render the idea of memetic illusion fallible.

### 3.2 The Unexplained Emergence of Qualia

Moreover, it still remains elusive why and how memeplexes would give rise to the illusion of consciousness at all. If consciousness really were merely a false construction, what would be the evolutionary and cultural advantage to produce such a persistent and vivid illusion? The memetic framework often suggests that successful memes are those that enhance replication, but it is far from clear how the illusion of a continuous inner subject is serving this role. Unlike the spread of language, technology, or religions, the experience of consciousness does not straightforwardly guide the succession of species, nor the spread of other memes. To propose that memes created the illusion of consciousness is to leave unanswered why this specific illusion would be selected for, and what functional role, if any, it plays in memetic competition [19]. From a memetic account, the spread of mind-body dualism in Cartesian philosophy can be reduced to the transmission of a philosophical memeplex, just as the very concept of qualia might be treated as a meme that "infected" human discourse. Centuries later, this same idea inspires rival memeplexes in the philosophy of mind, with some defending the authenticity of qualia, others insisting they are illusions. On this view, the ongoing debate itself is merely the competition of memeplexes, but this observation does not resolve the problem of consciousness. If both the claim that qualia exist and the claim that they do not are just memes, then the assertion that "all is memes" is also just another meme competing for influence. In this way, memetics brackets the question of truth rather than settling it, while not responding to the deeper question of why there is anything to experience in the first place. Consequently, this account is not self-explanatory, and raises further questions about the origins and persistence of conscious experience.

If there's really no "I" inside the body, then it becomes difficult to explain the persistence of the first-person standpoint, of the capacity for reflexive evaluation, and the sense of continuity of subjectivity. Similarly, the ability of introspection (to reflect upon one's own states, to report experiences, and to recognize oneself as the subject) would also be reduced to mechanisms of memplex. To declare that the self is a mere illusion risks erasing the very faculty that makes such claims possible. For it is precisely the "self," however understood, that distinguishes between "my" experience and "someone else's," and between inner states and outward behaviors [20]. Even if the notion of a substantial Cartesian ego is abandoned (or interpreted as memes), some conception of self remains necessary to account for the unique way conscious beings access their own experiences. To dismiss the self entirely is not only to deny the persistent human intuition, but also to deconstruct the parts of memetics itself, since memes cannot replicate or be contested without a "host" who "adopts, resists, or modifies them" [1].

#### 4. Conclusion

Strong meme theory presents itself as a radical account of mind, reducing the self and consciousness to the play of replicating memplexes. Yet as this essay has argued, such a view faces certain difficulties. In the domain of normativity, memetics may describe how ideas spread, but it cannot explain how they are judged. To ignore this evaluative dimension is to reduce cultural change to arbitrary succession, collapsing the distinction between persuasion by evidence and persuasion by force. In the domain of consciousness, the theory seems inconsistent in branding consciousness as an illusionary byproduct of memplexes. As a framework for cultural transmission, memetics remains powerful, but cannot by itself serve as a complete theory of the mind. An adequate account should go beyond memes to include both our capacity for normativity and our irreducible experience of consciousness.

This essay has focused on memetics as the central framework in explaining cultural transmission, but it should be noted that memetics is not the only model available. Popular Alternatives like "dual inheritance theory" or socio-cultural leaning models provide different accounts of how ideas and practices propagate and evolve without the use of "meme."

The analysis offered in this paper is only conceptual, rather than empirical, and does not aim to discredit memetics

as a cultural framework, but only to question its adequacy as a theory of mind (self). In this paper, the analogy between memes and cognitive structures also involves interpretive risk, and some of the argument depends on philosophical rather than empirical reasoning. While in principle, more empirical studies could be done to further investigate these claims, developing such principles would lie within the scope of cognitive science, psychology, or neural science. Moreover, the discussion of perception, memory, or rational reference and other psychological mechanisms isn't covered in detail, which might complicate the picture. Future studies might therefore focus on integrating memetic insights with more robust theories of cognition, normativity, and consciousness. Interdisciplinary approaches drawing from cognitive science and philosophy of mind might be needed to reveal how memes interact with the evaluative and experiential dimensions of human life.

#### References

- [1] R. Dawkins, *The Selfish Gene*, London: Oxford University Press, 2006.
- [2] D. Dennet, „Darwin's Dangerous Idea,“ *Behavior and Philosophy*, vol. 24, no. 2, pp. 169-174, 1994.
- [3] S. Blackmore, „Consciousness in meme machines,“ *Journal of Consciousness Studies*, vol. 10, no. 4-5, pp. 19-30, 2003.
- [4] S. Blackmore, „Memes and the malign user illusion,“ *Consciousness and Cognition*, vol. 9, no. 2, pp. S49 - S49, 2000.
- [5] Y. Zheng, „Memes, mind and normativity,“ in *Culture, Nature, Memes*, 191-201.
- [6] F. Jackson, „Epiphenomenal qualia,“ in *Consciousness and emotion in cognitive science*, Routledge, 1998, pp. 197-206.
- [7] K. Distin, *The selfish meme; A critical reassessment*, Cambridge University Press, 2005.
- [8] N. Chomsky, *Aspects of the Theory of Syntax*, MIT press, 2014.
- [9] D. Richard, *Worldviews: An introduction to the history and philosophy of sciences*, John Wiley & Sons, 2018.
- [10] G. Gillett, „Dennet, Foucault, and the selection of memes,“ *Inquiry*, vol. 42, no. 1, pp. 3-23, 1999.
- [11] N. Daniels, „Reflective equilibrium and Archimedean points,“ *Canadian Journal of Philosophy*, vol. 10, no. 1, pp. 83-103, 1980.
- [12] D. Ermakov and A. Ermakov, „Memetic approach to cultural evolution,“ *Biosystems*, vol. 204, p. 104378, 2021.
- [13] T. Nagel, „What is it like to be a bat?,“ in *The language and thought series*, Harvard University Press, 1980, pp. 159-168.



- [14] S. Blackmore, „There is no stream of consciousness,“ *Journal of Consciousness Studies*, vol. 9, no. 5-6, pp. 17-28, 2002.
- [15] D. Dennett, *Consciousness Explained*, Penguin Books, 1991.
- [16] D. J. Chalmers, *The conscious mind: In search of a fundamental theory*, Oxford Paperbacks, 1997.
- [17] D. M. Wegner, *The illusion of conscious will*, MIT press, 2017.
- [18] E. Nahmias, „When consciousness matters: A critical review of Daniel Wegner’s *The illusion of conscious will*,“ *Philosophical Psychology*, vol. 15, no. 4, pp. 527-541, 2002.
- [19] D. A. R. Sperber, „An objection to the memetic approach to culture,“ 2000.
- [20] E. Davidson, „Memetic Intention and The Implications For Agency,“ *The Oracle*, no. 13, pp. 32-43, 2019.
- [21] N. Davis, *An Analysis of Richard Dawkins’s *The Selfish Gene**, Macat Library, 2017.