Artigos integrais em acesso aberto

Cópias integrais autorizadas por licença aberta + acesso oficial às demais fontes

O que está incluído

Este arquivo incorpora os textos completos abaixo quando a licença permite redistribuição. O artigo de Zador et al. (2023), licenciado em Creative Commons Attribution 4.0, será anexado na íntegra em PDF publicado pela revista. Para as outras fontes da lista da pesquisa, indicamos acesso ao texto oficial, mas não redistribuímos cópias integrais sem uma licença de reutilização clara. Assim, este caderno não é uma coleção pirateada nem uma síntese substituta.

CamScanner

Preuss (2006) — Who’s afraid of Homo sapiens?

CC BY; texto integral via PubMed Central. Repositório: https://pmc.ncbi.nlm.nih.gov/articles/PMC1764430/

Licença: Creative Commons Attribution (CC BY). A atribuição original acompanha este texto; consulte a página do artigo para metadados completos e avisos de correção.

Who's afraid of Homo sapiens?

Understanding how humans differ from other animals, as well as how we are like them, requires comparative investigations. For the purpose of documenting the distinctive features of humans, the most informative research involves comparing humans to our closest relatives–the chimpanzees and other great apes. Psychology and anthropology have maintained a tradition of empirical comparative research on human specializations of cognition. The neurosciences, by contrast, have been dominated by the model-animal research paradigm, which presupposes the commonality of "basic" features of brain organization across species and discourages serious treatment of species differences. As a result, the neurosciences have made little progress in understanding human brain specializations. Recent developments in neuroimaging, genomics, and other non-invasive techniques make it possible to directly compare humans and nonhuman species at levels of organization that were previously inaccessible, offering the hope of gaining a better understanding of the species-specific features of the human brain. This hope will be dashed, however, if chimpanzees and other great ape species become unavailable for even non-invasive research.

Introduction

What a pleasure it is to read Karl Pribram's 1970 James Arthur Lecture on human specializations of the brain and cognition. It takes us back to a time of bold scientific conjectures about human nature, an era that saw the publication of Konrad Lorenz's On Aggression (1966; first German edition, 1963), B.F. Skinner's Beyond Freedom and Dignity (1971), and E.O. Wilson's Sociobiology (1975), to take a few examples, all books that intrigued me (and sometimes infuriated me) as a student. Pribram displays an impressively broad knowledge of psychology and neuroscience, achieving a synthesis that few scientists at the time would have been capable of and that even fewer scientists could credibly attempt today. The work is also impressive because the author is reasonably specific about what he considers to be the distinctive characteristics of human beings. In the realm of cognition, he emphasizes meaning, and in particular the ability of humans to construct propositions. He doesn't quite claim that language is a signature specialization of human cognition, but this is implicit in his argument. He considers, for example, that chimpanzees can use signs, but is not impressed with their syntactic capacity. In the realm of neuroscience, Pribram is less specific, but he does suggest that the distinctiveness of the human brain involves not the organization of cortico-cortical connections, as one might suppose, but rather the organization of cortical connections with subcortical motor structures, consistent with his view that meaning is grounded in intention and action.

How do Pribram's conjectures hold up in the light of modern research? It must be acknowledged that Pribram's lecture did not spark a new set of investigations into human cognitive or neurological specializations. In the 1970s, comparative human-ape research was dominated by the ape-language project, which involved teaching apes (mainly chimpanzees) to communicate using systems of manual gesture, such as American Sign Language, or with physical tokens. Popular culture today takes it as given that apes share with humans the capacity for language. Students of language, however, have for the most part drawn the opposite lesson from the ape-language project, concluding that apes demonstrate little productive language capacity, even after intensive training, and that

CamScanner

language is a human specialization (e.g.,). Recent years have brought new ideas about human cognitive specializations. For example, Povinelli has proposed that unlike humans, chimpanzees do not form explicit representations of abstract, unobservable variables, such as mental states (in the case of behavioral causation) or mass force (in the case of physical interactions). These proposals have prompted much debate in the comparative cognition community, along with new experiments (to get the flavor of the debate, see Tomasello et al. and accompanying commentaries). It is noteworthy that Povinelli's "reinterpretation" hypothesis, which holds that humans create new causal interpretations of ancient behaviors, recalls Pribram's claim that humans create and represent meaning through propositions. Additionally, anthropologists have revived the classical idea that culture is a human adaptation, and have begun to characterize cognitive specializations involved in the acquisition and transmission of cultural knowledge (e.g.,).

What are the evolutionary specializations of the human brain?

While a thin but vital thread of research on human cognitive specializations runs through psychology and anthropology from the 1970s to the present, in the neurosciences the thread has been drawn so fine as to be nearly invisible. Consider Pribram's list of human brain specializations: brain enlargement, hemispheric dominance and specialization, and "somewhat more generalized" cytoarchitecture. Of these, brain enlargement (encephalization) remains well accepted, although the nature of this enlargement–was there a general enlargement or a more specific enlargement of association cortex?–is controversial. Whether humans show greater hemispheric specialization than do apes and other nonhuman primates is also controversial: there is some evidence that our closest relatives, the chimpanzees and other apes, possess homologues of Wernicke's and Broca's areas, and that these are lateralized in ways that are at least qualitatively similar to humans. There are, however, those who maintain that humans show extreme hemispheric asymmetry and functional laterality, and that these are among the defining features of humans.

A few human specializations have been identified at finer levels of structural organization. For example, the histological organization of human primary visual cortex displays some striking differences with apes and monkeys. Humans have several populations of morphologically and biochemically distinctive pyramidal cells in anterior cortex. The spacing and cell density of minicolumns differ in Wernicke's area of the left and right hemispheres in humans, but not in chimpanzees or monkeys. With the extension of functional imaging techniques to nonhuman primates has come evidence that the functions of homologous visual areas differ to some extent between humans and macaque monkeys (the animals most commonly used as models of the human visual system), and also raise the possibility that humans have several higher-order visual areas not present in monkeys. Unfortunately, there have not yet been functional imaging studies of ape visual cortex, so it is not possible to say whether these differences represent true human specializations or specializations of the larger group (the Hominoidea) to which both humans and apes belong; the possibility that differences represent specializations of macaques also needs to be considered. One can point to additional differences between humans and various "model" species, such as human-macaque or human-mouse differences–some of them quite remarkable –but as interesting and important as these differences undoubtedly are, we cannot safely conclude that these are human specializations without studies that directly compare humans to apes, as well as to other species.

What slim pickings! It seems extraordinary that neuroscience has so little to offer on a matter so fundamental as what is it about our brains that makes us human. No less disturbing is the understanding that if we don't have the information to say much about how our brains differ from those of other animals, we probably know less than we suppose about how humans resemble other animals, also.

Why do we know so little about human brain specializations?

How is it that after more than 125 years of experimental neuroscience we know so little about how the human brain differs from that of other species? Consider the kind of research required to understand the place of humans among animals: one would need studies that examine human brain structure and function in detail, and that compare humans to other species using comparable techniques. The most informative comparisons, if one's goal is to understand human specializations, are with the great apes, and especially with chimpanzees (Pan troglodytes) and bonobos ("pygmy" chimpanzees; Pan paniscus), as these are the animals most closely related to us. If one's main goal is to understand what humans share with other animals, comparative studies would still be essential, although in this case one would need to study a broad range of species, of varying degrees of relationship to humans.

We do not have these kinds of studies. The obvious reason for this is it has been difficult to study the human brain directly in much detail. Historically, the most valuable investigative techniques in the neurosciences have required invasive and terminal techniques, or other manipulations (such as genetic manipulations) that we regard as unethical in humans. So, for the most part, neuroscientists study not real humans but surrogate humans, nonhuman species that we believe are similar enough to humans to be informative about the human condition, but not so similar as to be ethically problematic (or at least not prohibitively so). This approach–the model-animal paradigm–has been widely employed in the experimental biomedical sciences, particularly, in preference to the broader, comparative approach that has flourished in some other biological disciplines.

The adoption of the model-animal paradigm brought about important changes in the way scientists understand the role of animals in research. In the early part of the 20th century, before the establishment of currently favored model-animal species such as rats, mice, and rhesus monkeys, experimental biologists sought to identify features

CamScanner

of biological organization that are shared by a wide variety of species. To identify these features, it was felt necessary to actually compare a wide variety of species. Attitudes changed as the favored model animals became entrenched as research resources: increasingly it was assumed that the experimentally identified characteristics of these species have broad generality and the imperative to empirically demonstrate generality faded. As historians of science have observed, the establishment of breeding colonies of rodents transformed these animals from subjects of research into something more like standardized chemical reagents. From this perspective, cross-species variation in biological organization is a problem, a potential threat to the standing of one's particular model animal within the larger research community. Species differences, if they are acknowledged at all, tend to be soft-peddled and relegated to the status of noise: it's the commonalities that matter. In this context, the word "basic" has come to mean "common" or "widely shared." As a result, "basic" keeps close company with "same," so that one often hears the expression "the same basic X," where X can be structure, function, organization, or any other attribute of organisms.

This analysis does not imply that the use of nonhuman species as research models of human biological systems or specific diseases is misguided; after all, humans do share many features in common with other animals. What is problematic is the mindset fostered by the model-animal paradigm: viewing the biological world primarily through this lens has serious and negative consequences for the scientific enterprise. For one thing, because the model-animal paradigm discourages systematic, rigorous treatment of the similarities and differences across species, we remain in the position of assuming generality, rather than demonstrating it empirically. This is a particular problem for the neurosciences–mammalian cerebral cortex, for example, has proven to be far more variable across species than believed a decade or two ago, fundamentally compromising the idea that there is a "basic uniformity" of cortical design. In addition, by failing to take differences seriously, we have largely ignored the correlated variations between biological organization and function produced by evolution, and in so doing have ignored a very rich source of information about structure-function relationships (for examples of how such variations can be exploited, see). Finally, the model-animal paradigm has little room for human specializations, for if we consider to be "basic" only those characteristics of biological organization that are shared among species, then the features that distinguish human brains from those of other species don't count as "basic." These would include some of the most interesting and important things we would like to know about human beings, such as what features of our brains support our distinctive cognitive capacities and what features render humans (alone among primates) vulnerable to Alzheimer's disease.

Making room for humans: the critical need for human-chimpanzee comparisons

We can begin to redress some of the deficiencies in our scientific knowledge stemming from the primacy of the model-animal paradigm, with its very indirect approach to understanding human nature, by investing more in comparative studies. The time is especially propitious for studies that address human specializations through direct comparisons of humans and closely related species, because new techniques are available that make it possible to directly study the human brain in great detail, and in ethically acceptable ways, and we can apply these same techniques to the study of our closest relatives.

As the spatial resolution of neuroimaging techniques has improved, it has become practical to use them to compare the structure and function of human brains to those of other primates. While comparisons with macaque monkeys–the favored model nonhuman primates–have been emphasized, these approaches can be applied to the study of chimpanzees and other ape species as well, opening the door to new explorations of human-specific brain organization. In addition, histological studies of humans, using tissue obtained postmortem, have undergone something of a rebirth, driven in part by the need to provide a better understanding of the regional organization of human cortex for the correct interpretation of functional imaging results than can be obtained from the hundred-year-old cortical map of Brodmann. It is clear that one can obtain reliable and informative results with postmortem tissue derived not only from humans, as well as from apes and other nonhuman primates (e.g.,).

The growth of knowledge about the genomic organization of humans and nonhuman primates, including chimpanzees, creates additional opportunities for understanding human-specific brain organization. A number of studies have identified gene-expression or gene-sequence differences between humans, chimpanzees, and other nonhuman primates (reviewed in). Large-scale genomic changes in human evolution have also been documented, involving duplications and rearrangements of DNA, which in some cases have even resulted in the creation of novel, human-specific genes (e.g.,).

Comparative genomics research on human-ape differences has generated enormous interest in the popular media, which is not surprising given the strong interest of the public in understanding what makes us human. One important point that is not widely appreciated, however, is that the identification of genetic differences that distinguish humans from other animals is ultimately of little value if we cannot connect the genetic differences to phenotypic differences. At the present time, the wealth of information about human-chimpanzee genetic differences stands in stark contrast to the poverty of our understanding of human-chimpanzee differences in brain organization. My colleagues and I have suggested that we can use the information from comparative genetic studies as clues to identify previously unknown phenotypic specializations of the human brain, for example, by following the trail from genes to mRNA and protein expression in tissue. While we have focused on comparative histological studies, genomics-driven "phenotype discovery" could employ any of the growing array of non-invasive, non-terminal techniques (biochemical, proteomic, imaging, behavioral) we have for comparing humans to other animals.

CamScanner

Will there be chimpanzees to study?

The advent of technologies like genomics and neuroimaging gives me hope that neuroscience can adopt a more direct approach to the study of human nature and once again deal with fundamental questions of the kind addressed in Karl Pribram's lecture. My optimism is tempered, however, by the knowledge that we might soon lose a resource vital for this pursuit. Understanding what makes our brain distinctively human requires comparing humans to chimpanzees, our closest relatives. In a world where the value of animals in research is measured mainly in terms of their utility as "models," chimpanzees don't fit in very well, as they are more expensive and difficult to maintain than, say, rodents, and as with humans, their use in invasive research is restricted. So, despite the obvious value of comparative studies of chimpanzees and humans for understanding the human brain (including its distinctive vulnerability to neurodegenerative disease), and despite the availability of powerful new techniques for comparing human and chimpanzee brains, it may soon become impossible to pursue this essential research. In 1997, faced with a larger population of captive chimpanzees than could reasonably be supported, NIH imposed a moratorium on breeding chimpanzees. Consequently, the population is not being replaced as animals die of old age and the number of chimpanzees available for the kind of non-invasive research described above–the kind of research we conduct with humans–is rapidly shrinking. Given the highly endangered status of chimpanzees, which seem destined for extinction in the wild within a few decades, it is unlikely that such a resource, once lost, would be reconstituted. In 2007, NIH will decide whether or not to lift the moratorium. As someone committed to understanding the structure, function, and diseases of the human nervous system, I hope the moratorium will be ended. It seems prudent, now that the population has been reduced, to allow enough breeding to maintain a viable population. To be sure, it would be simpler to continue the current policy, avoiding the political heat that would result from choosing to maintain the chimpanzee population, and instead devote our limited resources to supporting mice and a few other favored model-animal species. But, after all, the mouse we will always have with us, whereas chimpanzees will be with us for only a short time longer, if we continue on our current course. If we fail to preserve the means to understand what makes us human, in health and in disease, future generations will surely ask: What were they thinking?

Competing interests

The author(s) declare that they have no competing interests.

Authors' contributions

TMP is the sole author of this paper.

Recent use of signs by chimpanzees (Pan troglodytes) in interactions with humans

Signing behavior in apes: A critical review

Reconstructing the evolution of mind

Understanding and sharing intentions: the origins of cultural cognition

Reinterpreting behavior: a human specialization?

The mentality of apes revisited

The evolution of cultural evolution

The frontal cortex: does size matter?

What is it like to be a human?

Human and nonhuman primate brains: are they allometrically scaled versions of the same design?

Prefrontal white matter volume is disproportionately larger in humans than in other primates

Humans and great apes share a large frontal cortex

Asymmetry of chimpanzee planum temporale: humanlike pattern of Wernicke's brain language area homolog

Planum parietale of chimpanzees and orangutans: a comparative resonance of human-like planum temporale asymmetry

Planum temporale asymmetries in great apes as revealed by magnetic resonance imaging(MRI)

Chimpanzee right-handedness: internal and external validity in the assessment of hand use

Schizophrenia as the price that Homo sapiens pays for language: a resolution of the central paradox in the origin of the species

Human-specific organization of primary visual cortex: alternating compartments of dense Cat-301 and calbindin immunoreactivity in layer 4A

An unusual population of pyramidal neurons in the anterior cingulate cortex of hominids contains the calcium-binding protein calretinin

A neuronal morphologic type unique to humans and great apes

Intuition and autism: a possible role for Von Economo neurons

Lateralization of minicolumns in human planum temporale is absent in nonhuman primate cortex

CamScanner

Comparative mapping of higher visual areas in monkeys and humans

Specializations of the human visual system: The monkey model meets human reality

From monkeys to humans: what do we now know about brain homologies?

Origin of GABAergic neurons in the human neocortex

The altered rationale for the choice of a standard animal in experimental psychology: Henry H. Donaldson, Adolf Meyer and 'the' albino rat

Before there were standards: The role of test animals in the production of empirical generality in physiology

The argument from animals to humans in cognitive neuroscience

Taking the measure of diversity: comparative alternatives to the model-animal paradigm in cortical neuroscience

Microstructure of the neocortex: comparative aspects

Neurochemical and cellular specializations in the mammalian neocortex reflect phylogenetic relationships: evidence from primates, cetaceans, and artiodactyls

Heterogeneity in the columnar number of neurons in different neocortical areas in the rat

The discovery of cerebral diversity: An unwelcome scientific revolution

Microsatellite instability generates diversity in brain and sociobehavioral traits

Species differences in central oxytocin receptor gene expression: comparative analysis of promoter sequences

A chimpanzee genome project is a biomedical imperative

The neurobiology of aging in nonhuman primates

Visual activation in prefrontal cortex is stronger in monkeys than in humans

Mapping the parietal cortex of human and non-human primates

Quantitative investigation of connections of the prefrontal cortex in the human and macaque using probabilistic diffusion tractography

The evolution of prefrontal inputs to the cortico-pontine system: diffusion imaging evidence from Macaque monkeys and humans

Comparing resting brain activity in monkeys, apes and humans

A comparative diffusion tensor imaging (DTI) study of the arcuate fasciculus language pathway in humans, chimpanzees and rhesus macaques

Quantitative architectural analysis: a new approach to cortical mapping

Scaling of inhibitory interneurons in areas V1 and V2 of anthropoid primates as revealed by calcium-binding protein immunohistochemistry

Genetics and the making of Homo sapiens

Human brain evolution: insights from microarrays

Comparative primate genomics: the year of the chimpanzee

Evolutionary mechanisms shaping the genomic structure of the Williams-Beuren syndrome chromosomal region at human 7qll.23

Divergent origins and concerted expansion of two segmental duplications on chromosome 16

A human-specific gene in microglia

Great ape phenome project?

A Pan-oramic view: insights into hominoid evolution through the chimpanzee genome

Increased cortical expression of two synaptogenic thrombospondins in human brain evolution

CamScanner

Kazu et al. (2014) — Cellular scaling rules for the brain of Artiodactyla include a highly folded cortex with few neurons

CC BY; texto integral via PubMed Central. Repositório: https://pmc.ncbi.nlm.nih.gov/articles/PMC4228855/

Licença: Creative Commons Attribution (CC BY). A atribuição original acompanha este texto; consulte a página do artigo para metadados completos e avisos de correção.

Cellular scaling rules for the brain of Artiodactyla include a highly folded cortex with few neurons

Quantitative analysis of the cellular composition of rodent, primate, insectivore, and afrotherian brains has shown that non-neuronal scaling rules are similar across these mammalian orders that diverged about 95 million years ago, and therefore appear to be conserved in evolution, while neuronal scaling rules appear to be free to vary in a clade-specific manner. Here we analyze the cellular scaling rules that apply to the brain of artiodactyls, a group within the order Cetartiodactyla, believed to be a relatively recent radiation from the common Eutherian ancestor. We find that artiodactyls share non-neuronal scaling rules with all groups analyzed previously. Artiodactyls share with afrotherians and rodents, but not with primates, the neuronal scaling rules that apply to the cerebral cortex and cerebellum. The neuronal scaling rules that apply to the remaining brain areas are, however, distinct in artiodactyls. Importantly, we show that the folding index of the cerebral cortex scales with the number of neurons in the cerebral cortex in distinct fashions across artiodactyls, afrotherians, rodents, and primates, such that the artiodactyl cerebral cortex is more convoluted than primate cortices of similar numbers of neurons. Our findings suggest that the scaling rules found to be shared across modern afrotherians, glires, and artiodactyls applied to the common Eutherian ancestor, such as the relationship between the mass of the cerebral cortex as a whole and its number of neurons. In turn, the distribution of neurons along the surface of the cerebral cortex, which is related to its degree of gyrification, appears to be a clade-specific characteristic. If the neuronal scaling rules for artiodactyls extend to all cetartiodactyls, we predict that the large cerebral cortex of cetaceans will still have fewer neurons than the human cerebral cortex.

INTRODUCTION

Artiodactyls are even-toed hoofed mammals, such as pigs, antelopes, and giraffes, of large body and brain size, with a highly convoluted cerebral cortex. Artiodactyls group together with cetaceans under the order Cetartiodactyla, which consists of 290 extant species. Despite the morphological distinctions between extant cetaceans and artiodactyls, the order Cetartiodactyla is evolutionarily monophyletic in origin, having diverged from the shared common ancestor with the nearest groups (Perissodactyla and Carnivora) around 90 million years ago.

CamScanner

Artiodactyl brains are of great interest from the standpoint of comparative neuroanatomy, amongst other reasons for being comparable in size to the largest non-human primate brains. Despite this, few artiodactyl brains have been investigated so far: early studies of neuronal density in the cerebral cortex occasionally included cow and sheep, if they investigated artiodactyls at all, although their highly folded cerebral cortex drew attention already 40 years ago. Artiodactyl brains vary over 10 times in mass within a range that overlaps both with the largest non-human primates and with large carnivores, and the artiodactyl cerebral cortex is even more folded than that of primates and carnivores for a similar brain mass. Yet, artiodactyls, primates, and carnivores of similar brain sizes have very distinct behavioral repertoires. While artiodactyls are mostly grazing herbivores, carnivores exhibit complex hunting behaviors, and primates are usually considered to exhibit an even more complex and flexible range of feeding and social behaviors than artiodactyls and carnivores. Considering that the number of brain neurons must be a major limiting factor to computational capacity, the similarity in brain size in the face of major behavioral differences raises the possibility that the brains of artiodactyls and primates of similar brain size do not contain similar numbers of neurons – that is, that they differ in their cellular composition, contrary to what was the dominant view in the past century (reviewed in).

Using a novel method that allows the rapid quantification of the cellular composition of brain structures (the isotropic fractionator), we have previously shown that primates differ from rodents, lagomorphs (Glires, collectively), and afrotherians in the relationship between the mass of the cerebral cortex and the number of neurons that constitute it. Primates stand out in having a much larger number of neurons in the cerebral cortex than rodents of a similar cortical mass, as well as a larger number of neurons in the cerebellum than predicted for glires or afrotherians with a similar cerebellar mass. Given the behavioral differences between artiodactyls and primates, and assuming that neurons are the functional information-processing units of the brain, we predict artiodactyls to have fewer neurons than primates in both the cerebral cortex and cerebellum, given structures of a similar mass.

We refer to the mathematical relationship that describes how the mass of a brain structure varies as a function of the number of neurons in the structure across species as the neuronal scaling rule that applies to that structure. Similarly, the mathematical relationship that describes how the mass of a brain structure varies as a function of its number of non-neuronal cells across species is referred to as the non-neuronal scaling rule for that structure. The picture emerging from the isotropic fractionator studies is one where neuronal scaling rules are variable across brain structures and mammalian orders, while the non-neuronal scaling rules are shared across all brain structures and across all species of the different orders examined so far: Afrotheria, Glires, Scandentia, Primata, and Eulipotyphla (reviewed in).

We have found that the neuronal scaling rules that apply to the cerebral cortex are shared amongst afrotherians, glires, and eulipotyphlans, but differ from those that apply to primates; the neuronal scaling rules that apply to the cerebellum are shared amongst afrotherians and glires, while eulipotyphlans and primates have their own separate scaling rules. In contrast, the neuronal scaling rules that apply to the “rest of brain” (brainstem plus diencephalon and basal ganglia) appear to be shared across afrotherians, glires, eulipotyphlans, and primates, although there is a larger spread of data points than for the other structures. These findings raise the possibility that, while mammalian orders may have characteristic neuronal scaling rules that apply to some brain structures, those neuronal scaling rules shared by afrotherians and glires, and also eulipotyphlans for the cerebellum and rest of brain, also applied to building the brains of the common mammalian ancestor to these modern groups.

Here we determine whether different neuronal scaling rules apply to Artiodactyla than to other mammalian orders, or whether artiodactyls have retained the neuronal scaling rules that we found to be shared by glires and afrotherians. We also examine whether non-neuronal scaling rules remained conserved across brain structures and species in the evolution of artiodactyls. We test these hypotheses by determining the cellular composition of the brain of five artiodactyl species of a wide range of brain sizes: the domestic pig, the springbok, the blesbok, the greater kudu, and the giraffe, by means of the isotropic fractionator, a non-stereological method that allows for the absolute numbers of neurons and non-neuronal cells to be readily quantified in anatomically defined regions of the brain. The isotropic fractionator yields results that are comparable to those obtained with stereology.

In addition to investigating how the cellular scaling rules for artiodactyls relate to previously found scaling rules, we compare the cellular composition of their brains specifically to those of primates, which have a similar range of brain sizes and cortical gyrification. We have previously shown that cortical folding scales differently across rodents and primates with numbers of cortical neurons, and proposed that folding is a consequence of tension related to connectivity through the white matter rather than a function of the simple addition of neurons to the cortical surface. Here we extend to the highly convoluted artiodactyl cerebral cortex our test of the hypotheses that the degree of cortical folding is (1) a similar function of whole brain mass, (2) a shared function of the number of neurons in the cerebral cortex (reviewed in), and (3) inversely correlated with cortical thickness. Finally, we examine whether variations in the degree of folding along the anteroposterior axis of the cerebral cortex are also related to local variations in the number of cortical neurons or in cortical thickness, and whether the artiodactyl cerebral cortex, like the human cerebral cortex, has separate zones that differ in their neuronal densities.

MATERIALS AND METHODS

ANIMALS

CamScanner

One specimen each of domestic pig (Sus scrofa domesticus, 100 kg), blesbok (Damaliscus dorcas phillipsi, 60 kg), springbok (Antidorcas marsupialis, 25 kg), greater kudu (Tragelaphus strepsiceros, 218 kg), and giraffe (Giraffa camelopardalis, 470 kg) was used in the current study. The specimen of the domestic pig was obtained from local breeders in South Africa, whereas all other specimens were obtained from animals caught from field populations living in their natural environment, and thus their precise age could not be ascertained. All animals were healthy males with no obvious pathologies upon veterinary examination, and no neuropathologies visible, and had the typical body mass of adults, except for the giraffe, which was a juvenile (the typical adult body mass for giraffes is over 1,000 kg). All animals were treated and used in accordance with the University of the Witwatersrand Animal Ethics Committee Guidelines (clearance number 2008/36/1) which parallel those of the NIH for the care and use of animals in scientific experiments.

DISSECTION

All animals were euthanized (overdose of sodium pentobarbital, 100 mg/kg, i.v.) and the head was perfused through the carotids with 0.9% saline, followed by 4% paraformaldehyde in 0.1 M phosphate buffer (PB, pH 7.4). Following perfusion, the brains were removed, weighed, and post-fixed in 4% paraformaldehyde in 0.1 M PB overnight, cryoprotected in 30% sucrose in 0.1 M PB at 4∘C and stored in an antifreeze solution at -20∘C until processing. The brains were divided into two halves along the mid-sagittal fissure and one hemisphere of each brain processed. The olfactory bulbs, when available, were dissected and weighed individually. The cerebellum was dissected by cutting the cerebellar peduncles at the surface of the brainstem. The cerebrum was separated from the brainstem by cutting at a plane anterior to the colliculi and posterior to the thalamus and mammillary bodies of the hypothalamus. The brainstem was divided into pons + medulla and mesencephalon by an axial transection anterior to the basilar pons and posterior to the inferior colliculus (Figure 1A). The cerebrum was then cut manually into 2 mm coronal sections in order to allow the removal of the ensemble of diencephalon and striatum, of the hippocampus, and the dissection of the remaining cerebral cortex into gray and white matter, which had their numbers of cells counted separately. Numbers of cells obtained separately for the pons + medulla, mesencephalon, and diencephalon + striatum were later pooled together and are reported as “rest of brain,” for comparison with data obtained previously in other species.

(A) Main structures analyzed. Shown is the hemisphere of Damaliscus dorcas, illustrating the separation of pons + medulla from the mesencephalon, and the cut that separated the latter from the diencephalon. The ensemble of diencephalon + striatum were later separated from the cerebrum in the individual coronal sections cut from the hemisphere already stripped of the mesencephalon. (B) Surface areas and volumes shown on a coronal section of Giraffa camelopardalis. Total pial surface perimeter (PG) is shown in blue; exposed perimeter, in yellow; perimeter of the white–gray matter interface (PW) in green (solid line); and area of subcortical white matter in the coronal plane also in green, including the dashed line.

Coronal sections of the cerebral cortex were pooled in groups of three (amounting to 6 mm sections) for analysis of the distribution of neurons along the anterior–posterior axis. Each set of coronal sections was then dissected into gray and white matter prior to cell counting. This procedure was followed in all species except for the blesbok, whose cerebral cortex was processed separately as gray and white matter only.

Since only one hemisphere of each brain was available for analysis, values reported here are multiplied by 2 to give estimates for the whole brain that are comparable with our previously published data. While this ignores possible asymmetries between the hemispheres, such asymmetries would have no influence on the results reported here, as the latter span several orders of magnitude whereas any asymmetries would be of the order of a few percentual points between the hemispheres. For the sake of consistency with our previous studies, and because the olfactory bulb was not available for all specimens, whole brain values used in the analysis exclude the olfactory bulb.

MORPHOMETRY

All coronal sections of the cerebral cortex were scanned at 300 dpi and imported into NeuroLucida software (MBF Bioscience, Vermont, USA) for tracing and reconstruction of cortical gray and white matter surface areas and volumes, as described previously. Briefly, for each 2 mm section, we determined the total pial surface perimeter (PG) and the area of gray matter in the coronal plane (SG); the perimeter of the white–gray matter interface (PW, which does not include the interface with the striatum); and the area of subcortical white matter external to the striatum in the coronal plane (SW; Figure 1B).

Determining cortical surface area and volumes by multiplying coronal perimeters and areas by section thickness (2 mm) would underestimate areas and volumes at both extremities of the cerebral cortex, due to its curved shape. Instead, we used formulas that calculate area and volume from a series of sections in a manner that is sensitive to the inclination of the cerebral cortical surface. These formulas are based on the measurements of the gray or white matter perimeters in the coronal plane, P, and of the coronal area occupied by the gray or white matter in the section, S. The formulas that calculate total gray and white matter surface areas AG and AW and volumes VG and VW for each section in a series of coronal areas and perimeters S0, S1, S2, ... and P0, P1, P2, ... obtained as above are the following:

where h = 2 mm (section thickness).

CamScanner

We calculated the local cortical thickness (T) of the gray matter as the ratio VG/AG. The local folding index (FI), a measure of the degree of cortical gyrification of a coronal section, was calculated by first tracing the exposed surface of the gray matter (the smallest surface of gray matter that did not enter sulci) to determine AE in a section, then determining the ratio AG/AE for that section. The total FI for each species is the average of the folding indices for all sections.

After scanning, each section had its cortical white and gray matter separated with a scalpel, using the small amount of resistance that the white matter offers to being pulled from the gray matter exactly at the gray/white matter interface. The hippocampus was removed from all sections. We recorded the mass (g) of the gray and white matter of each analyzed structure to determine cell densities in a manner that is comparable with our previous studies. Sections were then pooled in groups of three for quantification of their cellular composition. Surface areas and volumes were correspondingly added across the groups of three sections for analysis. To determine the total number of neuronal and other (mostly glial) cells in each structure, we then applied the Isotropic Fractionator, which transforms the anisotropic cerebral tissue into an isotropic suspension of cell nuclei whose density in the suspension of known volume can promptly be determined.

ISOTROPIC FRACTIONATION

Total numbers of cells, neurons, and non-neuronal (other cells) were estimated as described previously using the isotropic fractionator method. Briefly, this method turns each dissected brain division into an isotropic suspension of isolated nuclei of known, defined volume, kept homogeneous by agitation. The total number of nuclei in suspension – and therefore the total number of cells in the original tissue – is estimated by determining the density of nuclei in small aliquots stained with the fluorescent DNA marker DAPI (4′-6-diamidino-2-phenylindole dihydrochloride, Invitrogen, USA) under the microscope.

For each structure, at least four samples of the nuclear suspension were counted independently, in different chambers of the hemocytometer, to determine the number of nuclei/ml of the suspension. The reported values for total number of cells refer to the average nuclei/ml of the samples taken multiplied by the total volume of the suspension. This consistently yields a coefficient of variation of 0.10 or less, and never more than 0.15, across samples from the same structure.

Once the total cell number in a structure is known, the proportion of neurons is determined by immunocytochemical detection of neuronal nuclear antigen (NeuN), expressed in the nuclei of most neuronal cell types and not in non-neuronal cells. We used a polyclonal primary antibody against NeuN that is labeled with Cy3 (Millipore, ABN78), which gives results similar to those obtained with the unlabeled monoclonal primary (Millipore, MAB377) followed by a secondary antibody reaction, but yields a much stronger signal in a single step reaction. Estimates of the proportion of NeuN-positive nuclei are considered reliable since the coefficient of variation among animals of the same species is typically below 0.15. Numbers of other cells are derived by subtraction.

DATA ANALYSIS

All statistical analyses and regressions were performed in JMP 9.0 (SAS Institute, Cary, NC, USA). Regressions to power and linear functions were performed to find the best fit for each distribution.

For the comparison with cellular scaling rules reported previously, we used the equations that apply to the average structure size and cellular composition for the species of the groups described earlier: Primata, excluding tree shrews; Glires (; excluding the naked mole-rat, the only fossorial animal in our sample, which is an outlier amongst Glires – see), Eulipotyphla, and Afrotheria. Because we found the artiodactyl data points to overlap with the distributions for glires and afrotherians, we also tested their alignment with the power function that applies jointly to glires and afrotherians.

The hand dissection of the cerebral cortex into gray and white matter yields small numbers of neurons in the “white matter.” Checks performed in cores of white matter dissected with no risk of contaminating gray matter yielded no significant numbers of neurons in the tissue, which suggests that neurons found in white matter preparations are contaminants from the gray matter, and not resident neurons in the white matter, although interstitial neurons are known to exist in other species. For consistency, all analyses that refer to scaling with cortical surface area or volume use total numbers of neurons (gray and white combined), except values reported explicitly for the gray matter, such as neuronal density in the gray matter, which compute only the numbers of neurons found in the dissected tissue (Table 1).

Cellular composition of Artiodactyla brains.

Sus scrofa domesticus Antidorcas marsupialis Damaliscus dorcas phillipsi Tragelaphus stripceros Giraffa camelopardalis MBD, kg ~100 25 60 218 470 MBR, g 57.758 102.102 155.066 308.522 528.026 MCXT, g 35.780 62.814 109.044 212.004 389.616 MHP, g 1.928 3.434 n.a. 10.936 7.486 MCB, g 8.128 11.458 13.402 31.776 67.730 MRoB, g 13.850 27.830 32.620 64.742 70.680 MD+BG, g 6.728 13.814 n.a. 30.418 33.322 MMES, g 2.338 5.304 n.a. 15.928 15.928 MP+M, g 4.784 8.312 n.a. 18.396 21.430 MOB, g 0.822 1.200 n.a. 5.546 2.052 NBR 2.22 × 109 3.06 × 109 3.90 × 109 6.09 × 109 11.21 × 109 NCXT 292.96 × 106 375.49 × 106 548.58 × 106 793.07 × 106 1.67 × 109 NGM 207.75 × 106 293.77 × 106 361.95 × 106 615.87 × 106 1.33 × 109 NHP 12.91 × 106 20.48 × 106 n.a. 28.36 × 106 58.59 × 106 NCB 1.86 × 109 2.26 × 109 2.40 × 109 4.04 × 109 8.88 × 109 NRoB 58.71 × 106 70.48 × 106 92.24 × 106 106.28 × 106 150.65 × 106 ND+BG 34.40 × 106 40.12 × 106 n.a. 58.88 × 106 34.32 × 106 NMES 12.43 × 106 7.52 × 106 n.a. 26.07 × 106 26.63 × 106 NP+M 11.88 × 106 22.84 × 106 n.a.

CamScanner

21.64 × 106 47.44 × 106 NOB 9.20 × 106 16.00 × 106 58.71 × 106 38.33 × 106 24.68 × 106 DNCXT 8,118 5,978 5,031 3,741 4,283 DNGM 7,375 7,051 5,142 4,798 5,882 DNHP 6,695 5,965 n.a. 2,594 7,826 DNCB 228,632 196,999 179,206 127,218 131,080 DNRoB 4,238 2,532 2,828 1,642 2,131 DND+BG 5,113 2,904 n.a. 1,936 1,030 DNMES 5,317 1,418 n.a. 1,637 1,672 DNP+M 2,483 2,748 n.a. 1,176 2,214 DNOB 11,187 13,332 n.a. 6,912 12,026 O/NBR 2.076 3.251 3.989 2.481 3.252 O/NCXT 10.361 10.834 11.956 14.625 15.949 O/NGM 8.544 7.239 8.356 8.443 7.764 O/NHP 10.334 10.111 n.a. 17.868 10.628 O/NCB 0.188 0.207 0.184 0.313 0.622 O/NRoB 18.682 18.710 24.842 32.076 32.333 O/ND+BG 17.779 19.408 n.a. 31.841 38.841 O/NMES 15.667 30.250 n.a. 30.546 41.017 O/NP+M 24.452 13.706 n.a. 34.088 23.631 O/NOB 8.434 6.576 n.a. 8.523 9.417

RESULTS

Across the five species analyzed here (Figure 2), average body mass varies 18.8-fold (from 25 kg in the springbok to 470 kg in the giraffe), while brain mass varies 9.1-fold, and total number of brain neurons varies only 5.0-fold (Table 1). Brain mass increases as a power function of body mass with a small exponent of 0.555 ± 0.029 ($p = 0.0027$; Figure 3A, excluding the pig, which is an obvious outlier in the relationship, with a much larger body than expected for its brain mass, a probable consequence of domestication – see Figure 1 in). The relationship between brain mass and body mass for artiodactyls does not overlap with any of those found previously for glires, insectivores, afrotherians, or primates (Figure 3A), but it does overlap with the relationship found for an independent dataset of 22 artiodactyl species, with a similarly small exponent of 0.596 ± 0.031 ($p < 0.0001$, Figure 3B), where again the domestic pig is an obvious outlier. Although the giraffe individual we analyze was a juvenile, with a smaller body and brain mass than in adult individuals, its body and brain mass still fit the relationship found in the dataset of; Figure 3B). The total number of brain neurons (not including the olfactory bulb) also increases as a power function of body mass, with an even smaller exponent of 0.425 ± 0.069 ($p = 0.00254$; Figure 3C), and this relationship also does not overlap with that found previously for other mammalian clades, although the wider 95% CI includes most other species analyzed previously (Figure 3C).

Phylogenetic relationship and appearance of the artiodactyl brains used in this study. (A) Phylogenetic relationship across the artiodactyl species examined here and other mammals investigated previously. (B) Appearance of the artiodactyl brains used in this study. Scale bar, 2 cm. The springbok brain is shown in a dorsal view, while the lateral view of the left hemisphere is shown for the other species.

Artiodactyl brains gain mass with a different relationship to body mass compared to other mammals. Scaling relationships for artiodactyl brains (in black) are plotted along with rules found previously for other groups (primates in red, glires in green, afrotherians in blue, insectivores in orange, scandentia in gray). Graphs relate (A,B) total brain mass (excluding the olfactory bulb) and (C) total numbers of neurons in the brain (excluding the olfactory bulb) to body mass. (A,C) our dataset; (B) our dataset (filled circles), and dataset of; open circles). Each point represents the average values for each species. Only the power functions for Artiodactyla are plotted (excluding the pig), along with the 95% confidence interval (dotted line). Exponents are 0.555 ± 0.029 (A, brain mass × body mass), 0.596 ± 0.031 (B, brain mass × body mass, for the Boddy dataset exclusively), and 0.425 ± 0.069 (brain neurons × body mass). Data from),, and.

RELATIVE DISTRIBUTION OF MASS AND NEURONS

On average, we find that the cerebral cortex as a whole (with the hippocampus) corresponds to 67.2 ± 2.4% of brain mass, the cerebellum amounts to 11.4 ± 0.9% of brain mass, and the rest of the brain to 21.3 ± 2.3% of brain mass, values in the same range as found for other mammals by ourselves (reviewed in) and others. The relative size of the cerebral cortex and of the cerebellum (expressed as a percentage of total brain mass) does not vary significantly with total brain mass across species studied (Spearman correlation, $p = 0.1041$ and 0.6238, respectively), while the relative mass of the rest of brain decreases significantly with increasing brain mass (Spearman correlation, $\rho = -0.900$, $p = 0.0374$).

The relative number of brain neurons found in each structure does not vary significantly with brain mass across the five artiodactyls (all correlations, $p > 0.18$). As in other mammalian species, the cerebellum of these artiodactyls concentrates most of the brain neurons, with on average 82.3 ± 0.9% of all brain neurons despite its much smaller relative mass of 11.4 ± 0.9%. Conversely, the cerebral cortex has only 15.4 ± 0.8% of all brain neurons, despite representing 67.2 ± 2.4% of brain mass, and the rest of the brain, which accounts for 21.3 ± 2.3% of brain mass, has only 2.4 ± 0.3% of all brain neurons. Across species, variations in the relative mass of the cerebral cortex do not correlate with variations in relative number of brain neurons in the cerebral cortex ($p = 0.2848$), nor does a larger relative mass of the rest of brain correlate with a larger relative number of brain neurons in the rest of brain ($p = 0.2848$), although relatively larger cerebella contain relatively more of brain neurons ($\rho = 0.900$, $p = 0.0374$). The hippocampus is relatively small in artiodactyls, representing between 1.9% (in the giraffe) and 5.5% (in the springbok) of total cortical mass, and containing only between 3.5% (in the giraffe) and 5.4% (in the springbok) of all cortical neurons (Table 1).

NEURONAL SCALING RULES

We find that the relationship between brain size and number of neurons is shared between artiodactyls, glires, and afrotherians, such that brains of a similar mass across these orders, if they existed, would have a similar number of neurons (Figure 4A). Across these five artiodactyls, total brain mass varies as a power function of its number of neurons with an exponent of 1.373 ± 0.120 ($r2 = 0.977$, $p = 0.0015$), indicating that the brain as a whole gains mass more rapidly than it gains neurons (Figure 4A). The distribution of brain mass and total number of neurons in

CamScanner

artiodactyls falls within the 95% confidence interval of the joint distribution for glires (rodents and lagomorphs) and afrotherians, which, moreover, has a similar exponent of 1.456 ± 0.092 ($r2 = 0.950$, $p < 0.0001$; Figure 4A). In contrast, artiodactyls have much larger brains than primates with similar numbers of neurons (Figure 4A, compare black and red points).

Neuronal scaling rules for artiodactyl brains and brain structures. Average brain structure mass for each species is plotted as a function of its total number of neurons. (A) Whole brain (excluding olfactory bulb); (B) Cerebral cortex (including hippocampus); (C) cerebellum; (D) rest of brain; (E) olfactory bulb; (F) hippocampus. (A) Artiodactyls overlap with the power function relating brain mass and number of brain neurons for glires and afrotherians together (green line; exponent, 1.456 ± 0.092), and the power function that applies to artiodactyls (black line, with the giraffe; exponent, 1.373 ± 0.120) also includes glires and afrotherians. (B) Artiodactyls overlap with the power function relating the mass of the cerebral cortex (including the hippocampus) and number of cortical neurons for glires and afrotherians together (green line; exponent, 1.645 ± 0.090), and the power function that applies to artiodactyls (black line, excluding the giraffe; exponent, 1.739 ± 0.092) also includes glires, afrotherians, and eulipotyphlans, but excludes primates. (C) The power function relating cerebellar mass and number of cerebellar neurons for artiodactyls (black line; exponent, 1.351 ± 0.118) includes other non-primate mammalian clades. (D) The power function relating mass and number of neurons in the rest of brain mass (brainstem to basal ganglia) for glires and afrotherians (green line; exponent, 1.481 ± 0.180) excludes four of five artiodactyl species. While all but one non-artiodactyl species fall below the power function that applies to artiodactyls (black line; exponent, 1.706 ± 0.376), the 95% CI is wide enough that they are still included. (E) The power function relating mass and number of neurons in the olfactory bulb for artiodactyls (black line; exponent, 1.309 ± 0.257) excludes all other mammalian clades. (F) The artiodactyl hippocampus falls within the 95% CI of the power functions relating hippocampal mass and number of neurons for afrotherians (blue line; exponent, 1.211 ± 0.312, $p = 0.0304$) and for eulipotyphlans (yellow line; exponent, 1.041 ± 0.422; $p = 0.0903$). Artiodactyl species in black, primates in red, glires in green, afrotherians in blue, insectivores in orange, scandentia in gray. Data from),, and.

The relationship between the mass of the cerebral cortex and its number of neurons for the five artiodactyl species has an exponent of 1.364 ± 0.150 ($p = 0.0028$). This, however, includes the giraffe, which was a juvenile, with a cerebral cortex that was possibly not yet fully grown, thus artificially decreasing the exponent (notice that report a total brain mass of 700 g for the giraffe, while our specimen had a cortical mass of only 528 g). When the giraffe is excluded from the relationship, the exponent that applies to the cerebral cortex of artiodactyls (1.739 ± 0.092, $r2 = 0.994$, $p = 0.0028$; Figure 4B, black) becomes indistinguishable from the exponent that applies to this distribution across glires and afrotherians together, of 1.645 ± 0.090 ($r2 = 0.962$, $p < 0.0001$; Figure 4B, green). The mass of the cerebral cortex of artiodactyls is therefore well predicted from its number of neurons by the neuronal scaling rules that apply to glires and afrotherians. In contrast, the artiodactyl cerebral cortex is much larger than a primate cerebral cortex of a similar number of neurons (Figure 4B, compare black and red data points). Thus, the relationship between cerebral cortical mass and number of neurons that applies to glires and afrotherians also applies to artiodactyls, and to eulipotyphlans as well, but excludes the primates (Figure 4B, red data points).

The relationship between cerebellar mass and number of neurons across the five artiodactyl species is described by a power function with exponent 1.351 ± 0.118 ($r2 = 0.978$, $p = 0.0014$) that overlaps with the distribution that applies to glires and afrotherians, of exponent 1.306 ± 0.070 ($r2 = 0.964$, $p < 0.0001$) and includes data points for those species (Figure 4C). This relationship is such that, for a similar number of cerebellar neurons, artiodactyls have a much larger cerebellar mass than primates (Figure 4C).

The mass of the rest of brain scales as a power function of its number of neurons across artiodactyls with an exponent of 1.706 ± 0.376 ($r2 = 0.872$, $p = 0.0201$; Figure 4D, black). Although this exponent is not significantly different from the exponent that applies jointly to afrotherians and glires, of 1.481 ± 0.180 ($r2 = 0.838$, $p < 0.0001$; Figure 4D, green), four of the five artiodactyl data points fall outside of the 95% confidence interval for afrotherians and glires. This suggests that a different neuronal scaling rule applies to artiodactyls than to afrotherians and glires, and also to primates, such that the mass of the rest of brain is larger in artiodactyls than in afrotherians, glires, and also primates, for a similar number of neurons in the rest of brain. Within the rest of brain, we find larger numbers of neurons in the combined diencephalon and basal ganglia than in the mesencephalon or the combined pons and medulla in the smaller artiodactyls (pig, springbok, and kudu), but not in the giraffe (Table 1).

The mass of the olfactory bulb scales across artiodactyls as a function of its number of neurons of exponent 1.309 ± 0.257 ($r2 = 0.928$, $p = 0.0364$), similar to the exponent that applies to afrotherians and glires together (1.348 ± 0.238, $r2 = 0.762$, $p = 0.0002$; Figure 4E). However, the two distributions are shifted, so that the olfactory bulb of artiodactyls has a significantly larger mass than the olfactory bulb of afrotherians and glires with a similar number of neurons, falling outside of the 95% CI for those clades. For the hippocampus, in contrast, the power function that describes the mass of the structure with number of neurons does not reach significance across artiodactyls ($p = 0.2254$), and the distribution of hippocampal mass as a function of number of neurons cannot be distinguished from the distribution across afrotherians (Figure 4F).

NON-NEURONAL SCALING RULES

CamScanner

While the mass of each brain structure varies as a different power function of its number of neurons across artiodactyl species, the relationship between mass of each brain structure (cerebral cortex, cerebellum, and rest of brain) and number of other (non-neuronal) cells can be described as a single power function of exponent 0.956 ± 0.045 ($p < 0.0001$). Besides the overlap across brain structures, this distribution also overlaps with that across afrotherians, glires, eulipotyphlans as well as primates (joint exponent of 1.034 ± 0.020, $p< 0.0001$). The addition of artiodactyls to the analysis does not alter the exponent significantly (1.035 ± 0.017, $p < 0.0001$; Figure 5A). As a result, whole brain mass varies as a similar functions of numbers of other cells across artiodactyls (exponent, 1.104 ± 0.216, $p = 0.0146$), the ensemble of glires, primates, afrotherians, and eulipotyphlans (exponent, 1.040 ± 0.032, $p < 0.0001$), and all clades together (exponent, 1.020 ± 0.026, $p < 0.0001$; Figure 5B). Thus, brains of a similar size are composed of similar numbers of non-neuronal cells across different mammalian clades, from afrotherians to artiodactyls.

Artiodactyls fit the non-neuronal scaling rules for all brain structures and mammalian clades. Plots show the relationship between structure mass and number of non-neuronal (other) cells for (A) each brain structure or (B) the brain as a whole. Circles, cerebral cortex (A) or whole brain (B); squares, cerebellum; triangles, rest of brain. Functions plotted apply to the ensemble of species, including artiodactyls. Exponents are 1.035 ± 0.017 (A) and 1.020 ± 0.026 (B). Artiodactyl species in black, primates in red, glires in green, afrotherians in blue, insectivores in orange, scandentia in gray. Data from),, and.

CELL DENSITIES

In artiodactyls, as in other groups of mammals, neuronal density varies considerably more than other cell density across structures (Figures 6A,B). Neuronal density in the cerebral cortex (gray + white matter) varies between 3,741 neurons/mg in the greater kudu to 8,118 neurons/mg in the pig; in the cerebellum, it ranges between 127,218 neurons/mg in the greater kudu to 228,632 neurons/mg in the pig; and in the rest of brain, from 1,642 neurons/mg in the greater kudu to 4,238 neurons/mg in the pig (Table 1). Neuronal densities in the cerebral cortex and cerebellum are lower in artiodactyl than in primates for a similar structure mass (Figure 6A). In contrast, other cell densities vary across all structures only between 33,123 other cells/mg in the blesbok cerebellum and 81,505 other cells/mg in the giraffe cerebellum, and overlap between artiodactyl and primate structures of similar mass (Figure 6B). There is no systematic correlation between variations in other cell density and in structure mass across artiodactyls species (Cx, $p = 0.5046$; Hp, $p = 0.6000$; Cb, $p = 0.8729$; RoB, $p = 0.6238$; all structures: $p = 0.2810$), as found in other clades (Figure 6B).

Cellular density variation across structures and species. (A) Average neuronal density (in number of neurons per mg of tissue) in each structure and species. (B) Average density of other cells (in number of other, non-neuronal cells per mg of tissue) in each structure and each species. Both graphs are plotted on a similar Y-axis for comparison. For clarity, only the significant power functions for Artiodactyla are plotted (cerebral cortex with hippocampus, exponent -0.424 ± 0.028; cerebellum, exponent -0.276 ± 0.063; rest of brain, -0.488 ± 0.113). No brain structure in any mammalian order examined here exhibits a significant correlation between other cell density and structure mass. Notice in (A) that the function for neuronal densities in artiodactyl cerebral cortex (minus the giraffe; circles) includes glires, afrotherians and eulipotyphlans, while the function for artiodactyl cerebellum (with the giraffe; squares) includes all other non-primate mammalian clades. Artiodactyl species in black, primates in red, glires in green, afrotherians in blue, insectivores in orange, scandentia in gray. Data from),, and.

Consistently with the faster increase in structure mass than in number of neurons in the artiodactyl cerebral cortex in a manner that is similar across artiodactyls, afrotherians, glires, and eulipotyphlans, but not primates, neuronal density in the artiodactyl cerebral cortex (minus the giraffe) decreases with increasing cortical mass, as a power function of exponent -0.424 ± 0.028 ($p = 0.0044$; Figure 6A, circles). The 95% confidence interval for this function includes afrotherians, glires and eulipotyphlans, but excludes primates (Figure 6A), indicating that the scaling of neuronal densities in the cerebral cortex is shared among the former clades, but is distinct in primates. Similarly, in the artiodactyl rest of brain, neuronal density also decreases significantly as a power function of increasing structure mass (exponent, -0.488 ± 0.113, $p = 0.0227$), in a manner that cannot be distinguished from the trend in the ensemble of other clades (Figure 6A, triangles). Neuronal density in the artiodactyl cerebellum (including the giraffe) scales in a manner that includes most afrotherians and glires, but excludes primates, varying as a power function of cerebellar mass with exponent -0.276 ± 0.063 ($p = 0.0222$; Figure 6A, squares).

The ratio between the numbers of other cells and neurons in each structure (the O/N ratio), which approximates the glia/neuron ratio, varies between 0.184 (in the blesbok cerebellum) and 32.333 (in the giraffe rest of brain) across structures and species in these artiodactyls (Figure 7). The O/N ratio in the artiodactyl cerebral cortex exceeds 10 when both gray and white matter are included, and within the cortical gray matter alone it varies between 7.2 and 8.5 across species (Table 1). The O/N ratio varies widely across structures and species as a function of structure mass, with no single relationship evident (Figure 7A). In contrast, the O/N ratio varies as a common power function of neuronal density across all artiodactyl structures with an exponent of -1.082 ± 0.035 ($p < 0.0001$), in a distribution that overlaps with the variation of O/N as a function of neuronal density across non-artiodactyl species (exponent, -0.917 ± 0.028, $p < 0.0001$; Figure 7B). The addition of artiodactyl structures to the distribution does not change the exponent significantly (-0.935 ± 0.024, $p < 0.0001$).

Uniform variation in the O/N ratio with neuronal density, but not structure mass. (A) Ratio between numbers of other (non-neuronal) and neuronal cells, O/N, varies in no uniform fashion across structures (cerebral cortex,

CamScanner

circles; cerebellum, squares; rest of brain, triangles) and species (each symbol) as a function of the mass of each structure. (B) In contrast, the O/N ratio varies uniformly across brain structures and species as a function of average neuronal density in the structure. The power function plotted in b applies to the non-artiodactyl species in the sample, and has and exponent of -0.917 ± 0.028. Artiodactyl species in black, primates in red, glires in green, afrotherians in blue, insectivores in orange, scandentia in gray. Data from), and.

GYRIFICATION

We next examined how cortical folding and the distribution of neurons along the cortical surface compare across artiodactyls and primates species of a similar range of brain sizes (Table 2). We find that the cerebral cortex of artiodactyls appears to be as folded as primate cortices of a similar brain mass (Figure 8A), and with a distribution of folding along the anteroposterior axis of artiodactyl cortices that is similar to that observed in primates (; see Figure 14A). However, we find that the cerebral cortex of artiodactyls is much more folded than primate or rodent cortices of similar numbers of neurons (Figure 8B). For instance, while the kudu and the pig-tailed macaque monkey both have around 400 million neurons in a single cortical hemisphere, the former has a larger FI of 2.009, compared to 1.652 in the macaque (Figure 8B). This larger degree of folding is related to the spreading of similar numbers of neurons over a much larger cortical surface area in artiodactyls than in primates (Figure 8C): for instance, the ca. 400 million neurons are spread over 22,203 mm2 in the kudu, but over only 9,381 mm2 in the pig-tailed monkey.

Cortical areas, volumes and folding index for artiodactyls.

Species AGM, mm2 VGM, mm3 VWM, mm3 FI Sus scrofa domesticus 6,594 8,807 5,834 1.889 Antidorcas marsupialis 9,623 16,063 10,809 2.018 Damaliscus dorcas phillipsi 16,915 36,692 20,330 2.007 Tragelaphus strepsiceros 22,203 52,041 25,691 2.010 Giraffa camelopardalis 40,128 96,048 60,309 2.408

The artiodactyl cortex is more folded than that of rodents and primates. (A) Folding index (FI) of the cerebral cortex (ratio between the total surface area and the exposed surface area of the cortex) plotted as a function of total brain mass. (B) FI of the cerebral cortex plotted as a function of the number of cortical neurons in each species. Notice that for similar numbers of cortical neurons, the cortex of artiodactyls is more folded that the cortex of both primates and rodents. (C) Variation in total surface area of the cerebral cortical gray matter, in mm2, plotted as a function of the total number of neurons in the cortical gray matter. Power functions for artiodactyls (minus the giraffe), rodents, and primates are plotted and have exponents 1.230 ± 0.169, 1.174 ± 0.051, and 0.884 ± 0.071, respectively. (D) Variation in average number of neurons per mm2 of cortical surface area plotted as a function of the total volume of the cerebral cortex (gray and white matter combined). Artiodactyls in black, afrotherians in blue, glires in green, primates in red. Data from, and.

As a consequence, the average number of neurons underneath 1 mm2 of cortical surface, N/A, is 2–6 times smaller in artiodactyls (18,706 ± 1,176 neurons/mm2) than in primates (between 40,000 and 120,000 neurons/mm2; Figure 8D). The ratio N/A in artiodactyls does not scale with brain size, number of cortical neurons, cortical volume or surface area (Spearman correlation, $p = 0.6238$; Figure 8D).

Cortical surface area increases with numbers of neurons raised to exponents of 1.230 ± 0.169 across artiodactyls (minus the giraffe) and 1.174 ± 0.051 across rodents, but only 0.884 ± 0.071 across primates (Figure 8C). Although the relationship between cortical surface area and numbers of neurons seems overlapping in a log-log scale between artiodactyls and glires (Figure 8C), the strikingly different folding of the cortical surface suggests a different distribution of the cortical volume and its neurons along the surface. Indeed, there are three distinct patterns of surface distribution of the cortical volume in rodents, primates, and artiodactyls, visible in the relationship between cortical gray matter volume and gray matter surface area (Figure 9A). Gray matter volume increases with gray matter surface area raised to the power of 1.466 ± 0.036 in artiodactyls (minus the giraffe; $p = 0.0006$), 1.165 ± 0.026 in primates ($p \leq 0.0001$), and 1.350 ± 0.037 in rodents ($p < 0.0001$; Figure 9A). The different surface spreading of the cortical volume is more clearly apparent in the relationships between cortical gray matter thickness and surface area (Figure 9B) and between cortical thickness and number of cortical neurons (Figure 9C). Gray matter thickness increases with cortical surface area raised to the power of 0.466 ± 0.036 in artiodactyls (minus the giraffe; $p = 0.0059$), 0.350 ± 0.037 in rodents ($p = 0.0025$), and 0.165 ± 0.026 in primates ($p = 0.0001$). Gray matter thickness increases with numbers of cortical neurons raised to the power of 0.562 ± 0.118 in artiodactyls (minus the giraffe; $p = 0.0413$), 0.413 ± 0.040 in rodents ($p = 0.0019$), and 0.149 ± 0.024 in primates ($p = 0.0001$). These analyses show that the cortical volume is spread more thinly in artiodactyls than in rodents, with a thinner cortex in artiodactyls for a similar number of neurons (Figure 9C) or surface area (Figure 9B). Although artiodactyls and primates have some overlap between cortical thickness and surface area (Figure 9B), similar numbers of cortical neurons are spread more thinly in primate than in artiodactyls (Figure 9C). Yet, the thicker artiodactyl cortices (for a similar number of neurons; Figure 9C) are more folded than primate cortices of either similar numbers of neurons (Figure 8B) or similar thickness (this is the case for the pig, blesbok, and giraffe; Figure 9D), which indicates that the differential thickness of the cerebral cortex is not determinant of the degree of cortical folding.

Scaling of the cortical gray matter. (A) Variation in the volume of the cortical gray matter as a function of the total surface area of the gray matter. Power functions for artiodactyls (minus the giraffe), rodents, and primates are plotted and have exponents 1.466 ± 0.036, 1.350 ± 0.037, and 1.165 ± 0.026, respectively. Notice that the artiodactyl cortex has a larger surface area than rodents for a similar gray matter volume. (B) Variation in the

CamScanner

thickness of the cortical gray matter as a function of the total surface area of the gray matter. Power functions for artiodactyls (minus the giraffe), rodents, and primates are plotted and have exponents 0.466 ± 0.036, 0.350 ± 0.037, and 0.165 ± 0.026, respectively. For a similar surface area, the artiodactyl cortex is thinner than the rodent cortex. (C) Variation in the thickness of the cortical gray matter as a function of the total number of cortical neurons. Power functions for artiodactyls (minus the giraffe), rodents, and primates are plotted and have exponents 0.562 ± 0.118, 0.413 ± 0.040, and 0.149 ± 0.024, respectively. Again, for a similar number of cortical neurons, the artiodactyl cortex is thinner than the rodent cortex. (D) FI of the cerebral cortex plotted as a function of the average thickness of the cortical gray matter. Notice that for similar cortical thicknesses, both the artiodactyl and the primate cortex are more folded than the rodent cortex. Artiodactyls in black, afrotherians in blue, glires in green, primates in red. Data from, and.

DISTRIBUTION OF NEURONS WITHIN THE CEREBRAL CORTEX

If neurons were distributed evenly across the cortical mass in the five artiodactyl species, then the accumulated number of neurons along the anteroposterior axis should vary linearly with the accumulated mass of the cortical gray matter along the same axis and this distribution should be similar across species. In contrast, while in the kudu the distribution of cortical neurons is mirrored in the distribution of cortical mass (Figure 10A, gray points), in the pig, in the springbok, and in the giraffe, there is a slower accumulation of neurons along the anterior two thirds of the cortex, and a faster accumulation of neurons along the posterior third of the cortex. This suggests that, as was observed in the human cortex, the cerebral cortex of artiodactyls (although not of the kudu) has a larger density of neurons in the posterior regions. Indeed, the neuronal density in the cortical gray matter increases significantly from the anterior to the posterior pole across the four species combined (Spearman correlation, $\rho = 0.3121$, $p = 0.0171$; Figure 10B). Within each species, however, the correlation is only significant for the giraffe ($\rho = 0.6211$, $p = 0.0035$), and it is only in the giraffe that a one-way ANOVA shows a significant difference in neuronal density between the anterior two thirds and the posterior third of the cortex (giraffe, $p = 0.0136$; pig, $p = 0.0969$; springbok, $p = 0.1516$; kudu, $p = 0.7337$; Figure 10C). Still, in the pig and springbok, despite the non-significant correlation between neuronal density and position along the anteroposterior axis (pig, $p = 0.1175$; springbok, $p = 0.3911$), the posteriormost block of cerebral cortex is the portion that exhibits the largest neuronal density of all, with values about 4× larger than in the remainder of the cortex (Figure 10B). In the giraffe, the block with the largest neuronal density is situated 70% along the anteroposterior axis (Figure 10B).

Distribution of neurons along the anteroposterior axis of the cerebral cortex. Each artiodactyl species is represented in a different color. (A) Cumulative percentage of neurons along the anteroposterior axis of the cerebral cortex plotted as a function of the cumulative mass of the gray matter of the cerebral cortex along the same axis. Anterior is to the left. Each point represents the values for one 6 mm block of cortex along the axis. The plotted line indicates the linear relationship expected for a uniform distribution of neurons in the cortical mass along the anteroposterior axis. While the kudu fits the uniform distribution, the springbok, pig, and giraffe deviate in such a manner that the cumulative distribution of neurons along the cortical mass is better described by a different function (plotted) in the anterior 2/3 of the cortical mass. (B) Variation in neuronal density in each 6 mm block of cerebral cortex along the anteroposterior (A-P) axis. Anterior is to the left. While there is little variation in neuronal density along the cortex of the kudu, densities vary by 4–5× among cortical blocks in the pig, springbok and giraffe. (C) Difference in neuronal densities found in the anterior 2/3 and in the posterior 1/3 of the cerebral cortex in each species. Averages in each group are shown for each species. The largest neuronal densities are found in the posterior 1/3 of the cortex in the giraffe, pig, and springbok. (D) Variation in the surface density of neurons in the cortex (that is, neurons/mm2) in each 6 mm block of cerebral cortex along the anteroposterior axis. Despite the large variation along the cortex of each species, there is no systematic difference in surface density of neurons along the A-P axis.

In contrast to the trend toward larger neuronal densities in the posterior cortex, there is no correlation between the distribution of neurons per cortical surface area (neurons/mm2) and position along the anteroposterior axis ($\rho = 0.6211$, $p = 0.0035$; Figure 10D). While this does not imply that neurons are distributed evenly across the cortical surface, the lack of correlation suggests that there is no systematic variation in the distribution of neurons under the cortical surface along the anteroposterior axis.

In these four artiodactyl species, the degree of folding of the cortical surface varies about 2× along the anteroposterior axis (Figure 11A), in the same manner found in the human cerebral cortex, with less folding in the poles than in the center. Across all four species, variations in the degree of folding along the anteroposterior axis are correlated best with variations in the surface area of the white–gray matter interface ($\rho = 0.7459$, $p < 0.0001$; Figure 11B), but also with variations in the surface area of the gray matter ($\rho = 0.6822$, $p < 0.0001$; Figure 11C), in numbers of cortical neurons ($\rho = 0.6721$, $p < 0.0001$; Figure 11D) and in cortical thickness ($\rho = 0.5361$, $p < 0.0001$; Figure 11E). Notice that the correlation with cortical thickness is positive, that is, regions where the cortical gray matter is thicker tend to be more folded. Similarly, within each species separately, the best predictor of variations in FI along the anteroposterior axis is the surface area of the white–gray matter interface (all species, $\rho > 0.6$, $p < 0.02$; Figure 11B); gray matter surface area is not correlated with FI in the giraffe ($p = 0.1906$), numbers of cortical neurons are not correlated with FI in the springbok ($p = 0.1121$), and cortical thickness is not correlated with FI in the giraffe ($p = 0.1844$) and in the pig ($p = 0.5772$). Thus, variations in cortical folding along the cortical surface appear to be best predicted by variations in the extent of the white–gray matter surface – although the relationship is not a single one for the four species of artiodactyls (Figure 11B).

CamScanner

Variation in FI within the cerebral cortex. Each artiodactyl species is represented in a different color. (A) Variation in FI of the cortex along the anteroposterior axis of the cerebral cortex. Anterior is to the left. Each point represents the values for one 6 mm block of cortex along the axis. The smallest folding indices in the poles are a trivial consequence of the shape of the cortex in relation to the axis of measurement (that is, of the surface becoming orthogonal to the axis at the poles), but the overall shape of the distribution indicates a larger folding of the cortex in the intermediate positions along the A-P axis in artiodactyls, as in primates. (B–E) Variation in FI amongst blocks of cerebral cortex along the A-P axis for each species as a function of the surface area of the white matter in the block (B), surface area of the gray matter in the block (C), number of cortical neurons in the block (D), or average thickness of the gray matter in the block (E).

PREDICTIONS FOR CETACEANS

Artiodactyls are members of the order Cetartiodactyla, which includes the closely related cetaceans, large animals with large brains. So far, we have found that neuronal scaling rules apply to all species within a given order (primates, rodents) or superorder (afrotherians), and even across orders, as in the case of the cerebral cortex of rodents, lagomorphs, afrotherians, and artiodactyls. Because the total mass of the artiodactyl cerebral cortex conforms to the neuronal scaling rules that apply to the ensemble of artiodactyls (minus the giraffe), glires, and afrotherians, and given that cetaceans and artiodactyls belong to the same mammalian order, the rules that apply to the ensemble of glires, afrotherians, and artiodactyls can be used to predict the neuronal composition of the cetacean cerebral cortex. The prediction is given by the equation NCXT = e17.327±0.045. MCXT0.590±0.018, where NCXT is the total number of neurons in the cerebral cortex and MCXT it the total mass of the cerebral cortex (gray + white matter, in g). This equation is obtained for the relationship between cortical mass and number of neurons across glires, afrotherians, and artiodactyls (minus the giraffe) written to have number of cortical neurons as a dependent variable and cortical mass as the independent variable. Using the cortical volumes given in cm3 by, which can be approximated to cortical mass in grams (gray + white matter combined), we predict the cerebral cortex of Phocoena phocoena, Tursiops truncatus, Grampus griseus, and Globicephala macrorhyncha, at 340, 815, 1,127, and 2,045 cm3, to be composed of 1.04, 1.75, 2.11, and 3.01 billion neurons, respectively.

DISCUSSION

Here we determined the numbers of neuronal and non-neuronal cells that compose the brains of five artiodactyl species. We find that the artiodactyl cerebral cortex and cerebellum conform to the neuronal scaling rules that apply to afrotherians, glires, and eulipotyphlans, which we have suggested also apply to the last common ancestor of modern eutherians. Although the volume of the cerebral cortex is a shared function of its number of neurons across artiodactyls, afrotherians, rodents, and eulipotyphlans, the surface distribution of this volume differs across clades, resulting in widely different numbers of neurons per unit of surface area across the species, and in cortices that fold to different extents in different species, even when composed of similar numbers of neurons. These dissimilarities underline one common principle in brain evolution that becomes more and more clear: that there is no single common principle regarding how neurons build the brain, such that all mammalian brains are not the same, although patterns are beginning to emerge that illuminate the ways in which mammalian brains have evolved.

We will next discuss how the simple determination of the cellular composition of the brain of five artiodactyl species, in comparison to a total of 31 other mammalian species, illuminates several of the most basic aspects of brain organization, development, and evolution.

NEURONAL SCALING RULES

Upon realizing that even closely related clades such as rodents and primates had brain structures formed according to strikingly different neuronal scaling rules, we had initially supposed that each mammalian order might have its own characteristic set of neuronal scaling rules dictating how the size of each brain structure scales with number of neurons in the structure. While we then found eulipotyphlans to share neuronal scaling rules for the cerebral cortex, but not for the cerebellum, with rodents and lagomorphs (that is, glires), we recently found that the afrotherian neuronal scaling rules for both the cerebral cortex and the cerebellum are also shared with glires. Given the position of Afrotheria closer to the base of the Eutheria evolutionary tree, the sharing of neuronal scaling rules for the cerebral cortex by afrotherians, glies, and eulipotyphlans raised the possibility that those were the ancestral rules that applied to the formation of the cerebral cortex of mammals. The present finding that the neuronal scaling rules for the cerebral cortex and cerebellum of afrotherians and glires are also shared by artiodactyls strengthens this possibility and suggests that primates and eulipotyphlans diverged from the ancestral rules, which however still apply to Artiodactyla.

The existence of ancestral neuronal scaling rules that remained conserved in the evolution of some clades is further supported by the finding that the rules that apply to the rest of the brain (brainstem plus diencephalon and basal ganglia), which are evolutionarily older structures than the cerebral cortex, are shared by afrotherians, glires, eulipotyphlans, and even primates – although possibly not by artiodactyls, a more recent group than the others.

Unfortunately, the giraffe brain specimen that we had available for analysis was not from a fully grown adult, but only a juvenile, of about half the expected adult body mass and subadult brain mass (470 and 528 g compared to 1,200 and 700g, respectively). Although we do not know how artiodactyl development compares to rodents and primates, we presume that the final stages of development that expand brain mass, with extension of arbors, resulting in increased neuronal cell mass, and with addition of large numbers of non-neuronal cells, were not yet

CamScanner

completed, although the mechanisms that determine final numbers of neurons were possibly already completed, such that the average total number of neurons in the giraffe brain is possibly well represented by the numbers found here. Such an incomplete developmental brain expansion would provide a simple explanation for our findings of structure masses that are well below the expected values for the numbers of neurons in the structures in comparison to the other artiodactyls in the sample. It was based on this rationale that we excluded the giraffe data from the analyses, although they were presented for comparison.

NON-NEURONAL SCALING RULES

We show that artiodactyls share non-neuronal scaling rules not only with afrotherians and glires, but also with eulipotyphlans and even with primates, and across all brain structures. “Non-neuronal cells” are not only glial cells (of all different types) but also endothelial cells and pericytes. However, given that the vascular volume has been measured at less than 3% of the cerebral cortex of the human brain as well as in rodents and other primates (reviewed in), we presume that glial cells are the vast majority of non-neuronal cells. Thus, the shared relationship between the mass of brain structures and their numbers of non-neuronal cells indicates that glial cells are added to different brain structures in different species following similar rules and mechanisms that have been conserved in evolution. This is illustrated by the finding that primate and artiodactyl brains of similar sizes are composed of strikingly different numbers of neurons, but similar numbers of non-neuronal (mostly glial) cells. For example, the pig brain, at 68.2 g, has 2.2 billion neurons, almost half than the slightly smaller bonnet monkey brain, at 62 g, which has 3.8 billion neurons, but both have fairly similar numbers of other cells: 4.6 billion in the pig, 4.9 billion in the bonnet monkey. Such evolutionary conservation of glial scaling rules, especially in the face of diversity in the neuronal scaling rules, suggests that physiological constraints are in place that limit diversification in how glial cells, of all glial subtypes, are added to the neuronal parenchyma in development.

One consequence of the evolutionarily conserved, shared glial scaling rules is the uniform increase in the O/N ratio, and thus in the glia/neuron ratio, with decreasing neuronal density. We have shown that the glia/neuron ratio increases as a consequence of increasing average neuronal cell mass following the same rules across brain structures and species. In the present artiodactyls, we find glia/neuron ratios that exceed 7–8 within the cortical gray matter, 10 in the cerebral cortex as a whole, and 20 or even 30 in the brainstem. Notice that these artiodactyl cortices reach at most 1/3 of the mass of the human cerebral cortex, and yet have glia/neuron ratios in the cortical gray matter that are as much as 5× larger than in the human gray matter. This difference illustrates that the glia/neuron ratio is not a function of brain mass or cortical mass, but rather of neuronal density, which may or may not vary with cortical mass, as we have pointed out before.

COMPARISON WITH PRIMATES

One result of the scaling of the artiodactyl cerebral cortex conforming to the neuronal scaling rules that relate structure mass to numbers of neurons in afrotherians and glires, but not primates, is that the cerebral cortex is much larger in artiodactyls than in primates with similar numbers of neurons. For instance, the springbok and the owl monkey have cerebral cortices with around 400 million neurons (springbok, 375 million; owl monkey, 442 million), but the springbok cortex, at 62.8 g, is nearly 6× larger than the owl monkey cortex, which weighs only 10.6 g. Similarly, the kudu and the pig-tailed monkey both contain ca. 800 million neurons in the cerebral cortex, but the structure weighs 212 g in the kudu, in contrast to only 36.2 g in the macaque; and the giraffe has as many cortical neurons as a rhesus monkey (1.7 billion), but in a structure that weighs 390 g in a juvenile, versus only 70 g in an adult rhesus monkey. The gorilla, the primate species with a comparable cortical mass to the giraffe, has, in contrast, an estimated 8.9 billion cortical neurons, or 5× more neurons in the cerebral cortex than the giraffe. We speculate that these large differences in numbers of cortical neurons result in very different cognitive abilities between artiodactyl and primate species of similar cortical mass such as the giraffe and the gorilla (at ~400 g), and the pig and the pig-tailed monkey (at ~36 g), potentially endowing primates with more processing power, behavioral complexity, and flexibility than their artiodactyl counterparts.

Moreover, non-sensory, non-motor and thus presumably associative cortical areas in artiodactyls are limited to the very frontal pole of the cortex, while in great apes (and human) they amount to 14% of the cortical gray matter volume. Thus, the distribution of similar numbers of neurons between sensorimotor and associative areas is also likely to contribute to the different cognitive abilities of artiodactyls and primates with similar numbers of cortical neurons. When compounded with the larger average size of neurons in artiodactyls than in primates for a similar number of cortical neurons, we predict that the gorilla not only has 5× more cortical neurons overall than the giraffe, for a similar cortical mass, but it has even more than 5× more neurons available for associative (prefrontal) processing, rather than purely sensorimotor processing, than the giraffe. We are now examining these differences across clades in numbers of associative × sensorimotor neurons directly.

Our data show that the larger cortical volume in artiodactyls than in primates for a similar number of neurons is not simply attributable to a disproportionately larger white matter. While a similar number of cortical neurons is indeed accompanied by a nearly 10× larger volume of white matter in artiodactyls than in primates, those neurons reside in a cortical gray matter that is also nearly 10× larger in volume in artiodactyls than in primates, such that the relationship between gray and white matter volumes is overlapping across both clades. The difference is that the same volumes of gray or white matter comprise far fewer neurons in artiodactyls than in primates. Given that non-neuronal densities are comparable across the clades, it can be inferred that artiodactyl neurons, with cell parts in both gray and white matters, are about 10× larger in artiodactyls than in primates. Our results thus demonstrate

CamScanner

that there is not a single way of building a large cerebral cortex in nature. Still, the fairly overlapping distributions of gray and white matter volumes across clades suggest the existence of common principles underlying the organization of the two portions of the cortical matter, that is, principles that link the volume of white matter to the volume of the gray matter. As we demonstrate here, however, such common principles are not related simply to sheer numbers of cortical neurons; instead, we suggest that they are related directly to the volume of the gray matter, which in turn results from variations in both number and average size of the neurons therein.

DISTRIBUTION OF CORTICAL NEURONS WITHIN THE CORTICAL SURFACE AND IMPLICATIONS FOR CORTICAL EXPANSION IN EVOLUTION

We have previously shown that the visual cortex of mouse and man has a different neuronal composition that the rest of the cerebral cortex, in agreement with previous findings, which prompted us to perform a similar systematic analysis of the neuronal composition of the cortical volume in artiodactyls. Our current analysis, however, was hampered by the shortage of data on functional localization in the cerebral cortex of artiodactyls and by the disagreement between the few sources available regarding the position of the visual cortex. Using microelectrode recordings in the llama, placed its visual cortex medial to the suprasylvian fissure, along the dorsalmost cortical surface; the auditory cortex lateral to the suprasylvian fissure; and somatosensory areas rostrally in the cortex, although not extending all the way to the frontal pole. In contrast,, studying the sheep, placed the visual cortex in the occipitalmost areas, as found in primates, much more posteriorly than described by.

Our findings of a trend toward larger neuronal densities in the posterior third than in the anterior two thirds of the cortex is compatible with an occipital location of a visual cortex with a distinctly high neuronal density, at least in some species, such as the pig and springbok. However, the largest neuronal densities occur in the posteriormost cortex of the pig and springbok, but at about 70% of the giraffe, while there is no section with a noticeably larger neuronal density in the kudu. While these differences may be due to a lack of resolution in our sampling, for instance because a dorsal (rather than occipital) visual cortex would have been mixed with non-visual areas in the same sections, they are compatible with a scenario in which the location of the visual cortex is not shared across artiodactyl species, being occipital in the pig and springbok (as in the sheep) but more rostral in the giraffe (and llama). Thus, while we show evidence of heterogeneity in the distribution of neurons along the cortical surface in three of four species of artiodactyls, with zones of much higher neuronal densities than in the remainder of the cortex, there is not enough evidence at present to assign these regions of higher neuronal density to the visual cortex.

In contrast, we do show that there is no systematic variation in the average number of neurons per surface area of cerebral cortex along the anteroposterior axis, contrary to what has been found in the cortex of human and non-human primates, and to what has been proposed by Barbara Finlay’s group to be the basis of the expansion of the cortical surface: a front-to-back gradient in the start of neurogenesis, which would lead to a front-to-back gradient in the surface density of neurons. If such a gradient in neurogenesis does indeed lead to a matching gradient in the surface density of neurons, then our data suggest that there is no gradient in the timing of neurogenesis in the formation of the artiodactyl cortex. Alternatively, the existence of such a gradient in the development of the artiodactyl cortex would provide evidence that a rostrocaudal gradient in the timing of neurogenesis does not necessarily lead to a gradient in the surface density of neurons along the cortex.

It has been suggested that the extension of the developmental interval required to give rise to larger total numbers of cortical neurons across species would occur through a combination of (1) an increase in the difference in timing of neurogenesis in rostral versus caudal cortex and (2) late-born supragranular neurons becoming disproportionately more numerous in the caudal than in the rostral pole relative to early born infragranular neurons. Such a combination would supposedly result necessarily in steeper gradients in the surface density of neurons in those cortices with more neurons, such as in primates compared to rodents. Thus, if such a mechanism in which the extension of neurogenesis necessarily leads to steep gradients in the surface density of neurons were the universal mechanism whereby cortical expansion occurs across mammalian species, then cerebral cortices with larger numbers of neurons should necessarily exhibit steeper gradients than cortices with fewer neurons. In contrast, here we find that the pig and the springbok, with numbers of cortical neurons in the range between a marmoset and an owl monkey, and the kudu, with more cortical neurons than the owl monkey, exhibit no significant variation in surface density of neurons along the rostrocaudal axis of the cerebral cortex, in contrast to the primate species with similar numbers of cortical neurons, which do exhibit sharp gradients. The discrepancy indicates that, contrary to the proposition of, there is no single mechanism for expanding the cerebral cortex across mammals; an overall larger number of neurons may (as in primates) or may not (as in artiodactyls) occur in the presence of a rostrocaudal gradient in the surface density of neurons.

Instead, we suggest that such gradients result from variations in connectivity-related parameters that determine the extent of cortical surface formed by a given number of neurons, depending on their lateral spread, such as the fraction of gray matter neurons that are connected through the white matter and the average caliber of those fibers. Thus, given a different pattern of connectivity of cortical neurons through the white matter (amongst other variables, such as the average size of the neurons in the gray matter), similar numbers of cortical neurons may give rise to smaller or larger cortical surfaces in different species, with or without establishing a gradient in surface density along the rostrocaudal axis.

CamScanner

Finally, the increasing degree of gyrification that is found to accompany increasing cortical volume is usually credited to the expansion of the number of neurons in the cerebral cortex, whether homogeneously, through the lateral addition of columnar modules or for instance via a conical expansion of the outer subventricular zone. In contrast, here we show that the artiodactyl cerebral cortices are more folded than primate cerebral cortices of similar numbers of neurons. Thus, the degree of cortical gyrification is not necessarily proportional to the number of neurons in the cortex, and for this reason, the mechanisms driving cortical folding must be sought in events other than simply the regulation of neurogenesis. This will be the subject of a separate investigation.

PREDICTIONS FOR CETACEANS

Here we examined only the artiodactyl branch of the Cetartiodactyla; determining whether the present scaling rules that apply to artiodactyls extend to the whole of Cetartiodactyla will require examining cetacean brains, which we are undertaking now. Still, given our previous findings that species in the same order (or even superorder, as is the case of Afrotheria) share the same neuronal scaling rules, it is plausible that the rules that we found here for artiodactyls will also apply to cetacean brains.

If this is the case, we predict that the large cerebral cortex of several cetacean species, such as the pilot whale Globicephala macrorhyncha, which is about twice larger than the human cerebral cortex, is composed of only around 3 billion neurons. This is at odds with a previous stereological estimate of 13 billion neurons in the cerebral cortex of the minke whale, Balaenoptera acutorostrata, at a density of about 8,000 neurons/mm3 of its 1,622 cm3 of gray matter. The same group later estimated an even larger total of 15 billion neurons in the smaller cerebral cortex of the harbor porpoise. Those studies, however, may have been biased by undersampling, given that the authors sampled only 12–13 sections out of over 3,000 sections of the minke whale cerebral cortex, counting a total of only about 100 disectors per entire cortex, with a total of around 200 cells sampled from the entire cortex (; the study on the harbor porpoise does not provide information on the sampling fraction and other parameters used). In contrast, our method samples the entire cerebral cortex, which removes any concerns of bias due to undersampling. The neuronal density that report in the cortical gray matter of the minke whale, 8,000 neurons/mm3, is even higher than those densities we find in much smaller cerebral cortices of other cetartiodactyls (see Table 1), whereas, given the rapid decrease in neuronal density with increasing cortical mass, we would expect densities of no more than 5,000 neurons/mm3, and probably fewer, in cetacean cortices. Thus, there are two possible scenarios: that estimate is inflated by undersampling, which would leave open the possibility that cetaceans do fit the rules reported here for other Cetartiodactyls; or that our estimates are underestimates, and cetaceans have larger neuronal densities and total numbers of neurons than can be predicted from the scaling rules reported here. However, given that we sample the entire cerebral cortex using a method that has been shown to be as accurate as stereology, while they use stereology to sample only a handful of sections out of over 3,000 sections with a small number of disectors, we feel the first scenario is likely. The direct examination of cetacean brains using our method will show whether or not the scaling rules for artiodactyl brains do extend to Cetartiodactyla as a whole. Such a study will also illuminate what happened in evolution when modern cetaceans branched off from the common ancestor with modern artiodactyls.

In the meantime, the three billion neurons that we predict to compose the cerebral cortex of the pilot whale is about one-fifth of the number of neurons found in the human cerebral cortex, even though the pilot whale cortex is twice larger than the human cortex. We predict, therefore, that even the largest cetacean cerebral cortex still has fewer neurons than the human cerebral cortex, supporting our hypothesis that the remarkable cognitive abilities of the human brain compared to even larger brains are related to the remarkable number of neurons in its brain, despite its modest size.

Conflict of Interest Statement

The Associate Editor Chet C. Sherwood declares that, despite having collaborated with one of the authors (Paul R. Manger), the review process was handled objectively. José Maldonado is the representative in Latin America for MBF Bioscience, the company that markets the NeuroLucida software employed in this study. The other authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

REFERENCES

Equal numbers of neuronal and non-neuronal cells make the human brain an isometrically scaled-up primate brain.

Validation of the isotropic fractionator: comparison with unbiased stereology and DNA extraction for quantification of glial cells.

Changing numbers of neuronal and non-neuronal cells underlie postnatal brain growth in the rat.

Comparative analysis of encephalization in mammals reveals relaxed constraints on anthropoid primate and cetacean brain scaling.

Systematic, balancing gradients in neuron density and number across the primate isocortex.

Systematic, cross-cortex variation in neuron numbers in rodents and primates.

Interstitial cells of the adult neocortical white matter are the remnant of the early generated subplate neuron population.

Scalable architecture in mammalian brains.

CamScanner

The cortical visual areas of the sheep.

Total neocortical cell number in the mysticete brain.

Cellular scaling rules for the brains of an extended number of primate species.

Neuronal density, size and shape in the human anterior cingulate cortex: a comparison of Nissl and NeuN staining.

Brain sizes, surfaces, and neuronal sizes of the cortex cerebri: a stereological investigation of man and his variability and a comparison with some mammals (primates, whales, marsupials, insectivores, and one elephant).

The human brain in numbers: a linearly scaled-up primate brain.

Not all brains are made the same: new views on brain scaling in evolution.

Neuronal scaling rules for primate brains: the primate advantage.

The glia/neuron ratio: how it varies uniformly across brain structures and species and what that means for brain physiology and evolution.

Cellular scaling rules for primate brains.

Gorilla and orangutan brains conform to the primate scaling rules: implications for hominin evolution.

Isotropic fractionator: a simple, rapid method for the quantification of total cell and neuron numbers in the brain.

Brain scaling in mammalian evolution as a consequence of concerted and mosaic changes in numbers of neurons and average neuronal cell size.

Cellular scaling rules for rodent brains.

Connectivity-driven white matter scaling and folding in primate cerebral cortex.

Updated neuronal scaling rules for the brains of Glires (rodents/lagomorphs).

Distribution of neurons in functional areas of the mouse cerebral cortex reveals quantitatively different cortical zones.

Size and shape of the cerebral cortex in mammals. I. The cortical surface.

“The effects of domestication on brain size,” in

Morphometry of the human cerebral cortex microcirculation: general characteristics and space-related profiles.

Conical expansion of the outer subventricular zone and the role of neocortical folding in evolution and development.

Acquisition of brains from the African elephant (Loxodonta africana): perfusion-fixation and dissection.

Quantitative analysis of neocortical gyrencephaly in African elephants (Loxodonta africana) and six species of cetaceans: comparison with other mammals.

Three counting methods agree on cell and neuron number in chimpanzee primary visual cortex.

How the cortex gets its folds: a connectivity-based universal model of mammalian gyrification.

All brains are made of this: a fundamental building block of brain matter with matching neuronal and glial masses.

NeuN, a neuronal specific nuclear protein in vertebrates.

Mammalian phylogenomics comes of age.

Cellular scaling rules for the brains of Afrotheria.

Order-specific quantitative patterns of cortical gyrification.

Human and animal cognition: continuity and discontinuity.

Specification of cerebral cortical areas.

The human cerebral cortex is neither one nor many: neuronal distribution reveals two quantitatively different zones in the grey matter, three in the white matter, and explains local variations in cortical folding.

The basic uniformity in structure of the neocortex.

Evolution of the brain and intelligence.

Cellular scaling rules of insectivore brains.

Volumen- und Oberflächenmessungen an Gehirnen verschiedener Säugetiere im Vergleich zu einem errechneten Modell.

Prefrontal white matter volume is disproportionately larger in humans than in other primates.

Evolution of increased glia-neuron ratios in the human frontal cortex.

Activity of acetylcholine system in cerebral cortex of various unanesthetized mammals.

Correlations of neuronal and microvascular densities in murine cortex revealed by direct counting and colocalization of nuclei and vessels.

CamScanner

Different scaling of white matter volume, cortical connectivity, and gyrification across rodent and primate brains.

A neurological comparative study of the harp seal (Pagophilus groenlandicus) and harbor porpoise (Phocoena phocoena) brain.

Somatic sensory cortex of Llama (Lama glama).

The control of neuron number.

The human pattern of gyrification in the cerebral cortex.

CamScanner

Jonas & Kording (2017) — Could a Neuroscientist Understand a Microprocessor?

CC BY; texto integral via PubMed Central. Repositório: https://pmc.ncbi.nlm.nih.gov/articles/PMC5230747/

Licença: Creative Commons Attribution (CC BY). A atribuição original acompanha este texto; consulte a página do artigo para metadados completos e avisos de correção.

Could a Neuroscientist Understand a Microprocessor?

There is a popular belief in neuroscience that we are primarily data limited, and that producing large, multimodal, and complex datasets will, with the help of advanced data analysis algorithms, lead to fundamental insights into the way the brain processes information. These datasets do not yet exist, and if they did we would have no way of evaluating whether or not the algorithmically-generated insights were sufficient or even correct. To address this, here we take a classical microprocessor as a model organism, and use our ability to perform arbitrary experiments on it to see if popular data analysis methods from neuroscience can elucidate the way it processes information. Microprocessors are among those artificial information processing systems that are both complex and that we understand at all levels, from the overall logical flow, via logical gates, to the dynamics of transistors. We show that the approaches reveal interesting structure in the data but do not meaningfully describe the hierarchy of information processing in the microprocessor. This suggests current analytic approaches in neuroscience may fall short of producing meaningful understanding of neural systems, regardless of the amount of data. Additionally, we argue for scientists using complex non-linear dynamical systems with known ground truth, such as the microprocessor as a validation platform for time-series and structure discovery methods.

Author Summary

Neuroscience is held back by the fact that it is hard to evaluate if a conclusion is correct; the complexity of the systems under study and their experimental inaccessability make the assessment of algorithmic and data analytic technqiues challenging at best. We thus argue for testing approaches using known artifacts, where the correct interpretation is known. Here we present a microprocessor platform as one such test case. We find that many approaches in neuroscience, when used naïvely, fall short of producing a meaningful understanding.

Introduction

CamScanner

The development of high-throughput techniques for studying neural systems is bringing about an era of big-data neuroscience. Scientists are beginning to reconstruct connectivity, record activity, and simulate computation at unprecedented scales. However, even state-of-the-art neuroscientific studies are still quite limited in organism complexity and spatiotemporal resolution. It is hard to evaluate how much scaling these techniques will help us understand the brain.

In neuroscience it can be difficult to evaluate the quality of a particular model or analysis method, especially in the absence of known truth. However, there are other systems, in particular man made ones that we do understand. As such, one can take a human-engineered system and ask if the methods used for studying biological systems would allow understanding the artificial system. In this way, we take as inspiration Yuri Lazbnick's well-known 2002 critique of modeling in molecular biology, "Could a biologist fix a radio?". However, a radio is clearly much simpler than the nervous system, leading us to seek out a more complex, yet still well-understood engineered system. The microprocessors in early computing systems can serve this function.

Here we will try to understand a known artificial system, a classical microprocessor by applying data analysis methods from neuroscience. We want to see what kind of an understanding would emerge from using a broad range of currently popular data analysis methods. To do so, we will analyze the connections on the chip, the effects of destroying individual transistors, single-unit tuning curves, the joint statistics across transistors, local activities, estimated connections, and whole-device recordings. For each of these, we will use standard techniques that are popular in the field of neuroscience. We find that many measures are surprisingly similar between the brain and the processor but that our results do not lead to a meaningful understanding of the processor. The analysis can not produce the hierarchical understanding of information processing that most students of electrical engineering obtain. It suggests that the availability of unlimited data, as we have for the processor, is in no way sufficient to allow a real understanding of the brain. We argue that when studying a complex system like the brain, methods and approaches should first be sanity checked on complex man-made systems that share many of the violations of modeling assumptions of the real system.

An engineered model organism

The MOS 6502 (and the virtually identical MOS 6507) were the processors in the Apple I, the Commodore 64, and the Atari Video Game System (VCS) (see for a comprehensive review). The Visual6502 team reverse-engineered the 6507 from physical integrated circuits by chemically removing the epoxy layer and imaging the silicon die with a light microscope. Much like with current connectomics work, a combination of algorithmic and human-based approaches were used to label regions, identify circuit structures, and ultimately produce a transistor-accurate netlist (a full connectome) for this processor consisting of 3510 enhancement-mode transistors. Several other support chips, including the Television Interface Adaptor (TIA) were also reverse-engineered and a cycle-accurate simulator was written that can simulate the voltage on every wire and the state of every transistor. The reconstruction has sufficient fidelity to run a variety of classic video games, which we will detail below. The simulation generates roughly 1.5GB/sec of state information, allowing a real big-data analysis of the processor.

The simplicity of early video games has led to their use as model systems for reinforcement learning and computational complexity research. The video game system ("whole animal") has a well defined output in each of the three behavioral conditions (games). It produces an input-dependent output that is dynamic, and, in the opinion of the authors, quite exciting. It can be seen as a more complex version of the Mus Silicium project. It is also a concrete implementation of a thought experiment that has been mentioned on and off in the literature. The richness of the dynamics and our knowledge about its inner workings makes it an attractive test case for approaches in neuroscience.

Here we will examine three different "behaviors", that is, three different games: Donkey Kong (1981), Space Invaders (1978), and Pitfall (1981). Obviously these "behaviors" are qualitatively different from those of animals and may seem more complicated. However, even the simple behaviors that are studied in neuroscience still involve a plethora of components, typically including the allocation of attention, cognitive processing, and multiple modalities of inputs and outputs. As such, the breadth of ongoing computation in the processor may actually be simpler than those in the brain.

The objective of clever experimental design in neuroscience often is to find behaviors that only engage one kind of computation in the brain. In the same way, all our experiments on the chip will be limited by us only using these games to probe it. As much as more neuroscience is interested in naturalistic behaviors, here we analyze a naturalistic behavior of the chip. In the future it may be possible to excute simpler, custom code on the processor to tease apart aspects of computation, but we currently lack such capability in biological organisms.

Much has been written about the differences between computation in silico and computation in vivo —the stochasticity, redundancy, and robustness present in biological systems seems dramatically different from that of a microprocessor. But there are many parallels we can draw between the two types of systems. Both systems consist of interconnections of a large number of simpler, stereotyped computing units. They operate on multiple timescales. They consist of somewhat specialized modules organized hierarchically. They can flexibly route information and retain memory over time. Despite many differences there are also many similarities. We do not wish to overstate this case—in many ways, the functional specialization present in a large mammalian brain far eclipses that present in the processor. Indeed, the processor's scale and specialization share more in common with C. elegans than a mouse.

CamScanner

Yet many of the differences should make analysing the chip easier than analyzing the brain. For example, it has a clearer architecture and far fewer modules. The human brain has hundreds of different types of neurons and a similar diversity of proteins at each individual synapse, whereas our model microprocessor has only one type of transistor (which has only three terminals). The processor is deterministic while neurons exhibit various sources of randomness. With just a couple thousand transistors it is also far smaller. And, above all, in the simulation it is fully accessible to any and all experimental manipulations that we might want to do on it.

What does it mean to understand a system

Importantly, the processor allows us to ask “do we really understand this system?” Most scientists have at least behavioral-level experience with these classical video game systems, and many in our community, including some electrophysiologists and computational neuroscientists, have formal training in computer science, electrical engineering, computer architecture, and software engineering. As such, we believe that most neuroscientists may have better intuitions about the workings of a processor than about the workings of the brain.

What constitutes an understanding of a system? Lazbnick’s original paper argued that understanding was achieved when one could “fix” a broken implementation. Understanding of a particular region or part of a system would occur when one could describe so accurately the inputs, the transformation, and the outputs that one brain region could be replaced with an entirely synthetic component. Indeed, some neuroengineers are following this path for sensory and memory systems. Alternatively, we could seek to understand a system at differing, complementary levels of analysis, as David Marr and Tomaso Poggio outlined in 1982. First, we can ask if we understand what the system does at the computational level: what is the problem it is seeking to solve via computation? We can ask how the system performs this task algorithmically: what processes does it employ to manipulate internal representations? Finally, we can seek to understand how the system implements the above algorithms at a physical level. What are the characteristics of the underlying implementation (in the case of neurons, ion channels, synaptic conductances, neural connectivity, and so on) that give rise to the execution of the algorithm? Ultimately, we want to understand the brain at all these levels.

In this paper, much as in systems neuroscience, we consider the quest to gain an understanding of how circuit elements give rise to computation. Computer architecture studies how small circuit elements, like registers and adders, give rise to a system capable of performing general-purpose computation. When it comes to the processor, we understand this level extremely well, as it is taught to most computer science undergraduates. Knowing what a satisfying answer to “how does a processor compute?” looks like makes it easy to evaluate how much we learn from an experiment or an analysis.

What would a satisfying understanding of the processor look like?

We can draw from our understanding of computer architecture to firmly ground what a full understanding of a processor would look like (Fig 1). The processor is used to implement a computing machine. It implements a finite state machine which sequentially reads in an instruction from memory (Fig 1a, green) and then either modifies its internal state or interacts with the world. The internal state is stored in a collection of byte-wide registers (Fig 1a, red). As an example, the processor might read an instruction from memory telling it to add the contents of register A to the contents of register B. It then decodes this instruction, enabling the arithmetic logic unit (ALU, Fig 1a, blue) to add those registers, storing the output. Optionally, the next instruction might save the result back out to RAM (Fig 1a, yellow). It is this repeated cycle that gives rise to the complex series of behaviors we can observe in this system. Note that this description in many ways ignores the functions of the individual transistors, focusing instead on circuits modules like “registers” which are composed of many transistors, much as a systems neuroscientist might focus on a cytoarchitecturally-distinct area like hipppocampus as opposed to individual neurons.

A microprocessor is understood at all levels.

(A) The instruction fetcher obtains the next instruction from memory. This then gets converted into electrical signals by the instruction decoder, and these signals enable and disable various internal parts of the processor, such as registers and the arithmetic logic unit (ALU). The ALU performs mathematical operations such as addition and subtraction. The results of these computations can then be written back to the registers or memory. (B) Within the ALU there are well-known circuits, such as this one-bit adder, which sums two one-bit signals and computes the result and a carry signal. (C) Each logic gate in (B) has a known truth table and is implemented by a small number of transistors. (D) A single NAND gate is comprised of transistors, each transistor having three terminals (E). We know (F) the precise silicon layout of each transistor.

Each of the functions within the processor contains algorithms and a specific implementation. Within the arithmetic logic unit, there is a byte wide adder, which is in part made of binary adders (Fig 1b), which are made out of AND/NAND gates, which are made of transistors. This is in a similar way as the brain consists of regions, circuits, microcircuits, neurons, and synapses.

If we were to analyze a processor using techniques from systems neuroscience we would hope that it helps guide us towards the descriptions that we used above. In the rest of the paper we will apply neuroscience techniques to data from the processor. We will finally discuss how neuroscience can work towards techniques that will make real progress at moving us closer to a satisfying understanding of computation, in the chip, and in our brains.

Results

CamScanner

Validating our understanding of complex systems is incredibly difficult when we do not know the actual ground truth. Thus we use an engineered system, the MOS6502, where we understand every aspect of its behavior at many levels. We will examine the processor at increasingly-fine spatial and temporal resolutions, eventually achieving true “big-data” scale: a “processor activity map”, with every transistor state and every wire voltage. As we apply the various techniques that are currently used in neuroscience we will ask how the analyses bring us closer to an understanding of the microprocessor (Fig 2). We will use this well defined comparison to ask questions about the validity of current approaches to studying information processing in the brain.

Optical reconstruction of the microprocessor to obtain its connectome.

In, the (A) MOS 6502 silicon die was examined under a visible light microscope (B) to build up an image mosaic (C) of the chip surface. Computer vision algorithms were used to identify metal and silicon regions (E) to detect transistors (F), (G) ultimately producing a complete accurate netlist of the processor (D).

Connectomics

The earliest investigations of neural systems were in-depth anatomical inquiries. Fortunately, through large scale microscopy (Fig 2a) we have available the full 3d connectome of the system. In other words, we know how each transistor is connected to all the others. The reconstruction is so good, that we can now simulate this processor perfectly—indeed, were it not for the presence of the processor’s connectome, this paper would not have been possible. This process is aided by the fact that we know a transistor’s deterministic input-output function, whereas neurons are both stochastic and vastly more complex.

Recently several graph analysis methods ranging from classic to modern approaches have been applied to neural connectomes. The approach in was also applied to a region of this processor, attempting to identify both circuit motifs as well as transistor “types” (analogous to cell types) in the transistor wiring diagram. Fig 3 (adapted from) shows the results of the analysis. We see that one identified transistor type contains the “clocked” transistors, which retain digital state. Two other types contain transistors with pins C1 or C2 connected to ground, mostly serving as inverters. An additional identified type controls the behavior of the three registers of interest (X, Y, and S) with respect to the SB data bus, either allowing them to latch or drive data from the bus. The repeat patterns of spatial connectivity are visible in Fig 3a, showing the man-made horizontal and vertical layout of the same types of transistors.

Discovering connectivity and cell type.

Reproduced from. (A) The spatial distribution of the transistors in each cluster show a clear pattern (B) The clusters and connectivity versus distance for connections between Gate and C1, Gate and C2, and C1 and C2 terminals on a transistor. Purple and yellow types have a terminal pulled down to ground and mostly function as inverters. The blue types are clocked, stateful transistors, green control the ALU and orange control the special data bus (SDB).

While superficially impressive, based on the results of these algorithms we still can not get anywhere near an understanding of the way the processor really works. Indeed, we know that for this processor there is only one physical “type” of transistor, and that the structure we recover is a complex combination of local and global circuitry.

In neuroscience, reconstructing all neurons and their connections perfectly is the dream of a large community studying connectomics. Current connectomics approaches are limited in their accuracy and ability to definitively identify synapses, Unfortunately, we do not yet have the techniques to also reconstruct the i/o function–neurotransmitter type, ion channel type, I/V curve of each synapse, etc.—of each neuron. But even if we did, just as in the case of the processor, we would face the problem of understanding the brain based on its connectome. As we do not have algorithms that go from anatomy to function at the moment that go considerably beyond cell-type clustering it is far from obvious how a connectome would allow an understanding of the brain.

Note we are not suggesting connectomics is useless, quite the contrary–in the case of the processor the connectome was the first crucial step in enabling reliable, whole-brain-scale simulation. But even with the whole-brain connectome, extracting hierarchical organization and understanding the nature of the underlying computation is incredibly difficult.

Lesion a single transistor at a time

Lesions studies allow us to study the causal effect of removing a part of the system. We thus chose a number of transistors and asked if they are necessary for each of the behaviors of the processor (Fig 4. In other words, we asked if removed each transistor, if the processor would then still boot the game. Indeed, we found a subset of transistors that makes one of the behaviors (games) impossible. We can thus conclude they are uniquely necessary for the game—perhaps there is a Donkey Kong transistor or a Space Invaders transistor. Even if we can lesion each individual transistor, we do not get much closer to an understanding of how the processor really works.

Lesioning every single transistor to identify function.

We identify transistors whose elimination disrupts behavior analogous to lethal alleles or lesioned brain areas. These are transistors whose elimination results in the processor failing to render the game. (A) Transistors which impact only one behavior, colored by behavior. (B) Breakdown of the impact of transistor lesion by behavioral state. The elimination of 1565 transistors have no impact, and 1560 inhibit all behaviors.

CamScanner

This finding of course is grossly misleading. The transistors are not specific to any one behavior or game but rather implement simple functions, like full adders. The finding that some of them are important while others are not for a given game is only indirectly indicative of the transistor's role and is unlikely to generalize to other games. Lazebnik made similar observations about this approach in molecular biology, suggesting biologists would obtain a large number of identical radios and shoot them with metal particles at short range, attempting to identify which damaged components gave rise to which broken phenotype.

This example nicely highlights the importance of isolating individual behaviors to understand the contribution of parts to the overall function. If we had been able to isolate a single function, maybe by having the processor produce the same math operation every single step, then the lesioning experiments could have produced more meaningful results. However, the same problem exists in neuroscience. It is extremely difficult or technically impossible to produce behaviors that only require a single aspect of the brain.

Beyond behavioral choices, we have equivalent problems in neuroscience that make the interpretation of lesioning data complicated. In many ways the chip can be lesioned in a cleaner way than the brain: we can individually abolish every single transistor (this is only now becoming possible with neurons in simple systems). Even without this problem, finding that a lesion in a given area abolishes a function is hard to interpret in terms of the role of the area for general computation. And this ignores the tremendous plasticity in neural systems which can allow regions to take over for damaged areas. In addition to the statistical problems that arise from multiple hypothesis testing, it is obvious that the "causal relationship" we are learning is incredibly superficial: a given transistor is obviously not specialized for Donkey Kong or Space Invaders.

While in most organisms individual transistors are not vital, for many less-complex systems they are. Lesion individual interneurons in C. elegans or the H1 neuron in the fly can have marked behavioral impacts. And while lesioning larger pieces of circuitry, such as the entire TIA graphics chip, might allow for gross segregation of function, we take issue with this constituting "understanding". Simply knowing functional localization, at any spatial scale, is only the most nacent step to the sorts of understanding we have outlined above.

Analyzing tuning properties of individual transistors

We may want to try to understand the processor by understanding the activity of each individual transistor. We study the "off-to-on" transition, or "spike", produced by each individual transistor. Each transistor will be activated at multiple points in time. Indeed, these transitions look surprisingly similar to the spike trains of neurons (Fig 5). Following the standards in neuroscience we may then quantify the tuning selectivity of each transistor. For each of our transistors we can plot the spike rate as a function of the luminance of the most recently displayed pixel (Fig 6). For a small number of transistors we find a strong tuning to the luminance of the most recently displayed pixel, which we can classify into simple (Fig 6a) and (Fig 6b) complex curves. Interestingly, however, we know for each of the five displayed transistors that they are not directly related to the luminance of the pixel to be written, despite their strong tuning. The transistors relate in a highly nonlinear way to the ultimate brightness of the screen. As such their apparent tuning is not really insightful about their role. In our case, it probably is related to differences across game stages. In the brain a neuron can calculate something, or be upstream or downstream of the calculation and still show apparent tuning making the inference of a neurons role from observational data very difficult. This shows how obtaining an understanding of the processor from tuning curves is difficult.

Analyzing the spikes to understand their statistics.

(A) 10 identified transistors and (B) their spiking (rising edge) behavior over a short time window during behavior DK.

Quantifying tuning curves to understand function.

Mean transistor response as a function of output pixel luminance. (A) Some transistors exhibit simple unimodal tuning curves. (B) More complex tuning curves. (C) Transistor location on chip.

Much of neuroscience is focused on understanding tuning properties of neurons, circuits, and brain areas. Arguably this approach is more justified for the nervous system because brain areas are more strongly modular. However, this may well be an illusion and many studies that have looked carefully at brain areas have revealed a dazzling heterogeneity of responses. Even if brain areas are grouped by function, examining the individual units within may not allow for conclusive insight into the nature of computation.

The correlational structure exhibits weak pairwise and strong global correlations

Moving beyond correlating single units with behavior, we can examine the correlations present between individual transistors. We thus perform a spike-word analysis by looking at "spike words" across 64 transistors in the processor. We find little to very weak correlation among most pairs of transistors (Fig 7a). This weak correlation suggests modeling the transistors' activities as independent, but as we see from shuffle analysis (Fig 7b), this assumption fails disastrously at predicting correlations across many transistors.

Spike-word analysis to understand synchronous states.

(A) Pairs of transistors show very weak pairwise correlations during behavior SI, suggesting independence. (B) If transistors were independent, shuffling transistor labels (blue) would have no impact on the distribution of spikes per word, which is not the case (red).

CamScanner

In neuroscience, it is known that pairwise correlations in neural systems can be incredibly weak, while still reflecting strong underlying coordinated activity. This is often assumed to lead to insights into the nature of interactions between neurons. However, the processor has a very simple nature of interactions and yet produces remarkably similar spike word statistics. This again highlights how hard it is to derive functional insights from activity data using standard measures.

Analyzing local field potentials

The activity of the entire chip may be high dimensional, yet we know that the chip, just like the brain, has some functional modularity. As such, we may be able to understand aspects of its function by analyzing the average activity within localized regions, in a way analogous to the local field potentials or the BOLD signals from functional magnetic imaging that are used in neuroscience. We thus analyzed data in spatially localized areas (Fig 8a). Interestingly, these average activities look quite a bit like real brain signals (Fig 8b). Indeed, they show a rather similar frequency power relation of roughly power-law behavior. This is often seen as a strong sign of self-organized criticality. Spectral analysis of the time-series reveals region-specific oscillations or “rhythms” that have been suggested to provide a clue to both local computation and overall inter-region communication. In the chip we know that while the oscillations may reflect underlying periodicity of activity, the specific frequencies and locations are epiphenomena. They arise as an artifact of the computation and tell us little about the underlying flow of information. And it is very hard to attribute (self-organized) criticality to the processor.

Examining local field potentials to understand network properties.

We recorded from the processor during behavior DK. (A) Transistor switching is integrated and low-pass filtered over the indicated region. (B) local-field potential measurements from the indicated areas. (C) Spectral analysis of the indicated LFP regions identifies varying region-specific oscillations or “rhythms”.

In neuroscience there is a rich tradition of analyzing the rhythms in brain regions, the distribution of power across frequencies as a function of the task, and the relation of oscillatory activity across space and time. However, the example of the processor shows that the relation of such measures to underlying function can be extremely complicated. In fact, the authors of this paper would have expected far more peaked frequency distributions for the chip. Moreover, the distribution of frequencies in the brain is often seen as indicative about the underlying biophysics. In our case, there is only one element, the transistor, and not multiple neurotransmitters. And yet, we see a similarly rich distribution of power in the frequency domain. This shows that complex multi-frequency behavior can emerge from the combination of many simple elements. Analyzing the frequency spectra of artifacts thus leads us to be careful about the interpretation of those occurring in the brain. Modeling the processor as a bunch of coupled oscillators, as is common in neuroscience, would make little sense.

Granger causality to describe functional connectivity

Granger causality has emerged as a method of assessing putative causal relationships between brain regions based on LFP data. Granger causality assesses the relationship between two timeseries X and Y by comparing the predictive power of two different time-series models to predict future values of Y. The first model uses only past values of Y, whereas the second uses the history of X and Y. The additon of X allows one to assess the putative “causality” (really, the predictive power) of X.

To see if we can understand information transmission pathways in the chip based on such techniques, we perform conditional Granger causality analysis on the above-indicated LFP regions for all three behavioral tasks, and plot the resulting inferences of causal interactions (Fig 9). We find that the decoders affect the status bits. We also find that the registers are affected by the decoder, and that the accumulator is affected by the registers. We also find communication between the two parts of the decoder for Donkey Kong, and a lack of communication from the accumulator to the registers in Pitfall. Some of these findings are true, registers really affect the accumulator and decoders really affect the status bits. Other insights are less true, e.g. decoding is independent and the accumulator obviously affects the registers. While some high level insights may be possible, the insight into the actual function of the processor is limited.

Analyzing conditional Granger causality to understand functional connectivity.

Each of the recordings come from a well defined functional subcircuit. Green and blue are two parts of the decoder circuit. Red includes the status bits. Violet are part of the registers and yellow includes parts of the accumulator. We estimated for each behavioral state from LFP sites indicated in Fig 8. Arrows indicate direction of Granger-causal relationship, arrow thickness indicates effect magnitude.

The analysis that we did is very similar to the situation in neuroscience. In neuroscience as well, the signals come from a number of local sources. Moreover, there are also lots of connections but we hope that the methods will inform us about the relevant ones. It is hard to interpret the results—what exactly does the Granger causality model tell us about. Granger causality tells us how activity in the past are predictive of activity in the future, and the link from there to causal interactions is tentative at best and yet such methods are extensively used across large subfields of neuroscience. Even if such methods would reliably tell us about large scale influences, it is very hard to get from a coarse resolution network to the microscopic computations.

Dimensionality reduction reveals global dynamics independent of behavior

CamScanner

In line with recent advances in whole-animal recordings, we measure the activity across all 3510 transistors simultaneously for all three behavioral states (Fig 10) and plot normalized activity for each transistor versus time. Much as in neural systems, some transistors are relatively quiet and some are quite active, with a clear behaviorally-specific periodicity visible in overall activity.

The processor activity map.

For each of three behavioral states we plotted all the activities. Each transistor's activity is normalized to zero-mean and unit variance and plotted as a function of time.

While whole-brain recording may facilitate identification of putative areas involved in particular behaviors, ultimately the spike-level activity at this scale is difficult to interpret. Thus scientists turn to dimensionality reduction techniques, which seek to explain high-dimensional data in terms of a low-dimensional representation of state. We use non-negative matrix factorization to identify constituent signal parts across all time-varying transistor activity. We are thus, for the first time in the paper, taking advantage of all transistors simultaneously.

Non-negative matrix factorization assumes each recovered timeseries of transistor activity is a linear combination of a small number of underlying nonnegative time-varying signals (dimensions). Analogous with we plot the recovered dimensions as a function of time (Fig 11a) and the transistor activity profile of each component (Fig 11b). We can also examine a map of transistor-component activity both statically (Fig 11c) and dynamically (S1–S3 Videos available in online supplementary materials). Clearly there is a lot of structure in this spatiotemporal dataset.

Dimensionality Reduction to understand the roles of transistors.

We apply non-negative matrix factorization (NMF) to the space invaders (SI) task. (A) shows the six reduced dimensions as a function of time showing clear stereotyped activity. (B) the learned transistor state vectors for each dimension (C) Map of total activity—color indicates the dimension where the transistor has maximum value, and both saturation and point size indicate the magnitude of that value.

To derive insight into recovered dimensions, we can try and relate parts of the low-dimensional time series to known signals or variables we know are important (Fig 12a). Indeed, we find that some components relate to both the onset and offset (rise and fall) of the clock signal(Fig 12b and 12c). This is quite interesting as we know that the processor uses a two-phase clock. We also find that a component relates strongly to the processors read-write signal (Fig 12d). Thus, we find that variables of interest are indeed encoded by the population activity in the processor.

Relating dimensions to known signals to understanding the population code.

(A) For each of the recovered dimensions in Fig 11 we compute the correlation in time with 25 known signals inside the process. As we know the purpose of these signals we can measure how well the dimensions explain true underlying function. (B) Dimension 1 is strongly correlated with the processor clock CLK0, whereas (C) dimension 4 is correlated with the 180-degree out of phase CLK1OUT signal. (D) dimension 0 is strongly correlated with signal RW, indicating the processor switching between reading and writing memory.

In neuroscience, it is also frequently found that components from dimensionality reduction relate to variables of interest. This is usually then seen as an indication that the brain cares about these variables. However, clearly, the link to the read-write signal and the clock does not lead to an overly important insight into the way the processor actually processes information. Similar questions arise in neuroscience where scientists ask if signals, such as synchrony, are a central part of information processing or if they are an irrelevant byproduct. We should be careful at evaluating how much we understand and how much we are aided by more data.

Pondering the results of the processor analysis we can obtain some insights into the developments needed to better utilize dimensionality reduction towards an understanding. The narrow range of games that we considered and the narrow range of their internal states (we just simulated booting), means that many aspects of computation will not be reflected by the activities and hence not in the dimensionality reduction results. Moreover, the fact that we used linear reduction only allows for linear dependencies and transistors, just like neurons, have important nonlinear dependencies. Lastly, there is clearly a hierarchy in function in the processor and we would need to do it justice using hierarchical analysis approaches. The results of dimensionality reduction should be meaningful for guiding new experiments, necessitating transfer across chips in the same way as neuroscience experiments should transfer across animals. Importantly, the chip can work as a test case while we develop such methods.

Discussion

Here we have taken a reconstructed and simulated processor and treated the data "recorded" from it in the same way we have been trained to analyze brain data. We have used it as a test case to check the naïve use of various approaches used in neuroscience. We have found that the standard data analysis techniques produce results that are surprisingly similar to the results found about real brains. However, in the case of the processor we know its function and structure and our results stayed well short of what we would call a satisfying understanding.

Obviously the brain is not a processor, and a tremendous amount of effort and time have been spent characterizing these differences over the past century. Neural systems are analog and and biophysically complex, they operate at temporal scales vastly slower than this classical processor but with far greater parallelism than is available in state of the art processors. Typical neurons also have several orders of magnitude more inputs than a transistor.

CamScanner

Moreover, the design process for the brain (evolution) is dramatically different from that of the processor (the MOS6502 was designed by a small team of people over a few years). As such, we should be skeptical about generalizing from processors to the brain.

However, we cannot write off the failure of the methods we used on the processor simply because processors are different from neural systems. After all, the brain also consists of a large number of modules that can equally switch their input and output properties. It also has prominent oscillations, which may act as clock signals as well. Similarly, a small number of relevant connections can produce drivers that are more important than those of the bulk of the activity. Also, the localization of function that is often assumed to simplify models of the brain is only a very rough approximation. This is true even in an area like V1 where a great diversity of co-localized cells can be found. Altogether, there seems to be little reason to assume that any of the methods we used should be more meaningful on brains than on the processor.

To analyze our simulations we needed to convert the binary transistor state of the processor into spike trains so that we could apply methods from neuroscience to (see Methods). While this may be artefactual, we want to remind the reader that in neuroscience the idea of an action potential is also only an approximate description of the effects of a cell's activity. For example, there are known effects based on the extrasynaptic diffusion of neurotransmitters and it is believed that active conductances in dendrites may be crucial to computation.

Our behavioral mechanisms are entirely passive as both the transistor based simulator is too slow to play the game for any reasonable duration and the hardware for game input/output has yet to be reconstructed. Even if we could "play" the game, the dimensionality of the input space would consist at best of a few digital switches and a simple joystick. One is reminded of the reaching tasks which dominate a large fraction of movement research. Tasks that isolate one kind of computation would be needed so that interference studies would be really interpretable.

If we had a way of hypothesizing the right structure, then it would be reasonably easy to test. Indeed, there are a number of large scale theories of the brain. However, the set of potential models of the brain is unbelievably large. Our data about the brain from all the experiments so far, is very limited and based on the techniques that we reviewed above. As such, it would be quite impressive if any of these high level models would actually match the human brain to a reasonable degree. Still, they provide beautiful inspiration for a lot of ongoing neuroscience research and are starting to exhibit some human-like behaviors. If the brain is actually simple, then a human can guess a model, and through hypothesis generation and falsification we may eventually obtain that model. If the brain is not actually simple, then this approach may not ever converge. Simpler models might yield more insight—specifically seeking out an "adder" circuit might be possible, if we had a strong understanding of binary encoding and could tease apart the system to specifically control inputs and outputs of a subregion—examine it in slice, if you will.

The analytic tools we have adopted are in many ways "classic", and are taught to graduate students in neuroinformatics courses. Recent progress in methods for dimensionality reduction, subspace identification, time-series analysis, and tools for building rich probabilistic models may provide some additional insight, assuming the challenges of scale can be overcome. Culturally, applying these methods to real data, and rewarding those who innovate methodologically, may become more important. We can look at the rise of bioinformatics as an independent field with its own funding streams. Neuroscience needs strong neuroinformatics to make sense of the emerging datasets and known artificial systems can serve as a sanity check and a way of understanding failure modes.

We also want to suggest that it may be an important intermediate step for neuroscience to develop methods that allow understanding a processor. Because they can be simulated in any computer and arbitrarily perturbed, they are a great testbed to ask how useful the methods are that we are using in neuroscience on a daily basis. Scientific fields often work well in situations where we can measure how well a project is doing. In the case of processors we know their function and we can know if our algorithms discover it. Unless our methods can deal with a simple processor, how could we expect it to work on our own brain? Machine learning and statistics currently lack good high-dimensional datasets with complex underlying dynamics and known ground truth. While not a perfect match, the dynamics of a processor may provide a compelling intermediate step. Additionally, most neural datasets are still "small data"—hundreds of cells over tens of minutes. The processor enables the generation of arbitrary complexity and arbitrarially-long timeseries, enabling a focus on scalable algorithms. We must be careful to not over-fit, but neuroscience is rife with examples of adopting analytic tools from vary different domains (linear system theory, stochastic process theory, kalman filtering) to understand neural systems.

In the case of the processor, we really understand how it works. We have a name for each of the modules on the chip and we know which area is covered by each of them (Fig 13a). Moreover, for each of these modules we know how its outputs depend on its inputs and many students of electrical engineering would know multiple ways of implementing the same function. In the case of the brain, we also have a way of dividing it into regions (Fig 13b, adopted from). However, we only use anatomy to divide into modules and even among specialists there is a lot of disagreement about the division. Most importantly though, we do not generally know how the output relates to the inputs. As we reviewed in this paper, we may even want to be careful about the conclusions about the modules that neuroscience has drawn so far, after all, much of our insights come from small datasets, with analysis methods that make questionable assumptions.

Understanding the processor.

CamScanner

(A) For the processor we understand its hierarchical organization as well as which part of the silicon implements which function. For each of these “functional modules” we know how the outputs depend on the inputs. (B) For the brain, it is harder to be sure. The primate visual system is often depicted in a similar way, such as this diagram adapted from the classic Felleman and vanEssen diagram. These areas are primarially divided according to anatomy, but there is extensive debate about the ideal way of dividing the brain into functional areas. Moreover, we currently have little of an understanding how each area’s outputs depend on its inputs.

There are other computing systems that scientists are trying to reverse engineer. One particularly relevant one are artificial neural networks. A plethora of methods are being developed to ask how they work. This includes ways of letting the networks paint images and ways of plotting the optimal stimuli for various areas. While progress has been made on understanding the mechanisms and architecture for networks performing image classification, more complex systems are still completely opaque. Thus a true understanding even for these comparatively simple, human-engineered systems remains elusive, and sometimes they can even surprise us by having truly surprising properties. The brain is clearly far more complicated and our difficulty at understanding deep learning may suggest that the brain is hard to understand if it uses anything like gradient descent on a cost function.

What kind of developments would make understanding the processor, and ultimately the brain, more tractable? While we can offer no definitive conclusion, we see multiple ways in which we could have better understood the processor. If we had experiments that would more cleanly separate one computation then results would be more meaningful. For example, lesion studies would be far more meaningful if we could also simultaneously control the exact code the processor was executing at a given moment. Better theories could most obviously have helped; if we had known that the microprocessor has adders we could have searched for them. Lastly, better data analysis methods, e.g. those that can explicitly search for hierarchical structure or utilize information across multiple processors. Development in these areas seems particularly promising. The microprocessor may help us by being a sieve for ideas: good ideas for understanding the brain should also help us understand the processor. Ultimately, the problem is not that neuroscientists could not understand a microprocessor, the problem is that they would not understand it given the approaches they are currently taking.

Methods

Netlist acquisition

All acquisition and development of the initial simulation was performed in James. 200°F sulfuric acid was used to decap multiple 6502D ICs. Nikon LV150n and Nikon Optiphot 220 light microscopes were used to capture 72 tiled visible-light images of the die, resulting in 342 Mpix of data. Computational methods and human manual annotation used developed to reconstruct the metal, polysilicon, via, and interconnect layers. 3510 active enhancement-mode transistors were captured this way. The authors inferred 1018 depletion-mode transistors (serving as pullups) from the circuit topology as they were unable to capture the depletion mask layer.

Simulation and behaviors

An optimized C++ simulator was constructed to enable simulation at the rate of 1000 processor clock cycles per wallclock second. We evaluated the four provided ROMs (Donkey Kong, Space Invaders, Pitfall, and Asteroids) ultimately choosing the first three as they reliably drove the TIA and subsequently produced image frames. 10 seconds of behavior were simulated for each game, resulting in over 250 frames per game.

Lesion studies

Whole-circuit simulation enables high-throughput targeted manipulation of the underlying circuit. We systematically perturb each transistor in the processor by forcing its input high, thus leaving it in an “on” state. We measure the impact of a lesion by whether or not the system advances far enough to draw the first frame of the game. Failure to produce the first frame constitutes as a loss of function. We identified 1560 transistors which resulted in loss of function across all games, 200 transistors which resulted in loss of function across two games, and 186 transistors which resulted in loss of function for a single game. We plot those single-behavior lesion transistors by game in Fig 4.

Connectomic analysis

Using the acquired netlist, we implement the authors method from on the region of the processor consisting of the X, Y, and S registers. A nonparametric distance-dependent stochastic block model is jointly fit to six connectivitiy matrices: G → C1, G → C2, C1 → C2 C2 → C1, C1 → G, C2 → G, and via Markov-chain Monte Carlo, seeks the maximum a posteriori estmate for the observed connectivity.

Spiking

We chose to focus on transistor switching as this is the closest in spirit to discrete action potentials of the sort readily available to neuroscientific analysis. The alternative, performing analysis with the signals on internal wires, would be analogous to measuring transmembrane voltage. Rasters were plotted from 10 example transistors which showed sufficient variance in spiking rate.

Tuning curves

We compute luminance from the RGB output value of the simulator for each output pixel to the TIA. We then look at the transistor rasters and sum activity for 100 previous timesteps and call this the “mean rate”. For each transistor we then compute a tuning curve of mean rate versus luminance, normalized by the frequency of

CamScanner

occurrence of that luminance value. Note that each game outputs only a small number of discrete colors and thus discrete luminance values. We used SI as it gave the most equal sampling of luminance space. We then evaluate the degree of fit to a unimodial Gaussian for each resulting tuning curve and classify tuning curves by eye into simple and complex responses, of which Fig 4 contains representative examples.

Spike-word analysis

For the SI behavior we took spiking activity from the first 100ms of SI and performed spike word analysis on a random subset of 64 transistors close to the mean firing rate of all 3510.

Local field potential

To derive “local field potentials” we spatially integrate transistor switching over a region with a Gaussian weighting of $\sigma = 500\mu m$ and low-pass filter the result using a window with a width of 4 timesteps.

We compute periodograms using Welch’s method with 256-sample long windows with no overlap and a Hanning window.

Granger causality

We adopt methods for assessing conditional Granger causality as outlined in. We take the LFP generated using methods in section and create 100 1ms-long trials for each behavioral experiment. We then compute the conditional Granger causality for model orders ranging from 1 to 31. We compute BIC for all behaviors and select a model order of 20 as this is where BIC plateaus.

Whole brain recording

The transistor switching state for the first 106 timestamps for each behavioral state is acquired, and binned in 100-timestep increments. The activity of each transistor is converted into a z-score by subtracting mean and normalizing to unit variance.

Dimensionality reduction

We perform dimensionality reduction on the first 100,000 timesteps of the 3510-element transistor state vectors for each behavioral condition. We use non-negative matrix factorization, which attempts to find two matrices, W and H, whose product WH approximates the observed data matrix X. This is equivalent to minimizing the objective .

The Scikit-Learn implementation initialized via nonnegative double singular value decomposition solved via coordinate descent, as is the default. We use a latent dimensionality of 6 as it was found by hand to provide the most interpretable results. When plotting, the intensity of each transistor in a latent dimension is indicated by the saturation and size of point.

To interpret the latent structure we first compute the signed correlation between the latent dimension and each of the 25 known signals. We show particularly interpretable results.

Supporting Information

References

Putting big data to good use in neuroscience

Mapping brain activity at scale with cluster computing

Charting the Brain’s Networks

The Brain Activity Map Project and the Challenge of Functional Connectomics

The human brain project

Brain-wide neuronal dynamics during motor adaptation in zebrafish

Simultaneous whole-animal 3D imaging of neuronal activity using light-field microscopy

Whole-brain calcium imaging with cellular resolution in freely behaving Caenorhabditis elegans

Can a biologist fix a radio?—Or, what I learned while studying apoptosis

A visual motion detection circuit suggested by Drosophila connectomics

Connectomic reconstruction of the inner plexiform layer in the mouse retina

Human-level control through deep reinforcement learning

Aloupis G, Demaine ED, Guo A, Viglietta G. Classic Nintendo Games are (Computationally) Hard. In: Proceedings of the 7th International Conference on Fun with Algorithms (FUN 2014),. Lipari Island, Italy; 2014. p. 41–50. Available from: http://arxiv.org/abs/1203.1895.

What is a moment? Transient synchrony as a collective mechanism for spatiotemporal integration

From circuits to behavior: a bridge too far?

On the Precarious Path of Reverse Neuro-Engineering

In the brain, the model is the goal

CamScanner

The tale of the neuroscientists and the computer: Why mechanistic theory matters
Processing of complex stimuli and natural scenes in the visual cortex
The atoms of neural computation
Variability, compensation and homeostasis in neuron and network function
Deep molecular diversity of mammalian synapses: why it matters and how to measure it
An Analog VLSI Saccadic Eye Movement System
A cortical neural prosthesis for restoring and enhancing memory
Neuroanatomy: Cajal and after Cajal
Stochastic Blockmodeling of the Modules and Core of the Caenorhabditis elegans Connectome
Automatic discovery of cell types and microcircuitry from neural connectomics
The Rich Club of the C. elegans Neuronal Connectome
Towards neural circuit reconstruction with volume electron microscopy techniques
Ome sweet ome: what can the genome tell us about the connectome?
Structural properties of the Caenorhabditis elegans neuronal network
Unbiased classification of sensory neuron types by large-scale single-cell RNA sequencing
Using human brain lesions to infer function: a relic from a past era in the fMRI age?
A GAL4-Driver Line Resource for Drosophila Neurobiology
The neuronal architecture of the mushroom body provides a logic for associative learning
Dissociated functional significance of choice-related activity across the primate dorsal stream
Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex
The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat
Microstructure of a spatial map in the entorhinal cortex
The fusiform face area: a module in human extrastriate cortex specialized for face perception
Neural responses to polar, hyperbolic, and cartesian grating in area V4 of the macaque monkey
Classifying simple and complex cells on the basis of response modulation
Invariant visual representation by single neurons in the human brain
Weak pairwise correlations imply strongly correlated network states in a neural population
Self-organized criticality as a fundamental property of neural systems
Granger Causality Analysis in Neuroscience and Neuroimaging
On the Similarity of Functional Connectivity between Neurons Estimated across Timescales
Dimensionality reduction for large-scale neural recordings
Neural population dynamics during reaching
Learning the parts of objects by non-negative matrix factorization
Spatially Distributed Local Fields in the Hippocampus Encode Rat Position
Neuromechanics of muscle synergies for posture and movement
Neuronal synchrony does not correlate with motion coherence in cortical area MT
Signal-processing machines at the postsynaptic density
Neuronal Oscillations in Cortical Networks
Orientation selectivity in macaque V1: diversity and laminar dependence
Cellular, synaptic and network effects of neuromodulation
Dendritic Computation
A Large-Scale Model of the Functioning Brain
ACT-R: A Theory of Higher Level Cognition and its Relation to Visual Attention
Distributed hierarchical processing in the primate cerebral cortex
Yosinski J, Clune J, Nguyen A, Fuchs T, Lipson H. Understanding Neural Networks Through Deep Visualization. International Conference on Machine Learning—Deep Learning Workshop 2015. 2015; p. 12.
Visualizing and understanding convolutional networks
Lipton RJ, Regan KW. Magic To Do; 2016. Available from: https://rjlipton.wordpress.com/2016/02/07/magic-to-do/.

CamScanner

Intriguing properties of neural networks
Granger Causality: Basic Theory and Application to Neuroscience
Scikit-learn: Machine Learning in Python

CamScanner

nature communications

Perspective

https://doi.org/10.1038/s41467-023-37180-x

Catalyzing next-generation Artificial Intelligence through NeuroAI

Received: 11 September 2022

Accepted: 3 March 2023

Published online: 22 March 2023

Check for updates

Anthony Zador[1,29] ✉, Sean Escola[2,29], Blake Richards[3,4,5,6,7], Bence Ölveczky[8], Yoshua Bengio[3], Kwabena Boahen[9], Matthew Botvinick[10], Dmitri Chklovskii[11], Anne Churchland[12], Claudia Clopath[13], James DiCarlo[14], Surya Ganguli[15], Jeff Hawkins[16], Konrad Körding[17], Alexei Koulakov[1], Yann LeCun[18,19], Timothy Lillicrap[10], Adam Marblestone[20], Bruno Olshausen[21], Alexandre Pouget[22], Cristina Savin[23], Terrence Sejnowski[24], Eero Simoncelli[25], Sara Solla[26], David Sussillo[18,27], Andreas S. Tolias[28] & Doris Tsao[21]

Neuroscience has long been an essential driver of progress in artificial intelligence (AI). We propose that to accelerate progress in AI, we must invest in fundamental research in NeuroAI. A core component of this is the embodied Turing test, which challenges AI animal models to interact with the sensorimotor world at skill levels akin to their living counterparts. The embodied Turing test shifts the focus from those capabilities like game playing and language that are especially well-developed or uniquely human to those capabilities – inherited from over 500 million years of evolution – that are shared with all animals. Building models that can pass the embodied Turing test will provide a roadmap for the next generation of AI.

Over the coming decades, Artificial Intelligence (AI) will transform society and the world economy in ways that are as profound as the computer revolution of the last half century and likely at an even faster pace. This AI revolution presents tremendous opportunities to unleash human creativity and catalyze economic growth, relieving workers from performing the most dangerous and menial jobs. However, to reach this potential, we still require advances that will make AI more human-like in its capabilities. Historically, neuroscience has been a critical driver and source of inspiration for improvements in AI, particularly those that made AI more proficient in areas that humans and other animals excel at, such as vision, reward-based learning, interacting with the physical world, and language[1,2]. It can still play this role. To accelerate progress in AI and realize its vast potential, we must invest in fundamental research in "NeuroAI."

[1]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA. [2]Department of Psychiatry, Columbia University, New York, NY 10027, USA. [3]Mila, Montréal, QC H2S 3H1, Canada. [4]School of Computer Science, McGill University, Montreal, Canada. [5]Montreal Neurological Institute, McGill University, Montreal, Canada. [6]Department of Neurology & Neurosurgery, McGill University, Montreal, Canada. [7]Learning in Machines and Brains Program, CIFAR, Toronto, Canada. [8]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA. [9]Department of Bioengineering, Stanford University, Stanford, CA 94305, USA. [10]Google Deepmind, London N1C 4AG, UK. [11]Flatiron Institute, Simons Foundation, New York, NY 10010, USA. [12]Department of Neurobiology, University of California Los Angeles, Los Angeles, CA 90095, USA. [13]Department of Bioengineering, Imperial College London, London SW7 2BW, UK. [14]Department of Brain and Cognitive Sciences, MIT, Cambridge, MA 02139, USA. [15]Department of Applied Physics, Stanford University, Stanford, CA 94305, USA. [16]Numenta, Redwood City, CA 94063, USA. [17]Department of Neuroscience, University of Pennsylvania, Philadelphia, PA 19104, USA. [18]Meta, Menlo Park, CA 94025, USA. [19]Department of Electrical and Computer Engineering, NYU, Brooklyn, NY 11201, USA. [20]Media Lab, MIT, Cambridge, MA 02140, USA. [21]Helen Wills Neuroscience Institute, University of California Berkeley, Berkeley, CA 94720, USA. [22]Department of Basic Neurosciences, University of Geneva, Genève 1211, Switzerland. [23]Center for Neural Science, NYU, New York, NY 10003, USA. [24]Salk Institute for Biological Studies, La Jolla, CA 92037, USA. [25]Departments of Neural Science, Mathematics, and Psychology, NYU, New York, NY 10003, USA. [26]Department of Physiology, Northwestern University, Chicago, IL 60611, USA. [27]Department of Electrical Engineering, Stanford University, Stanford, CA 94305, USA. [28]Department of Neuroscience, Baylor College of Medicine, Houston, TX 77030, USA. [29]These authors contributed equally: Anthony M. Zador, Sean Escola. ✉e-mail: zador@cshl.edu

CamScanner

The seeds of the current AI revolution were planted decades ago, mainly by researchers attempting to understand how brains compute[3]. Indeed, the earliest efforts to build an "artificial brain" led to the invention of the modern "von Neumann computer architecture," for which John von Neumann explicitly drew upon the very limited knowledge of the brain available to him in the 1940s[4,5]. Later, the Nobel-prize winning work of David Hubel and Torsten Wiesel on visual processing circuits in the cat neocortex inspired the deep convolutional networks that have catalyzed the recent revolution in modern AI[6–8]. Similarly, the development of reinforcement learning was directly inspired by insights into animal behavior and neural activity during learning[9–15]. Now, decades later, applications of ANNs and RL are coming so quickly that many observers assume that the long-elusive goal of human-level intelligence—sometimes referred to as "artificial general intelligence"—is within our grasp. However, in contrast to the optimism of those outside the field, many front-line AI researchers believe that major breakthroughs are needed before we can build artificial systems capable of doing all that a human, or even a much simpler animal like a mouse, can do.

Although AI systems can easily defeat any human opponent in games such as chess[16] and Go[17], they are not robust and often struggle when faced with novel situations. Moreover, we have yet to build effective systems that can walk to the shelf, take down the chess set, set up the pieces, and move them around during a game, although recent progress is encouraging[18]. Similarly, no machine can build a nest, forage for berries, or care for young. Today's AI systems cannot compete with the sensorimotor capabilities of a four-year old child or even simple animals. Many basic capacities required to navigate new situations—capacities that animals have or acquire effortlessly—turn out to be deceptively challenging for AI, partly because AI systems lack even the basic abilities to interact with an unpredictable world. A growing number of AI researchers doubt that merely scaling up current approaches will overcome these limitations. Given the need to achieve more natural intelligence in AI, it is quite likely that new inspiration from naturally intelligent systems is needed[19].

Historically, many key AI advances, such as convolutional ANNs and reinforcement learning, were inspired by neuroscience. Neuroscience continues to provide guidance—e.g., attention-based neural networks were loosely inspired by attention mechanisms in the brain[20–23]—but this is often based on findings that are decades old. The fact that such cross-pollination between AI and neuroscience is far less common than in the past represents a missed opportunity. Over the last decades, through efforts such as the NIH BRAIN initiative and others, we have amassed an enormous amount of knowledge about the brain. *The emerging field of NeuroAI, at the intersection of neuroscience and AI, is based on the premise that a better understanding of neural computation will reveal fundamental ingredients of intelligence and catalyze the next revolution in AI.* This will eventually lead to artificial agents with capabilities that match those of humans. The NeuroAI program we advocate is driven by the recognition that AI historically owes much to neuroscience and the promise that AI will continue to learn from it–but only if there is a large enough community of researchers fluent in both domains. We believe the time is right for a large-scale effort to identify and understand the principles of biological intelligence and abstract those for application in computer and robotic systems.

It is tempting to focus on the most characteristically human aspects of intelligent behavior, such as abstract thought and reasoning. However, the basic ingredients of intelligence—adaptability, flexibility, and the ability to make general inferences from sparse observations—are already present in some form in basic sensorimotor circuits, which have been evolving for hundreds of millions of years. As AI pioneer Hans Moravec[24] put it, abstract thought "is a new trick, perhaps less than 100 thousand years old….effective only because it is supported by this much older and much more powerful, though usually unconscious, sensorimotor knowledge." This implies that the bulk of the work in developing general AI can be achieved by building systems that match the perceptual and motor abilities of animals and that the subsequent step to human-level intelligence would be considerably smaller. This is good news because progress on the first goal can rely on the favored subjects of neuroscience research—rats, mice, and non-human primates—for which extensive and rapidly expanding behavioral and neural datasets can guide the way. Thus, we believe that the NeuroAI path will lead to necessary advances if we figure out the core capabilities that all animals possess in embodied sensorimotor interaction with the world.

NeuroAI grand challenge: the embodied turing test

In 1950, Alan Turing proposed the "imitation game"[25] as a test of a machine's ability to exhibit intelligent behavior indistinguishable from that of a human (Fig. 1, left). In that game, now known as the Turing test, a human judge evaluates natural language conversations between a real human and a machine trained to mimic human responses. By focusing on conversational abilities, Turing evaded asking whether a machine could "think," a question he considered impossible to answer. The Turing test is based on the implicit belief that language represents the pinnacle of human intelligence and that a machine capable of conversation must surely be intelligent.

Until recently, no artificial system could come close to passing the Turing test. However, a class of modern AI systems called "large language models" can now engage in surprisingly cogent conversations[26]. In part, their success reveals how easily we can be tricked into imputing intelligence, agency, and even consciousness to our interlocutor[27]. Impressive though these systems are, because they are not grounded in real-world experiences, they nonetheless continue to struggle with many basic aspects of causal reasoning and physical common-sense. Thus, the Turing test does not probe our prodigious perceptual and motor abilities to interact with and reason about the physical world, abilities shared with animals and honed through countless generations of natural selection.

We therefore propose an expanded "embodied Turing test," one that includes advanced sensorimotor abilities (Fig. 1, right). The spirit of the original Turing test was to establish a simple qualitative standard against which our progress toward building artificially intelligent machines can be judged. This embodied Turing test would benchmark and compare the interactions with the world of artificial systems versus humans and other animals. Similar ideas have been proposed previously[28–32]. However, in light of recent advances enabling large-scale behavioral and neural measurements, as well as large-scale simulations of embodied agents in silico, we believe the time is ripe to instantiate a major research effort in this direction. As each animal has its own unique set of abilities, each animal defines its own embodied

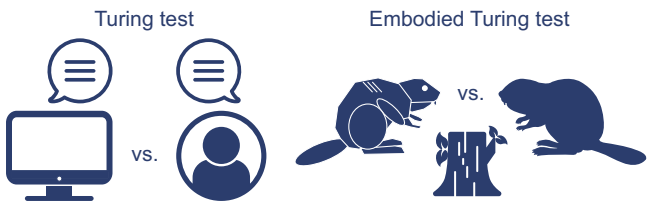


Fig. 1 | Turing tests: comparisons between the performance of AI systems and their living counterparts. *Left*: The original Turing test as proposed by Alan Turing[25]. If a human tester cannot determine whether their interlocutor is an AI system or another human, the AI passes the test. Modern large language models have made substantial progress towards passing this test[26]. *Right*: The embodied Turing test. An AI animal model—whether robotic or in simulation—passes the test if its behavior is indistinguishable from that of its living counterpart. No AI systems are close to passing this test. Here, an artificial beaver is tested on the species-specific behavior of dam construction.

CamScanner

Turing test: An artificial beaver might be tested on its ability to build a dam, and an artificial squirrel on its ability to jump through trees. Nonetheless, many core sensorimotor capabilities are shared by almost all animals, and the ability of animals to rapidly evolve the sensorimotor skills needed to adapt to new environments suggests that these core skills provide a solid foundation. This implies that after developing an AI system to faithfully reproduce the behavior of one species, the adaptation of this system to other species—and even to humans—may be straightforward. Below we highlight a few of the characteristics that are shared across species.

Animals engage their environments

The defining feature of animals is their ability to move around and interact with their environment in purposeful ways. Despite recent advances in optimal control, reinforcement learning, and imitation learning, robotics is still far from achieving animal-level abilities in controlling their bodies and manipulating objects, even in simulation. Of course, neuroscience can provide guidance about the kinds of modular and hierarchical architectures that could be adapted to artificial systems to give them these capabilities[33]. It can also provide us with design principles like partial autonomy (how low-level modules in a hierarchy act semi-autonomously in the absence of input from high-level modules) and amortized control (how movements generated at first by a slow planning process are eventually transferred to a fast reflexive system). These principles could guide the design of systems for perception, action selection, locomotion, and fine-grained control of limbs, hands, and fingers. Understanding how specific neural circuits participate in different tasks could also inspire solutions for other forms of 'intelligence,' including in more cognitive realms. For example, we speculate that incorporating principles of circuitry for low-level motor control could help provide a better basis for higher-level motor planning in AI systems.

Animals behave flexibly

Another goal is to develop AI systems that can engage a large repertoire of flexible and diverse tasks in a manner that echoes the incredible range of behaviors that individual animals can generate. Modern AI can easily learn to outperform humans at video games like *Breakout* using nothing more than pixels on a screen and game scores[34]. However, these systems, unlike human players, are brittle and highly sensitive to small perturbations: changing the rules of the game slightly, or even a few pixels on the input, can lead to catastrophically poor performance[35]. This is because these systems learn a mapping from pixels to actions that need not involve an understanding of the agents and objects in the game and the physics that governs them. Similarly, a self-driving car does not inherently know about the danger of a crate falling off a truck in front of it unless it has literally seen examples of crates falling off trucks leading to bad outcomes. And even if it has been trained on the dangers of falling crates, the system might consider an empty plastic bag being blown out of the car in front of it as an obstacle to avoid at all costs rather than an irritant, again, because it doesn't actually understand what a plastic bag is or how unthreatening it is physically. This inability to handle scenarios that have not appeared in the training data is a significant challenge to widespread reliance on AI systems.

To be successful in an unpredictable and changing world, an agent must be flexible and master novel situations by using its general knowledge about how such situations are likely to unfold. This is arguably what animals do. Animals are born with most of the skills needed to thrive or can rapidly acquire them from limited experience, thanks to their strong foundation in real-world interaction, courtesy of evolution and development[36]. Thus, it is clear that training from scratch for a specific task is not how animals obtain their impressive skills; animals do not arrive into the world *tabula rasa* and then rely on large labeled training sets to learn. Although machine learning has been pursuing approaches for sidestepping this *tabula rasa* limitation, including self-supervised learning, transfer learning, continual learning, meta learning, one-shot learning and imitation learning[37], none of these approaches comes close to achieving the flexibility found in most animals. Thus, we argue that understanding the neural circuit-level principles that provide the foundation for behavioral flexibility in the real-world, even in simple animals, has the potential to greatly increase the flexibility and utility of AI systems. Put another way, we can greatly accelerate our search for general-purpose circuits for real-world interaction by taking advantage of the optimization process that evolution has already engaged in[38–45].

Animals compute efficiently

One important challenge for modern AI—that our brains have overcome—is energy efficiency. Training a neural network requires enormous amounts of energy. For example, training a large language model such as GPT-3 requires over 1000 megawatts-hours, enough to power a small town for a day[46]. Biological systems are, by contrast, much more energy efficient: The human brain uses about 20 watts[47]. The difference in energy requirement between brains and computers derives from differences in information processing. First, at an algorithmic level, modern large-scale ANNs, such as large language models[26], rely on very large feedforward architectures with self-attention to process sequences over time[23], ignoring the potential power of recurrence for processing sequential information. One reason for this is that currently we do not have efficient mechanisms for credit assignment calculations in recurrent networks. In contrast, brains utilize flexible recurrent architectures that can solve the temporal credit assignment problem with great efficiency. Uncovering the mechanisms by which this happens could potentially enable us to increase the energy efficiency of artificial systems. Alternatively, it has been proposed that the synaptic dynamics within adjacent dendritic spines could serve as a mechanism for learning sequential structure, a scheme that could potentially be efficiently implemented in hardware[48]. Second, at an implementation level, neural circuits differ from digital computers. Neural circuits compute effectively despite the presence of unreliable or "noisy" components. For example, synaptic release, the primary means of communication between neurons, can be so unreliable that only one in every ten messages is transmitted[49]. Furthermore, neurons interact mainly by transmitting action potentials (spikes), an asynchronous communication protocol. Like the interactions between conventional digital elements, the output of a neuron can be viewed as a string of 0 s and 1 s; but unlike a digital computer, the energy cost of a "1" (i.e., of a spike) is several orders of magnitude higher than that of a "0"[50]. As biological circuits operate in a regime where spikes are sparse—even very active neurons rarely fire at more than 100 spikes per second and typical cortical firing rates may be less than 1 spike/second—they are much more energy efficient[51]. Spike-based computation has also been shown to be orders of magnitude faster and more energy efficient in recent hardware implementation[52].

A roadmap for solving the embodied Turing test

How might artificial systems that pass the embodied Turing test be developed? One natural approach would be to do so incrementally, guided by our evolutionary history. For example, almost all animals engage in goal-directed locomotion; they move toward some stimuli (e.g., food sources) and away from others (e.g., threats). Layered on top of these foundational abilities are more sophisticated skills, such the ability to combine different streams of sensory information (e.g., visual and olfactory), to use this sensory information to distinguish food sources and threats, to navigate to previous locations, to weigh possible rewards and threats to achieve goals, and to interact with the world in precise ways in service of these goals. Most of these—and many other—sophisticated abilities are found to some extent in even very simple organisms, such as worms. In more complex animals, such

CamScanner

as fish and mammals, these abilities are elaborated and combined with new strategies to enable more powerful behavioral strategies.

This evolutionary perspective suggests a strategy for passing the embodied Turing test by breaking it down into a series of incrementally challenging ones that build on each other, and iteratively optimizing on this series[53]. Specifically, the embodied Turing test comprises challenges that include a wide range of organisms used in neuroscience research, including worms, flies, fish, rodents and primates. This would enable us to deploy the vast amount of knowledge we have accumulated about the behavior, biomechanics, and neural mechanisms of these model organisms to both precisely define each species-specific embodied Turing test and serve as strong inductive biases to guide the development of robust AI controllers that can pass it.

The performance of these artificial agents could be compared with that of animals. Rich behavioral datasets representing a large swath of a species' ethogram have now been collected and can be deployed to benchmark performance on species-specific embodied Turing tests. Furthermore, these datasets are being rapidly expanded given new tools in 3D videography[54–57]. Additionally, detailed biomechanical measurements support highly realistic animal body models, complete with skeletal constraints, muscles, tendons, and paw features[58]. Combined with the open-sourcing of powerful, fast physics simulators and virtual environments[59,60], these models will afford the opportunity for embodied Turing test research to be performed in silico at scale[33]. Finally, existing extensive neural datasets with simultaneous neural recordings across multiple brain regions during behavior, combined with increasingly detailed neural anatomy and connectomics, provide a powerful roadmap for the design of AI systems that can control virtual animals to recapitulate the behaviors of their in vivo counterparts and thus pass the embodied Turing test.

Importantly, the specifics of the embodied Turing test for each species can be tuned to the needs of different groups of researchers. We can test the capacity of AI systems in terms of sensorimotor control, self-supervised and continual learning, generalization, memory-guided behavior on both short and life-long timescales, and social interactions. Despite these potentially different areas of interest, the challenges that compose the embodied Turing test can be standardized to permit the quantification of progress and comparison between research efforts. Standardization can be fostered by stakeholders including government and private funders, large research organizations such as the Allen Institute, and major collaborations like the International Brain Lab, with an eye toward the development of common APIs and support for competitions as has been an important impetus for much progress in machine learning and robotics[61,62]. Ultimately, virtual organisms that demonstrate successful recapitulation of behaviors of interest can be adapted to the physical world with additional efforts in robotics and deployed to solve real-world problems.

What we need

Achieving these goals will require significant resources deployed in three main areas. *First, we must train a new generation of AI researchers who are equally at home in engineering/computational science and neuroscience*. These researchers will chart fundamentally new directions in AI research by drawing on decades of progress in neuroscience. The greatest challenge will be in determining how to exploit the synergies and overlaps in neuroscience, computational science, and other relevant fields to advance our quest: identifying what details of the brain's circuitry, biophysics, and chemistry are important and what can be disregarded in the application to AI. There is thus a critical need for researchers with dual training in AI and neuroscience to apply insights from neuroscience to advance AI and to help design experiments that generate new insights relevant to AI. Although there is already some research of this type, it exists largely at the margins of mainstream neuroscience; training in neuroscience has thus far been motivated and funded mainly by the goal of improving human health and of understanding the brain as such. This lack of alignment between fields might explain, e.g., the multi-decade gap between Hubel and Wiesel's discovery of the structure of the visual system[6] and the development and application of convolutional neural networks in modern machine learning[8]. Thus, the success of a NeuroAI research program depends on the formation of a community of researchers for whom the *raison d'être* of their training is to exploit synergies between neuroscience and AI. Explicit design of new training programs can ensure that the NeuroAI research community reflects the demographics of society as a whole and is equipped with the ethical tools needed to ensure that the development of AI benefits society[63].

Second, we must create a shared platform capable of developing and testing these virtual agents. One of the greatest technical challenges that we will face in creating an iterative, embodied Turing test and evolving artificial organisms to pass it is the amount of computational power required. Currently, training just one large neural network model on a single embodied task (e.g. control of a body in 3-dimensional space) can take days on specialized distributed hardware[64]. For multiple research groups to iteratively work together to optimize and evaluate a large number of agents over multiple generations on increasingly complex embodied Turing tests, a large investment in a shared computational platform will be required. Much like a particle accelerator in physics or large telescope in astronomy, this sort of large-scale shared resource will be essential for moving the brain-inspired AI research agenda forward. It will require a major organizational effort, with government and ideally also industry support, that has as its central goal scientific progress on animal and human-like intelligence.

Third, *we must support fundamental theoretical and experimental research on neural computation*. We have learned a tremendous amount about the brain over the last decades, through the efforts of the NIH, in no small measure due to the BRAIN Initiative, and other major funders, and we are now reaching an understanding of the vast diversity of the brain's individual cellular elements, neurons, and how they function as parts of simple circuits. With these building blocks in place, we are poised to shift our focus toward understanding how the brain functions as an integrated intelligent system. This will require insight into how a hundred billion neurons of a thousand different types, each one communicating with thousands of other neurons, with variable, adaptable connections, are wired together, and the computational capabilities—the intelligence—that emerges. We must *reverse engineer the brain* to abstract the underlying principles. Taking advantage of the powerful synergies between neuroscience and AI will require program and infrastructure support to organize and enable research across the disciplines at a large scale.

Fortunately, there is now broad political agreement that investments in AI research are essential to humanity's technological future. Indeed, IARPA (Intelligence Advanced Research Projects Activity) was a pioneer in this field, launching the Machine Intelligence from Cortical Networks (MICrONS) project. This project spearheaded the collection of an unprecedented data set consisting of a portion of a mouse connectome and associated functional responses with the specific goal of catalyzing the development of next-generation AI algorithms[65]. Nonetheless, community-wide efforts to bridge the fields of neuroscience and AI will require robust investments from government resources, as well as oversight of project milestones, commercialization support, ethics, and big bets on innovative ideas. In the U.S., there are currently some lines of federal resourcing, such as the NSF's National Artificial Intelligence Research Institutes, explicitly dedicated to driving innovation and discovery in AI from neuroscience research, but these are largely designed to support a traditional academic model with different groups investigating different questions, rather than the creation of a centralized effort that could create something like the

CamScanner

embodied Turing test. Likewise, AI support grants in the U.S. are predominantly ancillary programs through the NIH, NSF, DoD, and even the EPA—each of which have their own directives and goals—and this pattern is shared by funding agencies globally. This leaves a significant funding gap for technology development as an end in itself. The creation of overarching directives either through existing entities, or as a stand-alone agency, to support NeuroAI and AI research would drive this mission and consequently unlock the potential for AI to benefit humanity.

Conclusions

Despite the long history of neuroscience driving advances in AI and the tremendous potential for future advances, most engineers and computational scientists in the field are unaware of the history and opportunities. The influence of neuroscience on shaping the thinking of von Neumann, Turing and other giants of computational theory are rarely mentioned in a typical computer science curriculum. Leading AI conferences such as NeurIPS, which once served to showcase the latest advances in both computational neuroscience and machine learning, now focus almost exclusively on the latter. Even some researchers aware of the historical importance of neuroscience in shaping the field often argue that it has lost its relevance. "Engineers don't study birds to build better planes" is the usual refrain. However, the analogy fails, in part because pioneers of aviation did indeed study birds[66,67], and some still do[68,69]. Moreover, the analogy fails also at a more fundamental level: The goal of modern aeronautical engineering is not to achieve "bird-level" flight, whereas a major goal of AI is indeed to achieve (or exceed) "human-level" intelligence. Just as computers exceed humans in many respects, such as the ability to compute prime numbers, so too do planes exceed birds in characteristics such as speed, range and cargo capacity. However, if the goal of aeronautical engineers were indeed to build a machine with the "bird-level" ability to fly through dense forest foliage and alight gently on a branch, they would be well-advised to pay very close attention to how birds do it. Similarly, if AI aims to achieve animal-level common-sense sensorimotor intelligence, researchers would be well-advised to learn from animals and the solutions they evolved to behave in an unpredictable world.

References

1. Hassabis, D., Kumaran, D., Summerfield, C. & Botvinick, M. Neuroscience-inspired artificial intelligence. *Neuron* **95**, 245–258 (2017).
2. Macpherson, T. et al. Natural and artificial intelligence: a brief introduction to the interplay between AI and neuroscience research. *Neural Netw.* **144**, 603–613 (2021).
3. McCulloch, W. S. & Pitts, W. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.* **5**, 115–133 (1943).
4. von Neumann, J. *First Draft of a Report on the EDVAC*. https://doi.org/10.5479/sil.538961.39088011475779 (1945).
5. von Neumann, J. *The Computer and the Brain* (Yale University Press, 2012).
6. Hubel, D. H. & Wiesel, T. N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.* **160**, 106–154 https://doi.org/10.1113/jphysiol.1962.sp006837 (1962).
7. Fukushima, K. Neocognitron: a self organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol. Cybern.* **36**, 193–202 (1980).
8. LeCun, Y. & Bengio, Y. Convolutional networks for images, speech, and time series. In: *The Handbook of Brain Theory and Neural*. 255–258 (ACM, 1995).
9. Thorndike, E. L. Animal intelligence: an experimental study of the associative processes in animals. https://doi.org/10.1037/10780-000 (1898).
10. Thorndike, E. L. The law of effect. *The Am. J. Psychol.* **39**, 212 https://doi.org/10.2307/1415413 (1927).
11. Thorndike, E. L. The fundamentals of learning. https://doi.org/10.1037/10976-000 (1932).
12. Crow, T. J. Cortical synapses and reinforcement: a hypothesis. *Nature* **219**, 736–737 (1968).
13. Rescorla, R. A. A theory of pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: Black, A. H. & Prokasy, W. F. (eds.) *Classical Conditioning II: Current Research and Theory*. 64–99 (Century-Crofts, 1972).
14. Klopf, A. H. *Brain Function and Adaptive Systems: A Heterostatic Theory* (AIR FORCE CAMBRIDGE RESEARCH LABS HANSCOM AFB MA, 1972).
15. Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science* **275**, 1593–1599 (1997).
16. Campbell, M., Hoane, A. J. & Hsu, F.-H. Deep blue. *Artif. Intell.* **134**, 57–83 (2002).
17. Silver, D. et al. Mastering the game of Go with deep neural networks and tree search. *Nature* **529**, 484–489 (2016).
18. Reed, S. et al. A generalist agent. https://arxiv.org/abs/2205.06175 (2022).
19. Sinz, F. H., Pitkow, X., Reimer, J., Bethge, M. & Tolias, A. S. Engineering a less artificial intelligence. *Neuron* **103**, 967–979 (2019).
20. Itti, L., Koch, C. & Niebur, E. A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **20**, 1254–1259 (1998).
21. Larochelle, H. & Hinton, G. Learning to combine foveal glimpses with a third-order Boltzmann machine. *Adv. Neural Inform. Process. Syst.* **23**, 1243–1251 (2010).
22. Xu, K. et al. Show, attend and tell: neural image caption generation with visual attention. In: *Proceedings of the 32nd International Conference on Machine Learning* (eds. Bach, F. & Blei, D.) vol. 37, 2048–2057 (PMLR, 2015).
23. Vaswani, A. et al. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **30**, 6000–6010 (2017).
24. Moravec, H. *Mind Children: The Future of Robot and Human Intelligence* (Harvard University Press, 1988).
25. Turing, A. M. I.—Computing machinery and intelligence. *Mind* **LIX**, 433–460 (1950).
26. Brown, T. et al. Language models are few-shot learners. *Adv. Neural Inf. Process. Syst.* **33**, 1877–1901 (2020).
27. Sejnowski, T. Large language models and the reverse turing test. https://arxiv.org/abs/2207.14382 (2022).
28. Brooks, R. A. Intelligence without representation. *Artificial Intelligence*. **47**, 139–159 https://doi.org/10.1016/0004-3702(91)90053-m (1991).
29. Meyer, J.-A. & Wilson, S. W. *From Animals to Animats: Proceedings of the First International Conference on Simulation of Adaptive Behavior* (Bradford Books, 1991).
30. Pfeifer, R. & Scheier, C. Understanding intelligence. https://doi.org/10.7551/mitpress/6979.001.0001 (2001).
31. Pfeifer, R. & Bongard, J. *How the Body Shapes the Way We Think: A New View of Intelligence* (MIT Press, 2006).
32. Ortiz, C. L. Why we need a physically embodied turing test and what it might look like. *AI Magazine*. vol. 37, 55–62 https://doi.org/10.1609/aimag.v37i1.2645 (2016).
33. Merel, J., Botvinick, M. & Wayne, G. Hierarchical motor control in mammals and machines. *Nat. Commun.* **10**, 5489 (2019).
34. Mnih, V. et al. Human-level control through deep reinforcement learning. *Nature* **518**, 529–533 (2015).
35. Huang, S., Papernot, N., Goodfellow, I., Duan, Y. & Abbeel, P. Adversarial attacks on neural network policies. https://arxiv.org/abs/1702.02284 (2017).
36. Zador, A. M. A critique of pure learning and what artificial neural networks can learn from animal brains. *Nat. Commun.* **10**, 3770 (2019).

CamScanner

37. Bommasani, R. et al. On the opportunities and risks of foundation models. https://arxiv.org/abs/2108.07258 (2021).
38. Elman, J. L. Learning and development in neural networks: the importance of starting small. *Cognition* **48**, 71–99 (1993).
39. Lake, B. M., Ullman, T. D., Tenenbaum, J. B. & Gershman, S. J. Building machines that learn and think like people. *Behav. Brain Sci.* **40**, e253 (2017).
40. Doya, K. & Taniguchi, T. Toward evolutionary and developmental intelligence. *Curr. Opin. Behav. Sci.* **29**, 91–96 https://doi.org/10.1016/j.cobeha.2019.04.006 (2019).
41. Pehlevan, C. & Chklovskii, D. B. Neuroscience-inspired online unsupervised learning algorithms: artificial neural networks. *IEEE Signal Process. Mag.* **36**, 88–96 (2019).
42. Stanley, K. O., Clune, J., Lehman, J. & Miikkulainen, R. Designing neural networks through neuroevolution. *Nat. Mach. Intell.* **1**, 24–35 (2019).
43. Gupta, A., Savarese, S., Ganguli, S. & Fei-Fei, L. Embodied intelligence via learning and evolution. *Nat. Commun.* **12**, 5721 (2021).
44. Stöckl, C., Lang, D. & Maass, W. Structure induces computational function in networks with diverse types of spiking neurons. *bioRxiv*. https://doi.org/10.1101/2021.05.18.444689 (2022).
45. Koulakov, A., Shuvaev, S., Lachi, D. & Zador, A. Encoding innate ability through a genomic bottleneck. *bioRxiv*. https://doi.org/10.1101/2021.03.16.435261 (2022).
46. Patterson, D. et al. Carbon emissions and large neural network training. https://arxiv.org/abs/2104.10350 (2021).
47. Sokoloff, L. The metabolism of the central nervous system in vivo. *Handb. Physiol. Sect. I Neurophysiol.* **3**, 1843–1864 (1960).
48. Boahen, K. Dendrocentric learning for synthetic intelligence. *Nature* **612**, 43–50 (2022).
49. Dobrunz, L. E. & Stevens, C. F. Heterogeneity of release probability, facilitation, and depletion at central synapses. *Neuron* **18**, 995–1008 (1997).
50. Attwell, D. & Laughlin, S. B. An energy budget for signaling in the grey matter of the brain. *J. Cereb. Blood Flow. Metab.* **21**, 1133–1145 (2001).
51. Lennie, P. The cost of cortical computation. *Curr. Biol.* **13**, 493–497 (2003).
52. Davies, M. et al. Advancing neuromorphic computing with loihi: a survey of results and outlook. *Proc. IEEE Inst. Electr. Electron. Eng.* **109**, 911–934 (2021).
53. Cisek, P. & Hayden, B. Y. Neuroscience needs evolution. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **377**, 20200518 (2022).
54. Mathis, A. et al. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* **21**, 1281–1289 (2018).
55. Wu, A. et al. Deep Graph Pose: a semi-supervised deep graphical model for improved animal pose tracking. *Adv. Neural Inf. Process. Syst.* **33**, 6040–6052 (2020).
56. Marshall, J. D. et al. Continuous whole-body 3D kinematic recordings across the rodent behavioral repertoire. *Neuron* **109**, 420–437.e8 (2021).
57. Pereira, T. D. et al. Publisher Correction: SLEAP: A deep learning system for multi-animal pose tracking. *Nat. Methods* **19**, 628 (2022).
58. Merel, J. et al. Deep neuroethology of a virtual rodent. in *International Conference on Learning Representations* (Association for Computing Machinery, 2020).
59. Todorov, E., Erez, T. & Tassa, Y. MuJoCo: A physics engine for model-based control. in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems* (IEEE, 2012).
60. Brockman, G. et al. OpenAI Gym. (2016) https://doi.org/10.48550/arXiv.1606.01540.
61. Kitano, H., Asada, M., Kuniyoshi, Y., Noda, I. & Osawa, E. RoboCup: The Robot World Cup Initiative. in: *Proceedings of the first international conference on Autonomous Agents*. 340–347 (Association for Computing Machinery, 1997).
62. Bell, R. M. & Koren, Y. Lessons from the Netflix prize challenge. *ACM SIGKDD Explorations Newsletter*. vol. 9, 75–79 https://doi.org/10.1145/1345448.1345465 (2007).
63. Doya, K., Ema, A., Kitano, H., Sakagami, M. & Russell, S. Social impact and governance of AI and neurotechnologies. *Neural Netw.* **152**, 542–554 (2022).
64. Liu, S. et al. From motor control to team play in simulated humanoid football. https://arxiv.org/abs/2105.12196 (2021).
65. MICrONS Consortium et al. Functional connectomics spanning multiple areas of mouse visual cortex. *bioRxiv* https://doi.org/10.1101/2021.07.28.454025 (2021).
66. Lilienthal, O. *Birdflight as the Basis of Aviation: A Contribution Towards a System of Aviation, Compiled from the Results of Numerous Experiments Made by O and G Lilienthal*. (Longmans, Green, 1911).
67. Culick, F. What the Wright Brothers did and did not understand about flight mechanics-In modern terms. in *37th Joint Propulsion Conference and Exhibit* (American Institute of Aeronautics and Astronautics, 2001).
68. Shyy, W., Lian, Y., Tang, J., Viieru, D. & Liu, H. *Aerodynamics of Low Reynolds Number Flyers*. (Cambridge University Press, 2008).
69. Akos, Z., Nagy, M., Leven, S. & Vicsek, T. Thermal soaring flight of birds and unmanned aerial vehicles. *Bioinspir. Biomim.* **5**, 045003 (2010).

Acknowledgements

This Perspective grew out of discussions following the first conference From Neuroscience to Artificially Intelligent Systems (NAISys), held in CSHL in 2020, and includes many of the invited speakers from that meeting. A.Z. would like to thank Cat Donaldson for convincing him that this manuscript would be worth writing, and for contributing to early versions. B.Ö. would like to acknowledge valuable contributions on the manuscript from Diego Aldorando. The authors would like to acknowledge the following funding sources—A.Z.: Schmidt Foundation, Eleanor Schwartz Foundation, Robert Lourie Foundation; S.E.: NIH 5U19NS104649, 5R01NS105349; R.B.: CIFAR, NSERC RGPIN-2020-05105, RGPAS-2020-00031; B.Ö.: NIH R01NS099323, R01GM136972; Y.B.: CIFAR; K.B.: Stanford Institute for Human-Centered Artificial Intelligence, NSF 2223827; A.C.: NIH U19NS123716; J.D.: Semiconductor Research Corporation, DARPA, Office of Naval Research MURI-114407, MURI-N00014-21-1-2801, N00014-20-1-2589, NSF CCF-1231216, NSF 2124136, Simons SCGB 542965; S.G.: Simons Foundation, James S. McDonnell Foundation, NSF CAREER Award; A.M.: Federation of American Scientists; B.O.: NSF IIS-1718991; C.S.: NIH 1R01MH125571-01, R01NS127122, NSF 1922658, Google faculty award; E.S.: Simons Foundation, NIH EY022428; S.S.: NIH NINDS NS053603; D.S.: Simons 543049; A.T.: IARPA D16PC00003, DARPA HR0011-18-2-0025, NIH R01EY026927, R01MH109556, NSF NeuroNex DBI-1707400; D.T.: HHMI.

Author contributions

A.Z. conceived the project and wrote the first draft. A.Z., S.E., B.R., B.Ö. provided extensive editing to successive drafts. Y.B., K.B., M.B., D.C., A.C., C.C., J.D., S.G., J.H., K.K., A.K., Y.L., T.L., A.M., B.O., A.P., C.S., T.S., E.S., S.S., D.S., A.T., and D.T. provided additional guidance and edits at various stages.

Competing interests

The authors declare their engagements with the following relevant for-profit entities—A.Z.: Cajal Neuroscience; S.E.: Herophilus; R.B.: Google DeepMind; B.Ö.: Blackbird Neuroscience; K.B.: Femtosense Inc., Radical

CamScanner

Semiconductor, Neurovigil; C.C.: Google DeepMind; S.G.: Meta; J.H.: Numenta; K.K.: Paradromics; T.L.: Google DeepMind; A.M.: Google DeepMind, Kernel; D.S.: Meta; A.T.: Vathes Inc., Upload AI LLC, BioAvatar LLC. The remaining authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Anthony Zador.

Peer review information *Nature Communications* thanks Wolfgang Maass and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2023

CamScanner