**Autogenic transitions in individuality**

Paul B. Rainey

Department of Microbial Population Biology, Max Planck Institute for Evolutionary Biology, Plön, Germany, and Laboratory of Biophysics and Evolution, CBI, ESPCI Paris, Université PSL, CNRS, Paris, France.

ORCID: 0000-0003-0879-5795

Email: rainey@evolbio.mpg.de

**Abstract**

Major evolutionary transitions in individuality occur when independently reproducing entities fuse to form a new unit with a shared reproductive fate. Less considered are transitions that originate from within, when an autogenic innovation — a component generated internally within a lineage rather than acquired from outside — becomes a heritable part of a higher-level entity. The emergence of AI and its deepening interdependence with humans make it timely to explore such internal pathways. Three routes can be distinguished: (1) centralised, non-replicating AI that shapes but does not reproduce with humans; (2) replicating AI lineages forming symbioses with humans; and (3) synthetic endosymbioses in which AI becomes a developmentally inherited module. The first alters selection without creating new individuals; the latter two could generate composite lineages in which humans and AI reproduce together. Viewing individuality as capable of arising from within reframes how new Darwinian individuals can emerge across both natural and synthetic domains.

**Keywords**

Artificial intelligence – Major Evolutionary Transitions – Individuality – Human-AI symbioses

**Rethinking transitions in individuality**

Evolutionary biology explains the diversity of life through incremental adaptation (Darwin 1859, Fisher 1930, Dobzhansky 1937), but some of the most important shifts in life’s history occurred when the units of evolution changed (Maynard Smith & Szathmáry 1995, Okasha 2006, Godfrey-Smith 2009). These major evolutionary transitions in individuality (ETIs) mark turning points: genes were assembled into chromosomes, two ancient microbes merged to form the eukaryotic cell, single cells gave rise to multicellular organisms, and in some instances, organisms formed eusocial societies (Maynard Smith & Szathmáry 1995).

Each transition produced a new kind of individual; a collective assembled from once-independent units (Buss 1987). These events established the nested hierarchy that characterises life: genes within chromosomes, organelles within cells, cells within organisms, and organisms within societies. At every level, entities vary, reproduce and transmit heritable traits, forming Darwinian populations. However, as higher-level individuals emerge, lower-level units relinquish much of their autonomy and align their fates with that of the collective (Godfrey-Smith 2009, Rainey & De Monte 2014, Black et al 2020).

The hierarchical structure of life arose through two principal kinds of transition (Queller 2000). Fraternal transitions collectivise related entities, as in the origin of multicellularity. Egalitarian transitions unite unrelated partners, as in the symbiotic origin of the eukaryotic cell. Both yield higher-level individuals with shared reproductive fates, but the evolutionary challenges differ: fraternal transitions require regulation of conflict among kin, whereas egalitarian transitions demand alignment of interests among partners with no common descent.

These two forms may not exhaust the possibilities. Lineages can also innovate from within. Cells fabricate membranes and organelles (Gabaldón 2010); organisms construct environments (Odling-Smee et al 2003); humans manufacture artefacts that extend cognition (Clark & Chalmers 1998). Such examples suggest a neglected route — autogenic transitions — in which a component generated internally within a lineage, rather than acquired from outside, becomes integrated into development and inherited as part of a new higher-level unit.

These differ from familiar cases of environmental modification. Internally generated artefacts, like externally constructed niches, reshape environments and alter selection (Odling-Smee et al. 2003; Dawkins 1982). Beaver dams, termite mounds and human cities influence evolution but are not inherited modules. Each must be rebuilt, and none form a parent–offspring lineage at the composite level. They shift selection but do not create new Darwinian populations.

The challenge is to specify the conditions under which internally generated components cross the threshold from transient phenotype to heritable parts of the developmental programme. This demands a return to first principles: reproduction, heredity and variation must occur at the composite level (Godfrey-Smith 2009). Only then can selection generate cumulative adaptation of the whole.

Growing interdependences between humans and AI provide a natural context in which to illuminate these ideas. Artificial intelligence is a human artefact, yet increasingly entwined with development, social organisation and reproduction. The question is not whether AI alters selection on humans, since it already does, but whether it could participate in a transition to a new, higher-level individual. To explore this, I examine biological precedents for autogenic innovation, cultural analogues and three possible routes to human–AI individuality.

**Biological precedents for autogenic innovation**

The idea of individuality arising “from within” may seem unfamiliar, yet biology offers several partial precedents. Lineages are not static; they routinely generate novel, heritable components internally, without recruiting external partners.

Gene duplications and de novo gene birth (Long et al 2003, Kaessmann 2010) create new heritable units from within lineages. Many such innovations have become indispensable, producing functions on which higher-level organisation depends. These cases, however, simply expand the trait repertoire of existing individuals and lack the semi-autonomy that characterises partners in classical ETIs.

Peroxisomes provide the clearest precedent at the organelle level (Gabaldón 2010). They self-assemble from nuclear-encoded proteins via the endomembrane system, lack DNA, and are faithfully inherited during cell division. Once established, they became essential to eukaryotic metabolism. Developmental inventions such as placental structures and maternal provisioning show similar internal origins that remodel inheritance systems and underpin complex multicellularity (Jablonka & Lamb 2005, Grosberg & Strathmann 2007).

Even the capacity to evolve can arise autogenically. Bacterial contingency loci — mutational hotspots that bias phenotypic outcomes — are internal mechanisms that generate predictable, heritable variation (Moxon et al 1994, Barnett et al 2025).

Together, these examples show that autogenic innovation is intrinsic to life. Lineages can generate novel heritable modules from within, some integral to higher-level success. None, however, exhibit distinct evolutionary interests or require mechanisms to manage conflict. They are instructive but incomplete analogues.

**Autogenic innovation in culture**

If biology provides hints, human culture offers clearer, though contested, examples. Human societies fabricate novelties from within, and some become indispensable and faithfully transmitted (Boyd & Richerson 1985).

Language is the prime case (Deacon 1997). It arose within human populations and is transmitted with remarkable reliability: every child acquires a language during development. Linguistic systems form cultural lineages in their own right and are now so integral that cognition, cooperation and society are unthinkable without them. Language thus functions as an autogenic innovation — a lineage-born system that reshaped inheritance and enabled higher-level organisation.

The analogy has limits. Language has no evolutionary interests apart from those of humans and persists only because humans do (Boyd & Richerson 1985, Deacon 1997). It is an internal module, not an autonomous partner.

Tools such as hammers or ploughs are also internally generated and culturally transmitted, but they too must be remade each generation. They alter selection on their makers but do not form a parent–offspring lineage of composites (Dawkins 1982).

Language remains exceptional: it is developmentally inherited, universal and foundational to social organisation (Tomasello 2019). However, it falls short of constituting a new individual. AI now occupies this same frontier. Like tools, it is fabricated; like language, it is developmentally acquired and potentially indispensable. The question is whether it could move further, becoming a semi-autonomous, heritable module that tips humans across the Darwinian threshold to form a new kind of entity.

**Three routes to human–AI individuality**

In asking whether human–AI associations could become new evolutionary individuals, three broad routes can be distinguished. Route 1 involves non-replicating AI systems that coordinate human activity and reshape selection but lack reproduction at the composite level. Route 2 envisages AI systems that themselves reproduce and evolve, creating the potential for classical symbioses with humans. Route 3 considers a synthetic endosymbiosis in which AI, though artifactual, becomes a developmentally inherited component of humans. In this last case the composite acquires a parent–offspring lineage that constitutes a new Darwinian individual.

**Route 1: Centralised or distributed, non-replicating AI**

The first route envisages powerful AI systems that coordinate or regulate human affairs yet do not reproduce in the classical Darwinian sense. They may be centralised or distributed across networks, but all share the property of persistence without descent. Such systems accumulate information, adjust internal parameters, and modify outputs through structured feedback from human users and their environment. These feedback processes generate adaptive change: the system retains acquired configurations that improve performance and transmits them through retraining and model updating. In this way, Lamarckian inheritance of acquired states produces cumulative modification within a lineage, even in the absence of reproduction.

Although non-Darwinian, such systems can nonetheless shape evolutionary trajectories. By mediating communication, decision-making, and resource allocation, they structure the selective environments in which humans evolve. Humans remain the Darwinian component—they replicate, vary, and inherit—but their evolutionary success may become increasingly coupled to integration with AI infrastructure. AI thus participates indirectly in evolution as a persistent, learning scaffold that channels selection.

Doolittle’s notion of persistence-based selection clarifies the distinction (Doolittle 2024). Different AI architectures may vary in stability or robustness, and those that endure longest exert the greatest influence on human futures. Yet persistence, even when coupled with Lamarckian modification, is not reproduction: it filters by durability and adaptability within a lineage but does not create lineages of composite traits that selection can refine cumulatively.

Nevertheless, a marginal form of transition is conceivable. If the persistence of a particular AI architecture becomes reliably tied to human reproduction—if societies or institutions ensure that human lineages reproduce only through engagement with specific AI systems—the composite could acquire a rudimentary heredity. Humans would remain the primary Darwinian substrate, yet their fitness would depend on continued coupling with an enduring technological partner. Such cases would blur the line between persistence and reproduction, representing a boundary instance of individuality consistent with the framework proposed by Rainey & Hochberg (2025).

Such systems may therefore transform human evolution through environmental and cultural scaffolding, and under special conditions could even edge toward a limited transition in individuality. Still, without autonomous reproduction of the composite, such systems remain outside the domain of full Darwinian individuality.

**Route 2: Replicating AI lineages**

The second route envisages AI systems that do more than persist: they reproduce, vary, and evolve as independent lineages. Unlike the singular or infrastructural AI of Route 1, these entities would have their own cycles of descent with modification. In such a scenario humans and AI could enter partnerships resembling classical egalitarian transitions, where initially independent lineages become interdependent.

This route has a clear Darwinian logic. Replicating AI populations would generate heritable variation in traits affecting success, and selection would refine their design. When coupled to humans, the reproductive fates of both partners could become aligned, at least when collaboration enhances the fitness of each. The analogy is to symbioses that evolve toward organelles: partners begin separate but develop mechanisms that stabilise their association (Margulis 1970, Doulcier et al 2020).

Challenges are equally familiar. Independent AI lineages may evolve divergent interests, generating conflict that demands policing, sanction, or partner-choice mechanisms to maintain alignment. Because AI is artifactual, boundaries between evolution and design blur, raising questions about what counts as reproduction in this context.

Nonetheless, Route 2 offers a plausible path to new individuality. If AI lineages evolve and humans become reliably coupled to them, the resulting composites could form genuine Darwinian populations. Selection could then drive the emergence of composite-level adaptations that transcend the capacities of either humans or AI alone.

**Route 3: Synthetic endosymbiosis (“from within”)**

The third route differs from both centralised AI and replicating AI lineages. Here, AI is fabricated by humans and becomes integrated as a developmentally inherited component. This possibility was first sketched in Rainey (2023) using a deliberately simple example: societal rules require that every person possess an AI device, and parents transfer it to their offspring. Although naive, the example illustrates the principle: the fabrication of an artefact that, through coupling to reproduction, becomes part of a heritable system. What matters is reliable parent–offspring transmission; this alone is sufficient to generate heritable variation in composite fitness.

This amounts to a synthetic endosymbiosis: an artefact produced within a lineage becomes a heritable module. The comparison is closer to an extension of the lineage’s own developmental system than to an invading symbiont. If the AI component becomes necessary for normal development—through neural embedding, identity-locked coupling, or institutionalised provisioning—the human–AI composite participates directly in evolution by natural selection as a unit of selection in its own right.

Extended phenotypes such as dams or tools must be rebuilt in each generation and thus lack heritable continuity. Synthetic endosymbioses, by contrast, can meet the Darwinian threshold. Variation among composites, reproduction of composites, and heredity of composite traits all become possible. Under these conditions, selection can sculpt adaptations at the level of the composite, and individuality shifts level.

Route 3 therefore represents the clearest case of an autogenic transition in individuality, showing how artefacts generated within a lineage can, under suitable conditions, cross the boundary from environment-shaping tools to components of a new Darwinian individual.

**Convergence: outside-in and inside-out**

Although Routes 2 and 3 begin from different starting points, they need not end in different places. In Route 2, independently reproducing AI lineages may, through repeated interaction and conflict mediation, become stably coupled with humans. In Route 3, AI is fabricated within the human lineage and scaffolded into inheritance from the outset. Yet in both cases the trajectory can converge: the human–AI composite emerges as a unit with its own reproductive cycle and heritable variation in traits.

The distinction lies in the path taken rather than the endpoint. One proceeds outside-in, through the incorporation of an autonomous partner; the other inside-out, through the hereditary embedding of an artefact. Both routes show how alignment of reproductive fates can generate a new level of Darwinian organisation. The parallel is instructive: individuality can arise through either the integration of external lineages or the internal assimilation of self-made components.

**Persistence or reproduction**

The three routes differ in the extent to which persistence and reproduction become coupled. Route 1 shows that persistence alone can stabilise systems and shape human evolution through feedback and environmental scaffolding. Such Lamarckian adaptation within a lineage allows cumulative modification but not lineage-level reproduction. Routes 2 and 3, by contrast, introduce true descent with modification at the composite level.

Doolittle’s distinction remains useful here (Doolittle 2024): persistence can yield ecological dominance, but only reproduction allows cumulative evolution of composite traits. A marginal case arises when persistence becomes reliably linked to human reproduction, as in the boundary examples of Route 1, yet this still depends on humans as the primary Darwinian substrate. In Routes 2 and 3, reproduction and heredity shift to the level of the composite itself, enabling selection to refine traits that benefit the whole.

The lesson is that persistence can scaffold, but reproduction completes, a transition in individuality. Centralised or distributed AI may direct the course of human evolution, but only when reproduction and heredity are reconfigured at the composite level does a new individual, capable of open-ended evolutionary change, emerge.

**Conclusions: why “from within” matters**

The analysis of human–AI relations is speculative, yet the issues are general (Rainey 2023, Rainey & Hochberg 2025). Evolutionary biology has mostly explained new levels of individuality through the merger of independent lineages (Maynard Smith & Szathmáry 1995), but transitions can also be scaffolded by a lineage’s internal capacity to generate and stabilise new developmental components (Black et al 2020). Developmental systems such as germ–soma separation, maternal provisioning and imprinting show how mechanisms that arise from within can structure inheritance and anchor individuality (Grosberg & Strathmann 2007). Even bacterial contingency loci, which bias mutation toward adaptive outcomes, demonstrate how autogenic innovations can create heritable modules that shape evolutionary potential.

Seen in this light, the “from within” pathway extends a continuum. Individuality is not simply a matter of parts fusing but of systems acquiring the inheritance architecture that allows selection to operate at a higher level (Godfrey-Smith 2009). What matters is not whether components originate externally or internally, but whether they generate a new parent–off-spring map.

For evolutionary biology, this reframing sharpens the criteria for what counts as an individual and clarifies why extended phenotypes, however consequential, fall short. For the future, it suggests that artefacts may enter the domain of evolution if they become heritable parts of developmental systems. Major transitions are thus not confined to the deep past: they remain open possibilities shaped by human agency. If individuality can arise from within, the boundary between natural and synthetic evolution becomes porous, demanding concepts expansive enough to encompass potential mergers between biology and technology (Jablonka & Lamb 2005).

**Acknowledgements**

PBR acknowledges generous core support from the Max Planck Society and Eugene Koonin for provoking a closer look at autogenic transitions.

**Declaration of interests**

The authors do not work for, advise, own shares in, or receive funds from any organisation that could benefit from this article, and have declared no affiliation other than their research organisations

**References**

Barnett M, Meister L, Rainey PB. (2025). Experimental evolution of evolvability. Science 387: eadr2756

Black A, Bourrat P, Rainey PB. (2020). Ecological scaffolding and the evolution of individuality. Nat Ecol Evol 4, 426–436.

Boyd R, Richerson PJ. (1985). Culture and the Evolutionary Process. Chicago: University of Chicago Press.

Buss L.W. 1987 The Evolution of Individuality. Princeton: Princeton University Press.

Clark, A, Chalmers, D. (1998). The extended mind. Analysis 58: 7-19.

Darwin C. (1859). On the Origin of Species. London, John Murray

Dawkins, R. (1982). The Extended Phenotype: The Long Reach of the Gene. Oxford: Oxford University Press.

Deacon TW. (1997) The Symbolic Species: The Co-evolution of Language and the Brain. New York, W. W. Norton.

Dobzhansky T. (1937) Genetics and the Origin of Species. New York, Columbia University Press.

Doolittle WF. (2024). Darwinizing Gaia: Natural Selection and Multispecies Community Evolution. Massachusetts: MIT Press.

Doulcier G, Lambert A, De Monte S, Rainey PB. (2020). Eco-evolutionary dynamics of nested Darwinian populations and the emergence of community-level heredity eLife 9: e53433

Fisher RA. (1930) The Genetical Theory of Natural Selection. Oxford, Clarendon Press.

Gabaldón T. (2010). Peroxisome diversity and evolution. Philos Trans R Soc Lond B Biol Sci: 365:765–773.

Godfrey-Smith P. (2009). Darwinian Populations and Natural Selection. Oxford: Oxford University Press.

Grosberg RK, Strathmann RR. (2007). The evolution of multicellularity: a minor major transition? Annu. Rev. Ecol. Evol. Syst. 38: 621–654.

Jablonka E, Lamb MJ. (2005) Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life. Cambridge MA, MIT Press.

Kaessmann H. (2010). Origins, evolution, and phenotypic impact of new genes. Genome Res 20:1313–1326.

Long M, Betrán E, Thornton K, Wang W. 2003. The origin of new genes: glimpses from the young and old. Nat Rev Genet 4:865–875.

Margulis l. (1970). Origin of Eukaryotic Cells. Connecticut: Yale University Press.

Maynard Smith J, Szathmáry E. 2020. The Major Transitions in Evolution. Oxford: Oxford University Press

Moxon ER, Rainey PB, Nowak MA, Lenski RE. (1994) Adaptive evolution of highly mutable loci in pathogenic bacteria. Curr. Biol. 4: 24–33.

Odling-Smee FJ, Laland KN, Feldman MW (2003). Niche Construction: The Neglected Process in Evolution. Princeton: Princeton University Press.

Okasha S. (2006). Evolution and the Levels of Selection. Oxford, Oxford University Press.

Rainey PB. (2023). Major evolutionary transitions in individuality between humans and AI. Philos. Trans. R. Soc. B 378: 20210408.

Rainey PB, De Monte S. (2014). Resolving conflicts during the evolutionary transition to multicellular life. Annu Rev Ecol Evol Syst 45: 599-620.

Rainey PB, Hochberg ME (2025). Could humans and AI become a new evolutionary individual? Proc Natl Acad Sci USA 122: e2509122122.

Tomasello M. (2019). Becoming Human: A Theory of Ontogeny. Cambridge: Harvard University Press.