Beyond Gödelian Limits

Reflexivity as Non-Formal Self-Reference

Abstract

Gödel's incompleteness theorems demonstrate fundamental limitations of formal systems: no sufficiently powerful consistent system can prove its own consistency. This insight is often interpreted as limiting all forms of systematic self-reference. This article develops an alternative perspective by introducing a dimensional epistemology that distinguishes between different modes of cognitive processing. It is argued that Gödel's results pertain specifically to formal self-reference (Dimension 2), while reflexive metacognition (Dimension 3) represents a categorially different form of self-relation not subject to the same constraints. Reflexivity is understood not as a theorem within a formal system, but as a perspectival access to that system. This distinction has far-reaching consequences for debates about artificial intelligence, theories of consciousness, and the limits of formalizability.

Keywords: Gödelian incompleteness theorems, reflexivity, formal self-reference, dimensional epistemology, metacognition, artificial intelligence

1. Introduction

Kurt Gödel's incompleteness theorems of 1931 mark a turning point in the history of logic and mathematics. They demonstrate that every sufficiently powerful formal system must be either incomplete or inconsistent (Gödel, 1931). In particular, no such system, if consistent, can prove its own consistency. This insight has been interpreted in various ways, from mathematical and logical implications to far-reaching philosophical conclusions about the nature of mind and the limits of mechanical computation.

Particularly influential were interpretations that drew conclusions about the relationship between human thought and formal systems from Gödel's results. John Lucas (1961) argued that Gödel's theorems show that the human mind cannot be simulated by a mechanical process. Roger Penrose (1989, 1994) renewed this argument with reference to newer developments in physics and cognitive science. Both positions were intensively criticized, not least by Stewart Shapiro (1998) and David Chalmers (1995), who pointed to conceptual unclarity in transferring formal results to cognitive systems.

This article takes a different path. Rather than asking whether the human mind is limited by Gödel's theorems or transcends them, we first develop a more precise distinction between different forms of self-reference. The thesis is that Gödel's results concern a specific form of self-reference, namely formal self-reference within an axiomatic system, while there are other forms, particularly reflexive metacognition, that are not subject to the same constraints. This distinction is systematized through a dimensional model of epistemic processing that distinguishes four fundamental levels: reactive processing (D1), systematic formalization (D2), reflexive metacognition (D3), and contextual embedding (D4).

The central argument is: Gödel's incompleteness is a D2-result that reveals the limits of formal systems. Reflexivity as a D3-phenomenon, however, operates on a different categorial level and is not constrained by the same limitations. Reflection is not a theorem, but a perspective. This distinction is not only conceptually relevant but has concrete consequences for current debates in philosophy of mind, artificial intelligence, and epistemology.

2. Gödel's Incompleteness Theorems: A Reconstruction

To develop the dimensional perspective, we first require a precise presentation of Gödel's results. The first incompleteness theorem states: Let S be a consistent formal system sufficiently powerful to formalize basic arithmetic (such as Peano arithmetic or stronger systems). Then there exists a sentence G in the language of S that is true but not provable in S (Gödel, 1931; Nagel & Newman, 1958).

The construction of this sentence G proceeds through an ingenious self-reference. Gödel developed a method of arithmetical coding, the so-called Gödel numbering. Each symbol of the formal language receives a unique number, formulas are encoded as products of prime powers, and proof sequences as number sequences. This allows metamathematical statements about the system (such as 'formula x is provable') to be formulated as arithmetical statements within the system.

With this encoding, Gödel constructs a sentence that, informally expressed, states: 'This sentence is not provable in S.' Formally, G is an arithmetical formula expressing its own unprovability. The argument then runs as follows: Suppose G were provable in S. Then G would be false (since G states precisely that it is not provable). But in a consistent system, only true sentences can be proved. Contradiction. Therefore G is not provable. But that is exactly what G states. Therefore G is true.

The second incompleteness theorem sharpens this insight: no sufficiently powerful consistent formal system can prove its own consistency. The consistency of S can be formulated as an arithmetical statement Con(S) ('there exists no number that is the Gödel number of a proof of 0=1'). Gödel showed that Con(S) is formally equivalent to his undecidable sentence G. Since G is not provable (assuming S is consistent), Con(S) is also not provable.

These results are not mere technical curiosities, but touch fundamental questions about the relationship between truth and provability, between semantic and syntactic properties of formal languages. Alfred Tarski (1935) showed in his parallel work on the truth definition that truth cannot be defined in the object language but requires a metalanguage. A similar hierarchization principle seems to be at work in Gödel: the truth of G requires a standpoint outside S.

The philosophical interpretation of these results is disputed. One reading, prominently represented by Douglas Hofstadter (1979), sees in Gödel's theorems a deep insight into the nature of self-reference and consciousness. Hofstadter argues that the capacity for self-reference is central to intelligence and that Gödel's construction represents a formal analogue to self-conscious reflection. A more skeptical reading, represented by Torkel Franzén (2005), emphasizes the specifically formal prerequisites of Gödel's proofs and warns against hasty analogies to cognitive processes.

3. Previous Interpretations and Their Limits

3.1 The Lucas-Penrose Argument

John Lucas (1961) argued that Gödel's incompleteness shows that human mathematicians cannot be simulated by formal systems. His argument can be reconstructed as follows: Suppose a mathematician M were completely describable by a formal system S. Then M could not recognize the consistency of S, because S cannot prove its own consistency. But in fact mathematicians can recognize the consistency of simple formal systems (for example through the construction of models or through metamathematical argument). Therefore M is not describable by S.

Roger Penrose (1989, 1994) renewed this argument and connected it with considerations about quantum mechanics and gravitational theory. Penrose argues that mathematical insight contains a non-algorithmic component demonstrated by Gödel's theorems. He speculates that these non-algorithmic aspects might be based on quantum mechanical processes in the brain, particularly in the microtubules of neuronal cells.

The criticism of these arguments is extensive and convincing (Shapiro, 1998; Chalmers, 1995; Feferman, 1996). First, the argument presupposes that we can recognize the consistency of our own cognitive processes, which is by no means clear. How can we be certain that our own thinking is consistent? Second, a system S could recognize its own Gödel formula G as true if it had access to a stronger metatheory S'. The mathematician would then be described not by S alone, but by a hierarchy of systems S, S', S'', ... Third, the argument confuses the claim 'S cannot prove Con(S)' with the claim 'no system can recognize Con(S)'. Unprovability in S does not exclude that another system can prove Con(S).

3.2 The Tarskian Hierarchy

Alfred Tarski's (1935) solution to the liar paradox and related semantic antinomies consists in introducing a language hierarchy. Truth for an object language L_0 is defined in a metalanguage L_1, truth for L_1 in L_2, and so on. This hierarchization prevents a language from containing its own truth predicate, which would lead to paradoxes. The liar sentence 'This sentence is false' is not well-formed in this framework, because 'is false' is a metalanguage predicate that cannot be applied to sentences of the same level.

The Tarskian hierarchy offers a formal solution for certain kinds of self-reference. It shows that consistent discourse about truth requires a level distinction. However, it remains unclear whether this formal hierarchy adequately captures how cognitive systems actually deal with self-reference. A thinking subject does not seem to rigidly distinguish between object and meta-level, but fluidly switches between different perspectives. We do not speak of $truth_0$, $truth_1$, $truth_2$, but of truth simpliciter.

Graham Priest (2002, 2006) has argued that the Tarskian solution is ultimately unsatisfactory because it destroys the unity of our truth concept. Priest develops paraconsistent logics that can tolerate certain contradictions (so-called dialetheism) to enable a unified treatment of self-reference. His approach is radical: he accepts that some sentences (like the liar) can be both true and false without this leading to logical collapse.

3.3 Deflationary Strategies

Another reaction to Gödel's results consists in minimizing their philosophical significance. This position argues that while Gödel's theorems are mathematically significant, they do not provide deep insights about mind, meaning, or truth. They merely concern the technical properties of certain formal systems. Torkel Franzén (2005) represents this position and criticizes the 'Gödel industry' of philosophical speculation.

This deflationary attitude is methodologically cautious but risks underestimating the conceptual depth of Gödel's discoveries. Gödel's proof shows not only technical limits but raises fundamental questions about the relationship between syntactic structure and semantic content, between provability and truth. A philosophical position that ignores these questions misses essential dimensions of the problem.

4. The Dimensional Perspective: An Alternative Framework

Previous interpretations operate within a framework that does not sufficiently distinguish between different modes of epistemic processing. To systematize these distinctions, we propose a dimensional model that distinguishes four fundamental levels of cognitive organization.

4.1 Dimension 1: Reactive Processing

The first dimension (D1) encompasses reactive, non-representational processes. A system at D1 reacts to environmental stimuli without representing or formalizing them. Classic examples are simple reflex arcs, thermostats, or basic sensorimotor couplings. At D1 there is no self-reference in the proper sense. While causal feedback loops can exist (the output of a system influences its input), the system has no representation of itself. A thermostat regulates temperature without grasping itself as regulator.

D1 systems are operationally closed in the sense that their dynamics is determined by internal structures, but they are epistemically open since they draw no boundary between themselves and their environment. The system/environment distinction is an observer distinction, not a system-internal one. This level corresponds to the sensorimotor phase in Piaget's developmental psychology (Piaget, 1952).

4.2 Dimension 2: Systematic Formalization

The second dimension (D2) is characterized by the capacity for representation and formalization. A D2 system can map states of the world (and potentially of itself) in symbolic structures and operate over these structures. Formal systems in the sense of mathematical logic are paradigmatic D2 systems. They have a syntax (rules for manipulating symbols) and a semantics (interpretation of these symbols).

At D2, formal self-reference becomes possible. A system can contain variables that refer to elements of the system itself. Gödel's construction is an example of D2 self-reference: through Gödel numbering, sentences about the formal system are encoded as sentences within the system. Crucially: this self-reference is a structure described from the outside. The system itself does not 'know' that it refers to itself. The Gödel formula G refers to itself only for an external observer who understands the encoding. For the system S, G is merely an arithmetical formula without recognizable self-reference.

Gödel's incompleteness theorems are D2-results. They show the limits of what can be achieved within a formal system through syntactic manipulation. A consistent D2 system cannot derive its own consistency as a theorem. This is not a weakness of particular systems but a structural property of formalization as such, as Solomon Feferman (1996) emphasizes in his analysis of Gödel's results.

4.3 Dimension 3: Reflexive Metacognition

The third dimension (D3) marks a qualitative leap: the capacity for reflexive metacognition. A D3 system can not only form representations but reflect on its own representations. It can thematize its own epistemic position, recognize its beliefs as beliefs, grasp its limitations as limitations. This is not simply another level of formal representation (which would lead us back to D2), but a different cognitive mode.

The crucial difference: while D2 self-reference is an objective structure (a formula that refers to itself), D3 reflexivity is a perspective. A reflexive system takes a stance toward itself. It grasps itself as limited, as partial, as situated. This reflection is not itself a theorem within a formal system, but a standpoint toward that system. It is not a sentence in the object language, but the act of perspective-taking itself.

A concrete example clarifies the difference: A D2 system can represent the formula 'System S cannot prove its consistency' as a sentence in a metalanguage S'. A D3 system can reflexively grasp: 'I (as instance of a cognitive system) am limited in my ability to guarantee my own consistency.' The first statement is an objective description, the second a perspectival insight. The first can be formalized, the second not completely, because it contains the indexical component 'I'.

Epistemic modal operators like 'I know that...' or 'I believe that...' presuppose D3. They express not merely propositional attitudes but a reflexive relation of the subject to its own cognitive states. The statement 'I know that p, but I don't know whether q' is not a mere conjunction of knowledge claims but a reflection on one's own epistemic limitation. This structure was investigated particularly by Jaakko Hintikka (1962) in his epistemic logic, although the question remains open whether formal epistemic logic can capture the full structure of reflexive self-relation.

4.4 Indexicality and Self-Locating Beliefs

Indexicality is a key phenomenon for D3. The statement 'I am here' is not equivalent to an objective description 'Wolfgang is in Leipzig on January 9, 2026'. The first is perspectival and self-locating, the second objective. David Lewis (1979) and John Perry (1979) have shown in their work on self-locating beliefs that such beliefs are irreducible: they cannot be completely translated into non-indexical descriptions.

A famous example is Perry's 'messy shopper': A philosopher in the supermarket notices a trail of sugar on the floor and thinks 'Someone is making a mess.' When he then realizes 'I am the one making the mess,' his behavior changes, although no new objective information has been added. The change lies in the self-locating perspective. This perspectivity is essential to D3 and cannot be reduced to D2 representations.

Another illuminating example is the Sleeping Beauty problem (Elga, 2000): Sleeping Beauty is put to sleep on Sunday. A fair coin is tossed. On heads, she is awakened on Monday and interviewed; on tails, on Monday and Tuesday (with no memory of Monday on Tuesday). When she awakes and is asked 'What is the probability that the coin showed heads?', two answers are defensible: 1/2 (Halfer position: the coin was fair) or 1/3 (Thirder position: of three possible awakening episodes, only one shows heads). The debate shows how self-locating beliefs ('I am now in this awakening episode') lead to non-trivial epistemic problems that cannot be solved purely through objective probabilities.

4.5 Dimension 4: Contextual Embedding

The fourth dimension (D4) concerns the embedding of cognitive processes in concrete contexts. While D3 thematizes the capacity for reflection, D4 shows that this reflection never occurs abstractly and context-free but is always situated. A reflexive system reflects within specific pragmatic, social, and historical frameworks (Kuhn, 1962).

For the present discussion, D4 is less central but becomes relevant when we think about applying formal results to cognitive systems in concrete situations. The question whether a particular cognitive system can be described by a formal system depends on the context and purposes of this description.

5. The Central Argument: Reflexivity Beyond Formal Limits

With the dimensional framework, we can now formulate the central argument precisely: Gödel's incompleteness theorems demonstrate limits of D2 systems. They show that formal self-reference has certain boundaries. But D3 reflexivity is categorially different from D2 self-reference and is not subject to the same constraints.

5.1 Theorem versus Perspective

The fundamental difference can be stated thus: A theorem is a sentence within a formal system derived from axioms according to certain rules. A perspective, however, is a stance toward a system. Gödel's sentence G is a theorem (or rather an underivable theorem) in S. The reflexive insight 'This system S is limited' is not a formula in S but a reflection about S.

One might object: Is this reflexive insight not simply a sentence in a metalanguage S', thus ultimately still a D2 phenomenon? The answer is: No, for two reasons. First, the construction of the metalanguage S' already presupposes the ability to take a perspective on S. The hierarchization of object and metalanguage is not itself enforced by formal rules but is a reflexive achievement. The transition from S to S' is not an algorithmic step but a perspective shift. Second, a reflexive system can shift its perspective without constructing a new formal level each time. Reflection is fluid, not rigidly hierarchical.

Formally, this can be expressed as follows: Let S be a formal system and G its Gödel sentence. Gödel shows that $S \nvdash G$ (S does not prove G), provided S is consistent. A metasystem S' that contains S as a model can, however, derive $S' \vdash G$. The classical interpretation: one needs the higher level S'. The dimensional interpretation: A D3 system can grasp G as true, not because it operates in S', but because it reflexively penetrates the structure of S. This penetration is not a derivation but an insight.

5.2 The Irreducibility of Perspectivity

The irreducibility of indexical statements supports this position. When I say 'I recognize that my cognitive system is limited,' I express something that cannot be completely formulated in the third person. 'The system S is limited' is an objective statement that can be formulated in a metasystem. 'I recognize that I (as instance of S) am limited' contains an essential indexical component that is not exhausted by any objective description.

This does not mean that reflexive insights are mysterious or unnaturalizable. On the contrary: they are physically realized in cognitive systems that have certain architectural properties. But their epistemic structure cannot be reduced to D2 formalisms. Reflexivity is an emergent phenomenon that opens new possibilities that do not exist at the level of mere formalization.

5.3 Self-Constitution and Autocatalysis

A related phenomenon is the self-constitution of autocatalytic systems. An autocatalytic process produces its own conditions. In biology, living cells are autocatalytic: they reproduce the molecular components necessary for their own existence (Maturana & Varela, 1980). This self-production is not formally paradoxical, although it appears circular.

Analogously, reflexive consciousness could be understood as a form of cognitive autocatalysis. The system constitutes itself as reflexive subject through the act of reflection. This self-constitution cannot be described by a formal system because it is the emergence of the perspective itself. Gödel's theorems concern the limits of what an already constituted formal system can state about itself. They do not concern the self-constitution of the reflexive standpoint.

6. Consequences for Current Debates

6.1 Artificial Intelligence and Machine Consciousness

The dimensional perspective has direct implications for debates about artificial intelligence. If reflexivity cannot be realized by D2 systems, then purely formal systems (in the strict sense) cannot become reflexive. This does not necessarily mean that artificial systems generally cannot be reflexive, but it shows that reflexivity requires more than just complex information processing.

Modern Large Language Models (LLMs) operate primarily at D2: they manipulate symbolic structures according to learned statistical patterns. They can generate sentences of the form 'I think that...' or 'I am uncertain about...', but it is questionable whether this represents genuine reflexivity in the D3 sense or merely the simulation of reflexive language. The difference is not trivial: genuine reflexivity would mean that the system takes a perspective on its own processes, not just reproducing patterns in training data.

An illuminating test would be the capacity for metacognitive calibration: a truly reflexive system should not only give answers but accurately assess its uncertainty about these answers. Empirical studies on metacognition (Fleming & Dolan, 2012; Fleming, 2021) show that humans have specialized neural networks (particularly in the anterior prefrontal cortex) that monitor the quality of their own judgments. The question is whether LLMs have analogous architectures or merely simulate uncertainty through statistical variance in their outputs.

6.2 Higher-Order Thought Theories

In philosophy of mind, Higher-Order Thought (HOT) theories postulate that phenomenal consciousness arises through higher-order thoughts about mental states (Rosenthal, 1997, 2005). A mental state becomes conscious when it is the object of a HOT. From a dimensional perspective, such theories remain at D2 insofar as they understand meta-representations as mere formal structures.

Criticism of HOT theories has often pointed out that a mere hierarchy of representations does not explain why and how phenomenal quality arises (Block, 1995; Kriegel, 2009). The dimensional perspective sharpens this criticism: HOTs at the D2 level (as formal meta-representations) would not suffice. D3 reflexivity would be needed, i.e., a genuine perspectival relation of the subject to its own states. This is more than another representational level. It is the constitution of a standpoint from which states are grasped as my states.

6.3 Epistemological Implications

For epistemology, an important insight emerges: the limits of formalizability are not identical with the limits of the knowable. Gödel's theorems show what is not provable in formal systems. They do not show what is epistemically inaccessible. A reflexive subject can have insights that cannot be encoded as theorems in a formal system.

This rehabilitates a form of rational intuition without becoming mystical. Mathematical or philosophical insight is not merely the mechanical derivation of theorems but can include reflexive grasping of structures. This grasping is intersubjectively comprehensible and rationally criticizable, even if it is not algorithmic. Charles Parsons (2008) has developed similar considerations in his work on mathematical intuition, without however making the dimensional differentiation explicit.

7. Discussion of Possible Objections

7.1 Is D3 Just Disguised Metalanguage?

The most obvious objection is: D3 reflexivity is merely a paraphrase of what in logic is called metalanguage. Every reflexive statement about a system S could be formalized as a statement in a metasystem S'. The distinction between D2 and D3 would then be conceptually superfluous.

The answer was already indicated: the construction of the metalanguage S' presupposes the ability to take a perspective on S. This ability is not itself formalizable in S. Moreover, a reflexive system can switch between different perspectives without constructing a new formal level each time. The Tarskian hierarchy is a formal reconstruction of reflexive capacities, but not their complete capture.

An analogy: physical models can describe human movements (kinematics, dynamics). But the phenomenal experience of movement, the feeling for one's own body, is not contained in these models. Similarly, one can formally model reflexive processes, but perspectivity itself is not in the model. It is the standpoint from which modeling occurs.

7.2 How Do We Know Humans Have D3?

A second objection: the argument presupposes that humans (or at least some cognitive systems) possess D3 capacities. But how do we know this? Perhaps we too are only complex D2 systems that generate the illusion of reflexivity.

This is a serious skeptical objection. The answer cannot consist in a proof that would convince skeptics, since they precisely dispute the possibility of reflexive insight. But one can argue: the phenomenal evidence of reflexive experience ('I am aware of my thoughts') is a datum that must be explained. The dimensional perspective offers an explanation. The skeptical alternative ('It's all D2') would have to explain why it seems as if there is genuine reflexivity. This explanation has not yet been convincingly achieved.

Moreover, there is empirical evidence for the special role of reflexive processes. Studies on metacognition show that the ability to judge one's own uncertainty correlates with specific neural structures and can be selectively impaired in lesions of certain brain regions (Fleming & Dolan, 2012). This suggests that reflexivity is not just an illusion but a really implemented function.

7.3 Is the Distinction Empirically Testable?

A third objection concerns empirical testability. If D3 is categorially different from D2, there should be empirical correlates. But what neural or functional properties characterize D3 as distinct from D2?

This is an important question for further research. Possible candidates would be: (a) global working memory processes that integrate different information sources, as postulated in Bernard Baars' Global Workspace Theory (1988, 2005); (b) recurrent processing loops between cortical areas, proposed in Victor Lamme's Recurrent Processing Theory (2006) as basis for consciousness; (c) metacognitive networks in the anterior prefrontal cortex that monitor the uncertainty of one's own judgments (Fleming & Dolan, 2012). These empirical questions are open, but the conceptual distinction can guide research.

8. Conclusion and Outlook

This article has developed a dimensional perspective on Gödel's incompleteness theorems. The central thesis is: Gödel's results concern specifically formal self-reference (D2), while reflexive metacognition (D3) represents a categorially different form of self-relation not subject to the same limitations.

This distinction is not only conceptually relevant but has consequences for current debates in philosophy of mind, artificial intelligence, and epistemology. It shows that the question of machine consciousness can be posed more precisely: Can artificial systems develop D3 capacities? It sharpens criticisms of Higher-Order Thought theories: mere meta-representations (D2) do not suffice for phenomenal consciousness. And it rehabilitates a form of rational intuition: insight is not reducible to algorithmic derivation.

The dimensional perspective leaves many questions open. The empirical correlates of D3 must be identified. The formal structure of reflexive operators must be further elaborated. The relationship between reflexivity and other aspects of consciousness (phenomenality, intentionality, temporality) requires clarification.

But the basic insight remains: the limits of formalizability that Gödel revealed are not the limits of thought. Reflexivity transcends formal systems not through magic but through a different form of cognitive organization. Understanding this organization is a central task for philosophy and cognitive science.

References

Baars, B. J. (1988). A Cognitive Theory of Consciousness. Cambridge: Cambridge University Press.

Baars, B. J. (2005). Global Workspace Theory of Consciousness: Toward a Cognitive Neuroscience of Human Experience. Progress in Brain Research, 150, 45-53.

Block, N. (1995). On a Confusion about a Function of Consciousness. Behavioral and Brain Sciences, 18(2), 227-247.

Chalmers, D. J. (1995). Minds, Machines, and Mathematics: A Review of Shadows of the Mind by Roger Penrose. Psyche, 2(9).

Elga, A. (2000). Self-Locating Belief and the Sleeping Beauty Problem. Analysis, 60(2), 143-147.

Feferman, S. (1996). Gödel's Program for New Axioms: Why, Where, How and What? In P. Hájek (Ed.), Gödel '96: Logical Foundations of Mathematics, Computer Science and Physics (pp. 3-22). Berlin: Springer.

Fleming, S. M. (2021). Know Thyself: The Science of Self-Awareness. New York: Basic Books.

Fleming, S. M., & Dolan, R. J. (2012). The Neural Basis of Metacognitive Ability. Philosophical Transactions of the Royal Society B, 367(1594), 1338-1349.

Franzén, T. (2005). Gödel's Theorem: An Incomplete Guide to Its Use and Abuse. Wellesley, MA: A K Peters.

Gödel, K. (1931). Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. Monatshefte für Mathematik und Physik, 38, 173-198.

Hintikka, J. (1962). Knowledge and Belief: An Introduction to the Logic of the Two Notions. Ithaca, NY: Cornell University Press.

Hofstadter, D. R. (1979). Gödel, Escher, Bach: An Eternal Golden Braid. New York: Basic Books.

Kriegel, U. (2009). Subjective Consciousness: A Self-Representational Theory. Oxford: Oxford University Press.

Kuhn, T. S. (1962). The Structure of Scientific Revolutions. Chicago: University of Chicago Press.

Lamme, V. A. F. (2006). Towards a True Neural Stance on Consciousness. Trends in Cognitive Sciences, 10(11), 494-501.

Lewis, D. (1979). Attitudes De Dicto and De Se. The Philosophical Review, 88(4), 513-543.

Lucas, J. R. (1961). Minds, Machines and Gödel. Philosophy, 36(137), 112-127.

Maturana, H., & Varela, F. (1980). Autopoiesis and Cognition: The Realization of the Living. Dordrecht: D. Reidel.

Nagel, E., & Newman, J. R. (1958). Gödel's Proof. New York: New York University Press.

Parsons, C. (2008). Mathematical Thought and Its Objects. Cambridge: Cambridge University Press.

Penrose, R. (1989). The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics. Oxford: Oxford University Press.

Penrose, R. (1994). Shadows of the Mind: A Search for the Missing Science of Consciousness. Oxford: Oxford University Press.

Perry, J. (1979). The Problem of the Essential Indexical. Noûs, 13(1), 3-21.

Piaget, J. (1952). The Origins of Intelligence in Children. New York: International Universities Press.

Priest, G. (2002). Beyond the Limits of Thought (2nd ed.). Oxford: Oxford University Press.

Priest, G. (2006). In Contradiction: A Study of the Transconsistent (2nd ed.). Oxford: Oxford University Press.

Rosenthal, D. M. (1997). A Theory of Consciousness. In N. Block, O. Flanagan, & G. Güzeldere (Eds.), The Nature of Consciousness: Philosophical Debates (pp. 729-753). Cambridge, MA: MIT Press.

Rosenthal, D. M. (2005). Consciousness and Mind. Oxford: Oxford University Press.

Shapiro, S. (1998). Incompleteness, Mechanism, and Optimism. The Bulletin of Symbolic Logic, 4(3), 273-302.

Stegemann, W. (2026). Dimensional Logic as a Formal Inference System. Zenodo.

Tarski, A. (1935). Der Wahrheitsbegriff in den formalisierten Sprachen. Studia Philosophica, 1, 261-405.