Tool evolution as a prerequisite for consciousness

Carsten Korth

Carsten Korth*

Tool evolution as a prerequisite for consciousness

https://doi.org/10.1515/revneuro-2024-0166
Received November 20, 2024; accepted January 18, 2025;
published online February 20, 2025

Abstract: Within the concept of the extended mind, the active modification of external objects, externalizations, is seen as an auxiliary means to adapt to the environment. Toolmaking and use are advanced stages of externalizations that evolve. All past or present tools can, theoretically, be precisely assigned a location in an evolutionary tree with predecessors and progeny. Tools are reliably replicated, modified, and selected by their ability to facilitate human needs. Tool evolution, therefore, fulfills Darwinian criteria where the material tool is the phenotype and the instruction to build it is the code. The ostensive triangle consisting of a pointing individual, an observing individual, and a pointed-at object or tool is the germ cell of social transmission of instructions. Tool-building instructions ultimately can be reduced to distinct sequences of motor acts that can be recombined and are socially transmitted. When executed, they replicate tools for the reward of convenience or improved fitness. Tools elicit affordances relating to their use that synchronize different individuals' perceptions, result in psychological "understanding," and thereby modify social networks. Massive tool fabrication as present today in the "tool-sphere" has, therefore, accelerated prosociality and over time led to the acquisition of an individual's third person perspective. The entangled biological evolution accelerated the ongoing cumulative cultural evolution by selecting traits facilitating social transmission. In this context, tool evolution and the corresponding acquired individual instructional content is a precondition to the emergence of higher cognition and "consciousness." A neuroscience investigating externalizations as the starting point of this process is urgently needed.

Keywords: tool evolution; extended mind; cognitive offloading; social transmission; consciousness

This paper provides a comprehensive concept for how advanced cognition emerged from the required social-borne acquisition of tool-making instructions during the coevolution of cumulative culture and biology. The particular angle and a crucial element of this proposed entanglement is the to this end underappreciated fact that most of our cognitive functions involve external objects and that these objects or tools are tightly interwoven in our cognition and thinking leading to the conclusion that the omnipresent interaction with tools was also a prerequisite for higher cognition.

As will be detailed, tool evolution as the crown event of the extended mind results from a cyclic, algorithmic process, which involves social transmission of instructions to make and use tools, modifications of those instructions, and ultimately selecting them resulting in an ever-repetitive Darwinian process of replication, variation, and selection. Yet the biological foundations of how the brain externalizes and integrates external objects into cognition, i.e., a neuroscience of externalization, is widely absent.

This paper aims to provide a theoretical scaffold supported by experimental data and yielding testable hypotheses rather than a definite explanation, with the goal of inspiring concrete experimental investigations into a future neuroscience of externalization.

The line of reasoning in this review is developed from the concept of the extended mind to basic communication mechanisms through ostension, to tool evolution and the social transmission of tool-building and making instructions. Consequences of the growing tool-sphere for intersubjective behavior, affordances, social synchronization, and the acceleration of prosociality are derived. Individual advantages through tool interactions are demonstrated within the theory of predictive processing and an intersubjective theory of understanding, as well as the emergence of a third person perspective and content consciousness. Finally, the entanglement of evolutionary mechanisms between biological and tool evolution are discussed and a neuroscience of externalizations is outlined. A glossary of term definitions can be found in Table 1.

1 The extended mind

1.1 Examples of the extended mind and externalizations

A comprehensive view of cognition must involve the integration of external objects that are necessary for externalizing memory or other cognitive functions, termed active externalism (Clark 2024; Clark and Chalmers 1998; Menary 2010; Noe 2010; O'Regan and Noe 2001). "External" is defined

***Corresponding author: Carsten Korth**, Heinrich Heine University Düsseldorf, Moorenstrasse 5, 40225 Düsseldorf, Germany,
E-mail: ckorth@hhu.de. https://orcid.org/0000-0003-1503-1822

Table 1: Glossary and abbreviations.

Cumulative cultural evolution	The evolution of tools with transgenerational social transmission, i.e., the accumulation of tools or culture beyond an individual's ability to reproduce them during life time
Deixis	Taking the relative perspective of the observed person by the observer
Extended mind	The idea that objects outside the brain fulfill a necessary, active function for cognition
Extended phenotype	The concept that the materialization of a genotype, the phenotype, is not restricted to a body but also extends to its environment, like nid building for birds
External	Is defined as being outside one's body's space as limited by its physical, discrete borders such as the skin
Externalism	The idea that objects outside the brain are actively sought for to support cognitive functions
Externalization	A cognition that is assigned to an object
Evolvability	The capacity to increase generating phenotypic variety in order to speed up the evolutionary process
fMRI	Functional magnetic resonance imaging – an imaging method enabling to measure blood oxygenation as a correlate for neural activity of distinct brain regions
Imitation	The social copying of a body topography
Ostension	Pointing to an object to designate it to someone else
Prosociality	The idea that (large) groups of humans or other species cooperate and readily form social bonds or networks rather than being hostile to each other
Social transmission	The transfer of information between individuals by communicative means
Synchronization	The adjusting of motor output (behavior), or internal molecular or electrophysiological events between two or more individuals in time, such that they become more similar
Tool	An object that has been modified in a characteristic way to serve a function for the convenience of a user

as being outside one's body's space as limited by its physical, discrete borders such as the skin. Active externalism as defined in the original publication by Clark and Chalmers (1998) (see also Menary 2010 for multiple reviews and discussion) is the concept that material objects outside the brain can be essential parts of cognitive processes in a constitutive, symmetrical, and causal, i.e., active way. The idea of active externalism is not that material objects merely somehow assist in cognition but that they are a necessary constituent of a cognitive process, meaning that a particular kind of cognition could not take place without external objects. This, however, does not mean that every cognitive process has an external component. In cognitive processes involving active externalism "the human organism is linked with an external entity in a two-way interaction, creating a coupled system that can be seen as a cognitive system in its own right" (Clark and Chalmers 1998). Consequently, if the external component is removed, cognition will collapse as if a part of the brain had been removed. There is a fundamental continuity between the brain and cognitive extensions into the external world (Clark 2024). This initial radical argument of Chalmers and Clark has subsequently been modified to allow a more malleable internal versus external distinction as "outsourced cognition" (Gerken 2014).

The high degree of how (even transiently) external objects are readily integrated into our perception is, for example, demonstrated by the rubber hand illusion (Botvinick and Cohen 1998) and its subsequent variations, including a whole body illusion (Lenggenhager et al. 2007) where external body (parts) can get "manipulated" into our body image. Along similar lines, external tools such as mirrors have been shown to be able to appease phantom limb pain through feeding back body images into our perception and thereby beneficially complementing aberrant brain-internal functional connectivity (Chan et al. 2007; Ramachandran and Rogers-Ramachandran 1996). Evidence for the deep anchoring of external objects into brain computation also comes from experiments with monkeys where it was demonstrated that training for tool use changes functional network representations (Iriki and Sakura 2008). Furthermore, elaborate forms of memory like autobiographical memory are externalized and dependent on social interactions (Heersmink 2022; Van Bergen and Sutton 2019).

A simple example for everyday's active externalism is a diary in which appointments are noted (Clark and Chalmers 1998): when looked up, these appointments can be fulfilled, i.e., the diary is causally active. Without the diary, one might not have performed this cognitive process. In this case, memory was externalized, and minimized: instead of prospectively memorizing a series of appointments for upcoming weeks by heart, it was reduced to the instruction: "look in the diary at the appropriate date, then read and short-memorize the time for this or that appointment."

Another example for progress through externalization is mathematics: beyond the limited individual's ability to perform mental arithmetics by heart, calculating is mainly done with the help of external tools, be it as simple as using a pen and a paper, in order to externalize memory. Clearly, beyond the number hundred or so, most individuals are not able to perform multiplications or divisions by mental arithmetics alone but have learned to use a pen and a paper to assist in a memorized, abstract algorithmic calculating procedure. A calculator as an advanced tool has now replaced these simple tools used until about 50 years ago. Since advances in the natural sciences necessarily involve the generation of quantitative data and rather complex statistical mathematics, it is fair to state that without cognitive externalism there would be no natural sciences as we know them today (see also Fabry 2019 for an extended perspective of the enculturated origins of mathematics). Mathematics historically emerging from external tool evolution as presented here (detailed in chapter 2d) is thus in opposition to philosophical views assuming mathematics as a given (Tegmark 1998).

While an externalization of memory into “exograms” (Donald 1991) has massively increased memory storage capacity, it also has made memory more vulnerable to manipulation (Risko et al. 2019). Brain-centered (internal) cognition, as opposed to external cognition, is an exception and certainly not as powerful as a cognition that is a hybrid of brain-internal and environmental (outside-brain) tools that are causally coupled and reciprocally active (Menary 2010; Noe 2010). The explicit use of (external) objects for external memory has also been investigated under the term “cognitive off-loading” (Heersmink 2017; Hu et al. 2019; Risko and Gilbert 2016) and has been defined as “the use of physical action to alter the information processing requirements of a task so as to reduce cognitive demand” across several psychological domains (Risko and Gilbert 2016).

1.2 The range of externalizations

The concept of the extended mind is related to the concept of the “extended phenotype” (Dawkins 1982), which states that a phenotype corresponding to a genotype, i.e., the coding genetic information, is not merely constituted in forming a phenotypic body but also equally, via behavioral enactments, in modifying the environment. Examples are beavers building a dam, birds building a nid, and spiders forming a net. Since these environmental modifications executed by individuals require motor and cognitive acts, these animals require primitive “extended minds,” too (Facchin and Leonetti 2024). These primitive externalizations seem rather inborn and not acquired and, therefore, are not subject to social processing or transmission. Thus, while features of externalization themselves are not human-specific, externalizations triggering a cumulative cultural/object-driven evolutionary process are, as will be explained below.

Exactly why the transition of externalizations to tools and foremost a more extensive tool evolution happened only in humans is unclear. Likely, an emergence of rudimentary prosociality setting the frame for enabling social transmission such as the ability for “joint attention” (O’Madagain and Tomasello 2022; Tomasello 2014) and primitive tool use were biological preconditions to spark a coevolution between cultural evolution and culture-conditioned biological fixations (see below chapter 8) facilitating further cultural evolvability, i.e., the facilitation of biological hardware to accelerate cultural uploading and evolution. As will become clearer when following the line of reasoning in this paper, one major claim here is that externalizations by tools promoted prosociality, which in turn further accelerated cultural evolution and this mutual, repetitive reinforcement thus contributed to the evolvability (Kirschner and Gerhart 1998) both of the biological and cultural realm.

Earliest excavated material externalizations are artifacts such as cave paintings or statues that date back to around 40,000 years BCE (Conard 2003; Heersmink 2024). This, however, does not mean that perishable artifacts from wood or bones could not have been generated much earlier. The start of cumulative cultural evolution has been dated to around 600,000 BCE (Paige and Perreault 2024). Language, which in its earliest forms is dated back to 300,000 BCE (Fitch 2010) – 500,000 BCE (Perreault and Mathew 2012), is an externalization too, if we allow a broad definition. Ultimately, through distinct motor acts of the more than 100 different muscles from tongue, lips, jaw, and pharynx, sounds are generated that leave a temporary print on air molecules, which is transmitted between close-by individuals.

The role of language for evolving social networks and prosociality is eventually more intuitive than that of tools and likely language and tool evolution went hand in hand rather than one preceding the other as newest brain functional magnetic resonance imaging (fMRI) findings in humans suggest that reported activation of similar brain structures by syntactic use of language and tools (Morgan et al. 2015; Thibault et al. 2021). Language is transmitted between individuals in a social network and learning how to speak is usually transmitted vertically from parents to children. Oral language is an externalization process already, even though its mere acoustic dimension makes it a fleeting, transient “object” or transient tool (see chapter 2c), an ultra-fast expiring movement of air molecules requiring temporal resolution abilities both for sending as well as for receiving information. Even though language evolution will not be an object of this paper (for extensive reviews see Dennett 2017; Dunbar 1996; Everett 2017; Fitch 2010), it is important to note that the fleeting character of external “word objects” offers both advantages and disadvantages compared to that of permanently present, modified external objects, i.e., tools, that led to different evolutionary trajectories and cannot be compared here in depth (Figure 1).

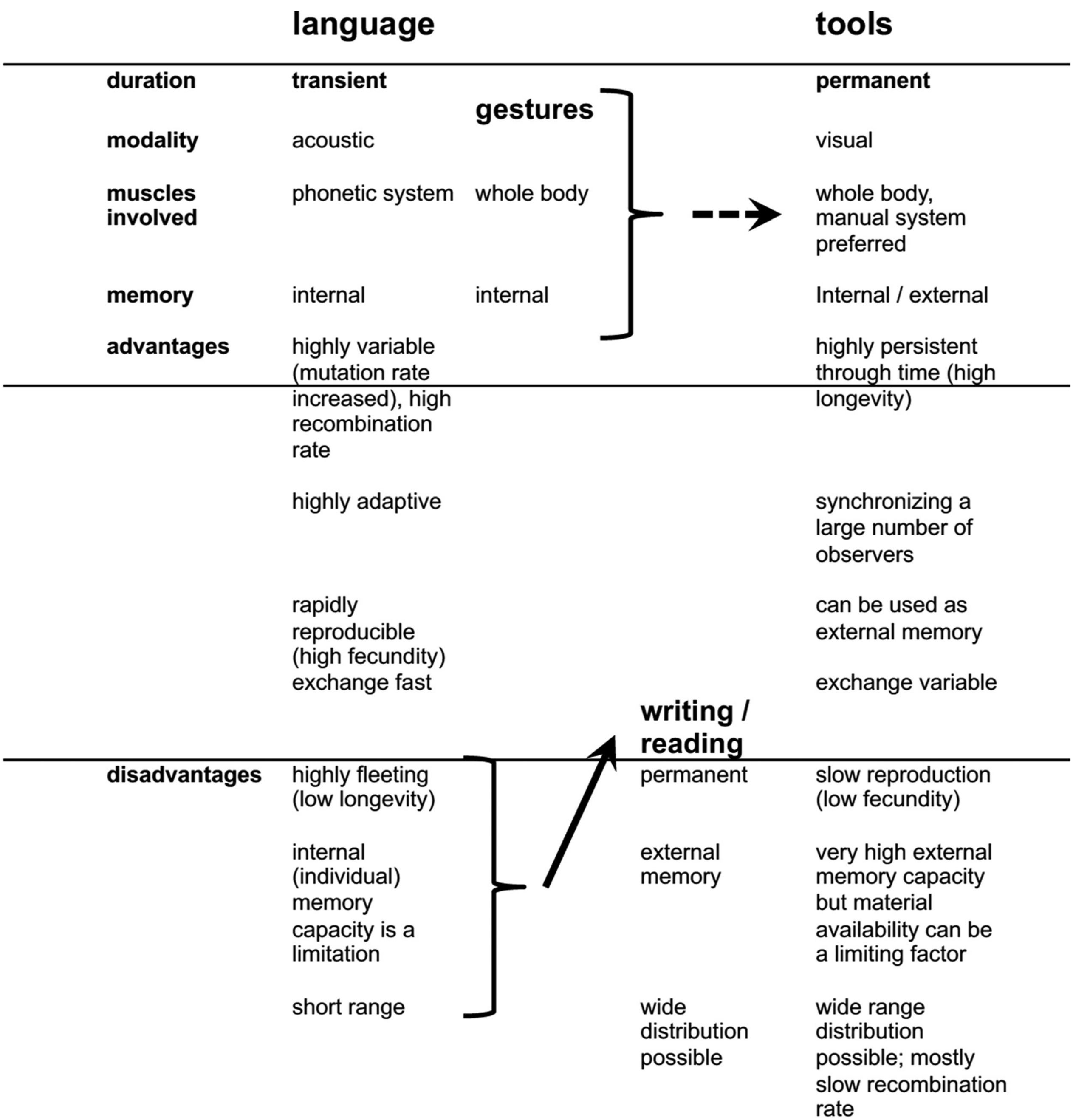

Figure 1: Comparing transient with permanent externalizations. Whereas language consists of transient externalizations, tools constitute relatively permanent externalizations – both with different advantages and disadvantages (the list is selective and incomplete). Initially, transient externalizations were repurposed or transformed to permanent tools at least twice: when motor acts such as gestures were used to execute instructions to make tools and when writing was invented as the permanent correlate to fleeting language (see arrows).

1.3 The ostensive triangle as the germ cell for igniting the evolution of externalizations

Given the likely parallel evolution of language and tool fabrication, it is instructive to analyze the beginnings of the evolution of language with regard to similarities with the beginnings of the evolution of tool use. Rolfe has given an excellent sketch on a plausible gradualistic, phylogenetic evolution of language acquisition, starting with profoundly simple interindividual interactions and ending in epistemic functions (Rolfe 1996) (Figure 2). The application of the ostensive triangle was developed for language but can easily be transposed for explaining tool evolution. The ostensive

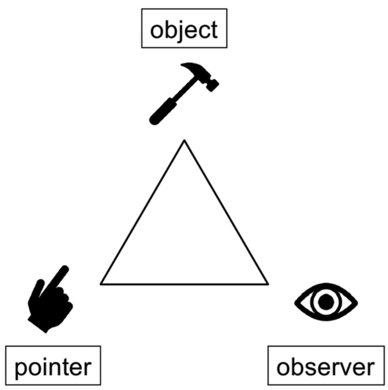

Figure 2: The ostensive triangle. The ostensive triangle as the germ cell of social transmission consists of the pointer, the observer, and the pointed-at subject or object. Whether a pointed-at subject preceded objects in anthropological evolution is discussed in the text.

triangle consists of three parties: a pointer (ostension), an observer, and a pointed-to object (or subject; see Figure 2, see also chapter 3).

A precondition of ostension and the first step for establishing social transmission is the solicitation between (at least) two individuals that share "joint attention" (see also chapter 3b), i.e., establish a joint frame of reference. The development of joint attention and the early apprehension of "organizing activity," i.e., basic dialogical abilities have been the topic of intensive research (O'Madagain and Tomasello 2022; Tomasello 2014) and are developed in humans as early as during breast feeding where, for example, already the sequential nature of chunks of social interaction are trained (Noe 2015).

Once joint attention of two individuals has been established through solicitation, ostension is executed by the active individual aiming to transmit an information: a distinct pointing to designate an object intending to draw the information-receiving individual's attention toward it so that the object can eventually be associated with a name by either a gesture or a sound (language), which will subsequently represent that object. The pointer sets a relative frame of reference centered on himself (deixis) that the receiving individual can relate to. This frame of reference requires a form of theory of mind since the observer has to adapt to the perspective of the pointer (O'Madagain and Tomasello 2022; Santiesteban et al. 2012). Pointing as a means of demonstrating intentionality seems to be superior to other attention-drawing or intentional gestures such as bodily expressions, gaze following, or attention getting (van Mazijk 2024). Active gestures executed by individuals have language-like syntax and their sequential structures are as fleeting as that of sound within language. Accordingly, it was hypothesized that "the slow-to-emerge precursor [human] from 5 million years ago to 2 million years ago may have built up a gesture language that derived from instrumental actions … it would have been an evolutionary track to pantomime" (Gärdenfors and Hochberg 2024; McNeill 2013). Sequential visuospatial integration of sight and hand movements was clearly an early requirement for pointing, and, later, for the generation of tools (Bruner and Iriki 2016).

The interaction in the above described ostensive triangle (pointer, receiver, object) is used to assign either a representation in the form of a name to an object (vocalization, gesture) or to perform a modification to the object as a first stage of tool evolution. The ostensive triangle is the germ cell of social transmission of externalizations setting in motion through its infinite repetition cumulative cultural evolution. The ostensive triangle tightly interweaves social communication and the transmission of object designations (with representational or instructional content), i.e., externalizations.

The ostensive triangle is similar to what has been termed triadic bodily mimesis, which has been emphasized as a main distinguishing factor in enabling cumulative cultural evolution, not present for example, in the great apes (Zlatev et al. 2005). Dyadic mimesis by contrast, according to this reasoning, just reflects a bipartite relationship that apes are in part capable of such as mirror self-recognition and the ability to copy bodily shapes on command but not fully implement shared intentionality (O'Madagain and Tomasello 2022; Zlatev et al. 2005).

Does externalization have roots in interpersonal recognition and interactions itself? For example, from a first-person perspective other highly mobile individuals could be regarded as (fleeting) external objects. It is conceivable, therefore, that first object externalizations may have been rooted in social individuals of kin or tribe members, as evidenced by the first symbolic artifacts as the "Löwenmensch" from 40,000 BCE (Conard 2003) or findings that go back even further in time (see Tattersall 2024 for a comprehensive review on the history of anthropologic evolution of symbolic use). The aim to externalize another person's representation, eventually by his/her own ostensive designation, in order to create a memory during that same social peer's absence as an autograph-like object could be conceived as a primordial motivation creating symbolic, i.e., representative artifacts or tools. Along these lines, representations of persons (dolls), just as in ontogenetic development, may have also preceded tools in anthropological tool evolution (a "doll-theory" of symbolic evolution; leftmost picture in Figure 3).

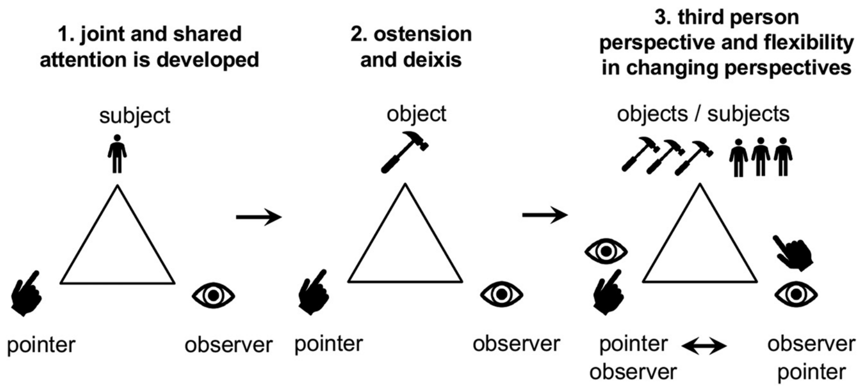

Figure 3: The evolution of a third person perspective from the ostensive triangle. Schematic drawing. 1. Joint attention was developed as an interpersonal ability in a primordial ostensive triangle (see Figure 2) where the attention was drawn (depicted by a pointing finger) to another individual. 2. In the ostensive triangle, objects were designated with representational content in the form of sounds (language) or instructions (tools). 3. As the tool-sphere grew, frequent changes in perspective between the pointer and the observer as well as an increasing number of objects led to social synchronization, which resulted in an extrapolation of multiple second person perspectives: the third person perspective, i.e., a joint minimum frame of reference of the material external world.

2 Tool evolution

2.1 Tool definition

What are tools? In this paper, tools are defined in the broadest sense, which includes all objects that have been modified by human individuals for further manipulative use on yet other objects but also those not used for further manipulation such as artwork, signs, or others. There is a spectrum of tool definitions from narrow (Fragaszy and Mangalam 2018) to broad (see also Mangalam et al. 2022; Ruck and Uomini 2024 for discussion). A classic functionalist definition is (Schubotz et al. 2023; Shumaker et al. 2011): 1. A tool is an object that can be manipulated; 2. During tool use, the tool is touched, held, or otherwise directly manipulated by the tool user; 3. The tool is used in an appropriate context, namely purposefully to achieve a goal; 4. The tool is used to alter the state of another object (including other organisms and the tool user); and 5. The tool extends the user's abilities to achieve certain goals. The latter point is important and clearly relates to the extended mind of the user.

A tool definition that takes an even more active externalist view is from Iriki and Sakura (2008) "external objects in their hands and moving these objects as extensions of their own body … this is the beginning of transitive movement, i.e., tool use." Iriki sees tool use as the end point of a gradual freeing of the limbs from movement and on a trajectory to support cognitive functions of which tool use is the most advanced one (Iriki and Sakura 2008). Tool use remaps the body scheme (extending boundaries) and thereby affects perception of the world (Noe 2010). Latour defines a tool as "a single element held directly in the hand of a man" and contrasts this with machines where single tools are linked to one another rather than to the hand of an individual (Latour 1987).

2.2 How tools are made and replicated

The more specific and functionalist tool definition given in this paper is as follows:

Tools are made from external materials that are modified in one or several steps and assigned to a particular use. The execution of these modifications requires defined sequences of precise motor acts. The protocol of sequential motor acts necessary to replicate a tool is the instructions that are transmitted from individual to individual (Figure 4).

Advanced tools or machines, such as robots, can also fabricate new tools, but this also follows distinct steps (programmed by humans) and the robots themselves as well as their components have been fabricated by humans through yet other motor act-requiring instructions. As such, ultimately, the antecedents of all present tools can be tracked back to a prehistory of components having been generated from raw materials or commodities that have been mined, harvested, or processed by humans, and so forth, extending to an extremely long, but not infinite, chain of linked events. Today's present tools are end points of branches in a large evolutionary tree where the thicker branches represent the predecessors of today's tools (thinner branches) similar to a tree of life (i.e., a pedigree of species) in biological evolution.

The important point is that all present or past tools have a position in this pedigree and thus ultimately can all be traced back to the simplest tools, such as handaxes, or similar. The existence of such an unambiguous pedigree

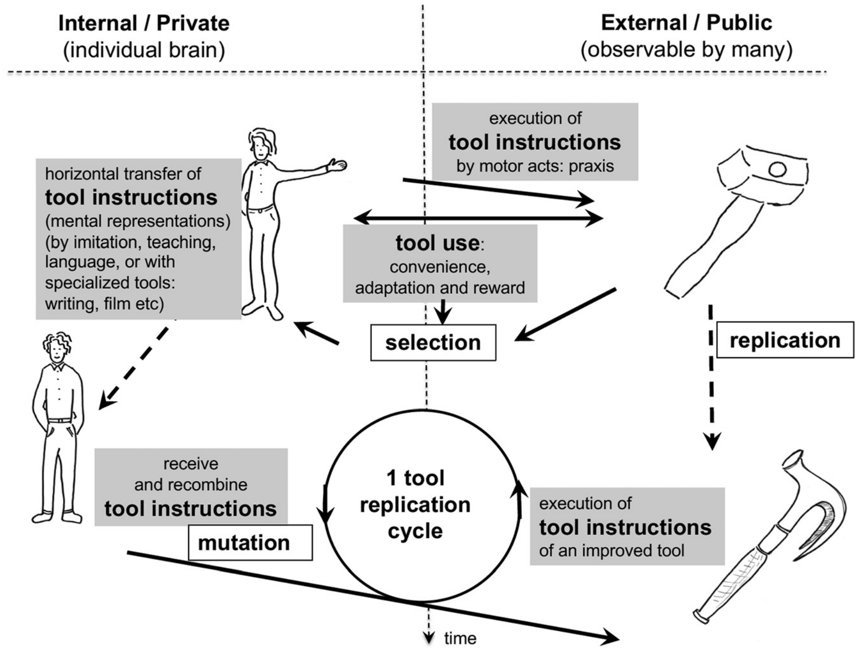

Figure 4: The tool replication cycle. Depicted is the flow between an individual's brain-internal tool instructions and the execution of those instructions during toolmaking. The modified object, the tool, is subsequently evaluated/ selected for convenience or fitness during tool use. When the instructions are socially transmitted, they might be modified resulting in a modified, novel tool (or not), and ending one tool replication cycle.

proves that there is Darwinian evolution in the tool-sphere, i.e., the "anthropogenic mass" (Elhacham et al. 2020), with its three characteristics replication (by transmitting the instructions tool fabrication), variation (by modifying these instructions), and selection (by deciding whether a modified tool is better suited for a given purpose than its predecessor; see Figure 4). Evolution of the tool-sphere was so successful that, as of today, it exceeds the biosphere by mass (Elhacham et al. 2020).

The three elements of Darwinian evolution: replication, variation, and selection can be abstracted from their exclusive application for modeling biological evolution to other replicating, material systems, provided that those reliably feature these three elements. The overarching concept is that Darwinian evolution is a functionalist model to explain the historical and causal trajectory of phenomena not restricted to genetic evolution (Skinner 1974). Dawkins and others (Lewontin 1970) stated that any system that replicates, induces variation and uses selection necessarily and unconditionally evolves along Darwinian characteristics. Dawkins called these functionalist principles "Universal Darwinism" (Dawkins 1983) (see also chapter 3d). The embodiment and kinetics of the replicating material within Universal Darwinism was never conceived to necessarily resemble genes but was conceived to feature any other plausible form. Darwin himself had proposed that cultural evolution demonstrated similarities to biological evolution in that a large number of gradual changes could amount to large changes over time that, for the uninformed observer who had missed the previous gradual changes, could present itself as inexplicable (Darwin 1871).

Dawkins distinguished the replicator and the vehicle as forming a Darwinian unit, where the replicator is ultimately the material that is replicated and modified during a Darwinian replication cycle of modification and replication, whereas the vehicle is the material into which the replicator and its modifications are translated to and, importantly, subjected to selection (Dawkins 1976/2016). The fitness of the vehicle, i.e., the convenience of how well it interacts with its ecological niche determines selection of the tightly physically connected replicator. In biological evolution, the replicator is the genetic material consisting of deoxy nucleic acid or ribonucleic acid, whereas the vehicles are cells, species, viruses, etc. Dawkins elaborated that the successful, surviving replicator has to display fecundity (a high replication rate), longevity (a long life), and fidelity (a high degree of similarity of the copy to the original) in order to outcompete other replicators. Darwinian evolution is fundamentally gradual rather than saltatory. Seeming gaps in a gradual sequence are explained with a "winner-takes-all" phenomenon: if the gradual development of an innovation is extremely successful (in terms of being adaptive, conveying advantages and being selected), all previous gradual stages are superseded and die out, as, for example, discussed for the evolutionary development of the eye enabling vision.

Could a similar trajectory be conceived for cumulative cultural evolution? Szathmary and Maynard Smith have rated the emergence of culture as one of the major transitions in evolution (Maynard Smith and Szathmary 1995). Clearly, several authors have proposed that cultural evolution follows Darwinian principles, too (Boyd and Richerson 1985; Cavalli-Sforza and Feldman 1981; Mesoudi et al. 2004).

Heyes and others have defined cumulative cultural evolution as being executed by social transmission in such a way that the behavior of a group of individuals is improved over successive generations (Heyes 2023). A "ratcheting the ratchet"-effect has been proposed to occur when previous success in tool achievements is used by successive generations (Tennie et al. 2009; Tomasello et al. 1993; Whiten 2022). In that respect, the term "cumulative" is important because it emphasizes the transgenerational nature of progress: progress that couldn't be achieved in individual's lifetime but that critically depends on prior knowledge accumulated by others (Tennie et al. 2017). Nonhuman primates such as chimpanzees have been demonstrated to exhibit primordial evolution of tools (see Arbib et al. 2023; Whiten 2005, 2022 for review). Recent evidence suggests that chimpanzees feature even primordial cumulative culture (Gunasekaram et al. 2024), which would mean that social transmission of tool use in these apes is indeed transgenerational (reviewed in Richerson and Boyd 2024). Intermediate or primordial forms of tool use are proof that cumulative cultural evolution evolved by gradual improvements and that Darwinian mechanisms are at work. Dennett (2017) has provided an illustrative example of cumulative tool evolution: during canoe building in Polynesia, subsequent generations of sailors built improved canoes that were tested for use on the ocean but only those designs were selected for replication that made it back from a life-threatening journey – the unsuccessful designs were selected out by the drowning of their builders, the carriers of the flawed toolmaking instructions.

More similarities exist between biological Darwinian and cumulative cultural evolution: innovation, for example, can also emerge from anastomosis, i.e., when previously separated branches of the evolutionary tree rejoin, such as in biological symbiosis. An analogy in tool evolution would be machines such as a car where several differently evolved tools are assembled to a novel tool with a novel purpose and reward (Dennett 2017; Latour 1987).

2.3 Language as a transient, fleeting "tool"

Language evolved in the acoustic modality where externalizations/signals are transient and limited to a moment, contrary to the visual modality where externalizations are usually permanent (to variable degrees; Figure 1). Language, as a highly specialized motor memory (of the phonetic muscles), is entirely stored in internal memory, which has to grow along with increasing linguistic skills and vocabulary. Pronounced single words are single externalizations just as the simplest tools, but these motor acts that are – like gestures – fleeting and do not lead to permanent modifications of an object, i.e., a tool or artifact. Complex arrays of words connected by syntactic structures are thus similar to more elaborated tools as shaped by multiple motor acts (instructions). A breakthrough discovery was the similarity of activated brain areas both during tool-related acts and language (Morgan et al. 2015; Thibault et al. 2021) suggesting that their biological correlates have a common origin. These findings suggest that language and tool evolution cannot be conceptualized separately, and it will be interesting to figure out the mutual communicative interactions between transient and permanent externalizations, i.e., language and tools, respectively, and the possible effects on their evolvability.

In language, the evolution of semiotic progression from indexes via icons as the earliest forms of intentional linkage to symbols required a normative convention (ostensive triangle, Figure 2, see also Everett (2017). Language allows to variably direct attention (Clark 2011) and heavily relies on the ability to internally memorize and process sequences of a more general nature. The domain general nature of language is supported by evidence that in dyslexics not only reading but also other "sequential functions" are impaired (Heyes 2018). Even though the cultural evolutionary origin of language was questioned for a long time (Chomsky 1965), increasingly, this view is replaced by a cultural, evolutionary view of language supported by a wealth of empirical data (for discussion see Dennett 2017; Everett 2017; Heyes 2018). In that context, for example, it has been claimed that grammatical construction is a unit of Darwinian-like selection itself (Steels and Szathmary 2018). Likely, alternating cycles of cultural evolution and genetic fixations during language evolution are at the origin of today's language (Henrich 2016). Language was subsequently "materialized" upon the invention of writing and thereby became a more permanent tool losing its fleeting character.

2.4 The language-writing transition: transforming externalizations from the transient to the material stage

Writing/reading recreates the transience of language by transposing phonetic movements to eye movements when tracking sequential letter strings. For simplicity, writing/reading will subsequently be referred to as "writing" only.

The analysis of the evolution from language to writing, or in the context of this paper, the evolution of language micro-instructions, i.e., motoric acts of the phonetic system with representational content to writing micro-tools, i.e., tools with representational content eliciting defined affordances, is an excellent and instructive example of how the materialization of previously fleeting internal instructions leads to a higher fidelity, higher longevity, and, eventually, higher fecundity of the replicated instructions.

Archeological evidence for the origins of writing starts to emerge from around 8,000 BC on in mesopotamia with the appearance of clay tokens in the wake of agriculture and the need to document transfer and property of commodities (Schmandt-Besserat 1996). Clearly, the initial purpose of writing was both economic and arithmetic in nature (for a detailed view on the enculturated practice of arithmetics see Fabry 2019; Menary 2015). Each clay token corresponded to one kind of good, had a unique geometric form, and was an entirely man-made artifact. By representing a kind of good, the token system allowed to simultaneously process data of different categories, allowed handling of a theoretically infinite number of goods without the risk of depending on human memory, and, importantly, in the absence of the real goods (Schmandt-Besserat 1996). In this initial first step, two events are remarkable: 1. There is an intention to externalize memory, which as a side effect allows to perform functional operations with those external, materialized memories, i.e., to recombine them in special or in random ways (which, "mutation-like" may lead to novel, originally unintended insights), and 2. The appearance of tokens as symbolic representations for goods including a numerical representation for documenting debts also intricately links the appearance of writing to the emergence of mathematics. The materialization of language certainly led to a massive expansion of external memory. The further evolution of the clay tokens allows a gradual reconstruction of steps leading from real goods and commodities to their abstract two-dimensional representations (Schmandt-Besserat 1996): 1. The original clay tokens representing each one a specific amount of commodities were 2. assembled in so-called envelopes, i.e., hollow vessels used for storing and grouping tokens, where the contents was marked upon, to 3. tablets, i.e., the former covers of envelope (vessels) where the contents was marked but without the prior content, to 4. the final step, the mere marking of signs with a pencil-like instrument on a two-dimensional, plain surface representing commodities achieved the full abstraction of representational content. Phonological representations, for example, for designating the names of the proprietor, came later (Malafouris 2013). Thus, gradual, "Darwinian" steps in the language-writing (tool) evolution can be precisely tracked.

Language which when orally transmitted has relatively low high variability thus got transformed into an external, permanent, and potentially lasting tool upon writing. The permanence of the material, i.e., written word increased longevity of the replicator, its fidelity (compare Chinese whispers with copy-writing of the same sentence), and its fecundity. Once paper was invented, subsequently, the book press, or ultimately the internet, the replicator fecundity of writing language rose exponentially. At the same time, the spontaneous drift in change of spoken languages particularly in small communities, due to the relatively low fidelity of the spoken word, was decelerated and the high fidelity of writing fed back to stabilize language variability (Noe 2015). Individual knowledge of writing also changes the attitude of speaking such as when the speaker is influenced by written texts to adopt an attitude of speaking as one would write (Noe 2015). What drives the replication of tools and how are tool-replicating instructions (Figure 4) transferred between humans?

3 Social transmission of instructions

Within the replicating cycle of tools (Figure 4), the transfer of the tool making or tool use instructions is executed by a process called social transmission. The ability for social transmission is known from animals (Harrison et al. 2024) but has advanced to an unmatched ability only in humans, likely reinforced by toolmaking itself. For example, social interaction is frequently used in humans for problem solving as compared to the great apes: while problem solving abilities are similar in adult great apes and small children, the ability to resort to others for help in cognitive tasks is unmatched in human children from the age of 5 years on, indicating that humans have evolved unique social skills (Tomasello 2014). Obviously, social transmission for toolmaking and use is today also transmitted by language, or writing (scripts) but the more interesting aspects are the anthropologically early events of social transmission.

Tool use of simple tools can also be acquired in the absence of social transmission but: 1. social transmission facilitates and accelerates toolmaking or use not unlike in biology where catalysts accelerate chemical events that would eventually take place but not within an affordable timeframe. 2. Complex tools such as machines require pedagogy or precise teacher–student instructions.

3.1 Copying motor acts

Social transmission of language is intuitive and a prototype of social transmission (Dennett 2017) because the fleeting sound is literally transmitted, i.e., a motor pattern of the phonetic muscles generates a sound, the sound travels instantly through oscillations of air molecules before it dissipates, which is long enough to activate a pattern in the brain of the receiver associating the sound with a designated representation. This kind of social transmission similarly applies to toolmaking and use, except that motor patterns in the latter involve a wider range of body muscles rather than only those of the phonetic system, and the product, the tool, is more permanent than the fleeting sound. Ultimately, all transmitted instructions can be reduced to performing a distinct sequence of motoric acts (Tamariz 2019). The specific transition of a simple nonintentional motor act to one with intentional content for a specific achievement has been termed praxis (Ruck and Uomini 2024). Praxis is "an action which is semantically rich, and a physical imposition of higher order (functional and social) internal states upon the external world" (Ruck and Uomini 2024). The instructions for motor acts could be transferred by language or text, which would be an indirect (but efficient) way for social transmission but copying is also possible without language through imitation of movement, similar to what is done during dancing (Hannay et al. 2024; Laland et al. 2016). The evolutionary accumulation of behaviors through an interaction with external memories has been termed a "non-genetically transmitted behaviourome" (Müller 2020).

How do sequences of motoric acts transition from more random events to less random ones, i.e., more significant intentional ones, such as praxis (Ruck and Uomini 2024) or mimesis (Donald 1991; Zlatev et al. 2005) (Figure 5)? Going back to active externalism, for an individual to access external cognition, motor acts have to be executed in order to access the external object, which usually involves the manipulation of this object. For example, this can consist in mere walking to somewhere for finding information or accessing external memory, browsing a book, or it can be more complex tool using. The important thing here is that the individual's interaction with an external material can be unambigously defined by a motor sequence, a sequential "organized act," executed by this individual with the intention of a particular achievement (Noe 2015; Ruck and Uomini 2024; Tamariz 2019).

The execution of a motor act with an external object by the first individual may draw the attention of a second individual. Any attended external material and the first individual's actions will then also become part of the second individual's passive cognition, initially merely drawing attentional resources. At the beginning, the aim or intentions of any actions of the first individual with the external material may not be fully "understood" by the second, observing individual (see chapter 6), but a "meaning" of the first individual's motor sequence, i.e., a change in the evaluation of the observed motor sequence may eventually unfold as further specific interactions of the first individual with the external material are recognized. This process is catalyzed by prior knowledge or experience of the second individual about the specific external material used by the first individual, or prior knowledge about specific motor sequences related to the external material. Prior knowledge will 1. prime specific motor sequence associations elicited by the specific external material memorized in similar situations in the past (by simple associative learning Birch et al.

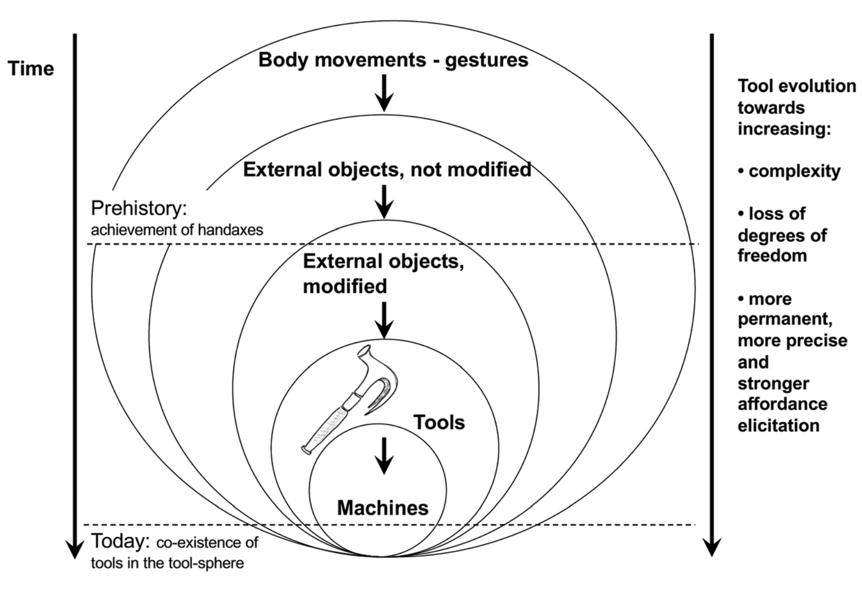

Figure 5: The continuum of externalizations from gestures to modified objects to tools and machines. Schematic drawing showing how simple externalizations such as motor movements evolve to defined sequences of movements such as gestures and, ultimately, lead to object modifications and tools.

2020) and thereby 2. greatly restrict the degrees of freedom for motor sequence associations performed in the context of that external object.

During the developmental ontogenesis of an individual, i.e., one's lifelong learning history starting in early childhood, specific external objects become associated to specific uses by mental representations of motor sequences or may become integrated into the child's body scheme (Bruner and Iriki 2016; Iriki 2006). While these initial associations occur in a social context through observations, imitation, or pedagogy, lateron, the specific external object will elicit those associations, or affordances (see chapter 4) also spontaneously in the absence of other individuals, or aid to predict potential actions of other individuals related to the use of this specific external object.

The passing on of an instruction leading to the making or use of a tool is thus transmitted within the ostensive triangle (Figure 2) where a pointer (or sender in semiotic terms) is in contact (by joint attention) with the observer (receiver in semiotic terms) to point at an object (signifier in semiotic terms) and eventually associates this object with a representative gesture or sound (sign in semiotic terms). When an observing, second individual reproduces the motor act observed from the pointer/sender, the motor act or instruction of the first individual has been replicated after passing through the cognition of the observing individual.

3.2 Joint attention

Joint attention in humans can be established in several domains: eye contact with gaze following to direct attention and perspective, language and sound in the acoustic realm, and gestures, or combinations thereof: all of these domains ultimately involve motor acts as outputs requiring perception as input between the sender and the recipient and, eventually, reverse interaction when sender and receiver are switched. “Our visible eyes provide information about eye direction and cues of mutual gaze that seem crucial to our unique forms of learning, cooperation, and communication” (Hare 2017). “Joint attention, which allows children to learn the association between labels and objects, relies on a child's ability to follow an adult's line of gaze” (Tomasello 2019; Tomasello et al. 2007). Experimental data report that when eye blinking is synchronized between individuals, they improve the coordination of their subsequent motor activities (Koike et al. 2016).

The question arises how the ostensive triangle and joint attention emerged in the first place. It has been pointed out that selection pressure has to act both on the sender and the receiver for a general benefit of social learning as compared to asocial learning (Turner et al. 2023). In that context, it is noteworthy that the achievement of joint attention itself may have evolved as a cognitive gadget (Heyes 2018), i.e., a culturally acquired ability that only later on got biologically fixed to ascertain its reliable availability. Ontogenetic development of pointing is seen as being rooted in grasping objects within the process of social shaping and has been hypothesized to develop from the simple transposing of objects towards the active referring to objects during joint activity (Liszkowski and Rüther 2024).

Along these lines are anthropological insights that during fire domestication in anthropological evolution (800–400,000 years ago), a cognitive prerequisite was not only the ability to ignite fire itself as a distinct sequence of a motor acts involving several objects, but also required prospective organizational abilities to maintain the fire (such as wood piling, fire supervision to prevent its premature extinction, etc.) (Twomey 2019). All these executions required complex social skills including joint attention and are evidence for an early evolution of social transmission of task and attention sharing itself.

An early externalized object might have been a peer or prey, i.e., another living individual, and pointing may have evolved for joint foraging or for other reproduction-related purposes. In this context, joint attention would have evolved due to a higher selection pressure on getting and staying in contact to align (“share”) intentions, for elementary needs. The transition within the ostensive triangle from a mobile living subject to a permanent, externalized object (Figure 3) may then have gradually occurred through, for example, associating the third individual in the ostensive triangle with doll-like objects as personal memory-artifacts representing real persons, and such practices may have occurred with the beginning of person-related memories when burial practices developed (Chapman 2013). Alternatively, the transition person-to-object may be related to the idea of “social tools” (Hare 2017), i.e., the reification of social peers and their assignment of roles in social interactions where under some circumstances individuals are handled as interchangeable tools.

3.3 Imitation and mimesis

There is a variety of mechanisms of social transmission and learning (see overview in Whiten 2022). A particularly important mechanism is imitation (Heyes 2016, 2023; Heyes and Frith 2014). One narrow definition of imitation states “more specifically, in imitation an observer learns and reproduces something about the topography of body movements” – which part of the body to arrange (e.g., head or foot), or how to move body parts relative to one another (e.g., to bring appendages closer to the trunk or to move them further away) (Heyes 2023). Imitation, i.e., the copying of another's body movements as precisely as possible, has invariably been considered key in social transmission. This is supported by the phenomenon of "overimitation," defined as including copying of causally irrelevant and perceivably unnecessary actions in relation to the given goal (Hoehl et al. 2019). Overimitation thus seems to have evolved to ascertain copying redundancy: motor patterns are first reproduced in overabundance without recognizing context-given termination points or borders, or any "understanding." This could mean that imitation is really a process that does not involve any (conscious) insight in line with that Dennett has called "competence without comprehension" (Dennett 2017). Other kinds of social learning include stimulus enhancement, observational conditioning, and many others (reviewed in Heyes 2018, 2023).

Mirroring movements or postures of other individuals clearly is at the center of imitation. Following that line, mirror self-recognition is considered to constitute a primordial form of self-representation (Anderson and Gallup 2015). A mirror-independent form of imitation is "movement mirroring" that is, for example, performed already in the early mother–child interaction where the mother imitates the child's spontaneous facial movements resulting in the association of self-movements and facial perceptions of others in "matching vertical associations" (Heyes 2019).

Mirror neurons, i.e., neurons originally discovered as being activated both on perceiving and executing hand mouth movements in monkeys (Rizzolatti and Craighero 2004) have been hypothesized to be a neural correlate for imitation. To this end, mirror neurons have been detected in nonhuman primates and humans (Bonini et al. 2022; Heyes and Catmur 2022). They are likely the result of extensive ontogenetic training rather than an innate endowment (Cook 2012; Heyes and Catmur 2022). A gradual evolution of imitation has been proposed, from grasp copying to complex imitation, gestural copying and protolanguage corresponding to distinct stages of early hominin toolmaking and use, and mediated by the mirror neuron system (Arbib 2011).

The report of tool-specific mirror neurons, i.e., cells that display highest activity when tool-manipulation is performed or observed suggests deep biological rooting of tool use (Ferrari et al. 2005; Rochat et al. 2010), conceptualized either as tools becoming part of the body scheme or, alternatively, hands or other body parts becoming objectified (Iriki 2006). What is certain today is that mirror neurons play a causal role in copying of body movement topography and contribute to processing of observed actions (such as grip) but they do not seem to play a role in inferring observed actors' intentions (Catmur et al. 2014; Cook et al. 2014; Heyes and Catmur 2022).

A precondition for multistep tools is the ability of memorizing a complex (motor) sequence. This requires episodic memory (Bevandic et al. 2024; Huston and Chao 2023; Tulving 1983), which is present in humans, apes, but also in many other species. The efficient process of social transmission of instructions was powered by a strong individual reward to replicate tool-related instructions. The benefit of tool use as increasing the convenience to achieve a goal was that without that tool, the benefit would have been costlier – thus, improving social transmission of tool-related instructions was rewarded and provided a strong reinforcing motivation to replicate tools (see Figure 4). Supporting the notion of tool-related rewards is the observation that nonhuman primates make profitable decisions relative to the reward quality in the context of preferred tool use choices (Laumer et al. 2019).

3.4 Universal Darwinism

On the assumption of Universal Darwinism, Dawkins coined the term "meme" to designate a replicator that, in analogy to a gene, could be the unit of cumulative cultural evolution. In the almost 50 years after its conception, however, and for reasons that become apparent below, the definition of a meme has undergone transformations that today make it too ambiguous to be used for more specific scientific definitions, such as the ones mentioned in this paper.

Dawkins originally defined memes as: "a unit of cultural transmission, or a unit of imitation" … "Examples of memes are tunes, ideas, catch-phrases, clothes fashions, ways of making pots or of building arches … Popular songs and stiletto heels are examples. Others, such as the Jewish religious laws …" (Dawkins 1976/2016). In these unfortunately fuzzy, attempted definitions, it is problematic that things of different complexity are mixed together: tunes (acoustic melodies), ideas (a thought most likely articulated in (representational) language), clothes (something material), and religious laws (written and nonwritten norms for group behavior). Dawkins attempted a more precise definition later (Dawkins 1982) mainly trying to clarify the replicator/phenotype distinction: "A meme should be regarded as a unit of information residing in a brain…" (Dawkins 1982). A comprehensive history of memes can be found in Daniel Dennett's recent book on memes (Dennett 2017). Blackmore has defined memes as "instructions for carrying out behavior, stored in brains (or other objects) and passed on by imitation" (Blackmore 1999), which introduces the term "instruction" but still limits meme transmission to imitation and locates units of selection in brains only. The idea of memes being related to instructions was also brought up by Delius "Memes are capable of instructing … behavior" (Delius 1991).

The concept of the meme as an algorithm or a software-like subroutine was extensively elaborated by Dennett (1995). In his latest publication, Dennett defined memes (Dennett 2017): "what are memes a kind of? They are a kind of way of behaving (roughly) that can be copied, brandished, ridiculed, parodied, censored, hallowed." Dennett (2017): "Memes are informational things. They are 'prescriptions' for ways of doing things that can be transmitted, stored, and mutated without being executed or expressed." Dennett tackles the problem by defining the units of memes as "the smallest elements that replicate themselves with reliability and fecundity" (Dennett 1995). The inability of previous authors to come to a crisp and clear definition of a meme ultimately limited its usefulness and left it mainly for colloquial use, for example, to denominate rapidly shared messages in the advertisement industry or in social media.

The here presented idea that tools and corresponding instructions are two different occurrences of the same underlying phenomenon has previous been termed "petrified memes" (Korth 2020) to emphasize that for tools – in contrast to the meme concept outlined above – a distinct material stage can be described (petrified = in stone). Both, the instructions transferred between social individuals on one side and the corresponding material tools on the other side were proposed to represent the same particular meme, existing in a duality, or dual states, similar to light that can be described both as a wave and as a particle. These dual states, embodied tools, and corresponding instructions are principally equal in significance and necessitate each other: the tool could not be built without the associated instructional mental representation, but the tool itself instructs the apprehending human individual about its making and use within a social context. Tool-related instructions, by general mechanisms of associative learning (Ginsburg and Jablonka 2019), can recombine with other instructions to yield novel instructions and thus fulfill the Darwinian imperative of variation (see Figure 4).

Powerful Darwinian evolutionary systems seem to have in common, that the replicator and the vehicle, i.e., the form that the corresponding gene is translated into, operate in different dimensions without or with rare functional contact. This strict division of labor between the coding replicator and its translation into the vehicle for selection has also been highlighted as a striking feature of evolution by Maynard Smith and Szathmary (Maynard Smith and Szathmary 1995, page 12f). Thus, a "clear cut" between genotype and phenotype in biology, where the genotype-encoded cells or bodies are under selection pressure rather than the physical form of the genotype, i.e., DNA or RNA themselves seem so lead to a particularly robust evolutionary system with high evolvability. This "clear cut" or separation between the encoding and the selected-for material offers several advantages for biological evolution: 1. one gene can code for several phenotypes without being selected out when one phenotype does not fit anymore (environmental flexibility), 2. redundancy in that several genes can code for a similar phenotype allowing the loss of one pathway without leading to a system-relevant loss, 3. evolvability in that a pool of genotypes can silently "evolve" (genetic drift) without being selected against but eventually becoming relevant in future changes in the environment.

The Darwinian evolutionary system for cumulative cultural evolution presented here as an alternating shuttle between the material stage of tools that are subject to selection on one side and on the other side the socially transmitted instructions for their making and use as sequences of motor acts (Figure 4) also constitutes such a "clear cut" between the encoding and selection dimensions indicating a stable evolutionary system. The analogy to biological Darwinian evolution is the genotype corresponding to the instructions/motor acts that are transmitted/replicated and the phenotype corresponding to material tools that are selected for their convenience or advantages for facilitating survival in the widest sense.

4 Shared tool-elicited affordances accelerate social synchronization

Tools elicit affordances that are related to their use. Affordances have been defined as "opportunities for action" (Gibson 1979). According to Gibson, things and features of the environment are said to afford possibilities or opportunities for action according to the subject's capabilities (Creem-Regehr et al. 2013; Veissiere et al. 2019). "An affordance is a relation between an agent's abilities and the physical states of its environment. For instance, water affords drinking, cups afford drinking-out-of, bridges afford crossing, axes cutting, handles holding, etc." (Veissiere et al. 2019). An affordance is also a decrease in the degree of freedom to act, because the affordance limits the possible actions in the context of an object (Fragaszy and Mangalam 2018). Depending on the elicited affordances, these are determined by a variety of acquired knowledge including expectations, enculturations, and patterns of social attention (Veissiere et al. 2019), and in some instances can be indirectly measured, for example, by gaze movements (Natraj et al. 2015).

The presence of a tool will thus spontaneously and automatically elicit two kinds of mental representations or affordances in the perceiving individual: 1. those regarding its use and 2. those regarding its making or aspects thereof. The associations regarding its making can be precise as in the case of simple objects, a hammer (…made out of wood and steel… , a shaft fixed on an iron top…) or a ladder (perpendicularly fixed bars…), or be more diffuse in complex objects like a computer (microprocessors made from silicon, integrated circuits, screen interfaces, etc…). The affordances elicited by tool use may be stereotypical (a hammer to beat a nail into a board, etc.), more complex, or even ambiguous. Both elicited affordances are sequential in nature: first do this, then this, etc. These affordances parallel the instructions, i.e., defined sequences of steps related to either their making or use and correspond to mental representations corresponding to a given tool. Obviously, one tool can elicit several affordances depending on the wider context (Borghi 2018), but the important point here is that at least one affordance is elicited and further affordances elicited by novel context may further narrow down the first affordance. Thus, importantly, tools, as opposed to simple unmodified objects of nature, evoke similar affordances, associations, or thoughts in different individuals thereby limiting the degrees of freedom of thoughts of these by constantly reiterating the same (or very similar) tool-related affordances.

The sharing of external objects or tools by two individuals in the ostensive triangle promotes the establishment of joint attention (Tomasello 2014, 2019) as a triadic situation in which two or more individuals attend to the same referent. Subsequently, this leads to an assimilation of affordances in these two individuals elicited by the external object, causing ultimately 1. an alignment of their brain states owing to both individuals' converging attention and affordances and 2. an increase in the fidelity of any social transmission, i.e., an increase in the redundancy of the communicated message by decreasing the degrees of freedom associated with a fixed motor sequence. The joint attention and perception of several individuals by external objects during social transmission thus acts as a synchronizer of their brain states, i.e., any prior drifting apart of those individuals' in terms of attention or perception will be constrained and even be converged onto the very same external object or tool. Synchronization is here defined as the adjusting of motor output (behavior), (body)-internal molecular or electrophysiological events of two or more individuals such that they become more similar.

The difference between a natural object such as a tree, a river, or a blackberry bush to tools is that the latter evoke stronger affordances (decreases of degrees of freedom to act). Their omnipresence in the abundant tool-sphere thus fundamentally structures our everyday life, and their simultaneous perception by individuals exposed to the same thousands of tools thus synchronizes the perception and, ultimately, brain events of many individuals – without these tools their brain states would be less aligned. Even if the tools are not actively used but only passively perceived, deeply rooted affordances act unconsciously such that their structuring effects are still be present.

It has been demonstrated that creating a shared point of reference, for example by mutual gaze, is facilitated by external objects (Koike et al. 2016). The resulting synchronization within the dyadic communication between two individuals has been conceptualized as the perceptual system of one brain to becoming coupled to the motor system of another through Hebbian association (Koike et al. 2016). In fact, measurable correlates of synchronization on the level of the electroencephalogram (EEG) during mother–child communication (Levy et al. 2017), on the level of brain region activation by functional magnetic resonance imaging (fMRI) (Stolk et al. 2014), or other methods (Stephenson et al. 2021) have been demonstrated. Brain synchronization between individuals, in turn, feeds back on individual's subjective experiences, for example pain (Goldstein et al. 2018). There are also historic examples of how tools fostered synchronization of whole societies: we are all aware that the setting of an absolute time scale through the invention of clocks and watches heavily contributed to interindividual synchronization of communication, work time partitioning, and many other individual activities (Noe 2010; Wajcman 2015).

In the evolution of externalizations from gestures to tools, it is therefore not surprising that the speaker–listener neural coupling in a dialogue (Stephens et al. 2010) leads to increased synchronization when gestures are added (Hasson et al. 2012). Ostensive cues, like, for example, gestures or pointing and contingent responsiveness accelerate social interaction since behavioral entrainment makes neural entrainment through phase-resetting of gaze and EEG more likely (Wass et al. 2020).

As externalizations become more complex leading to, ultimately, tools, the synchronization goes beyond shared attention: the joint elicitation of similar affordances even aligns conceptualizations leading to a synchronization of brain areas beyond those for simple signal perception (Stolk et al. 2014; Wheatley et al. 2024). Using a hyperscanning task, i.e., simultaneous real-time brain imaging of two interacting individuals, Stolk et al. (2016) showed that synchronization, measured by fMRI with combined coherence spectral-density analysis, increased upon multiple mutual communicative exchanges using tokens, i.e., simple tools. Social synchronization by tools feeds also back on social transmission itself in making imitation more efficient, for example, by the use of mirrors (Cook et al. 2013; Heyes 2018), and therefore contributes to the evolvability of social transmission. Through object permanence, tools thus act as scaffolds for social synchronization.

The synchronization-enhancing function of tools very much supports the idea that the relationship between social transmission and culture is bidirectional where imitation accelerates social transmission of instructions and, on the other hand, the products of those instructions, tools, align individuals' perceptive, and motoric worlds thus facilitating imitation skills themselves (Heyes 2023). Therefore, tool evolution ultimately promoted the emergence of prosociality.

5 Tools increase predictive processing

Tools, by eliciting similar affordances and thereby synchronizing individual brains within a group, make the social environment more predictable. This can be accommodated in the active inference/predictive processing framework (Clark 2016; Hohwy 2013). Active inference or predictive processing is a theory that proposes that every organism optimizes its internal model of the world via action-perception cycles with the goal of minimizing unexpected surprise, termed prediction error, and costs for building up an entirely novel reaction to the external environment (Friston and Stephan 2007). In the iterative interaction of the virtual model of the environment with the actual environment itself, both become more similar in order to fit into each other, or become predictive of one another via action-perception cycles (Constant et al. 2018; Vasil et al. 2019).

The error between the actual sensory signals and their predictions from internal estimates are minimized by either changing the internal estimates or sensory signals through an appropriate (motor) action. Thereby, the same internal representations that become active in perception are also activated to enable action, which leads to the main advantage of the active inference theory, namely that the representations underlying perceptions and motor actions are in the same data format (Clark 2023; Friston et al. 2012; Metzinger and Wiese 2017). "Active inference induces a circular causality that destroys conventional distinctions between sensory (consequence) and motor (cause) representations. This means that optimizing representations corresponds to perception or intention, i.e., forming percepts or intents" (Friston et al. 2011). Recent literature has suggested that the re-enactment or emulation of sensory patterns alone suffices for predictive processing (Clark 2023).

Tools are "active inference" (prediction)-devices themselves: by eliciting similar affordances they actively construct niches through toolmaking or use (Odling Smee et al. 2003). By their relative longevity (in evolutionary terms) and their increase of permanence in the visual modality in our everyday lives, tools also increase sensory steadiness in our immediate environment. Therefore, tools support predictive processing, particularly in the omnipresent social situations (Constant et al. 2018; Vasil et al. 2019), which likely was a strong reason why the evolution of the tool-sphere was advantageous for forming functional social networks. The executed motor acts of tool-related instructions foster prediction: as Clark puts it "Predictive brains constantly estimate the extent to which taking an action (such as using the stick to probe for depth) will reliably reduce uncertainty in ways that help us approach our goals" (Clark 2024). To resume the previous example of the use of a diary to memorize appointments: the tool (diary) usage decreases the future prediction error information (Clark 2024). Accelerated transmission of affordances by tools facilitates group coherence by providing an algorithm to align predictive processing and predictive error minimization for norm building. The selective allocation of attention has been conceptualized as a process of optimizing precision estimates (Friston and Stephan 2007; Hohwy 2012), i.e., decreasing variance of the generated models or hypotheses prior to perceptions. Translated to joint attention this means that the internal models within a social interaction are continuously improved. In summary, it is likely that the evolution of prosociality, joint attention, social predictive processing, and tools are entangled and mutually pushed themselves in steady reinforcement cycles with the goal to optimize predictive processing.

6 Intersubjective "understanding" as affordance reconstruction

Can the tool-catalyzed synchronization of affordances between individuals be meaningful for a psychological theory of understanding?

6.1 Understanding an object or a tool

To understand a tool is the ability to experience the same or very similar affordances as other individuals relating to the tool's making or use. It is the ability to deconstruct a tool into its instructions, to reverse-engineer the tool in efficient ways, i.e., to select a parsimonious best fitting deconstruction among many other possible ones. Teaching is "defined as the sender (or pointer) modifying their usual behavioral repertoire to actively facilitate social learning in others," which occurs when the benefit is larger than an action that involves only individual learning (Turner et al. 2023). The simultaneous perception of the tool by the teacher, the student (=receiver), and, eventually, a demonstration of its use will align the affordances of the teacher to those arising in the student and it is the gradual convergence of affordances related to a tool that is the equivalent of "understanding" the tool. The teacher optimizes relevance for the student regarding the tool in question such that inference of the most parsimonious tool-related instructions sufficiently reconstructing the affordances is possible (Goodman and Frank 2016). Since synchronization of affordances by a tool is also an alignment of predictive processing between teacher and student, in order to control apprehension of the receiver, the teacher checks arbitrarily the similarity of their affordances by observing cues, or evaluating answers of the receiver to the teacher's questions (verbal or nonverbal) that the receiver can only answer correctly when he/she has achieved a similar deconstruction.

This dyadic (or triadic if the tool is added) process with rapid bilateral communicative exchanges is similar to the process of the search for "hidden causes" representations within continuous prediction error minimization in the active inference paradigm (Metzinger and Wiese 2017). By recombining possible elementary instructions for best fitting a tool, i.e. the subject of understanding, teaching is thus like engineering a (virtual) tool itself. It is like building a fleeting novel cognitive gadget (Heyes 2018) of instructive steps that can be exchanged via motor representations (language) or a visualized motor execution with the tool. The importance of an organized teaching environment, i.e., a kind of permanent object or architecture, that includes tools for instructed pedagogy has been highlighted (Sterelny 2012). In this iterative process, both the teacher and the student are gradually improving to think through another mind (Veissiere et al. 2019). At the end, tool-inspired deconstruction or a reverse-engineering of affordances ultimately synchronizes teacher and receiver that ends with the teacher saying "the other individual understood." Having understood generates an experience of "familiarity" of the student relating to the novel object or tool. This process of iterative deconstruction has also been described as being "able to transform the new object into, so to speak, older objects" (Latour 1987).

6.2 Understanding an unfamiliar tool is similar to the process of cognitive archeology

The process of learning about novel tools and, ultimately, understanding them is similar to the method employed by cognitive archeology where the material remnants of ancient times are the only substrates left and the ancient world and human cognitions around it are reconstructed from these remains. The here-proposed tool-based theory of psychological understanding is similar to the method of cognitive archeology. In fact, cognitive archeology is the "study of past ways of thought as inferred from material remains" (Renfrew and Zubrow 1994), where starting from the external, material objects, cognition is inferred. Archeologists have to infer microscale processes, e.g., the making of a tool solely from the materials remnants (Overmann and Wynn 2019; Shennan 2011). For example, archeologists have investigated how handaxe knapping of different symmetries requires different levels of advanced cognitive and even neural brain functions (Hodgson 2019).

In cognitive archeology, this has led to a "the material engagement theory, a theoretical framework to set the sole substrate of archeology, material objects, as starting points for reconstructing past society's or individuals' cognitions" (Malafouris 2013; Renfrew and Zubrow 1994), where thinking with the help of tools (i.e., active externalism) or "things" has been termed "thinging" (Malafouris 2020). At its extreme, cognitive archeology even departs from a subject-centered view of cognition to adopt a completely decentralized view of cognition where materials are "cognitive scaffolds for thinking, material agency: materials have an effect on us, they make us do things" thus even reversing common agency dogmas (Malafouris 2013).

Obviously, there is a difference in time scale between psychological understanding and archeological deconstruction: whereas the process of deconstructing the very instructions that possibly led to the construction of an archeological object can take years, in interpersonal communication a deconstruction from a tool in the process of understanding eventually happens in minutes or even seconds – but both deconstructions still have a similar formal structure.

6.3 Understanding someone else

How do I understand someone else? The question is better rephrased as: How do I understand what someone else has in mind related to a specific topic? When detached from the context of tools, the explanation of such an understanding is more speculative but it follows a similar pattern: the questioner tries to achieve a deconstruction of the topic named by the person he/she wants to understand. Starting from a first guess that considers all possible contexts of the specific topic in question, an iterative process is started between questioner and respondent where an affordance alignment is the goal and its achievement is constantly confirmed or rejected by control questions (verbally or nonverbally). This process of alignment of mental states can be assisted efficiently by the presence of representations of tools (language), images, metaphors, or material evidence such as external objects to accelerate synchronization. Thus, when one's goal is to understand someone else by simple observation, searching for the most parsimonious deconstruction that explains this person's behavior or a situation is continuously ongoing and part of regulating behavior to establish common grounds (Vasil et al. 2019). It is similar to an engineering process similar to what has been described by Braitenberg as "synthetic psychology" (Braitenberg 1984): a behavior is most easily understood by reconstructing it from the most simple elements rather than by complicated speculations. In a sense, understanding is reverse-engineering of predictive processing. Affordance comparisons and mutual perspective taking are thus basic elements for inferring each other's goals (Creem-Regehr et al. 2013). Synchronization of brain activation is reported upon the mutual understanding of conceptual understanding (Stolk et al. 2014). The question "why" is a request for joint deconstruction in the act of understanding. A joint deconstruction is a first step to a theory of mind.

6.4 Competence without comprehension

Is understanding a process that depends on consciousness and insight? Not when a wide sense of the word is used and motor learning as well as the ability to replicate instructions are also comprised in the term "understanding," - not the deliberate exchange of representations by language alone. Understanding without rational insight has also been called "competence without comprehension" (Dennett 2017). An exchange of (tool-related) instructions can best explain understanding when defined as an interindividual alignment of affordances as proposed in this paper. It was recently shown in a reductionist experiment that technological advances can be made without any causal understanding of the phenomenon (Derex et al. 2019) referring to broader research that many cultural optimizations can evolve without any individual insight such as, for example, cultural instructions transmitted across generations in prescientific societies (Henrich 2016, 2021).

7 A trajectory from social super-synchronization to a third person perspective and consciousness

As developed in the previous sections, social synchronization can be conceived as an evolutionary late stage beginning with the ostensive triangle, and having evolved gradually (Figure 3). Social synchronization leads to the theory of mind, i.e. to the ability to attribute mental states to oneself and to others (Premack and Woodruff 1978). Theory of mind is not unique to humans but, depending on the narrowness of its definition, occurs also in chimpanzees, even though the latter are not able to solve false belief tasks, i.e. tasks that require temporal and spatial perspective taking. This is proof of the gradual nature of the theory of mind, and other higher cognitive abilities (Call and Tomasello 2008). The creation of a common affordance space and a social niche of affordances for "thinking through other minds" (Veissiere et al. 2019) is an advanced stage of social synchronization, a super-synchronization. The gradual nature and interindividual variation of a theory of mind becomes apparent in disorders like autism where impairment in neurodevelopment (i.e., hardware problem) leads to deficits in social cognition (Happe and Frith 2020) and to the inability to align affordances (Wadge et al. 2019). Individual, habitual tool use leads to "self-objectification" (Iriki 2006) through the construction of an individual tool-sphere that when merged with the individual tool-spheres of others fosters the construction of a third person perspective. Super-synchronization of many individual brains is, therefore, further accelerated by the massive increase of the tool-sphere in late cumulative cultural evolution (Elhacham et al. 2020). Could late-emergent properties of the brain such as consciousness also be fostered by tool evolution and social super-synchronization, hence be ultimately acquired both phylogenetically (during anthropological history) and ontogenetically (during individual development; Figure 6)?

There is no clear and undisputed definition of consciousness, which makes a serious scientific approach very difficult but it is undoubted that consciousness is also a form of higher cognition. "The problem of consciousness … is to understand what processes or mechanisms or events in the brain make certain contents phenomenally conscious" (O'Regan and Noe 2001). Clearly, state consciousness, i.e., the degree of awareness, and content consciousness, i.e., what a

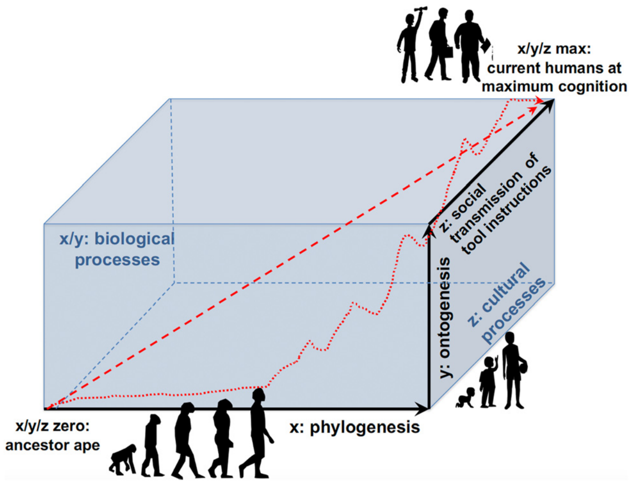

Figure 6: Higher cognition as the result of a coevolution of biological and cultural processes. The evolution of higher cognition is necessarily dependent on biological including genetic evolution (*x*-axis), an optimal ontogenetic development of a human individual (*y*-axis) in conjunction with the mandatory uploading of tool-accelerated content by social transmission (*z*-axis). This trajectory does not follow a straight line (broken red line) but has an irregular course halting, leaping, and changing direction depending on various genetic and tool-related modifiers of cognitive development.

phenomenal experience is about, can be distinguished. Consciousness will here be understood to mean content consciousness. The proposition made here is neither to explain state consciousness nor to explain on a detailed mechanistic level how phenomenal experiences are realized, but to draft an anthropological trajectory claiming that content consciousness could not have evolved without prior extensive cumulative cultural evolution executed by social transmission, tool evolution, and social synchronization (Figure 6). Higher cognition can only be achieved though the extended mind sparking tool evolution as the prerequisite of consciousness which then, by consequence, has to be considered as acquired or uploaded.

One of the major components of a developing content consciousness is the emergence of a third person perspective which can only be defined in conjunction with a co-emerging first person perspective. In fact, a third person perspective can be seen as an iconized, anonymous sum of a large number of second person perspectives. The gradual increase in perspective differentiation starts from the ostensive triangle (with deixis, i.e., the perspective taking of participating individuals) and resembles a progressive alignment or synchronization of individual affordances or the ability for understanding (Creem-Regehr et al. 2013) (Figure 3). In line with the above, consciousness has also been conceptualized to facilitate predictive processing (Hohwy 2012), which, as was presented here (chapter 5), is facilitated by tool evolution.

It is important to understand that, while consciousness in a mature adult individual mind does not require the immediate presence of other individuals, clearly, prior to the maturation of full consciousness, there is mandatory massive social transmission or uploading of tool-related instructions during child development necessary to arrive at a final cognitive state that comprises perspective taking, the notion of a "self," and, ultimately, consciousness (Cleeremans et al. 2020). Once uploaded, of course, consciousness is "ratcheted in" and can run off in an asocial mode, meaning no other direct presence of other individuals is required. Thus, even the notion of "pure consciousness" as achieved during mediation (Metzinger 2020) can neither be experienced nor conceptualized without taking into account the prior uploading process. "Cultural evolution shapes not just what we think but how we think it" (Heyes 2018). Consciousness is ultimately a social and acquired ability biologically anchored by dual inheritance mechanisms (see Section 8).

What are the consequences of the insight that consciousness, intuitively conceived as something very "private," is, ultimately, socially acquired? An immediate notion is that (content) consciousness was not ignited all the sudden but evolved gradually in parallel with tool evolution and social synchronization. Dissecting these gradual stages of consciousness in the future will be important. Another consequence is that consciousness during humankind's history in individuals 500, 2,000, 10,000 or 200,000 years ago was likely very different when compared to individuals today since prosociality, tool evolution, and social synchronization, as reviewed above, were not as advanced as today. Cognitive archeology, genetics, and other scientific methods will be needed to scientifically investigate this aspect of anthropological evolution.

It is conceivable that there are also large differences in interindividual content consciousness within present day individuals across different cultures. The cognitive bias of most academic consciousness researchers as being WEIRD (i.e. from a Western, educated, industrialized, rich, democratic background) (Henrich et al. 2010) has to be critically appreciated because among themselves, the academic community is likely more synchronized and thus experiences likely more similar content consciousness, as compared to the average nonacademic, non-WEIRD individuals, let alone illiterates. This is paralleled by a differential use and access to tools. Even the content consciousness experienced subjectively throughout one's individual's biography is likely to change during different life stages but the self-illusion (Blackmore 2016; Dennett 2016; Frankish 2016; Metzinger 2004) and our continuously recreated autobiographical memory spinning around our externalized self may prevent us from seeing this.

Today we live in a tool-overabundant environment where the mass of the tool-sphere already exceeds that of the biosphere (Elhacham et al. 2020). This is a very recent development in human history and has massive streamlining effects on increasing jointly shared affordances, social synchronization, conformity, and norm implementations, amounting to "super-synchronization." The consequence of an extremely prosocial environment led, for example, also to a decrease in major crime incidence in the last centuries (Pinker 2011). A tool/machine-free environment is almost impossible to imagine today. Even in plain nature we see paths, wear clothes, and are reminded of other human remains. We cannot "forget" language, i.e., "un-learn" it or not have phenomenal experiences because once ignited during ontogenetic development by social transmission, this uploaded linguistic content or content consciousness becomes biologically fixed (see Section 8).

The general idea of anthropologically acquired consciousness has also been proposed by Jaynes (1976), but his idea lost credibility through its exclusive linking of the idea with language evolution, the bicameral mind, the audacity of making precise predictions on its appearance to the second millenium B. C., and his opposition to Darwin's idea of evolution (Jaynes 1976). Nevertheless, it inspired Dennett to the term "software archeology" (Dennett 1986), i.e., the reconstruction of a cognitive evolution of mankind not unlike to what is called cognitive archeology today (Malafouris 2013; Renfrew and Zubrow 1994). Coming full circle, tool evolution within cumulative cultural evolution may after all provide a key to unravel the emergence of content consciousness.

8 Tool-conditioned biological and cultural coevolution

Do cumulative cultural evolution and biological evolution interact, i.e., coevolve, and if so, how? If biology is the hardware, i.e., consisting of a rather permanent, "automatic" and hard-wired response to an environmental selection pressure, then cultural adaptations executed by behavior responding to a given selection pressure are the software, i.e., consisting of more deliberate and flexible, but eventually slower responses. Of course, behavior is ultimately hard-wired, too, but, according to different species and behaviors, is more malleable and has more degrees of freedom. Evolutionary selection of human individuals has acted on both hardware and software, and selected whatever was more advantageous for ever changing environmental challenges. Sometimes, repetitive, stereotypical motoric behavioral patterns executed by individuals required for cumulative cultural evolution may have become fixed, meaning hard-wired, to ensure their reliability and fast response at the expense of flexibility. On the other hand, hard-wired, energy-expensive behavioral responses may have been built back during biological evolution after flexible behavioral responses fitted better to selection pressures. Scientific evidence for both scenarios is given below.

Both those simultaneously operating behavioral programs, the use of preset motor patterns as well as acquired behaviors for achieving best behavioral responses are realized simultaneously in individual brains. The brain is utilitarian: it selects those mechanisms whether inborn or flexible that serve best for adaptation to a varying environment. The individual behavioral repertoire at any moment is rooted in a mosaic of hardware and software-based processing solution with all possible graduations – all simultaneously present but each with different phylogenetic and ontogenetic histories. The major assertion of this paper is that both during phylogenesis and ontogenesis behaviors are acquired through interaction with tool evolution. During anthropological evolution, cognitive functions of individual brains were acquired by a coevolving tool evolution in a cumulative, transgenerational manner. During individual development in childhood, cognitive functions are acquired through a social environment that is structured by tool evolution within cumulative cultural evolution. The interaction is truly bi-directional, as will be shown in examples below, such that the presence and availability of tool making and using skills relieved pressure from maintaining biological hardware. The liberated energy expenditure of which could be diverted to other biological systems.

The **dual inheritance** theory provides abundant scientific evidence and a clear framework to explain mutual influencing within cultural and biological coevolution (Boyd and Richerson 1985; Cavalli-Sforza and Feldman 1981; Henrich 2016; Uchiyama et al. 2021). Dual inheritance posits that genes and culture continually feed back onto each other and that genetic changes can lead to changes in culture, which in turn modify genetic selection, and vice versa (Laland 2008).

Salient questions for the entanglement of cumulative cultural evolution and biological coevolution are as follows:

(1) Are there "evolvability hubs" connecting biology and cumulative cultural evolution for mutually accelerating each other's evolutionary speed?
(2) What genes or biological hardware favor social transmission and the establishment of the ostensive triangle, prerequisites of tool use, tool building, or cumulative cultural evolution?
(3) What biological hardware were needed or liberated by the cumulative success of behavioral/toolmaking solutions for arising selection pressures?

Here, only some of the most prominent features will be touched: evolvability mechanisms, prosociality, the role of assortative mating in accelerating social transmission, cognitive abilities enabling niche construction, the neural reuse model during cognitive evolution, and the Baldwin effect for biological fixation of successful behavioral circuitry.

8.1 Evolvability

Evolvability, in the original biological definition, is an organism's capacity to generate heritable phenotypic variation (Kirschner and Gerhart 1998). Translated into cumulative cultural evolution it means an individual's capacity to generate variations in tools that can be socially transmitted. In a narrower sense, evolvability may also be defined as a mechanism that increases heritable phenotypic variation, i.e., that speeds up evolutionary solutions to faster achieve adaptation. Based on the concept of evolvability, eight major transitions in Darwinian evolution were claimed from the chemical evolution of replicating molecules to the evolution of societies (Maynard Smith and Szathmary 1995). This concept was subsequently extended to the stepwise transitions in cognitive domains (Barron et al. 2023; Dennett 2017; Ginsburg and Jablonka 2019; Tomasello 2014). In each of these steps, novel mechanisms of selection were established opening novel phenotypic spaces.

8.2 Prosociality and consequences

Clearly, genes steering brain development must first have developed primordial brain hardware to enable the ostensive triangle and primordial social transmission of sequences of motor acts (instructions), but once ignited, abilities for social transmission were rapidly accelerated by evolvability. Social tolerance as early prosociality is a necessity for stabilizing populations of large overall size and high connectedness, which in turn are preconditions for generating, retaining, and diffusing cultural innovations (Cieri et al. 2014; Muthukrishna et al. 2014). Increased social tolerance promotes cultural transmission itself through the evolution of teaching and conformity (Tennie et al. 2009) as well as through increasing memory (Urban Levy et al. 2023).

Anthropological findings support the view that humans self-domesticated, i.e., selected themselves for prosociality (Cieri et al. 2014), as have done other species, e.g., possibly bonobos in contrast to chimpanzees (Hare et al. 2012). Biological features of the "domestication syndrome" such as craniofacial feminization and increased serotonergic and dopaminergic innervation of the amygdala coincide in the humanoid fossil record with the transition to settlements and increase in parallel to population density (Hare 2017). In a famous experiment recapitulating the domestication history of dogs, wild foxes were domesticated in only 45 years demonstrating the degree by which selective mating can accelerate the emergence of prosociality (Hare et al. 2005). Domesticated foxes shared many of the morphological and behavioral features (feminized skull, sensitivity to human social cues such as gestures, prolonged juvenile development, etc.) of dogs versus their wolf ancestors (Hare et al. 2005). It has been suggested that (self-) domestication genes are primarily found in neural crest development (Sanchez-Villagra et al. 2016), but others have argued that shared reproductive disruption would be sufficient to explain specific differentially expressed genes in domesticated animals (Gleeson and Wilson 2023).

Increasing prosociality led to the formation of very large groups and sedentariness, and together with the coemergence of domesticated animals, also guided coevolution of zoonotic pathogens and adaptive immune responses to these. In history, social groups with efficient immune responses to zoonotic pathogens came to dominate other groups that had not developed such an immunity as, for example, happened in the selection of prosocial life forms over hunter/gatherers (Diamond 1997; Kennedy 2024), thus further selecting and thereby accelerating prosociality as a human individual trait.

8.3 Assortative mating

The progression in cognitive and toolusing abilities in individuals likely increased the preference to company with individuals of similar cognitive abilities, leading to assortative mating, i.e., the nonrandom preference for mating partners. The increased synchronization of mates with improved communicative and social abilities consequently greatly accelerated fixation of genes favoring social transmission (Creanza et al. 2017; Zeng and Henrich 2022). As a consequence of assortative mating, social transmission-relevant but less frequent genes would have had a higher and faster chance of becoming penetrant and thereby guide biological evolution through the fixation of these traits. Socially biased learning and communication has been shown to select for intelligence (van Schaik and Pradhan 2003) and would thus accelerate assortative mating in a loop-like, self-reinforcing manner.

8.4 Baldwin effect

Acquired behaviors can be genetically fixed between generations by the "Baldwin effect" (Baldwin 1896). The Baldwin effect is a phenotypic change occurring in an individual as a successful adaptation to its environment that becomes gradually fixed in its developmental genetic or epigenetic repertoire (Baldwin 1896; Dennett 2003; Maynard Smith 1987). This means that a continued advantageous behavior of the same kind for which flexibility was not an advantage anymore, put biological selection pressure on genes that fixed this behavior to overcome its inherent instability and delayed reaction for the sake of a rapid and stable response. Such genetic fixations are gradual and concern domain-general abilities such as, for example, specific processing speed, episodic memory, social recognition, social transmission, and others, rather than whole complex behaviors. Alternatively, under a similar selection pressure, a rare random mutation is also conceivable to rather suddenly enable a useful adaptation.

8.5 Niche construction

Yet another mechanism how tool use shaped biological evolution is niche construction (Clark 2011; Laland et al. 2000; Odling Smee et al. 2003). Niche construction refers to "activities, choices, and metabolic processes of organisms, through which they define, choose modify, and partly create their own niches" and thereby actively changing selection pressure on the niche creator (Laland et al. 2000). Human cultural activities, including those involving tools, have fed back on genetic evolution by modifying natural selection and thereby altering genotype frequencies (Laland et al. 2000; Odling Smee et al. 2003). This has been abundantly reviewed and includes, for example, the persistence of lactose genes and light skin color to counter vitamin D deficiency when northern hemisphere habitats were colonized after the control of fire and the domestication of animals were achieved (Henrich 2016). In the process of niche construction, a cycle of mutual fitting between individual and niche is started that gets optimized and that includes both the natural and the social environment (Veissiere et al. 2019). Another example for a cumulative cultural solution replacing a hard-wired biological solution is the relative atrophy in digestive organs in response to the acquisition of controlling fire and cooking: the relative ease to obtain essential nutrients by the achievement of cooking led to smaller jaws and intestines and the saved energy needs of these organs could be expended on a growing brain (Wrangham 2017).

8.6 Neural reuse model

The neural reuse model assumes that existing brain structures are used for novel functional behaviors (Anderson 2014, 2016) and could explain how massive behavioral – ultimately sequential motor acts – could be stored and processed with increasing social transmission of language and tool instructions in a relatively short amount of time. According to the neural reuse model, "language works because it has developed to take advantage of and is fitted to our interactive sociality (and not because we evolved specialized, dedicated, modular neural machinery to support it)" (Anderson 2016). According to this concept, learning-initiated brain plasticity scaffolded by cultural practices drives the acquisition cognitive skills such as arithmetics, language, or writing (Fabry 2019).

The brain is extremely plastic, i.e., neural processing can be taken over by other regions. For example, it is a frequent clinical observation that when lesions occur at very young age, other parts of the brain can take over with little functional disability (de Schotten and Forkel 2022). In a landmark experiment, young ferrets were operated and rewired postnatally such that they could "see" with the auditory cortex (von Melchner et al. 2000). These observations clearly indicate that there are no inborn, i.e. predetermined specialized areas *per se* in the brain responsible for one function but rather that specialized areas are assigned during development and interaction with the environment. Along these lines, it has been shown that learning to read and write changes the structure of the embodied brain (Dehaene 2009). Significant rewiring of the brain upon externalizing large parts of computation through the extended mind is also believed to explain the factual shrinking of brain mass in *Homo sapiens* since the last ice age (10,000 years ago) by around 12 % (Tattersall 2024).

9 Toward a neuroscience of externalization and conclusion

The present paper started with the simple concept of the extended mind as the angle to bridge biological and tool/cumulative cultural evolution. This concept allows to conceive gradual increases in the complexities of externalizations, and ultimately of tools or machines. It is claimed that higher cognition in humans is the result of a necessary coevolution of biology and culture in the human brain as outlined, meaning that higher cognition as we see it today could not have happened without the routine social transmission of instructions regarding specific tool-making and use.

The fundamentally material nature of the present concept allows broad experimental approaches in humans and animals. Experimental approaches and evidence to many scientific claims made in this paper were cited and provide already starting points for further research. Other propositions and hypotheses made in this paper can be experimentally tested and subjected to the scientific process of confirmation and refutation.

It may be worthwhile to focus on a "neuroscience of externalization" as the simplest formulation of the present concepts. Research in this area has already started in cognitive neuroscience and psychology. Investigations have so far mostly focused on internal versus external memory management in humans. A "memory symbiosis-network" was proposed as an adapted rapid alternating use of both internal and external memory (Finley and Naaz 2023). In that context, it is remarkable that language is entirely stored in internal memory while tool use and building is much more externalized but (in many times requiring reading, i.e., written, externalized language). Memory modulation by tool use in humans has been extensively investigated in psychology (Pizzonia and Suhr 2022), as have been the developmental origins of cognitive offloading (i.e., externalizations) (Armitage et al. 2020) and mechanisms of social transmission in humans (Berger 2025). First psychological tests have been conceived able to quantify the balancing of individuals between internal and external memory sources to solve tasks (Bocanegra et al. 2019). Naturalistic tool-use in animal models can be correlated to novel brain circuitry (Pendergraft et al. 2023).

To this end, what is completely lacking is wet-lab neuroscience research elucidating the neural mechanisms of externalizations. The obvious difficulties are that advanced tool use and building as well as cumulative cultural evolution is exclusive to humans and monkey studies of tool use (Iriki and Sakura 2008) or theory of mind (Gallese et al. 1996) will give valuable hints but not the full picture of how biology is entangled with cumulative cultural evolution. While functional magnetic resonance imaging (fMRI) is widely used to investigate brain function, including in the context of tool use or cumulative cultural evolution (Corradi-Dell'acqua et al. 2008; Derex et al. 2019; Koike et al. 2016) and is a promising way to identify a functional neuroanatomy, is not molecular and, therefore, will not identify distinct molecular and evolutionary mechanisms. The exploration of external memory mechanisms is likely to be a fertile research ground since it can be scaled down to standard laboratory animals and this field is still completely uncharted territory.

How are externalizations of the simplest kind realized on the neuronal or cellular level? Certainly, the identification of mirror neurons (Rizzolatti and Craighero 2004) by meticulous and tedious wet-lab research promised a breakthrough in identifying correlates of the theory of mind, of imitation and of a kind of social transmission mechanism long sought after. At this point, not all hopes regarding mirror neurons have been fulfilled, but this area will provide more research opportunities (Heyes and Catmur 2022), particularly with the discovery of mutations impairing mirror neuron function (Yoshida et al. 2016), and the discovery or mirror neuron-related cells in rats (Viaro et al. 2021). Studies investigating cumulative cultural evolution in animals have recently been reviewed and will inspire novel experimental approaches (Arbib et al. 2023). Mechanisms of social transmission in animals (Harrison et al. 2024) and extended animal cognition (Facchin and Leonetti 2024) are further topics of investigations. Domestication has been a fertile ground for investigating the developmental and genetic origins of prosociality (Gleeson and Wilson 2023; Sanchez-Villagra et al. 2016).

Human diseases or possible animal models of these may also help to decipher mechanisms of externalizations. Even though no single disease impacts all aspects of toolmaking or use and the social transmission of instructions (Figure 4), distinct brain impairments, such as, for example, the apraxias, can be used to investigate procedural accuracy needed for tool handling (Baumard et al. 2016; Etcharry-Bouyx et al. 2017; Goldenberg and Randerath 2015; Osiurak and Badets 2016). Psychosis or delusions such as occurring during schizophrenia positive symptoms can be conceptualized as disorders of a third person perspective (Gambini et al. 2004; Hipolito 2016), i.e., the inability to establish a common frame of reference, while certainly many basic toolmaking and using abilities are conserved. The comparison of brain disorders helps to elucidate molecular and possible anatomical correlates of externalization interactions.

Large scale genetics can also be used to monitor the increase and decrease of selection pressure that acted on various genes during anthropological evolution, also termed "molecular archeology" (Kaczanowska et al. 2022; Paabo 2014). The dating of periods with increased or relaxed selection pressure on genes that can be correlated with cognitive traits allows the identification of periods of biological fixation or relaxation, respectively. These, in turn, may reveal important insights into biological circuitry critical for abilities required for externalizations or tool evolution.

In conclusion, what we call mind is the result of convergent and entangled evolution in biology and cumulative culture where an essential component is the acquired cognition through tool making and use. Conceptualizing the mind or higher cognition as such will be an advantage when investigating its origins.

Acknowledgments: I thank Thomas Metzinger for discussions on an early version of this manuscript.
Research ethics: Not applicable.
Informed consent: Not applicable.
Author contributions: The author has accepted responsibility for the entire content of this manuscript and approved its submission.
Use of Large Language Models, AI and Machine Learning Tools: None declared.
Conflict of interest: The author states no conflict of interest.
Research funding: None declared.
Data availability: Not applicable.

References

Anderson, J.R. and Gallup, G.G. (2015). Mirror self-recognition: a review and critique of attempts to promote and engineer self-recognition in primates. *Primates* 56: 317–326.

Anderson, M.L. (2014). *After phrenology: neural reuse and the interactive brain*. MIT Press, Cambridge, MA.

Anderson, M.L. (2016). Precis of after phrenology: neural reuse and the interactive brain. *Behav. Brain Sci.* 39: e120.

Arbib, M.A. (2011). From mirror neurons to complex imitation in the evolution of language and tool use. *Annu. Rev. Anthropol.* 40: 257–273.

Arbib, M.A., Fragaszy, D.M., Healy, S.D., and Stout, D. (2023). Tooling and construction: from nut-cracking and stone-tool making to bird nests and language. *Curr. Res. Behav. Sci.* 5, https://doi.org/10.1016/j.crbeha.2023.100121.

Armitage, K.L., Bulley, A., and Redshaw, J. (2020). Developmental origins of cognitive offloading. *Proc. Biol. Sci.* 287: 20192927.

Baldwin, J.M. (1896). A new factor in evolution. *Am. Nat.* 30: 441–451.

Barron, A.B., Halina, M., and Klein, C. (2023). Transitions in cognitive evolution. *Proc. Biol. Sci.* 290: 20230671.

Baumard, J., Lesourd, M., Jarry, C., Merck, C., Etcharry-Bouyx, F., Chauvire, V., Belliard, S., Moreaud, O., Croisile, B., Osiurak, F., et al. (2016). Tool use disorders in neurodegenerative diseases: roles of semantic memory and technical reasoning. *Cortex* 82: 119–132.

Berger, J. (2025). What gets shared, and why? Interpersonal communication and word of mouth. *Annu. Rev. Psychol.*, 76: 5.1–5.23.

Bevandic, J., Chareyron, L.J., Bachevalier, J., Cacucci, F., Genzel, L., Nwcombe, N.S., Vargha-Khadem, F., and Olafsdottir, H.F. (2024). Episodic memory development: bridging animal and human research. *Neuron* 112: 1060–1080.

Birch, J., Ginsburg, S., and Jablonka, E. (2020). Unlimited Associative Learning and the origins of consciousness: a primer and some predictions. *Biol. Philos.* 35, https://doi.org/10.1007/s10539-020-09772-0.

Blackmore, S. (1999). *The meme machine*. Oxford University Press, Oxford.

Blackmore, S. (2016). Delusions of consciousness. *J. Conscious. Stud.* 23: 52–64.

Bocanegra, B.R., Poletiek, F.H., Ftitache, B., and Clark, A. (2019). Intelligent problem-solvers externalize cognitive operations. *Nat. Human Behav.* 3: 136–142.

Bonini, L., Rotunno, C., Arcuri, E., and Gallese, V. (2022). Mirror neurons 30 years later: implications and applications. *Trends Cognit. Sci.* 26: 767–781.

Borghi, A.M. (2018). Affordances, context and sociality. *Synthese*, 199: 12485–12515.

Botvinick, M. and Cohen, J. (1998). Rubber hands 'feel' touch that eyes see. *Nature* 391: 756.

Boyd, R. and Richerson, P.J. (1985). *Culture and the evolutionary process*. University of Chicago Press, Chicago.

Braitenberg, V. (1984). Vehicles. In: *Experiments in synthetic psychology*. MIT Press, Cambridge, MA.

Bruner, E. and Iriki, A. (2016). Extending mind, visuospatial integration, and the evolution of the parietal lobes in the human genus. *Quat. Int.* 405: 98–110.

Call, J. and Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends Cognit. Sci.* 12: 187–192.

Catmur, C., Press, C., Cook, R., Bird, G., and Heyes, C. (2014). Authors' response: mirror neurons: tests and testability. *Behav. Brain Sci.* 37: 221–241.

Cavalli-Sforza, L.L. and Feldman, M.W. (1981). *Cultural transmission and evolution: a quantitative approach*. Princeton University Press, Princeton, USA.

Chan, B.L., Witt, R., Charrow, A.P., Magee, A., Howard, R., Pasquina, P.F., Heilman, K.M., and Tsao, J.W. (2007). Mirror therapy for phantom limb pain. *N. Engl. J. Med*. 357: 2206–2207.

Chapman, R. (2013). Death, burial, and social representation. In: Tarlow, S., and Nilsson Stutz, L. (Eds.), *The Oxford handbook of the archeology of death & burial*. Oxford University Press, Oxford, U.K.

Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press, Cambridge, MA.

Cieri, R.L., Churchill, S.E., Franciscus, R.G., Tan, J.Z., and Hare, B. (2014). Craniofacial feminization, social tolerance, and the origins of behavioral modernity. *Curr. Anthropol.* 55: 419–443.

Clark, A. (2011). *Supersizing the mind. Embodiment, action and cognitive extension*. Oxford University Press, Oxford, UK.

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

Clark, A. (2023). *The experience machine: how our minds predict and shape reality*. Penguin Random House, U.K.

Clark, A. (2024). Extending the predictive mind. *Australas. J. Philos.* 102: 119–130.

Clark, A. and Chalmers, D. (1998). The extended mind (Active externalism). *Analysis* 58: 7–19.

Cleeremans, A., Achoui, D., Beauny, A., Keuninckx, L., Martin, J.R., Munoz-Moldes, S., Vuillaume, L., and de Heering, A. (2020). Learning to be conscious. *Trends Cognit. Sci.* 24: 112–123.

Conard, N.J. (2003). Palaeolithic ivory sculptures from southwestern Germany and the origins of figurative art. *Nature* 426: 830–832.

Constant, A., Ramstead, M.J.D., Veissiere, S.P.L., Campbell, J.O., and Friston, K.J. (2018). A variational approach to niche construction. *J. R. Soc. Interface* 15, https://doi.org/10.1098/rsif.2017.0685.

Cook, R. (2012). The ontogenetic origins of mirror neurons: evidence from 'tool-use' and 'audiovisual' mirror neurons. *Biol. Lett.* 8: 856–859.

Cook, R., Johnston, A., and Heyes, C. (2013). Facial self-imitation: objective measurement reveals No improvement without visual feedback. *Psychol. Sci.* 24: 93–98.

Cook, R., Bird, G., Catmur, C., Press, C., and Heyes, C. (2014). Mirror neurons: from origin to function. *Behav. Brain Sci.* 37: 177–192.

Corradi-Dell'acqua, C., Ueno, K., Ogawa, A., Cheng, K., Rumiati, R.I., and Iriki, A. (2008). Effects of shifting perspective of the self: an fMRI study. *Neuroimage* 40: 1902–1911.

Creanza, N., Kolodny, O., and Feldman, M.W. (2017). Cultural evolutionary theory: how culture evolves and why it matters. *Proc. Natl. Acad. Sci. U. S. A.* 114: 7782–7789.

Creem-Regehr, S.H., Gagnon, K.T., Geuss, M.N., and Stefanucci, J.K. (2013). Relating spatial perspective taking to the perception of other's affordances: providing a foundation for predicting the future behavior of others. *Front. Hum. Neurosci.* 7: 596.

Darwin, C. (1871). *The descent of man and selection in relation to sex*. John Murray, London.

Dawkins, R. (1976/2016). *The selfish gene*, 40th anniversary edition. Oxford University Press, Oxford.

Dawkins, R. (1982). *The extended phenotype* Oxford University Press, Oxford.

Dawkins, R. (1983). In: Bendall, D.S. (Ed.), *Universal darwinism in evolution from molecules to men*. Cambridge University Press, Cambridge, UK, pp. 403–425.

de Schotten, M.T. and Forkel, S.J. (2022). The emergent properties of the connected brain. *Science* 378, https://doi.org/10.1126/science.abq2591.

Dehaene, S. (2009). *Reading in the brain: the new science of how we read*. Viking Books, New York.

Delius, J. (1991). The nature of culture. In: Dawkins, M.S., Halliday, T.R., and Dawkins, R. (Eds.), *The tinbergen legacy*. Chapman & Hall, London, UK, pp. 75–99.

Dennett, D. (1986). Julian Jaynes's software archeology. *Can. Psychol.* 27: 149–154.

Dennett, D. (1995). *Darwin's dangerous idea*. Touchstone, New York.

Dennett, D. (2003). The Baldwin effect: a crane, not a skyhook. In: Weber, B.H., and Depew, D.J. (Eds.), *Evolution and learning: the Baldwin effect reconsidered*. MIT Press, Cambridge, MA, pp. 69–106.

Dennett, D.C. (2016). Illusionism as the obvious default theory of consciousness. *J. Conscious. Stud.* 23: 65–72.

Dennett, D. (2017). *From bacteria to bach and back*. Norton, New York.

Derex, M., Bonnefon, J.F., Boyd, R., and Mesoudi, A. (2019). Causal understanding is not necessary for the improvement of culturally evolving technology. *Nat. Hum. Behav.* 5: 446–452.

Diamond, J. (1997). *Guns, germs and steel*. Norton, USA.

Donald, M. (1991). *Origins of the modern mind. Three stages in the evolution of culture and cognition*. Harvard University Press, Cambridge, MA.

Dunbar, R. (1996). *Grooming, gossip and the evolution of language*. Faber & Faber, London.

Elhacham, E., Ben-Uri, L., Grozovski, J., Bar-On, Y.M., and Milo, R. (2020). Global human-made mass exceeds all living biomass. *Nature* 588, https://doi.org/10.1038/s41586-020-3010-5.

Etcharry-Bouyx, F., Le Gall, D., Jarry, C., and Osiurak, F. (2017). Gestural apraxia. *Rev. Neurol.* 173: 430–439.

Everett, D. (2017). How language began. In: *The history of humanity's greatest invention*. Profile Books Ltd, London, UK.

Fabry, R.E. (2019). The cerebral, extra-cerebral bodily, and socio-cultural dimensions of enculturated arithmetical cognition. *Synthese* 197: 3685–3720.

Facchin, M. and Leonetti, G. (2024). Extended animal cognition. *Synthese* 203, https://doi.org/10.1007/s11229-024-04579-y.

Ferrari, P.F., Rozzi, S., and Fogassi, L. (2005). Mirror neurons responding to observation of actions made with tools in monkey ventral premotor cortex. *J. Cognit. Neurosci.* 17: 212–226.

Finley, J.R. and Naaz, F. (2023). Strategic use of internal and external memory in everyday life: episodic, semantic, procedural, and prospective purposes. *Memory* 31: 108–126.

Fitch, W.T. (2010). *The evolution of language*. Cambridge University Press, Cambridge, UK.

Fragaszy, D.M. and Mangalam, M. (2018). Tooling. In: Naguib, M., Barrett, L., Healy, S.D., Podos, J., Simmons, L.W., and Zuk, M. (Eds.), *Advances in the study of behavior*. Academic Press, Cambridge, MA, pp. 177–241.

Frankish, K. (2016). Illusionism as a theory of consciousness. *J. Conscious. Stud.* 23: 11–39.

Friston, K.J. and Stephan, K.E. (2007). Free-energy and the brain. *Synthese* 159: 417–458.

Friston, K., Mattout, J., and Kilner, J. (2011). Action understanding and active inference. *Biol. Cybern.* 104: 137–160.

Friston, K., Samothrakis, S., and Montague, R. (2012). Active inference and agency: optimal control without cost functions. *Biol. Cybern.* 106: 523–541.

Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain* 119: 593–609.

Gambini, O., Barbieri, V., and Scarone, S. (2004). Theory of mind in schizophrenia: first person vs third person perspective. *Conscious. Cognit.* 13: 39–46.

Gärdenfors, P. and Hochberg, A. (2024). The evolution of intentional teaching. In: Gontier, N., Lock, A., and Sinha, C. (Eds.), *The Oxford handbook of human symbolic evolution*. Oxford University Press, Oxford, U.K.

Gerken, M. (2014). Outsourced cognition. *Phil. Issues* 24: 127–158.

Gibson, J.J. (1979). The theory of affordances. In: *The ecological approach to visual perception*. Houghton Mifflin, Boston.

Ginsburg, S. and Jablonka, E. (2019). The transition to unlimited associative learning: how the dice became loaded. In: Evolution of the *s*ensitive *s*oul: *l*earning and the *o*rigins of *c*onsciousness: 347–403.

Gleeson, B.T. and Wilson, L.A.B. (2023). Shared reproductive disruption, not neural crest or tameness, explains the domestication syndrome. *Proc. Biol. Sci.* 290: 20222464.

Goldenberg, G. and Randerath, J. (2015). Shared neural substrates of apraxia and aphasia. *Neuropsychologia* 75: 40–49.

Goldstein, P., Weissman-Fogel, I., Dumas, G., and Shamay-Tsoory, S.G. (2018). Brain-to-brain coupling during handholding is associated with pain reduction. *Proc. Natl. Acad. Sci. U. S. A*. 115: E2528–E37.

Goodman, N.D. and Frank, M.C. (2016). Pragmatic Language interpretation as probabilistic inference. *Trends Cognit. Sci.* 20: 818–829.

Gunasekaram, C., Battiston, F., Sadekar, O., Padilla-Iglesias, C., van Noordwijk, M.A., Furrer, R., Manica, A., Bertranpetit, J., Whiten, A., van Schaik, C.P., et al. (2024). Population connectivity shapes the distribution and complexity of chimpanzee cumulative culture. *Science* 386: 920–925.

Hannay, P., McLeish, T., and Schofield, J. (2024). On the origins of dance. *World Archaeol.*, 1–10. https://doi.org/10.1080/00438243.2024.2370803.

Happe, F. and Frith, U. (2020). Annual research review: looking back to look forward – changes in the concept of autism and implications for future research. *J. Child Psychol. Psychiatry* 61: 218–232.

Hare, B. (2017). Survival of the friendliest: Homo sapiens evolved via selection for prosociality. *Annu. Rev. Psychol.* 68: 155–186.

Hare, B., Plyusnina, I., Ignacio, N., Schepina, O., Stepika, A., Wrangham, R., and Trut, L. (2005). Social cognitive evolution in captive foxes is a correlated by-product of experimental domestication. *Curr. Biol*. 15: 226–230.

Hare, B., Wobber, V., and Wrangham, R. (2012). The self-domestication hypothesis: evolution of bonobo psychology is due to selection against aggression. *Anim. Behav.* 83: 573–585.

Harrison, R.A., Dongre, P., van Schaik, C.P., and van de Waal, E. (2024). The forgotten adaptive social benefits of social learning in animals. *Biol. Rev*. 99: 1638–1651.

Hasson, U., Ghazanfar, A.A., Galantucci, B., Garrod, S., and Keysers, C. (2012). Brain-to-brain coupling: a mechanism for creating and sharing a social world. *Trends Cognit. Sci.* 16: 114–121.

Heersmink, R. (2017). Distributed selves: personal identity and extended memory systems. *Synthese* 194: 3135–3151.

Heersmink, R. (2022). Extended mind and artifactual autobiographical memory. *Mind Lang.* 37: 659–673.

Heersmink, R. (2024). In: Bietti, L.M., and Pogacar, M. (Eds.), *History of memory artifacts in the Palgrave Encyclopedia of memory studies*. Springer Nature, Switzerland.

Henrich, J. (2016). *The secret of our success*. Princeton University Press, New Jersey, USA.

Henrich, J. (2021). Cultural evolution: is causal inference the secret of our success? *Curr. Biol*. 31: R381–R83.

Henrich, J., Heine, S.J., and Norenzayan, A. (2010). Most people are not WEIRD. *Nature* 466: 29.

Heyes, C. (2016). Imitation: not in our genes. *Curr. Biol*. 26: R412–R414.

Heyes, C. (2018) Cognitive gadgets. In: *The cultural evolution of thinking*. Harvard University Press, Cambridge, MA.

Heyes, C. (2019). Precis of cognitive gadgets: the cultural evolution of thinking. *Behav. Brain Sci.* 42: 1–13.

Heyes, C. (2023). Imitation and culture: what gives? *Mind Lang.* 38: 42–63.

Heyes, C. and Catmur, C. (2022). What happened to mirror neurons? *Perspect. Psychol. Sci.* 17: 153–168.

Heyes, C.M. and Frith, C.D. (2014). The cultural evolution of mind reading. *Science* 344: 1243091.

Hipolito, I. (2016). The phenomenology of the intersubjective impairment. *J. Eval. Clin. Pract.* 22: 608–614.

Hodgson, D. (2019). The cognitive mechanisms deriving from the acheulean handaxe that gave rise to summetry, form, and pattern perception. In: Tracy, B., Henley, T.B., Rossano, M.J., and Kardas, E.P. (Eds.), *Handbook of cognitive archaeology. Psychology in prehistory*. Routledge, New York.

Hoehl, S., Keupp, S., Schleihauf, H., McGuigan, N., Buttelmann, D., and Whiten, A. (2019). Over-imitation': a review and appraisal of a decade of research. *Dev. Rev.* 51: 90–108.

Hohwy, J. (2012). Attention and conscious perception in the hypothesis testing brain. *Front. Psychol.* 3: 96.

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

Hu, X., Luo, L., and Fleming, S.M. (2019). A role for metamemory in cognitive offloading. *Cognition* 193: 104012.

Huston, J.P. and Chao, O.Y. (2023). Probing the nature of episodic memory in rodents. *Neurosci. Biobehav. Rev.* 144: 104930.

Iriki, A. (2006). The neural origins and implications of imitation, mirror neurons and tool use. *Curr. Opin. Neurobiol*. 16: 660–667.

Iriki, A. and Sakura, O. (2008). The neuroscience of primate intellectual evolution: natural selection and passive and intentional niche construction. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363: 2229–2241.

Jaynes, J. (1976). *The origins of consciousness in the breakdown of the bicameral mind*. Houghton Mifflin, New York.

Kaczanowska, J., Ganglberger, F., Chernomor, O., Kargl, D., Galik, B., Hess, A., Moodley, Y., von Haeseler, A., Buhler, K., and Haubensak, W. (2022). Molecular archaeology of human cognitive traits. *Cell Rep.* 40: 111287.

Kennedy, J. (2024). *Pathogenesis. How germs made history*. Random House, UK.

Kirschner, M. and Gerhart, J. (1998). Evolvability. *Proc. Natl. Acad. Sci. U. S. A.* 95: 8420–8427.

Koike, T., Tanabe, H.C., Okazaki, S., Nakagawa, E., Sasaki, A.T., Shimada, K., Sugawara, S.K., Takahashi, H.K., Yoshihara, K., Bosch-Bayard, J., et al. (2016). Neural substrates of shared attention as social memory: a hyperscanning functional magnetic resonance imaging study. *Neuroimage* 125: 401–412.

Korth, C. (2020). Tools as petrified memes: a duality. *Behav. Brain Sci.* 43: e169.

Laland, K., Wilkins, C., and Clayton, N. (2016). The evolution of dance. *Curr. Biol*. 26: R5–R9.

Laland, K.N. (2008). Exploring gene-culture interactions: insights from handedness, sexual selection and niche-construction case studies. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363: 3577–3589.

Laland, K.N., Odling-Smee, J., and Feldman, M.W. (2000). Niche construction, biological evolution, and cultural change. *Behav. Brain Sci.* 23: 131–146.

Latour, B. (1987). *Science in action. How to follow scientists and engineers through society*. Harvard University Press, Cambridge, MA.

Laumer, I.B., Auersperg, A.M.I., Bugnyar, T., and Call, J. (2019). Orangutans make flexible decisions relative to reward quality and tool functionality in a multi-dimensional tool-use task. *Plos One* 14, https://doi.org/10.1371/journal.pone.0211031.

Lenggenhager, B., Tadi, T., Metzinger, T., and Blanke, O. (2007). Video ergo sum: manipulating bodily self-consciousness. *Science* 317: 1096–1099.

Levy, J., Goldstein, A., and Feldman, R. (2017). Perception of social synchrony induces mother-child gamma coupling in the social brain. *Soc. Cognit. Affect Neurosci.* 12: 1036–1046.

Lewontin, R.C. (1970). The units of selection. *Annu. Rev. Ecol. Systemat.* 1: 1–18.

Liszkowski, U. and Rüther, J.M. (2024). Ontogenetic origins of infant pointing. In: Gontier, N., Lock, A., and Sinha, C. (Eds.), *The Oxford handbook of human symbolic evolution*. Oxford University Press, Oxford, U.K.

Malafouris, L. (2013). *How things shape the mind. A theory of material engagement*. MIT Press, Cambridge, MA.

Malafouris, L. (2020). Thinking as "thinging": psychology with things. *Curr. Dir. Psychol. Sci.* 29: 3–8.

Mangalam, M., Fragaszy, D.M., Wagman, J.B., Day, B.M., Kelty-Stephen, D.G., Bongers, R.M., Stout, D.W., and Osiurak, F. (2022). On the psychological origins of tool use. *Neurosci. Biobehav. Rev.* 134, https://doi.org/10.1016/j.neubiorev.2022.104521.

Maynard Smith, J. (1987). Natural selection: when learning guides evolution. *Nature* 329: 761–762.

Maynard Smith, J. and Szathmary, M. (1995). *The major transitions in evolution*. W. H. Freeman, Oxford.

McNeill, D. (2013). The co-evolution of gesture and speech, and downstream consequences. In: Müller, C., Cienki, E., Fricke, E., Ladewig, S., and McNeill, D. (Eds.), *Body-language communication. An international handbook on multimodality integration*. De Gruyter Mouton, Berlin.

Menary, R. (2010). *The extended mind*. MIT Press, Cambridge, MA.

Menary, R. (2015). Mathematical cognition: a case of enculturation. In: Metzinger, T., and Windt, J. (Eds), *Open MIND. MIND Group*, Frankfurt am Main, Germany. https://www.openmind.net.

Mesoudi, A., Whiten, A., and Laland, K.N. (2004). Perspective: is human cultural evolution Darwinian? Evidence reviewed from the perspective of the origin of species. *Evolution* 58: 1–11.

Metzinger, T. (2004). *Being no one*. MIT Press, Cambridge, MA.

Metzinger, T. (2020). Minimal phenomenal experience: meditation, tonic alertness, and the phenomenology of "pure" consciousness. *Philos. Mind Sci.* 1, https://doi.org/10.33735/phimisci.2020.i.46.

Metzinger, T. and Wiese, W. (2017). Vanilla PP for philosophers: a primer on predictive processing. In: Metzinger, T., and Wiese, W. (Eds.), *Philosophy and predictive processing*. MINF Group, Frankfurt am Main, Germany.

Morgan, T.J., Uomini, N.T., Rendell, L.E., Chouinard-Thuly, L., Street, S.E., Lewis, H.M., Cross, C.P., Evans, C., Kearney, R., de la Torre, I., et al. (2015). Experimental evidence for the co-evolution of hominin tool-making teaching and language. *Nat. Commun*. 6: 6029.

Müller, C.P. (2020). Mechanisms of a near-orthogonal ultra-fast evolution of human behaviour as a source of culture development. *Behav. Brain Res.* 384, https://doi.org/10.1016/j.bbr.2020.112521.

Muthukrishna, M., Shulman, B.W., Vasilescu, V., and Henrich, J. (2014). Sociality influences cultural complexity. *Proc. Biol. Sci.* 281: 20132511.

Natraj, N., Pella, Y.M., Borghi, A.M., and Wheaton, L.A. (2015). The visual encoding of tool-object affordances. *Neuroscience* 310: 512–527.

Noe, A. (2010). *Out of our heads: why you are not your brain, and other lessons from the biology of consciousness*. Hill and Wang, New York.

Noe, A. (2015). *Strange tools. Art and human nature*. Hill and Wang, New York.

O'Madagain, C. and Tomasello, M. (2022). Shared intentionality, reason-giving and the evolution of human culture. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 377: 20200320.

O'Regan, J.K. and Noe, A. (2001). A sensorimotor account of vision and visual consciousness. *Behav. Brain Sci.* 24: 939–973.

Odling Smee, J., Laland, K., and Feldman, M. (2003). *Niche construction: the neglected process in evolution*. Princeton University Press, Princeton, USA.

Osiurak, F. and Badets, A. (2016). Tool use and affordance: manipulation-based versus reasoning-based approaches. *Psychol. Rev.* 123: 534–568.

Overmann, K.A. and Wynn, T. (2019). On tools making minds: an archaeological perspective on human cognitive evolution. *J. Cognit. Cult.* 19: 39–58.

Paabo, S. (2014). The human condition-a molecular approach. *Cell* 157: 216–226.

Paige, J. and Perreault, C. (2024). 3.3 million years of stone tool complexity suggests that cumulative culture began during the Middle Pleistocene. *Proc. Natl. Acad. Sci. U. S. A*. 121: e2319175121.

Pendergraft, L.T., Marzluff, J.M., Cross, D.J., Shimizu, T., and Templeton, C.N. (2023). American crows that excel at tool use activate neural circuits distinct from less talented individuals. *Nat. Commun.* 14, https://doi.org/10.1038/s41467-023-42203-8.

Perreault, C. and Mathew, S. (2012). Dating the origin of language using phonemic diversity. *PLoS One* 7: e35289.

Pinker, S. (2011). *The better angels of our nature*. Viking Press., New York.

Pizzonia, K.L. and Suhr, J.A. (2022). Systematic review of correlates of internal and external memory strategy use in older adults. *J. Appl. Gerontol.* 41: 1491–1499.

Premack, D. and Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behav. Brain Sci.* 4: 515–526.

Ramachandran, V.S. and Rogers-Ramachandran, D. (1996). Synaesthesia in phantom limbs induced with mirrors. *Proc. Biol. Sci.* 263: 377–386.

Renfrew, C. and Zubrow, E. (1994). *The ancient mind. Elements of cognitive archeology*. Cambridge University Press, Cambridge, UK.

Richerson, P.J. and Boyd, R.T. (2024). Culture in humans and other animals. *Science* 386: 846–847.

Risko, E.F. and Gilbert, S.J. (2016). Cognitive offloading. *Trends Cognit. Sci.* 20: 676–688.

Risko, E.F., Kelly, M.O., Patel, P., and Gaspar, C. (2019). Offloading memory leaves us vulnerable to memory manipulation. *Cognition* 191: 103954.

Rizzolatti, G. and Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci*. 27: 169–192.

Rochat, M.J., Caruana, F., Jezzini, A., Escola, L., Intskirveli, I., Grammont, F., Gallese, V., Rizzolatti, G., and Umiltà, M.A. (2010). Responses of mirror neurons in area F5 to hand and tool grasping observation. *Exp. Brain Res.* 204: 605–616.

Rolfe, L. (1996). Theoretical stages in the prehistory of grammar. In: Lock, A., and Peters, C. (Eds.), *Handbook of symbolic evolution*. Oxford University Press, Oxford.

Ruck, L.M. and Uomini, N.T. (2024) Artifact, praxis, tool, and symbol. In: Gontier, N., Lock, A., and Sinha, C. (Eds.), *The Oxford handbook of human symbolic evolution*. Oxford University Press, Oxford, U.K.

Sanchez-Villagra, M.R., Geiger, M., and Schneider, R.A. (2016). The taming of the neural crest: a developmental perspective on the origins of morphological covariation in domesticated mammals. *R. Soc. Open Sci.* 3: 160107.

Santiesteban, I., White, S., Cook, J., Gilbert, S.J., Heyes, C., and Bird, G. (2012). Training social cognition: from imitation to theory of mind. *Cognition* 122: 228–235.

Schmandt-Besserat, D. (1996). *How writing came about*. University of Texas Press, Austin.

Schubotz, R.I., Ebel, S.J., Elsner, B., Weiss, P.H., and Wörgötter, F. (2023). Tool mastering today – an interdisciplinary perspective. *Front. Psychol.* 14, https://doi.org/10.3389/fpsyg.2023.1191792.

Shennan, S. (2011). Descent with modification and the archaeological record. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 366: 1070–1079.

Shumaker, R.W., Walkup, K.R., and Beck, B.B. (2011). *Animal tool behavior: the use and manufacture of tools by animals*. JHU Press, Baltimore, USA.

Skinner, B.F. (1974). *About behaviorism*. Knopf, New York.

Steels, L. and Szathmary, E. (2018). The evolutionary dynamics of language. *Biosystems* 164: 128–137.

Stephens, G.J., Silbert, L.J., and Hasson, U. (2010). Speaker-listener neural coupling underlies successful communication. *Proc. Natl. Acad. Sci. U. S. A*. 107: 14425–14430.

Stephenson, L.J., Edwards, S.G., and Bayliss, A.P. (2021). From gaze perception to social cognition: the shared-attention system. *Perspect. Psychol. Sci.* 16: 553–576.

Sterelny, K. (2012). *The evolved apprentice*.

Stolk, A., Noordzij, M.L., Verhagen, L., Volman, I., Schoffelen, J.M., Oostenveld, R., Hagoort, P., and Toni, I. (2014). Cerebral coherence between communicators marks the emergence of meaning. *Proc. Natl. Acad. Sci. U. S. A*. 111: 18183–18188.

Stolk, A., Verhagen, L., and Toni, I. (2016). Conceptual alignment: how brains achieve mutual understanding. *Trends Cognit. Sci.* 20: 180–191.

Tamariz, M. (2019). Replication and emergence in cultural transmission. *Phys. Life Rev.* 30: 47–71.

Tattersall, I. (2024). A timeline for the acquisition of symbolic cognition in the human lineage. In: Gontier, N., Lock, A., and Sinha, C. (Eds.), *The Oxford handbook of human symbolic evolution*. Oxford University Press, Oxford, U.K.

Tegmark, M. (1998). Is "the theory of everything" merely the ultimate ensemble theory? *Ann. Phys.* 270: 1–51.

Tennie, C., Call, J., and Tomasello, M. (2009). Ratcheting up the ratchet: on the evolution of cumulative culture. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364: 2405–2415.

Tennie, C., Premo, L.S., Braun, D.R., and McPherron, S.P. (2017). Resetting the null hypothesis: early stone tools and cultural transmission. *Curr. Anthropol.* 58: 652–672.

Thibault, S., Py, R., Gervasi, A.M., Salemme, R., Koun, E., Lovden, M., Boulenger, V., Roy, A.C., and Brozzoli, C. (2021). Tool use and language share syntactic processes and neural patterns in the basal ganglia. *Science* 374: eabe0874.

Tomasello, M. (2014). *A natural history of human thinking*. Harvard University Press, Cambridge, MA, London, UK.

Tomasello, M. (2019). *Becoming human. A theory of ontogeny*. The Belknap Press of Harvard University Press, Cambridge, MA.

Tomasello, M., Kruger, A.C., and Ratner, H.H. (1993). Cultural learning. *Behav. Brain Sci.* 16: 495–511.

Tomasello, M., Hare, B., Lehmann, H., and Call, J. (2007). Reliance on head versus eyes in the gaze following of great apes and human infants: the cooperative eye hypothesis. *J. Hum. Evol.* 52: 314–320.

Tulving, E. (1983). *Elements of episodic memory*. Oxford University Press, New York.

Turner, C.R., Mann, S.F., Spike, M., Magrath, R.D., and Sterelny, K. (2023). Joint evolution of traits for social learning. *Behav. Ecol. Sociobiol.* 77, https://doi.org/10.1007/s00265-023-03314-w.

Twomey, T. (2019) Domestic fire, domestic selves. How keeping fire facilitated the evolution of emotions and emotion regulation. In: Tracy, B., Henley, T.B., Rossano, M.J., and Kardas, E.P. (Eds.), *Handbook of cognitive archaeology. Psychology in prehistory*. Routledge, New York.

Uchiyama, R., Spicer, R., and Muthukrishna, M. (2021). Cultural evolution of genetic heritability. *Behav. Brain Sci.* 45: e152.

Urban Levy, P., Sklenar, A.M., Frankenstein, A.N., and Leshikar, E.D. (2023). Evidence for a memory advantage for prosocial behaviors. *Brain Behav.* 13: e3096.

Van Bergen, P. and Sutton, J. (2019). Sociocultural memory development research drives new directions in gadgetry science. *Behav. Brain Sci.* 42: e185.

van Mazijk, C. (2024). Intentionality, pointing, and early symbolic cognition. *Hum. Stud.* 47: 439–458.

van Schaik, C.P. and Pradhan, G.R. (2003). A model for tool-use traditions in primates: implications for the coevolution of culture and cognition. *J. Hum. Evol.* 44: 645–664.

Vasil, J., Badcock, P.B., Constant, A., Friston, K.J., and Ramstead, M.J.D. (2019). A world unto itself: human communication as active inference. *Front. Psychol.*, 11: 1–26, Art. No. 417.

Veissiere, S.P.L., Constant, A., Ramstead, M.J.D., Friston, K.J., and Kirmayer, L.J. (2019). Thinking through other minds: a variational approach to cognition and culture. *Behav. Brain Sci.*: 1–97, https://doi.org/10.1017/s0140525x19001213.

Viaro, R., Maggiolini, E., Farina, E., Canto, R., Iriki, A., D'Ausilio, A., and Fadiga, L. (2021). Neurons of rat motor cortex become active during both grasping execution and grasping observation. *Curr. Biol.* 31, https://doi.org/10.1016/j.cub.2021.07.054.

von Melchner, L., Pallas, S.L., and Sur, M. (2000). Visual behaviour mediated by retinal projections directed to the auditory pathway. *Nature* 404: 871–876.

Wadge, H., Brewer, R., Bird, G., Toni, I., and Stolk, A. (2019). Communicative misalignment in autism spectrum disorder. *Cortex* 115: 15–26.

Wajcman, J. (2015). Pressed for time. In: *The Acceleration of life in digital capitalism*. University of Chicago Press, Chicago.

Wass, S.V., Whitehorn, M., Haresign, M., Phillips, E., and Leong, V. (2020). Interpersonal neural entrainment during early social interaction. *Trends Cognit. Sci.* 4.

Wheatley, T., Thornton, M.A., Stolk, A., and Chang, L.J. (2024). The emerging science of interacting minds. *Perspect. Psychol. Sci.* 19: 355–373.

Whiten, A. (2005). The second inheritance system of chimpanzees and humans. *Nature* 437: 52–55.

Whiten, A. (2022). Blind alleys and fruitful pathways in the comparative study of cultural cognition. *Phys. Life Rev.* 43: 211–238.

Wrangham, R. (2017). Control of fire in the paleolithic: evaluating the cooking hypothesis. *Curr. Anthropol.* 58: S303–S13.

Yoshida, K., Go, Y., Kushima, I., Toyoda, A., Fujiyama, A., Imai, H., Saito, N., Iriki, A., Ozaki, N., and Isoda, M. (2016). Single-neuron and genetic correlates of autistic behavior in macaque. *Sci. Adv*. 2, https://doi.org/10.1126/sciadv.1600558.

Zeng, T.C. and Henrich, J. (2022). Cultural evolution may influence heritability by shaping assortative mating. *Behav. Brain Sci.* 45: e181.

Zlatev, J., Persson, T., and Gärdenfors, P. (2005). Triadic bodily mimesis is the difference. *Behav. Brain Sci.* 28: 720.