



The Prenatal Origins of Human Behavioral Modernity: A Speculative Theory

Richard Parncutt

Introduction

In anthropology and archeology, behavioral modernity refers to typical behaviors that distinguish humans from other animals. We humans are capable of planning future behaviors, based on memory of, and thinking about, the past. That ability allowed early humans to make sophisticated tools including knives, and to kill large, dangerous animals. We humans also assign arbitrary meanings to sounds (words) and objects (art). We perceive ourselves as members of society, identifying ourselves with clothes and ornaments.

The distinction between humans and other animals is not clearcut. Some of the cognitive skills that are characteristic of behavioral modernity can be observed in animals. The archeological record suggests that behavioral modernity emerged gradually, consistent with Darwinian biological evolution. Early humans started to make stone tools and kill large animals some 2 to 3 million years ago; behavioral modernity emerged during a period of roughly 100,000 years, starting with the appearance of anatomically modern humans roughly 200,000 years ago.

I will focus on language, art, music, religion and consciousness – five categories of behavioral modernity that can be considered either independently or interdependently. All have both universal and culture-specific aspects. Common to all five is *symbolic behavior*. A symbol is something that represents something else in an arbitrary relationship; for example, the sound of the word “dog” has nothing to do with a dog itself. Early human symbolic behaviors included assigning meaning to arbitrary patterns of sound and movement to create language. Consciousness emerged with symbols such as body decoration, suggesting that the people had personal names. Music, art, and religion are full of symbolic meanings; the meanings of melodies and rhythms, paintings and sculptures, gods and rituals

differ from one culture to the next. There are also interesting cross-cultural commonalities.

For centuries, researchers and thinkers have speculated about the origin of behavioral modernity. This paper presents an approach based on the behaviors and abilities of the preverbal human: the fetus and infant. The aim is not to replace existing approaches, but to complement them.

Existing theories

Existing theories of the origin of behavioral modernity have answered many interesting questions but left equally many unanswered. Discussion continues in the areas of uncertainty. Before entering that uncertain area, I will briefly present some existing theories.

Language is a complex system of sounds and meanings. It implies or entails abstract thinking, planning, and imagination. It includes words and syntax, both of which are learned in cultural contexts. Seen ecologically, language is based on speech, a complex behavior that includes the sound, articulation, and perception and cognition of words, phrases, and sentences. It also involves accompanying physical gestures (hand movements, facial expressions). An ecological approach implies that the origin of speech involves both vocalization and gestures. Non-human animals vocalize and gesticulate for various reasons, including emotion (expression of physical or emotional state in a social context), conflict (scaring off another animal or conspecific), or sexual selection (attracting or repelling a potential mate).

Speech sounds are *honest* in the sense of biological signaling theory if they communicate correct, reliable, or useful information about the speaker. An example of a dishonest signal is a low-pitched vocalization that suggests an animal is bigger or more dangerous than it really is. In an evolutionary approach, language can only develop if speakers trust each other to send honest signals. That makes communication between caregivers and infants a likely setting for the emergence of language. From an evolutionary viewpoint, caregivers and infants have the same aim: infant survival.

The prehistoric emergence of language presumably depended on parallel developments in art, music, and religion. Symbolic behaviors such as body painting may have helped early humans assign meanings to sounds produced by the vocal tract. Conversely, emerging linguistic abilities (such as the ability to imagine and describe something that one is planning)

could have enabled symbolic behaviors in other areas, including music, art, and religion.

Music involves entrainment (synchronized movement in performance and dance) and group singing (a different form of synchronization). It promotes social coherence, giving group members a feeling of belonging. It motivates group members to look after each other and fight for a common cause or against a common enemy. The idea of music as social glue is widely accepted, but it is unclear why music has that property. Neither entrainment nor singing necessarily implies cooperation. Music requires a lot of time and energy that from an evolutionary viewpoint is better spent directly promoting survival and reproduction.

Much the same can be said for religion, which involves beliefs in the supernatural that are expressed collectively in ritual. Religions include traditional narratives, moral rules or laws, special (holy) places, and social roles. Religious behaviors demand extraordinarily high amounts of time and effort, supported by high levels of devotion and commitment. What motivated people to make such enormous investments for unclear rewards?

A broad definition of religion includes the traditional religions of indigenous hunter-gatherer societies. Understanding them seems essential for understanding the origins of religion. Hunter-gatherer religions are animist, believing that most things in the world have a soul or spirit; that includes people, animals, plants, rocks, mountains, rivers, clouds, and so on. If the first religions were like that, a possible explanation involves the emergence of consciousness. When humans discovered themselves and their personal identity, they also discovered the souls and agency of other humans, and by extension everything else in their world.

Research on the origin of religion has answered many interesting questions, but from the (admittedly arbitrary) point of view of an atheist scientist, stubbornly difficult issues remain. Why do many modern, highly educated, scientifically aware people maintain beliefs in supernatural agency, in diverse cultures worldwide – despite the cognitive dissonance? Why did early societies regulate behavior with supernaturally controlled laws and punishments when human laws and punishments could have worked just as well – and often did?

Art often means decorating ritual objects or attracting attention to colors and shapes. Why did humans create the first abstract art? Did they enjoy looking at patterns of lines and colors? If so, why? If not, what motivated them to create the patterns? Much later, the first representational art appeared. Early cave art depicted the large animals that early humans hunted and ate. What did those artists expect to achieve with their paint-

ings? We guess that early art had symbolic or religious function, but what function exactly, we do not know. If early human behavior was determined by cost-benefit relationships, there would have been no rituals and no paintings: the artists would instead have been busy gathering food and hunting animals.

To be conscious is to be conscious of being a person: a member of society of conscious beings. Archeological evidence for the prehistoric emergence of a sense of self, and hence consciousness, includes body decoration: beads, bracelets, necklaces, pendants. Consciousness enables us to take the perspective of other people (empathy, theory of mind), or other times or places (mental travel). Mental time travel involves remembering what happened in the past or imagining what might happen in the future while simultaneously perceiving what is happening in the present.

Here, I will confine myself to aspects of consciousness that can be observed in empirical psychological research, such as the emergence of consciousness in infancy and childhood. I will assume that consciousness emerged in human prehistory as symbolic behaviors (language) helped humans create cognitive representations (imagine things) from different perspectives. I will avoid difficult philosophical questions about why we have subjective experiences at all, or the exact relationship between consciousness and brain states.

To define the boundaries of consciousness, it helps to consider the case of the domestic dog, a social species that has been adopted by another social species, the human. Pet dogs seem to share aspects of human consciousness, but empirical evidence suggests that humans are better at empathy, theory of mind, mental travel, and self-awareness. The difference is large, but it is also one of degree. Consciousness also depends on language: I can hardly know whether my dog has consciousness, in the sense of knowing that it is experiencing something, if it cannot talk and share its experience. That is an empirical problem, but it is also consistent with the (unprovable) assumption that language is a co-requisite for consciousness.

That raises interesting questions. Did the emergence of language enable the emergence of consciousness? Or did consciousness arise more directly from social interactions? Did early humans acquire consciousness in rituals that involved psychoactive substances, with altered states of consciousness attracting their attention to consciousness itself? Or did consciousness emerge naturally as the human mind became more complex?

Preverbal psychology

This paper presents a theory of the origin of behavioral modernity that is based on prenatal and infant (i.e., preverbal) psychology. The theory relies on general psychological foundations. One of them is the *schema*.

A psychological schema is a collection of typical perceptible characteristics of a given object, concept, or situation, and accompanying behaviors. Textbook examples include situations in which human participants usually behave (and therefore are expected to behave) in specific ways, such as railway stations, doctor's surgeries, classrooms, and teenage parties. A situational schema is activated when we perceive the characteristic features of the situation, and spontaneously respond with corresponding or appropriate behaviors.

Different schemas emerge at different developmental stages. There must, therefore, be a first schema in ontogeny, and it must be the first perceptible situation that evokes characteristic behavioral responses. Since the human fetus in the third trimester can perceive in all senses, and since it responds to what it perceives, the first schema must be prenatal. I will call it *prenatal schema*, on the assumption that it is unitary (although like other schemas it is possible to analyze it or divide it up). The prenatal schema is *situational* in that it refers to the prenatal situation, and *social* in that it incorporates early maternal attachment, the first relationship.

Motor activity is important for biological and neural development. The fetus makes breathing movements, opens and closes its mouth, yawns, sucks its fingers, and swallows amniotic fluid. It goes through wake-sleep cycles. Its behavior also depends on its situation: for example, it responds to a sudden loud noise.

In the third trimester, all sensory systems are supplying the fetal brain with constant streams of information that are part of perceptual and cognitive development. Whereas the fetus probably lacks consciousness as defined, it is capable of perception and attention, and it interacts with its environmental niche, like any other animal.

From the fetal perspective, the prenatal situation includes the following:

- vocalizations of the mother and (nearby, loud-voiced) others (low-pass filtered),
- the sounds and movements of maternal walking,
- the changing taste of amniotic fluid (depending on what the mother just ate),

- bent bodily postures (more so as the fetus grows),

The prenatal situation is characterized by proximal rather than distal perception. Proximal perception entails a perceptual focus on stimulus properties rather than the people or objects in the world that are producing them. For the fetus, that means perceiving the pitch-time patterns of the maternal voice, whole-body movements when the mother walks, the changing taste of amniotic fluid, and occasional visual stimulation – all without knowing how these patterns originated or what environmental objects or phenomena they represent.

Research in prenatal psychology suggests that these characteristics of the prenatal situation are cognitively processed and stored in prenatal memory. These processes can be seen through the lens of behavioral psychology, and independent of any form of consciousness. There is evidence for prenatal episodic memory, and hence for a form of prenatal consciousness, in psychotherapeutic practice; but that idea is unnecessary for the purpose of explaining the origin of behavioral modernity.

In an evolutionary approach, we expect the fetus to attend to stimuli that are important for postnatal survival. Anything that reflects the mother's physical or emotional state (voice, heartbeat, movements, footsteps, biochemical content of placental blood) is important in that evolutionary sense. In an ecological approach, those stimuli are the *affordances* of the prenatal situation.

Like other schemas, the prenatal schema includes characteristic emotions whose neurophysiological foundations are similar to those of adult human emotions. That foundation includes the limbic system, which develops throughout gestation, starting in the first trimester. After birth, the limbic system is relevant for motivation, emotion, and behaviors that promote survival and reproduction, including sex, childcare, and fight/flight.

Human emotions are associated with characteristic pitch-time-movement patterns or gestures that the fetus learns before birth. Note that here and elsewhere I use the word “learn” in the general psychological sense of acquiring information that changes later behavior. In everyday conversation, but not here, the word “learn” implies conscious learning.

The everyday life of an adult, whether pregnant or not, includes emotional ups and downs that correlate with internally audible sounds and movements. Changing hormonal concentrations in the maternal blood allow the fetus to learn how sound/movement patterns correlate with the mother's physical/emotional state. Each hormone has different physiolog-

ical and emotional correlates, and many are lipids that pass the lipid membranes of the placenta and blood-brain barrier.

In the absence of language and consciousness, fetal emotions cannot be described. But we might reasonably expect prenatally “experienced” emotions to correspond to (later) feelings of

- safety (protection, absence of fear),
- being inside an enclosed space,
- being part of, and dependent on, a larger living being, and
- floating, weightlessness, or being immersed in fluid.

There are two kinds of fetal learning. Both are quite intuitive and procedural – unlike learning in the everyday sense of learning a phone number or studying for an exam. *Habituation* is a simple form of familiarity; the response to a stimulus gets weaker when it is repeated. In *classical conditioning* or *statistical learning*, stimuli that happen together or in succession are associated with each other such that perceiving one of them evokes responses characteristic of the other. In this way, the fetus can acquire and store information about typical patterns of pitch, time, or movement in stimuli that are heard repeatedly before birth.

The fetus hears the maternal heartbeat for about 20 weeks before birth – half of the gestational period. The tempo of the heartbeat goes up and down in two characteristic ways that the fetus can learn: in faster cycles (lasting a few seconds) that are coupled to maternal breathing, and slower cycles (lasting minutes or hours) that depend on the changing physical and emotional state of the mother. The fetus also learns that heart-rate variability is higher when the mother is relaxed.

Footsteps are another example. Often, the fetus perceives footsteps and associated body movements. That enables it to learn about normal ranges of walking speed, the isochrony (equal spacing in time) of maternal walking by comparison to more temporally variable heartbeat sounds, the relation of maternal footsteps to fetal body movement, and relationships between their loudness and their tempo.

The fetus has many opportunities to learn about typical fundamental frequency contours of speech and other vocalizations, and their relation to physical and emotional state, and breathing. These natural patterns are similar to cultural patterns of musical melody and phrasing.

The third-trimester human fetus can open its eyes. It has a developed visual cortex, although there is little to see. Cones on the fetal retina are starting to differentiate colors. After birth, vision gradually becomes the dominant sense, taking over from hearing and demanding more neural

resources that other senses. Against this background, evolutionary theory predicts the fetus will, as far as possible, exercise and develop its visual abilities, to get a head-start on postnatal visual development.

The evolution of human infancy

For over a million years, the combination of upright walking and growing brain made the human birth process (parturition) increasingly difficult. Often, it was deadly for the infant or the mother or both. Therefore, infants were born earlier and became more fragile. Some were more likely to survive to reproductive age than others due to quasi-random variation in biological and behavioral traits. When infants became parents themselves, their successful genes were passed to future generations. Over many generations:

- The newborn's skull became softer.
- Parturition became more complex, including a series of fetal rotations.
- Midwifery emerged when groups of older women assisted younger women during childbirth – an early expression of human empathy (although midwifery has also been observed in bonobos).
- The newborn became more demanding, crying more often to motivate or manipulate the constant attentive presence of caregivers.
- Newborns forced caregivers to attend to them constantly in other ways, as demonstrated by the *still face experiment* in developmental psychology.
- Infants played in more complex ways and for longer periods of time, which developed their perceptual, motor, social, and cognitive skills. Play also maintained the joint attention of caregivers for longer periods.
- Infants evaluated the long-term behavior of their caregivers and responded to inappropriate behaviors in manipulative ways (cf. secure, avoidant, ambivalent, disorganized attachment).

In this way, genetically transmitted evolution had both biological and behavioral aspects, and behavioral aspects spilled over into cultural evolution.

According to the *baby schema* concept of Konrad Lorenz, human babies are perceived as cute if they have typical baby-like features like a round face and relatively large eyes. For the same reason, humans may perceive non-human animals as cute and then choose them as pets. In this way, the baby schema motivates infant care across different species. By analogy to

the caregiver's infant schema, human infants developed a *caregiver schema*, responding to adults who were perceived to be trustworthy caregivers (giving off the right signals) by behaving in cute ways that would in turn trigger the caregiver's baby schema. The interaction between baby schema and caregiver schema promoted the survival of fragile infants in a dangerous world.

The most dangerous time for any human infant is the first few days and weeks after birth. In preparation for that, an evolutionary approach predicts that the human fetus became increasingly sensitive to the changing physical and emotional state of the mother, as reflected by the internal sounds of her body. It increasingly perceived the mother as existentially important: all-loving and all-powerful. After birth, that prenatal preparation helped it monitor the changing emotional and physical states of caregivers and respond in appropriate ways to maximize its chance of survival.

Based on prenatal experience, infants increasingly detected signs of danger in caregiver behaviors such as fear in the maternal voice. They cried only if crying was likely to promote their survival in a given situation – not if crying might put the infant, caregiver, or tribe in danger. That can explain why the children of anxious mothers become anxious themselves, which affects their emotional self-regulation. Unconfirmed anecdotal evidence from the delivery room suggests that a healthy newborn can be remarkably quiet – crying less than other infants, or not at all – if its mother is experiencing a perinatal emergency such as delivery hemorrhage.

The survival of the fragile human infant increasingly depended on its ability to contribute actively to bonding relationships with older children and adults, including its biological mother. That happened in a context of *alloparenting*: in a typical ancestral tribe, probably several adults looked after several children, while others were busy with other tasks. Infants and children depended for their survival on the quality of their relationships with different caregivers. They increasingly adjusted their behavior to match the needs, possibilities, and expectations of caregivers.

The skills of caregivers evolved – biologically and behaviorally – with those of infants. It is not easy to guess infant needs from the sound of crying alone. As caregivers collectively became more skilled at that, using contextual information (e.g., when the infant was last fed), language also developed.

Preverbal origins of behavioral modernity

Preverbal psychology suggests how behavioral modernity might have emerged. In the following, this idea is pursued for each of the five mentioned aspects of behavioral modernity separately.

Language

Language probably emerged in everyday situations of one-to-one communication, and in honest (reliable, empathic, trusting) interactions such as those between caregivers and infants/children. The obstetric dilemma had made infants more fragile, forcing them to develop the survival-promoting skill of monitoring the physical/emotional states of caregivers. In this way, the fetus/infant learned to link an ever-growing number of complex patterns of sound and movement with meanings. Each pattern implicitly answered two questions: What is the current maternal/caregiver state/situation? What infant behavioral response would best promote infant survival?

In his book *The Language Instinct*, cognitive linguist Steven Pinker argued that humans have an innate capacity for language. This central claim is consistent with the present theory, but with some caveats. The word “innate” means “inborn;” it does not specify whether the capacity in question is prenatally learned (due to environmental interaction; “nurture”) or genetic (due to biological evolution: “nature”). Usually, complex skills are a combination of nature and nurture. I interpret Pinker’s “language instinct” as an ability, primarily genetically determined, and a motivation to associate large numbers of temporally and spectrally complex patterns of sound and movement with arbitrary meanings. Like any ability, the language ability can be trained – in this case, both prenatally and postnatally. This “instinct” enables children to quickly learn a large vocabulary of words, and on that basis to build language.

Syntax is central to language and hence to any theory of its origin. If the fetus gradually became more sensitive to complex, prenatally perceptible patterns of sound and movement, it also became sensitive to temporal order effects, which could have contributed to the development of linguistic syntax. For example, imagine a pregnant woman stumbling and falling. The fetus perceives a series of changes in maternal vocalizations, body movements, and heartrate. Later, there are hormonal changes in the blood. The changes happen in characteristic orders, to which the fetus can become sensitive by processes similar to classical conditioning. The complexity of

human behavior and physiology means that the fetus is constantly exposed to temporal orderings of perceptible stimulus combinations and their emotional implications.

Music

Early humans gathered in the evenings and cooked on fires, increasing the length of their active day. They experienced stimulus combinations such as talking (vocalizing) and rhythmic sounds from making and using tools.

The archeological record shows that this sometimes happened in caves. The cave's complex acoustic reflections (resonances) boosted acoustic energy at low frequencies and reduced clarity and intelligibility. The situation was dark and enclosed, and the participants felt protected from wild animals.

That resembled the prenatal situation in several respects. I propose that this experience therefore triggered the prenatal schema of several participants. In general, psychological schemas are unconsciously activated when several of their typical characteristics occur simultaneously. In this case, the schema was experienced as a mysterious but enjoyable (magical) feeling – similar to modern experiences of listening to music or participating in religious rituals.

The actors in this ancestral scenario were motivated by this special feeling to repeat the behaviors that seemed to have produced the feeling, according to well-known principles of operant conditioning. Those behaviors included producing early forms of melody and rhythm. Melody was produced by singing or imitating the singing voice with simple instruments. Rhythm meant creating repetitive percussive sounds, either with body percussion or objects (tools). Rhythmic sounds motivated rhythmic movement (dance), which further reinforced the magical feelings (think of how the fetus moves when the mother walks). Song, although mainly melodic, could also be rhythmic, if combined with rhythmic tapping or clapping. Percussive sound, while mainly rhythmic, could also be melodic if different percussion instruments produced different pitches.

Psychoactive drugs may have played a role, magnifying the magical feelings. But the effect was also possible without drugs. Therefore, drugs do not play a central role in the present theory, although in reality they could have been important.

Religion

Intrinsic to the prenatal schema is the mother as perceived by the fetus. The fetus starts to develop an attachment relationship with the mother, upon whom its postnatal survival will depend. Activation of the prenatal schema therefore co-activates emotions that we might expect to be associated with a large, moving agent or an ungendered, god-like protector.

Those emotions are diverse and range from devotion to fear. The emotion “experienced” by the newborn in the presence of the mother or another caregiver may be compared to the awe that participants in religious rituals experience in the presence of the divine. Such emotions become mysterious or magical when spontaneously re-awakened in childhood or adulthood by religious rituals that activate the prenatal schema.

Religion is a belief in supernatural agents that guide rituals and determine moral rules and human destiny. I propose that the first religions emerged when groups of early humans collectively experienced several features of the prenatal schema simultaneously. Those features include bent postures, enclosed spaces, dull light and sound (resonant caves or buildings), and music (song and/or dance). The activation of the prenatal schema evoked a feeling of supernatural presence that was based on the feeling of maternal presence before birth, as the fetus started to bond with the mother.

This origin-of-religion scenario overlaps with the above origin-of-music scenario. The emotion is linked to the sounds and movements in the musical case and the supernatural agent in the religious case. Early humans were motivated to enhance these feelings by repeating or intensifying the actions that seemed to have caused them. In the case of religion, the feelings could be enhanced by entering a dark, resonant space; by a leader speaking or vocalizing loudly; by song or dance; and/or by characteristic fetal body postures such as kneeling, prostration, or touching the face.

As religion developed, so too did other aspects of behavioral modernity, including language. That enabled early humans to explain the mysterious feelings and pass on their explanations from one generation to the next in a developing oral tradition, creating sacred stories and myths.

Art

The archeological record suggests that art emerged in two stages. About 100,000 years ago, early humans accidentally discovered the emotional effect of geometric patterns. In ritual situations, geometric patterns activated

the prenatal schema due to their similarity to prenatally perceived entoptic phenomena.

Entoptic phenomena are optical illusions that originate in the physiology of the eye. They may comprise moving dots (“floaters”), grids, zigzag lines, or other geometric shapes. In everyday life, we seldom notice them; instead, the visual system creates a model of the visible world based on the incomplete or distorted information provided in real time by the eyes.

I propose that the imitation of entoptic phenomena in early art – motivated by activation of the prenatal schema – enabled early humans to develop simple artistic skills. That was a prerequisite for the later emergence of representational art. About 50,000 years ago, art started to represent the supernatural agent(s) whose presence humans experienced in caves during early religious rituals.

Today, we humans can perceive the world in two different ways, *distal* and *proximal*. In everyday life, our perception focuses on distal stimuli: we perceive physical things (people, animals, trees and so on). But when we look at visual art (e.g., in a gallery), we adopt a more esthetic habitus, focusing on the proximal stimulus – the stimulus itself, with its colors and shapes. Probably, non-human animals do not do that.

But prenatal perception is also proximal. If the fetus is going to train its eyesight, just as it trains its other senses, it has no choice but to focus on the stimulus itself, which means looking at entoptic phenomena. After birth, perception suddenly becomes distal; proximal perception no longer has evolutionary value. That is a big change, and the infant takes several weeks or months to adjust.

I propose that, when early humans looked at geometric patterns in the context of ritual, the combination of stimuli activated their prenatal schema. In that way, they learned again to focus attention on the proximal stimulus – something that previous humans or animals had seldom done. They were emotionally rewarded for doing so by activation of the prenatal schema, so they did it repeatedly.

Representational art, including cave paintings of animals, humans, and mixtures of the two, emerged as early humans tried to visually represent the supernatural presence that they experienced in early religious rituals. The animals they represented tended to be large – consistent with the size of mother from the fetal viewpoint. That can explain why images of humans are rare in the earliest cave art.

Cave art is often found near acoustic resonances, high on cave walls or on the ceiling. That is consistent with the physical location of the mother’s vocal chords and heart relative to the fetus. Early in the third trimester, the

fetal organs of balance are functioning, and the fetus hears the mother's heartbeat and voice more or less loudly – depending on its orientation. The fetus can also move about, changing the position of its head and ears, which allows the fetus to perceive the mother's voice as above and her feet as below. The fetus might also be able to approximately locate other people in social situations, very roughly guessing direction and distance. These ideas are consistent with the universal idea of heaven or paradise being in the sky; the darkness and mystery of caves is consistent with hell or purgatory being under the ground (underworld).

Consciousness

Consciousness emerged in connection with other aspects of behavioral modernity. Perhaps that idea is sufficient to explain its origin. Prenatal theory nevertheless suggests some additional factors that may have promoted the emergence of consciousness.

The obstetric dilemma was resolved in part by social behaviors that helped and supported women giving birth. When such behaviors were selected for, midwifery became more sophisticated. That was an important milestone in the development of human empathy and cooperation.

As infants played, their increasing fragility put caregivers under pressure to predict and prevent potentially fatal accidents, developing their mental time travel – their ability to imagine and predict the future. That same ability was also used to create tools, kill dangerous animals, and design living areas.

The fragility of human infants also trained elements of consciousness in infants themselves. Carer-infant interactions trained their theory of mind as they learned to play with toys as if they had agency. Caregivers interacted with infants as if they were conscious language users, long before they had acquired this skill. That helped develop both language and consciousness.

A specific origin scenario

Hominins have used fire in a controlled way for at least a few hundred thousand years. Fire enabled early humans to stay awake and do things in the evening. That made the day effectively longer: social activity no longer stopped at sunset. Fires were sometimes lit in caves, which protected early humans from rain, cold, and dangerous animals. Fire also enabled cooking,

making it easier to digest large amounts of energy and protein. That led to a larger brain and smaller teeth.

The use of fire for light, warmth, and cooking allowed early humans to participate in social activities that could have strengthened the group in the struggle for survival. In that situation, we can imagine how music and religion might gradually have emerged in a process akin to operant conditioning: an animal (human ancestor) does something, gets an emotional reward for it, and then does it again in order to receive the reward again. For example, a rat in a behavioral psychology experiment may accidentally push a lever and be rewarded with sugar, motivating the rat to repeat the action that seemed to have caused the reward.

The sounds that early humans heard in caves had more energy at low frequencies than sounds heard in the open. A typical low tone sung by a bass singer is G₂, the lowest line of the bass clef. It has a fundamental frequency of about 100 Hz (100 cycles per second) and a wavelength of 3.4 m (since the speed of sound is 340 m/s). That is also the typical width or height of a cave that people can walk around in. In such a cave, low-pitched sounds are reinforced by acoustic resonances. That reduced the audibility of the higher harmonics of voiced speech sounds.

Sound in caves was also less clear, due to multiple reflections, which further reduced speech intelligibility. In addition, any people who were preparing food, making tools, or working with tools produced sounds that repeated periodically about one to two times per second – comparable with heartbeats and footsteps.

In these ways, the experiences of early humans in caves were similar to aspects of the prenatal schema described above – often, by accident. In schema theory, a schema is activated by presenting some (not all) of its characteristic stimuli, just as a pattern can be recognized even if part of the pattern is missing or occluded. Schema activation evokes the missing aspects. Here, they include fetal-maternal bonding from the fetal perspective and corresponding emotional states.

Music and religion are emotionally similar to each other. Both involve reverence (veneration, gratitude, surrender), transcendence (wonder, amazement, magic), awe (fear/dread combined with wonder), joy (bliss), power (activation), serenity (tranquility), and unconditional love (longing). Whereas this list initially seems biased toward Western culture or monotheism, it is also broadly consistent with diverse religious traditions including those of modern hunter-gatherers and their cosmologies, which typically include animism (spirits, totems) and afterlife (ancestor worship, shamanism).

Hunter-gatherer religions typically feature high gods with whom humans interact. In animist Native American mythology, the Great Spirit or Great Mystery, Wakan Tanka, created the Lakota Sioux. In the more totemic Australian Aboriginal cosmology, Baiame came down from the heavens to create the landscape and human culture. A theory of prenatal origins predicts that all religions involve activation of the prenatal schema; that the triggered emotions are strong (existentially important) and embodied (with a corporal foundation); and that the supernatural (the mother from the fetal viewpoint) is central.

New cultural behaviors did not emerge by accident, but were driven by an emotional reward. That reward could have been the emotion characteristic of the mother-fetus relationship from the fetal perspective. The inclusion of this behavioral mechanism in the present theory sets it apart from purely Darwinist evolutionary theories in which quasi-random variations in behavior led to selection for behaviors that better promote survival and reproduction. A theory in which specific behaviors are rewarded by specific emotions is a promising path toward understanding cultural developments such as art, music, and religion.

The theory being proposed is gradualist – also in the tradition of Darwin. The changes being proposed are understood to have happened incrementally over long periods. That enabled different aspects of behavioral modernity to interact with each other. For example, as the linguistic predisposition of the fetus grew – slowly, over hundreds or thousands of generations – the complexity of the acoustic communication of adults also improved, which in turn enriched the acoustical environment of the fetus. In this way, music and religion could have developed by many small steps, and the same applies to art and consciousness.

The theory is intended to be parsimonious and non-circular. Regarding parsimony, the aim is to minimize the number of controversial assumptions, making the resulting theory easier to falsify. That in turn makes the theory more convincing. Regarding non-circularity, it should be clear that the ultimate cause of behavioral modernity was not part of behavioral modernity itself. This criterion is satisfied if the cause was the obstetric dilemma, in combination with physiological prerequisites.

I am assuming that during infancy and childhood (for ontogeny), and in the period between 200,000 and 100,000 years ago (for phylogeny), all aspects of behavioral modernity (depending on definition) emerged in parallel, interacting with and reinforcing each other, while also taking advantage of developments that had happened before that time. In ontogeny, those previous developments included increasing sensitivity to the chang-

ing physical and emotional states of the mother from the fetal viewpoint; in phylogeny, they included physiological prerequisites for behavioral modernity such as upright posture, a large brain, a lower larynx, and behavioral prerequisites such as the controlled use of fire.

To ensure non-circularity, I am assuming in a first approximation that neither the fetus today, nor human ancestors before about 200,000 years ago (or any other animal), is/was significantly capable of any aspect of behavioral modernity. At that time – in a simplified first pass at the problem – there is/was no language, no music, no art, no religion, and no consciousness. In fact, all aspects of behavioral modernity emerge/d gradually and did not start from nothing.

Implications

The theory suggests that the obstetric dilemma triggered the emergence of human behavioral modernity. That generates a flood of questions and that could lead to new insights about the nature of art, music, religion, language, and consciousness. It challenges us to rethink who we are as humans.

Like other theories of the origin of behavioral modernity and its specific aspects, the present theory is difficult to test empirically. Evaluation might involve listing phenomena that are consistent with the theory or enumerating arguments for and against. Skepticism is justified if it is founded on scientific concerns or empirical data. The inseparability of a mother and her fetus makes it difficult to design experiments that clearly demonstrate causal relationships. But one can equally argue the opposite: a theory based on prenatal psychology is easier to test than one based on archeology or evolution, if prenatal behavior is happening in the here and now and can be observed increasingly accurately with emerging non-invasive technologies. Seen in that way, there is considerable potential for new research.

We live in a patriarchal society in which men are understood to be normal and women special. In addition, adults are considered normal and children special. If our existence as conscious, artistically and linguistically sophisticated humans depends on the obstetric dilemma, these relationships can be called into question. The theory suggests that the “normal” human – the one with the right to claim some kind of moral or political superiority – is a young woman. Whereas patriarchy will stubbornly resist such a conceptual revolution, implicit sexism and ageism could nevertheless explain why a theory of this kind is considered “speculative,” although

other theories of the origin of behavioral modernity are arguably no less so.

Acknowledgments. I am grateful for Annemarie Seither-Preisler, Markus Christiner, Klaus Evertz and Ludwig Janus for comments on a previous draft, and to participants at the 2023 Heidelberg conference of the International Society for Prenatal and Perinatal Psychology and Medicine (ISPPM) for discussion.