Quantifying the Relationship Between Cognitive Performance and Consciousness Based on IIT-Proxy Approach Across 15 Species

Primož Kocuvan
primoz.kocuvan@ijs.si
Jožef Stefan Institute
Ljubljana, Slovenia

Matjaž Gams
matjaz.gams@ijs.si
Jožef Stefan Institute
Ljubljana, Slovenia

Abstract

The relationship between intelligence and consciousness remains an open question across cognitive science, neuroscience, and artificial intelligence. This study explores a comparative scoring framework across 15 selected species. A cognitive-evidence index combines behavioral-test outcomes, tool use, and neuron-count categories, while a separate heuristic profile organizes comparative descriptions under five IIT-inspired dimensions, with large language models serving as analytical aids. Pearson and Spearman correlations indicate a positive association between the resulting indices. However, the cognitive index is not a validated measure of general intelligence, the IIT-inspired profile is not a direct measurement of consciousness or integrated information, and the two indices draw partly on overlapping evidence. The findings therefore describe exploratory concordance within the proposed scoring framework rather than an independently established relationship between intelligence and consciousness.

Keywords

cognitive performance, large language models, consciousness, intelligence, IIT

1 Introduction

The relationship between intelligence and consciousness is a fundamental but unresolved question in cognitive science, neuroscience, philosophy of mind, and artificial intelligence. Intelligence is generally associated with the capacity to learn, solve problems, adapt to changing conditions, and use information to achieve goals. Consciousness, by contrast, refers to the presence and organization of subjective experience. Although these capacities often appear together in humans and other animals, they are conceptually distinct, and their relationship cannot be assumed to be direct or causal.

Comparative research across species provides a useful framework for examining this relationship. Animals differ substantially in memory, behavioral flexibility, self-related processing, inhibitory control, tool use, and problem-solving ability. However, no single behavioral task provides a complete or species-neutral measure of intelligence, because performance may depend on sensory abilities, motivation, ecological specialization, prior experience, and experimental design. Brain size and neuron count indicate the biological resources available for information processing, but cognitive performance also depends on neural organization, connectivity, functional specialization, recurrent interactions, and information integration. Cognitive capacity is therefore better evaluated through a combination of behavioral and neurobiological indicators rather than biological scale alone.

The assessment of consciousness presents an even greater methodological challenge because subjective experience cannot be directly observed from an external perspective. Behavioral complexity may indicate advanced information processing, but it does not establish the presence, content, or degree of conscious experience. Comparative investigation therefore requires an explicit theoretical framework that identifies properties relevant to consciousness and translates them into criteria that can be evaluated using available evidence.

Integrated Information Theory provides one possible framework for such an analysis. It characterizes conscious experience. These properties correspond to the theory's phenomenological axioms. The theory emphasizes that consciousness depends not simply on the amount of processing performed by a system, but on how its elements form an integrated causal structure. Direct calculation of integrated information for the whole brains of different species is currently impractical because it would require detailed knowledge of their complete causal organization. Theory-informed proxy assessments can instead provide a structured means of examining how behavioral and neurobiological evidence corresponds to selected properties of conscious processing. Such assessments are not direct measurements of consciousness or formal calculations of integrated information. Large language models may assist in organizing heterogeneous evidence.

2 Related work

A broad scoping review identifies dozens of competing theories of consciousness, highlighting persistent disagreement over definitions, mechanisms, and explanatory scope [1, 2]. Research on neural correlates of consciousness (NCC) has identified candidate mechanisms associated with conscious perception, but their interpretation remains contested, complicating efforts to relate subjective experience to intelligent behavior [3]. In neuroscience and cognitive science, Global Workspace Theory (GWT) proposes that consciousness arises when information is globally broadcast across distributed systems, enabling flexible integration, cognitive control, reportable awareness, and coordination among otherwise modular processes [4]. Supporting evidence emphasizes recurrent processing and large-scale cortical integration, although GWT's neural implementation and explanatory sufficiency remain debated [5]. Higher-order theories (HOT) hold that conscious states require higher-order representations of lower-order mental states, linking awareness to metacognitive or representational capacity [6]; predictive-processing and re-entrant theories instead emphasize top-down and bottom-up interactions

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
Information Society 2026, Ljubljana, Slovenia
© 2026 Copyright held by the owner/author(s).
https://doi.org/10.70314/is.2026.cogni.26

that minimize prediction error and connect consciousness with attention and intelligence. By contrast, Integrated Information Theory (IIT) defines consciousness in terms of integrated causal structure, quantified by the informational measure Φ, and allows consciousness to dissociate from behavioral sophistication [7]. Recent adversarial and comparative studies have directly contrasted GWT-, HOT-, and IIT-based predictions, sharpening their differences while leaving the mechanisms of consciousness unresolved [8, 9].

Alongside these competing accounts, integrative perspectives seek to reconcile global broadcasting, predictive processing, attention schema approaches, and internal world-modeling, suggesting that consciousness and intelligence may rely on overlapping enabling mechanisms without standing in a simple identity relation [10, 5]. Although intelligent behavior can clearly occur without conscious awareness, many researchers agree that, in biological systems, consciousness often complements advanced cognition by supporting flexible integration, self-monitoring, and adaptive control, even as the precise relationship remains unresolved [5].

3 Methodology

According to Integrated Information Theory (IIT), as developed by Tononi and colleagues, consciousness is grounded in a set of phenomenological axioms—fundamental properties that every conscious experience must satisfy. These axioms are intended to be self-evident features of subjective experience itself and serve as the conceptual basis for corresponding physical postulates that a system must satisfy in order to be conscious. We distinguish intelligence as a broad capacity for learning, problem solving, and adaptation from the cognitive score used in this study. The latter is an operational index combining selected behavioral findings with a coarse indicator of neural resources. It therefore represents evidence relevant to some aspects of intelligence, rather than intelligence itself or a validated measure of general intelligence across species. Accordingly, a higher cognitive score indicates a higher aggregate score on the selected indicators; it does not imply a proportionally greater level of intelligence. The five axioms of IIT are as follows:

- **Intrinsic existence**: Conscious experience exists for itself. It is intrinsic to the experiencing subject and does not depend on observation by an external entity.
- **Information**: Every conscious experience is specific and differentiated; it is defined by being this particular experience rather than any of the many other possible experiences.
- **Integration**: Conscious experience is unitary and irreducible. It cannot be decomposed into independent components without losing the experience as a whole.
- **Exclusion**: Conscious experience is definite in both content and spatiotemporal extent. At any given moment, there is a single, determinate experience rather than multiple overlapping ones.
- **Composition**: Conscious experience is structured, consisting of multiple distinctions and relations that together form a unified experiential whole.

The IIT axioms describe properties of experience, whereas the corresponding physical postulates concern the causal organization of the substrate supporting that experience [11, 12]. Behavioral capacities are therefore not equivalent to the axioms. In this study, the axiom names provide an organizing framework for a heuristic comparative assessment, rather than independently measurable behavioral criteria.

The proposed connection is conceptual: information motivates attention to differentiated processing, integration to the coordination of information, and composition to structured relations among represented contents. Descriptions of behavioral flexibility, neural organization, and complex cognition can inform such a comparison, but they do not establish intrinsic information, causal irreducibility, or the causal structure required by IIT. Thus, these descriptions provide interpretive context rather than direct measurements of the corresponding postulates.

Two distinctions are especially important. Intrinsic existence does not mean reflective self-recognition: mirror-test performance does not establish that a system has intrinsic cause–effect power. Likewise, exclusion concerns the definite boundary and spatiotemporal grain of a conscious substrate, not inhibitory control or the selection of a single behavioral response. The behavioral tests included here do not directly assess either property.

Accordingly, self-recognition, theory of mind, metacognition, and language are considered aspects of the comparative evidence, not necessary or sufficient criteria for consciousness. The numerical ratings are IIT-inspired heuristic judgments, not measurements of how strongly an axiom is satisfied or calculations of Φ. LLMs served as analytical aids in this qualitative assessment, not as independent measurements of consciousness.

For each species, the three behavioral-test outcomes were scored as 2 for a clear pass, 1 for a partial or conditional pass, and 0 for a failed test or a test that had not been conducted. Tool use was scored as 2 for confirmed use, 1 for partial or limited evidence, and 0 for no documented evidence. Neuron counts were assigned to five ordered categories, from 0 for very low to 4 for very high, using the ranges specified in Table 2.

The difference in scale resolution reflects the distinction between categorical behavioral evidence and categorized numerical neuron-count estimates. Three levels capture the coarse evidence distinctions reported for behavioral tests and tool use, whereas five neuron-count categories retain a broader ordering of neural resources. These are pragmatic ordinal coding conventions, not empirically calibrated units of cognitive ability.

The cognitive score was calculated as

$$C_i = M_i + O_i + D_i + T_i + N_i,$$

where M_i, O_i, and D_i denote the mirror, novel-object-recognition, and delay-of-gratification scores, respectively, T_i denotes tool use, and N_i denotes the neuron-count category. Consequently, neuron count has twice the maximum point contribution of any individual behavioral component. This unequal contribution is a consequence of the chosen coding scheme, not an empirically established weighting of biological importance. Furthermore, assigning zero to an untested case indicates an absence of credited positive evidence, not demonstrated inability, making the index sensitive to differences in research coverage.

Separately, the five IIT-inspired proxy dimensions were each rated on a 1–10 scale and summed, giving a possible total of 5–50 (Table 1). To examine the association between the cognitive score and the IIT-inspired proxy total, we used Pearson's correlation coefficient (ρ) and Spearman's rank correlation coefficient (ρ_s).

$$\rho = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}},$$

$$\rho_s = \frac{\sum_{i=1}^{n}(\mathrm{R}[X_i] - \overline{\mathrm{R}[X]})\,(\mathrm{R}[Y_i] - \overline{\mathrm{R}[Y]})}{\sqrt{\sum_{i=1}^{n}(\mathrm{R}[X_i] - \overline{\mathrm{R}[X]})^2}\sqrt{\sum_{i=1}^{n}(\mathrm{R}[Y_i] - \overline{\mathrm{R}[Y]})^2}},$$

The two indices were constructed using partly overlapping evidence. The cognitive score assigns explicit points to behavioral-test outcomes, tool use, and neuron-count categories, whereas the IIT-inspired profile uses qualitative judgments informed by some of the same behavioral and neurobiological descriptions. For example, mirror-test evidence contributes to the cognitive score and also informs the discussion of self-recognition used in the proxy assessment. Neuron count enters the cognitive score directly and was also considered as contextual neurobiological information in the qualitative comparison.

The different scoring procedures therefore do not make the underlying evidence independent. Shared evidence may contribute to a positive association between the indices, although it does not mathematically determine the sign or magnitude of the resulting correlation. The analysis consequently evaluates agreement between partially dependent constructed indices, rather than an independent association between measured intelligence and measured consciousness.

4 Results

We calculated the scores with $n = 15$ species included in the comparison. The two scores showed a strong positive association across species:

$$\rho = 0.802, \qquad \rho_S = 0.792.$$

The Pearson correlation indicates a strong approximately linear relationship between the two variables: species with higher intelligence/cognitive-test scores tended to also have higher Tononi style proxy scores. The squared Pearson correlation is:

$$r^2 = 0.802^2 \approx 0.643,$$

which means that approximately 64% of the variance in the Tononi/IIT-style proxy score is associated with the linear variation in the intelligence/cognitive score. This should be interpreted as statistical association, not as a causal explanation.

The Spearman rank correlation was similarly high; this indicates that the rank ordering of species was also broadly consistent between the two scoring systems. In other words, species ranked highly by cognitive-test evidence also tended to rank highly by the IIT-inspired axiom profile.

Overall, the results indicate concordance between the cognitive score and the IIT-inspired heuristic profile under the selected scoring scheme. However, their partly shared evidence base means that this concordance is not an independent validation of either index. The present analysis does not separate the contribution of shared scoring evidence from any association between the underlying constructs. Similar Pearson and Spearman coefficients do not resolve this dependence. The findings should therefore be interpreted as exploratory agreement between constructed indices, not as evidence that intelligence determines consciousness or that the proxy total measures subjective experience.

A further limitation is that species are not statistically independent data points, because closely related species share evolutionary history. Therefore, the correlation is useful as an exploratory comparison, but a stronger comparative analysis would require phylogenetic correction and independently measured neurobiological variables.

The proxy scores in Table 1 were assigned through a two-step heuristic comparison. Human ratings were fixed at 10 for each dimension, giving a total of 50, and chimpanzee ratings were assigned a total of 30, equivalent to a mean of 6 across the five dimensions. These values are reference points within the scoring scheme, not experimentally established percentages of consciousness or integrated information.

The remaining species were positioned relative to these anchors using comparative descriptions of self-related processing, behavioral flexibility, neural organization, and cognitive performance. Such evidence informs the qualitative comparison but does not provide an independently validated conversion into numerical consciousness scores. In particular, a chimpanzee total of 30 relative to a human total of 50 does not establish that chimpanzees possess 60% of human consciousness. Numerical distances between species represent heuristic assignments rather than calibrated differences in subjective experience.

5 Future Work

Several directions for future research follow naturally from the present study. First, the proposed scoring framework for intelligence and consciousness should be refined by incorporating more fine-grained, empirically grounded measures. To explore alternative operationalizations of consciousness derived from competing theoretical frameworks beyond Integrated Information Theory, such as Global Neuronal Workspace or higher-order models. Second, the methodological dependence on large language models as analytical aids warrants systematic evaluation. Future studies should assess the sensitivity of the results to model choice, prompting strategy, and training data, as well as examine inter-model and inter-rater reliability when LLM-assisted assessments are compared with expert human judgments. This would help clarify the extent to which the observed correlations reflect underlying biological patterns rather than artifacts of linguistic or conceptual overlap. Third, the present analysis could be extended to a broader and more diverse set of biological species. Future work should explore alternative statistical analyses and robustness checks, including sensitivity analyses under different scoring scales, partial correlations controlling for neural or morphological variables, and longitudinal or developmental comparisons where data permit. Such extensions would strengthen the interpretability of the results and contribute to a more nuanced understanding of the relationship between intelligence and consciousness across natural systems.

A priority is to document evidence overlap at the level of individual scoring decisions and repeat the analysis using non-overlapping evidence sets. Independent assessors should assign proxy ratings without access to the cognitive totals, although blinding alone would not remove dependence if the underlying evidence remained shared. Sensitivity analyses should also examine alternative component weights, the treatment of untested

Table 1: IIT-inspired heuristic proxy profiles across selected species. Each proxy dimension is rated on a 1–10 scale, and the total is the sum of the five ratings. The dimension names refer to IIT-inspired interpretive categories, not direct tests of the axioms or calculations of Φ.

Species	Intrinsicality	Information	Integration	Exclusion	Composition	Sum
Human (*Homo sapiens*)	10.0	10.0	10.0	10.0	10.0	**50.0**
Chimpanzee (*Pan troglodytes*)	7.5	7.0	5.5	5.0	5.0	**30.0**
Bonobo (*Pan paniscus*)	7.3	6.8	5.3	4.8	4.8	**29.0**
Orangutan (*Pongo*)	7.2	6.7	5.2	4.7	4.7	**28.5**
Bottlenose dolphin (*Tursiops truncatus*)	7.0	6.5	5.0	4.5	4.5	**27.5**
Gorilla (*Gorilla gorilla*)	6.9	6.4	4.9	4.4	4.4	**27.0**
African elephant (*Loxodonta africana*)	6.6	6.1	4.6	4.1	4.1	**25.5**
Raven (*Corvus corax*)	5.8	5.5	4.2	3.7	3.8	**23.0**
African grey parrot (*Psittacus erithacus*)	5.6	5.3	4.0	3.5	3.6	**22.0**
Dog (*Canis familiaris*)	5.2	4.9	3.7	3.2	3.5	**20.5**
Cat (*Felis catus*)	4.7	4.4	3.3	2.8	3.3	**18.5**
Sheep (*Ovis aries*)	4.3	4.0	3.0	2.5	2.7	**16.5**
Rat (*Rattus norvegicus*)	3.6	3.4	2.5	2.0	2.5	**14.0**
Mouse (*Mus musculus*)	3.3	3.1	2.3	1.8	2.0	**12.5**
Rabbit (*Oryctolagus cuniculus*)	3.1	2.9	2.1	1.6	1.8	**11.5**

Scoring note. The ratings are heuristic comparative judgments, not calibrated measurements of consciousness. Human ratings are fixed at 10 for each dimension, giving a total of 50. Chimpanzee ratings sum to 30, corresponding to a mean of 6. These anchors are reference conventions within the scoring scheme.

cases as missing rather than zero, and the influence of the human and chimpanzee reference anchors.

A further methodological objective is to make the translation from theoretical concepts to proxy ratings explicit and reproducible. For each IIT-inspired dimension, future studies should specify the corresponding physical postulate, the empirical indicator proposed to approximate it, the supporting evidence, and the rule for assigning a numerical rating. Each proposed correspondence should be treated as a hypothesis requiring validation, rather than as an equivalence between a behavioral capacity and an IIT axiom. Where no defensible empirical indicator is available, the dimension should be reported as unassessed rather than inferred solely from general cognitive sophistication. Reporting dimension-specific profiles alongside aggregate totals would also allow researchers to examine whether summation obscures distinct patterns of capacities. This approach would make it possible to evaluate individual proxy dimensions separately, without assuming that all dimensions have equal empirical support.

Future comparative datasets should also prioritize the comparability and quality of evidence across species, rather than relying only on a larger species count. Behavioral tasks should accommodate species-specific sensory, motor, and ecological characteristics while retaining an explicitly defined cognitive target. The evidence base should distinguish direct species-specific observations from numerical estimates and findings extrapolated from related species, with uncertainty in these inputs propagated to the resulting scores and correlations. With a sufficiently broad sample, phylogenetically informed analyses could examine whether the association persists after accounting for shared evolutionary history, while leave-one-species-out analyses could assess its dependence on individual species. These checks would complement, rather than replace, the separation of evidence used to construct the two indices and would help establish the conditions under which the proposed comparative framework yields stable and interpretable results.

6 Conclusion

This study compared a cognitive-evidence index with an IIT-inspired heuristic profile across 15 selected species. Using comparative behavioral and neurobiological descriptions, with large language models as analytical aids, the analysis identified a positive association between the two constructed indices.

The interpretation is limited by the heuristic scoring scales, the selected reference anchors, and the partly overlapping evidence used in constructing the indices. The cognitive score should not be equated with general intelligence, and the IIT-inspired total should not be interpreted as a measured degree of consciousness or a calculation of Φ. Agreement between Pearson and Spearman coefficients describes the observed association but does not establish independence from these methodological choices.

The contribution is therefore an exploratory framework for organizing comparative evidence and making its assumptions explicit. Establishing whether a corresponding association persists beyond shared scoring information requires independently operationalized indicators, transparent evidence-to-score rules, and sensitivity analyses.

More broadly, the study highlights a methodological distinction between the evidence available for a species, the theoretical interpretation of that evidence, and the numerical representation chosen for comparison. Within the present framework, a behavioral finding contributes an observation, its relevance to an IIT-inspired dimension requires an interpretive argument, and its translation into a rating introduces a further modeling choice. Keeping these steps separate is essential for preventing apparent numerical precision from being mistaken for direct access to subjective experience. It also clarifies the role of LLMs: they may support the organization and synthesis of comparative descriptions, but the justification of a proxy remains dependent on the evidence and assumptions underlying it, rather than on the fluency or consistency of the generated assessment. The proposed comparison should therefore be used to formulate questions about the relationship between cognitive capacities and consciousness-related properties, not to establish a definitive hierarchy of conscious experience. In particular, the framework can help identify which proposed relationships warrant independent investigation and where the available evidence is insufficient to support a numerical distinction. Disagreement between indicators or between alternative interpretations should be retained as an informative feature of the comparison rather than resolved solely to produce consistent aggregate scores.

Table 2: Comparison of selected cognition-related findings across species.

Species	Mirror Test	Novel Object Recognition Test	Marshmallow / delay-of-gratification test	Tool use	Total brain neurons
Mouse (*Mus musculus*)	Passed, conditional [13]	Passed [14]	Passed [15]	Yes; proto-tool / string-pulling evidence [16]	Very low; approx. 71 million total brain neurons [17, 18]
Rat (*Rattus norvegicus*)	Not conducted / not found as formal visual mark test [19]	Passed [20]	Passed [21]	Yes; rake/tool-retrieval task [22]	Low; approx. 200 million total brain neurons [17]
Rabbit (*Oryctolagus cuniculus*)	Not conducted / not found as formal mark test [23]	Passed [24]	Not conducted / not found [25]	Not conducted / not found [26]	Low; approx. 494 million total brain neurons [18]
Cat (*Felis catus*)	Not conducted / not found as formal mark test [27]	Passed [28]	Not conducted / not found [25]	Yes; proto-tool / string-pulling evidence [29]	Middle; approx. 1.215 billion total brain neurons [30]
Dog (*Canis familiaris*)	Not passed in classic visual mirror test [31]	Passed in adapted toy/NOR task [32]	Passed, variable performance [33]	Yes; proto-tool / string-pulling evidence [34]	Middle; approx. 2.253 billion total brain neurons [30]
Sheep (*Ovis aries*)	Not conducted / not found as formal mark test [35]	Not conducted / not found as formal NOR memory test [36]	Passed [37]	Not conducted / not found [38]	Middle; approx. 3.7–3.9 billion total brain neurons, estimated from sheep brain mass and artiodactyl scaling [39, 40]
Chimpanzee (*Pan troglodytes*)	Passed [31]	Passed, conditional / great-ape group-level evidence [41]	Passed [42]	Yes; strong natural tool-use evidence [43]	High; approx. 28 billion total brain neurons [44]
Bonobo (*Pan paniscus*)	Passed, but less robust than chimpanzee/orangutan evidence [31]	Passed, conditional / great-ape group-level evidence [41]	Passed [45, 46]	Yes; rare wild evidence and stronger captive evidence [47]	High; approx. 25–29 billion total brain neurons, estimated from chimpanzee/great-ape scaling [44]
Gorilla (*Gorilla gorilla*)	Mixed / conditional; not a clean pass [31]	Passed, conditional / great-ape group-level evidence [41]	Passed, conditional / mixed [46]	Yes; rare natural evidence [48]	Very high; approx. 33.4 billion total brain neurons [44, 49]
Orangutan (*Pongo*)	Passed [31]	Passed, conditional / great-ape group-level evidence [41]	Passed [50]	Yes; strong natural evidence [51]	Very high; approx. 32.6 billion total brain neurons [44, 49]
Human (*Homo sapiens*)	Passed [31]	Passed in human analogue / visual-paired-comparison tasks [52]	Passed, with individual variation [53]	Yes; unequivocal archaeological and modern evidence [54]	Very high; approx. 86.1 billion total brain neurons [55]
African elephant (*Loxodonta africana*)	Not conducted / not found for African elephant; Asian elephant passed in related study [56]	Not conducted / not found as formal NOR memory test [57]	Not conducted / not found for African elephant; Asian elephant tested separately [58]	Yes [59]	Very high; approx. 257 billion total brain neurons [60]
Bottlenose dolphin (*Tursiops truncatus*)	Passed, with aquatic-species caveat [61]	Not conducted / not found as formal NOR memory test [62]	Not conducted / not found [25]	Yes; sponge-tool use in bottlenose dolphins, with species-label caution [63]	Very high, inferred; direct total-brain count not found. Cortex approx. 11.7–15.2 billion neurons, and the cerebellum is estimated to contain more neurons than the human cerebellum [64, 55]
Raven (*Corvus corax*)	Not passed [65]	Not conducted / not found as formal NOR memory test [66]	Passed, mainly for better-quality food rewards [67]	Yes [68]	Middle; approx. 2.17 billion total brain neurons; approx. 1.20 billion pallial neurons [69, 44]
Parrot: macaws and African grey parrots (*Ara*, *Psittacus*)	*Psittacus*: not passed; *Ara*: not conducted / not found [70]	Not conducted / not found as formal NOR memory test for *Ara* or *Psittacus* [71]	Passed in macaws and African grey parrots [72]	Yes; *Ara* source available in chat, *Psittacus* noted but not separately source-cited in chat [73]	Middle; approx. 1.56 billion total brain neurons in African grey parrot and up to approx. 3.1 billion in large macaws; pallium approx. 0.85–1.92 billion neurons [69]

Notes. Neuron categories here use approximate total brain neuron counts: very low < 0.1 billion; low 0.1–1 billion; middle 1–5 billion; high 5–30 billion; very high > 30 billion. Where the original source reports "whole brain" excluding olfactory bulbs, the reported value is retained. The sheep value is an estimate, not a direct species-specific neuron count. For bottlenose dolphins, the cited paper reports cortical neuron counts and cerebellar comparisons, but not a direct total-brain neuron count. Earlier values such as 16 billion for humans, 5–6 billion for African elephants, and 1–2 billion for birds were cortical or pallial values, not total brain values.

Acknowledgements

This research was supported by the European Union through the Horizon Europe programme, under the ChatMED project (Grant Agreement ID: 101159214). Additional support was provided by the Slovenian Research Agency through research core funding (Grant No. P2-0209)

References

[1] Davide Sattin et al. "Theoretical Models of Consciousness: A Scoping Review". In: *Frontiers in Psychology* (2021). URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC8146510/.

[2] Anil Seth and Tim Bayne. "Theories of consciousness". In: *Open Access Collection* (2021). Reviewed variety of consciousness theories including HOTs, GWTs, IIT and predictive processing. URL: https://sussex.figshare.com/articles/journal_contribution/Theories_of_consciousness/23488103/1/files/41196653.pdf.

[3] Thomas (Ed.) Metzinger. *Review of Neural Correlates of Consciousness*. Discussion of empirical neural correlates and conceptual challenges. 2025. URL: https://www.researchgate.net/publication/225110218_Review_of_Neural_Correlates_of_Consciousness_Empirical_and_Conceptual_Questions_Thomas_Metzinger_ed.

[4] Bernard J. Baars. *Global workspace theory*. Cognitive architecture for conscious access and integration. 2025. URL: https://en.wikipedia.org/wiki/Global_workspace_theory.

[5] Axel Cleeremans. "Consciousness science: where are we, where are we going?" In: *Frontiers in Science* (2025). URL: https://www.frontiersin.org/journals/science/articles/10.3389/fsci.2025.1546279/full.

[6] Peter Carruthers. *Higher-order theories of consciousness*. Philosophical entry on higher-order theories and mental-state consciousness. 2026. URL: https://plato.stanford.edu/entries/consciousness-higher/.

[7] Wikipedia Contributors. "Integrated information theory". In: *Wikipedia* (2025). URL: https://en.wikipedia.org/wiki/Integrated_information_theory.

[8] Cogitate Consortium. "Adversarial testing of global neuronal workspace and other theories". In: *Nature* (2025). URL: https://www.nature.com/articles/s41586-025-08888-1.

[9] Liad Mudrik et al. "Unpacking the complexities of consciousness: Theories compared". In: *ScienceDirect* (2025). Comparison of GNWT, HOT, IIT, Recurrent Processing, Predictive Processing. URL: https://www.sciencedirect.com/science/article/pii/S0149763425000533.

[10] Michael Graziano. *Attention Schema Theory*. Theory proposing internal model of attention as basis of subjective awareness. 2025. URL: https://en.wikipedia.org/wiki/Attention_schema_theory.

[11] Masafumi Oizumi, Larissa Albantakis, and Giulio Tononi. "From the Phenomenology to the Mechanisms of Consciousness: Integrated Information Theory 3.0". In: *PLOS Computational Biology* 10.5 (2014), e1003588. DOI: 10.1371/journal.pcbi.1003588.

[12] Larissa Albantakis et al. "Integrated information theory (IIT) 4.0: Formulating the properties of phenomenal existence in physical terms". In: *PLOS Computational Biology* 19.10 (2023), e1011465. DOI: 10.1371/journal.pcbi.1011465.

[13] J. Yokose, W. D. Marks, and T. Kitamura. "Visuotactile integration facilitates mirror-induced self-directed behavior through activation of hippocampal neuronal ensembles in mice". In: *Neuron* 112.2 (2024), 306–318.e8. DOI: 10.1016/j.neuron.2023.10.022. URL: https://doi.org/10.1016/j.neuron.2023.10.022.

[14] Lindsay M. Lueptow. "Novel Object Recognition Test for the Investigation of Learning and Memory in Mice". In: *Journal of Visualized Experiments* 126 (2017), p. 55718. DOI: 10.3791/55718. URL: https://doi.org/10.3791/55718.

[15] Z. Gao et al. "The neural basis of delayed gratification". In: *Science Advances* 7.49 (2021), eabg6611. DOI: 10.1126/sciadv.abg6611. URL: https://doi.org/10.1126/sciadv.abg6611.

[16] P. S. Sandhu et al. "Information-theory analysis of mouse string-pulling agrees with Fitts's Law: Increasing task difficulty engages multiple sensorimotor modalities in a dual oscillator behavior". In: *Behavioural Brain Research* 456 (2024), p. 114705. DOI: 10.1016/j.bbr.2023.114705. URL: https://doi.org/10.1016/j.bbr.2023.114705.

[17] Suzana Herculano-Houzel, Bruno Mota, and Roberto Lent. "Cellular scaling rules for rodent brains". In: *Proceedings of the National Academy of Sciences of the United States of America* 103.32 (2006), pp. 12138–12143. DOI: 10.1073/pnas.0604911103. URL: https://doi.org/10.1073/pnas.0604911103.

[18] Suzana Herculano-Houzel et al. "Updated Neuronal Scaling Rules for the Brains of Glires (Rodents/Lagomorphs)". In: *Brain, Behavior and Evolution* 78.4 (2011), pp. 302–314. DOI: 10.1159/000330825. URL: https://doi.org/10.1159/000330825.

[19] T. Yakura et al. "Visual recognition of mirror, video-recorded, and still images in rats". In: *PLOS ONE* 13.3 (2018), e0194215. DOI: 10.1371/journal.pone.0194215. URL: https://doi.org/10.1371/journal.pone.0194215.

[20] A. Ennaceur and J. Delacour. "A new one-trial test for neurobiological studies of memory in rats. 1: Behavioral data". In: *Behavioural Brain Research* 31.1 (1988), pp. 47–59. DOI: 10.1016/0166-4328(88)90157-X. URL: https://doi.org/10.1016/0166-4328(88)90157-X.

[21] D. R. Schuweiler et al. "Rats delay gratification during a time-based diminishing returns task". In: *Journal of Experimental Psychology: Animal Learning and Cognition* 47.4 (2021), pp. 420–428. DOI: 10.1037/xan0000305. URL: https://doi.org/10.1037/xan0000305.

[22] A. Nagano. "Rats' (Rattus norvegicus) tool manipulation ability exceeds simple patterned behavior". In: *PLOS ONE* 14.12 (2019), e0226569. DOI: 10.1371/journal.pone.0226569. URL: https://doi.org/10.1371/journal.pone.0226569.

[23] J. L. Edgar and S. C. Seaman. "The effect of mirrors on the behaviour of singly housed male and female laboratory rabbits". In: *Animal Welfare* 19.4 (2010), pp. 461–471. DOI: 10.1017/S0962728600001949. URL: https://doi.org/10.1017/S0962728600001949.

[24] K. L. Hoffman and E. Basurto. "One-trial object recognition memory in the domestic rabbit (Oryctolagus cuniculus) is disrupted by NMDA receptor antagonists". In: *Behavioural Brain Research* 250 (2013), pp. 62–73. DOI: 10.1016/j.bbr.2013.04.049. URL: https://doi.org/10.1016/j.bbr.2013.04.049.

[25] I. Susini et al. "Delay of gratification in non-human animals: A review of inter- and intra-specific variation in performance". In: *bioRxiv* (2020). Preprint. DOI: 10.1101/

2020.05.05.078659. URL: https://doi.org/10.1101/2020.05.05.078659.

[26] V. K. Bentley-Condit and E. O. Smith. "Animal tool use: Current definitions and an updated comprehensive catalog". In: *Behaviour* 147.2 (2010), pp. 185–221. DOI: 10.1163/000579509X12512865686555. URL: https://doi.org/10.1163/000579509X12512865686555.

[27] Y. Kim and P. Johns. "Using online media to study animal cognition: Domestic cat responses to reflective images". In: *Animal Behavior and Cognition* 12.2 (2025), pp. 258–270. DOI: 10.26451/abc.12.02.05.2025. URL: https://doi.org/10.26451/abc.12.02.05.2025.

[28] J. E. Fogle et al. "Improved neurocognitive performance in FIV infected cats following treatment with the p75 neurotrophin receptor ligand LM11A-31". In: *Journal of NeuroVirology* 27.2 (2021), pp. 302–324. DOI: 10.1007/s13365-021-00956-2. URL: https://doi.org/10.1007/s13365-021-00956-2.

[29] E. Whitt et al. "Domestic cats (Felis catus) do not show causal understanding in a string-pulling task". In: *Animal Cognition* 12.5 (2009), pp. 739–743. DOI: 10.1007/s10071-009-0228-x. URL: https://doi.org/10.1007/s10071-009-0228-x.

[30] Debora Jardim-Messeder et al. "Dogs Have the Most Neurons, Though Not the Largest Brain: Trade-Off between Body Mass and Number of Neurons in the Cerebral Cortex of Large Carnivoran Species". In: *Frontiers in Neuroanatomy* 11 (2017), p. 118. DOI: 10.3389/fnana.2017.00118. URL: https://doi.org/10.3389/fnana.2017.00118.

[31] Y. Lei. "Sociality and self-awareness in animals". In: *Frontiers in Psychology* 13 (2023), p. 1065638. DOI: 10.3389/fpsyg.2022.1065638. URL: https://doi.org/10.3389/fpsyg.2022.1065638.

[32] A. Sommese et al. "Toy exploration in gifted word learner dogs and typical dogs". In: *Animal Cognition* 29 (2026), p. 24. DOI: 10.1007/s10071-026-02047-3. URL: https://doi.org/10.1007/s10071-026-02047-3.

[33] D. Brucks et al. "Reward type and behavioural patterns predict dogs' success in a delay of gratification paradigm". In: *Scientific Reports* 7 (2017), p. 42459. DOI: 10.1038/srep42459. URL: https://doi.org/10.1038/srep42459.

[34] B. Osthaus, S. E. G. Lea, and A. M. Slater. "Dogs (Canis lupus familiaris) fail to show understanding of means-end connections in a string-pulling task". In: *Animal Cognition* 8.1 (2005), pp. 37–47. DOI: 10.1007/s10071-004-0230-2. URL: https://doi.org/10.1007/s10071-004-0230-2.

[35] S. D. McBride, N. Perentos, and A. J. Morton. "Understanding the concept of a reflective surface: Can sheep improve navigational ability through the use of a mirror?" In: *Animal Cognition* 18.1 (2015), pp. 361–371. DOI: 10.1007/s10071-014-0807-3. URL: https://doi.org/10.1007/s10071-014-0807-3.

[36] E. Stephenson and M. J. Haskell. "The use of a Go/Go cognitive bias task and response to a novel object to assess the effect of housing enrichment in sheep (Ovis aries)". In: *Journal of Applied Animal Welfare Science* 25.1 (2022), pp. 62–74. DOI: 10.1080/10888705.2020.1824786. URL: https://doi.org/10.1080/10888705.2020.1824786.

[37] L. Castro, R. Nowak, and V. Dufour. "Performance of sheep and goats in a delay of gratification task". In: *Animal Cognition* (2026). Advance online publication. DOI: 10.1007/s10071-026-02071-3. URL: https://doi.org/10.1007/s10071-026-02071-3.

[38] C. Nawroth et al. "Farm animal cognition: Linking behavior, welfare and ethics". In: *Frontiers in Veterinary Science* 6 (2019), p. 24. DOI: 10.3389/fvets.2019.00024. URL: https://doi.org/10.3389/fvets.2019.00024.

[39] Ashik Banstola and John N. J. Reynolds. "Mapping sheep to human brain: The need for a sheep brain atlas". In: *Frontiers in Veterinary Science* 9 (2022), p. 961413. DOI: 10.3389/fvets.2022.961413. URL: https://doi.org/10.3389/fvets.2022.961413.

[40] Rodrigo S. Kazu et al. "Cellular scaling rules for the brain of Artiodactyla include a highly folded cortex with few neurons". In: *Frontiers in Neuroanatomy* 8 (2014), p. 128. DOI: 10.3389/fnana.2014.00128. URL: https://doi.org/10.3389/fnana.2014.00128.

[41] M. Padberg et al. "Social attention increases object memory in adult but not younger great apes". In: *Animal Behaviour* 221 (2025), p. 123081. DOI: 10.1016/j.anbehav.2025.123081. URL: https://doi.org/10.1016/j.anbehav.2025.123081.

[42] M. J. Beran et al. "Delay of gratification in chimpanzees (Pan troglodytes)". In: *Developmental Psychobiology* 34.2 (1999), pp. 119–127. DOI: 10.1002/(SICI)1098-2302(199903)34:2<119::AID-DEV5>3.0.CO;2-P. URL: https://doi.org/10.1002/(SICI)1098-2302(199903)34:2%3C119::AID-DEV5%3E3.0.CO;2-P.

[43] J. Goodall. "Tool-using and aimed throwing in a community of free-living chimpanzees". In: *Nature* 201 (1964), pp. 1264–1266. DOI: 10.1038/2011264a0. URL: https://doi.org/10.1038/2011264a0.

[44] Andrew B. Barron and Faelan Mourmourakis. "The Relationship between Cognition and Brain Size or Neuron Number". In: *Brain, Behavior and Evolution* 99.2 (2024), pp. 109–122. DOI: 10.1159/000532013. URL: https://doi.org/10.1159/000532013.

[45] J. R. Stevens et al. "Waiting for grapes: Expectancy and delayed gratification in bonobos". In: *International Journal of Comparative Psychology* 24 (2011), pp. 99–111. URL: https://escholarship.org/uc/item/4km2r37j.

[46] L. Salas-Morellon, I. Palacios-Huerta, and J. Call. "Dynamic inconsistency in great apes". In: *Scientific Reports* 14 (2024), p. 18130. DOI: 10.1038/s41598-024-67771-7. URL: https://doi.org/10.1038/s41598-024-67771-7.

[47] S. Pashchevskaya, B. Fruth, and G. Hohmann. "Water scooping: Tool use by a wild bonobo (Pan paniscus) at LuiKotale, a case report". In: *Primates* 65.3 (2024), pp. 145–150. DOI: 10.1007/s10329-024-01121-z. URL: https://doi.org/10.1007/s10329-024-01121-z.

[48] T. Breuer, M. Ndoundou-Hockemba, and V. Fishlock. "First observation of tool use in wild gorillas". In: *PLOS Biology* 3.11 (2005), e380. DOI: 10.1371/journal.pbio.0030380. URL: https://doi.org/10.1371/journal.pbio.0030380.

[49] Suzana Herculano-Houzel and Jon H. Kaas. "Gorilla and orangutan brains conform to the primate cellular scaling rules: implications for human evolution". In: *Brain, Behavior and Evolution* 77.1 (2011), pp. 33–44. DOI: 10.1159/000322729. URL: https://doi.org/10.1159/000322729.

[50] A. E. Parrish et al. "Delay of gratification by orangutans (Pongo pygmaeus) in the accumulation task". In: *Journal of Comparative Psychology* 128.2 (2014), pp. 209–214. DOI: 10.1037/a0035660. URL: https://doi.org/10.1037/a0035660.

[51] C. P. van Schaik and C. D. Knott. "Geographic variation in tool use on Neesia fruits in orangutans". In: *American Journal of Physical Anthropology* 114.4 (2001), pp. 331–342. DOI: 10.1002/ajpa.1045. URL: https://doi.org/10.1002/ajpa.1045.

[52] M. H. Sivakumaran et al. "The discrimination ratio derived from novel object recognition tasks as a measure of recognition memory sensitivity, not bias". In: *Scientific Reports* 8 (2018), p. 11579. DOI: 10.1038/s41598-018-30030-7. URL: https://doi.org/10.1038/s41598-018-30030-7.

[53] W. Mischel and E. B. Ebbesen. "Attention in delay of gratification". In: *Journal of Personality and Social Psychology* 16.2 (1970), pp. 329–337. DOI: 10.1037/h0029815. URL: https://doi.org/10.1037/h0029815.

[54] S. Semaw et al. "2.5-million-year-old stone tools from Gona, Ethiopia". In: *Nature* 385.6614 (1997), pp. 333–336. DOI: 10.1038/385333a0. URL: https://doi.org/10.1038/385333a0.

[55] Frederico A. C. Azevedo et al. "Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain". In: *Journal of Comparative Neurology* 513.5 (2009), pp. 532–541. DOI: 10.1002/cne.21974. URL: https://doi.org/10.1002/cne.21974.

[56] J. M. Plotnik, F. B. M. de Waal, and D. Reiss. "Self-recognition in an Asian elephant". In: *Proceedings of the National Academy of Sciences of the United States of America* 103.45 (2006), pp. 17053–17057. DOI: 10.1073/pnas.0608062103. URL: https://doi.org/10.1073/pnas.0608062103.

[57] J. L. Webb et al. "Evaluating the reliability of non-specialist observers in the behavioural assessment of semi-captive Asian elephant welfare". In: *Animals* 10.1 (2020), p. 167. DOI: 10.3390/ani10010167. URL: https://doi.org/10.3390/ani10010167.

[58] H.-L. Jim, S. Yamamoto, and J. M. Plotnik. "Do Asian elephants (Elephas maximus) wait for a larger food reward in a delay of gratification task?" In: *32nd Annual International Conference on Comparative Cognition*. Conference presentation; no DOI available. Albuquerque, New Mexico, USA, 2025.

[59] S. Chevalier-Skolnikoff and J. Liska. "Tool use by wild and captive elephants". In: *Animal Behaviour* 46.2 (1993), pp. 209–219. DOI: 10.1006/anbe.1993.1183. URL: https://doi.org/10.1006/anbe.1993.1183.

[60] Suzana Herculano-Houzel et al. "The elephant brain in numbers". In: *Frontiers in Neuroanatomy* 8 (2014), p. 46. DOI: 10.3389/fnana.2014.00046. URL: https://doi.org/10.3389/fnana.2014.00046.

[61] D. Reiss and L. Marino. "Mirror self-recognition in the bottlenose dolphin: A case of cognitive convergence". In: *Proceedings of the National Academy of Sciences of the United States of America* 98.10 (2001), pp. 5937–5942. DOI: 10.1073/pnas.101086398. URL: https://doi.org/10.1073/pnas.101086398.

[62] C. M. DeLong et al. "Visual perception in a bottlenose dolphin (Tursiops truncatus): Successful recognition of 2-D objects rotated in the picture and depth planes". In: *Journal of Comparative Psychology* 134.2 (2020), pp. 180–196. DOI: 10.1037/com0000207. URL: https://doi.org/10.1037/com0000207.

[63] M. Kruetzen et al. "Cultural transmission of tool use in bottlenose dolphins". In: *Proceedings of the National Academy of Sciences of the United States of America* 102.25 (2005), pp. 8939–8943. DOI: 10.1073/pnas.0500232102. URL: https://doi.org/10.1073/pnas.0500232102.

[64] Sam H. Ridgway et al. "Higher neuron densities in the cerebral cortex and larger cerebellums may limit dive times of delphinids compared to deep-diving toothed whales". In: *PLOS ONE* 14.12 (2019), e0226206. DOI: 10.1371/journal.pone.0226206. URL: https://doi.org/10.1371/journal.pone.0226206.

[65] L.-C. Vanhooland et al. "A comparative study of mirror self-recognition in three corvid species". In: *Animal Cognition* 26.1 (2023), pp. 229–248. DOI: 10.1007/s10071-022-01696-4. URL: https://doi.org/10.1007/s10071-022-01696-4.

[66] M. Stoewe et al. "Effects of group size on approach to novel objects in ravens (Corvus corax)". In: *Ethology* 112.11 (2006), pp. 1079–1088. DOI: 10.1111/j.1439-0310.2006.01273.x. URL: https://doi.org/10.1111/j.1439-0310.2006.01273.x.

[67] F. Hillemann et al. "Waiting for better, not for more: Corvids respond to quality in two delay maintenance tasks". In: *Animal Behaviour* 90 (2014), pp. 1–10. DOI: 10.1016/j.anbehav.2014.01.007. URL: https://doi.org/10.1016/j.anbehav.2014.01.007.

[68] I. Jacobs and M. Osvath. "Tool use and tooling in ravens (Corvus corax): A review and novel observations". In: *Ethology* 129.3 (2023), pp. 169–181. DOI: 10.1111/eth.13352. URL: https://doi.org/10.1111/eth.13352.

[69] Seweryn Olkowicz et al. "Birds have primate-like numbers of neurons in the forebrain". In: *Proceedings of the National Academy of Sciences of the United States of America* 113.26 (2016), pp. 7255–7260. DOI: 10.1073/pnas.1517131113. URL: https://doi.org/10.1073/pnas.1517131113.

[70] Y. Lin et al. "Mirror responses in the common hill myna (Gracula religiosa) and the African grey parrot (Psittacus erithacus)". In: *Zoology* 147 (2021), p. 125942. DOI: 10.1016/j.zool.2021.125942. URL: https://doi.org/10.1016/j.zool.2021.125942.

[71] C. L. Meehan and J. A. Mench. "Environmental enrichment affects the fear and exploratory responses to novelty of young Amazon parrots". In: *Applied Animal Behaviour Science* 79.1 (2002), pp. 75–88. DOI: 10.1016/S0168-1591(02)00118-1. URL: https://doi.org/10.1016/S0168-1591(02)00118-1.

[72] D. Brucks et al. "Intra- and interspecific variation in self-control capacities of parrots in a delay of gratification task". In: *Animal Cognition* 25.2 (2022), pp. 473–491. DOI: 10.1007/s10071-021-01565-6. URL: https://doi.org/10.1007/s10071-021-01565-6.

[73] L. O'Neill et al. "Causal understanding of the stone dropping task in two species of macaw". In: *bioRxiv* (2020). Preprint. DOI: 10.1101/2020.08.24.264390. URL: https://doi.org/10.1101/2020.08.24.264390.