Leveling with Tinbergen: Four levels simplified to causes and consequences

Classic contributions: Tinbergen (1963) On Aims and Methods of Ethology

Thore J. Bergman[1,2], Jacinta C. Beehner[1,3]

Department of Psychology, University of Michigan, Ann Arbor, MI 48109

Department of Ecology & Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109

Department of Anthropology, University of Michigan, Ann Arbor, MI 48109

Abstract

In 1963, Niko Tinbergen published his foundational manuscript identifying the four questions we ask in animal behavior – how does the behavior emerge across the lifespan (development); how does it work (mechanism); how and why did it evolve (evolution); and why is it adaptive (function). Tinbergen clarified that these 'levels of analysis' are complementary, not competing, thereby avoiding many fruitless scientific debates. However, the relationship among the four levels was never established. Here, we propose 'leveling' Tinbergen's questions to a temporal timescale divided into causes (encompassing mechanism, development, and evolution) and consequences (encompassing function). Scientific advances now seamlessly link evolution, development, and mechanism into a continuum of 'causes'. The cause-consequence distinction separates the processes that precede (and lead to) a behavior, from the processes that come after (and result from) a behavior. Even for past behaviors, the functional outcomes are (historical) consequences of the causes that preceded them.

1 Introduction

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/evan.21931

Nearly 60 years ago, Nikolaas ('Niko') Tinbergen identified four different ways to correctly answer behavioral questions in his classic manuscript, *On Aims and Methods of Ethology* [1]. Most students of animal behavior today have internalized the foundational framework he established - variously called Tinbergen's four questions or Tinbergen's four levels of analysis (**Figure 1**). These levels give equal honor to the pursuit of four different but complementary approaches to ethology - each with its own methods and specialized knowledge - making larger questions more tractable. The *mechanism* level examines the immediate causes for a behavior; these are the physiological processes (e.g., neural, muscular, hormonal) that allow an animal to express a behavior. These physiological processes are often triggered by external stimuli and contexts, such as social and ecological cues. The *development* level examines the ontogenetic changes that cause a behavior; these are often the same (physiological) processes as the mechanism level but at an earlier stage in the individual's lifetime, laying down the architecture (neural or otherwise) for supporting later behaviors. The *evolutionary history* level examines the trajectory of a behavior across many generations and how that behavior has changed or has been maintained throughout a phylogenetic lineage. The *function* level examines the fitness consequences of a behavior; these consequences are primarily 'seen' by evolution when they contribute to differences in survival or reproductive success. Tinbergen taught us that a full understanding of each instance of behavior requires more than one type of answer. Not only are the molecular geneticists and the field biologists each doing valid science, their work synergistically can be brought to bear on questions about animal behavior - allowing us to answer old questions in new ways.

INSERT FIGURE 1 HERE

At the time of Tinbergen's publication (1963), ethology was struggling to both define the scope of the field and to seek theoretical connections between disparate research approaches. Although Tinbergen was focused on explaining behavior, the four questions help biologists explain any phenotypic trait. Modifying Julian Huxley's three major problems for biology (causation, survival value, and evolution) [1] and sprinkling in a bit of Ernst Mayr's differentiation of *how* versus *why* questions [2], Tinbergen set a broad agenda for ethology, defining it as a science that spans timescales from milliseconds to millennia and physical scales from molecules to biomes.

Tinbergen's framework took root and has been the organizing structure for studies of animal behavior ever since. The rise of integrative biology is a testament to the power of this type of thinking. By clarifying that hypotheses at different levels are complementary (not competing), this classic manuscript has averted fruitless debates. For example, we cannot ask: *do chimpanzees eat fruit because they find the sweet taste rewarding or because it provides energy for survival?* These are not mutually exclusive explanations. An explanation at one level cannot exclude a different explanation at another level.

2 Problems and Solutions

2.1 First problem - There remains conceptual ambiguity between 'how' and 'why' questions.

Tinbergen's four levels have aged remarkably well. They remain very influential, appearing in (and often organizing) most animal behavior textbooks [3,4]. However, there continues to be

ongoing debate and confusion about the relationship *among* the four levels. Tinbergen did not attempt to organize them in his original publication, and this has left the topic open for debate. Most commonly, the levels are grouped as *proximate* (mechanism, ontogeny) and *ultimate* (function, evolution) explanations for behavior [5,6], which often are equated to 'how' (proximate) and 'why' (ultimate) questions. Additionally, the four levels are often grouped as *historical sequences* (the short-term sequence of development and the long-term sequence of evolution) versus a *slice-in-time* (the underlying mechanism or function at the time of the behavior)[7]. Recently, Sapolsky narrated a scenario where we can identify different 'causes' of a specific behavior by zooming in (to identify specific neurons firing) or zooming out (to identify early life developmental effects on the individual) allowing us to view the causes of behavior at different timescales. In this way, we are able to blur the line between what counts as developmental and what counts as mechanistic into one continuum [8]. Similarly, the utility of the proximate/ultimate dichotomy has repeatedly been questioned [9,10].

More problematic, researchers continue to confuse explanations at different levels; they contribute an explanation at one level for a question posed at another (we detail an example of this at the end). This occurs most commonly between 'how' and 'why' questions, with people giving a how-answer (mechanism) to a why-question (function). Consider the following question: *Why did the chicken cross the road?* Because her legs carried her? Or, because she had to get away from the farmer? Why questions can be answered correctly with both proximate and ultimate explanations, which means the how/why distinction is not overly helpful. It is this very ambiguity that adds humor to the why-do-chickens-cross-roads jokes. Sapolsky uses this question to open his popular book *Behave* [8], and he answers this with a

narrative of explanations that span from the evolutionary to the mechanistic -- a narrative that inspired us to reconsider the four levels in the first place.

2.2 Second problem - Scientific discovery has broken down the boundaries between questions. Innovative technology and scientific advances have eroded the temporal boundaries between the evolution, development, and mechanism levels. At the time of Tinbergen's publication, separating evolution from development was justified, but our current understanding is far more sophisticated. Discoveries in evolutionary-development and gene expression (e.g., epigenetics) have made it clear that within- and between-generational processes overlap. Similarly, at shorter time scales, our increasing temporal resolution for measuring physiology and the brain has made it increasingly difficult to separate developmental processes from more-immediate mechanisms that cause a behavior to occur. We know that gene expression [11], developmental plasticity [12], and social experience [13] can produce permanent and irreversible brain organization that 'cause' behaviors once animals are adults [14]. We also know that other forms of brain plasticity and hormonal regulation continue well into adulthood [15] in temporary and reversible ways making it again difficult to separate development from mechanism.

2.3 The Solution - A two-level framework comprising cause and consequence. In the absence of a conceptual structure from Tinbergen, and to refine, integrate, and extend conceptual arrangements suggested by others, we propose a simplified framework. Both the conceptual ambiguity and the break-down of temporal boundaries can be solved by moving to a two-level

framework surrounding any single instance of a behavior (**Figure 2**): cause (encompassing mechanism, development, and evolution) and consequence (function). The cause and consequence distinction neatly separates the processes that precede (and, therefore, can lead to) a specific behavior, from the processes that come after (and could possibly result from) the behavior. The moment the specific behavior occurs separates cause from consequence. We have leveled the four questions to a temporal timescale (before and after the behavioral event). Indeed, Tinbergen himself proposed almost this exact scheme:

> *I have always found it helpful to think of biology as concerned...with two problems; that of causation and that of function in the sense of survival value. By this I mean that...we ask "what makes this happen?" and "how do the effects of what happens influence survival (including reproduction)?" The first question can be roughly divided into three separate questions, differing in the time scale involved* [16].

These “three separate questions, differing in the time scale involved” are three of the four levels combined (evolution, development, mechanism) in the yellow-orange arrow of **Figure 2**. This was published in the author notes with a volume of Tinbergen's articles and seems to have largely been lost (although see Shettleworth [17]). It is unfortunate that this conceptualization never caught on, while the much less clear proximate-ultimate grouping did [18]. As we argue below, ‘proximate’ explanations for behaviors may actually not be so proximate after all, especially when we consider that gene regulation in one generation can be implicated in the behavior of their grand-offspring [19]. Rather than four separate time points (i.e., evolutionary

past, developmental past, immediately-preceding-the-behavior past, and the following-the-behavior future), we can now think about a temporal continuum allowing researchers to zoom in or out for any single instance of a behavior to study the causes and consequences at different time scales. Note that **Figure 2** depicts a behavior once it has already occurred with hypothetical causes and consequences that have emerged; the reader should keep in mind that the causes that contribute to (and the consequences that emerge from) any particular behavior are probabilistic (not deterministic) in nature.

INSERT FIGURE 2 HERE

3 Causes and Consequences

3.1 Causes. The two-level framework breaks down the barriers, not just between the development and mechanism levels but also between an individual's lifetime and evolutionary history (**Figure 2**). Evolutionary history is the broadest timescale for how things came to be. Although it is cross-generational, it maintains a connection from one generation to the next via direct genetic transmission. The entire causal chain connects generations via genetic inheritance, connects neighboring generations via epigenetics [20], connects early life experiences to reactivity as an adult via developmental plasticity [21], connects something that happened that morning to a hormone state later that afternoon via regulatory changes [22], connects a hormone state that afternoon to a sensory neuron being more likely to fire [23], and so on - until we reach the shortest timescale (on the order of milliseconds) connecting a motor neuron firing as a chicken dodges a vehicle (**Figure 2**). This connection across timescales makes categorical

158 thinking obsolete. Our explanation for what caused the chicken to cross the road will simply
159 depend on how far back in time we, as scientists, are willing to look. A cause at one less-
160 proximate point in time is itself directly linked to a cause at a more-proximate point in time.
161
162 **3.2 Consequences**. The other side of the behavior - the aftermath - is a bit more difficult to
163 grasp conceptually (**Figure 2**, blue arrow). When we think about consequences, we are
164 concerned with the effect of the trait, usually in terms of how it relates to survival and
165 reproduction. Logically, consequences come after the behavior. And, they do. With current
166 behavior, the current consequences can be measured and analyzed in a relatively
167 straightforward manner. This is the primary level of inquiry in the field of animal behavior and is
168 captured in Tinbergen's 'function' level of analysis (**Figure 1**).
169 But, very often we are interested in how natural selection *in the past* shaped the trait
170 that we are observing now. This is *not* the process of documenting the sequence of
171 evolutionary steps that led to the trait, which is a simple, historical process that belongs
172 squarely on the *cause* side of the analysis. By contrast, if we are to understand how natural
173 selection operated in the past, we need to understand how the trait previously affected survival
174 and reproduction. This belongs squarely on the *consequence* side of the analysis. When a trait
175 evolves by natural selection, this means that, in previous generations, the trait had a net fitness
176 benefit, *a consequence of the trait at that point in time.* One of Tinbergen's greatest mistakes
177 was not making the distinction clear between these two processes within his *evolution* level.
178 Tinbergen (1963, p. 428) described evolution as encompassing both the evolutionary history
179 (how animals got their forms) *and* the selection that shaped the trait (why evolution proceeded

180 the way it did). In short, Tinbergen wanted his evolution level to 'do double-duty',
181 simultaneously calling upon this level to answer both cause and consequence questions about
182 the past. Mayr [2] attempts to reconcile this with his term, 'ultimate causation', to refer to (what
183 we call) the *historical consequences* of past selection. Ultimate causation is a problematic
184 concept that conflates past consequences with past causes [24].
185 To understand why this is problematic, an analogy might be useful. Imagine that
186 selection is a series of sieves with different-sized openings in the mesh. Each 'selection' sieve
187 only allows stones to move on to the next sieve if they can pass through the openings in the
188 previous one. Starting with a jumble of different-sized stones, the process ends up with a
189 relatively homogenous pile of similar-sized stones in each pile. If you reach into one of those
190 piles and grab a stone, you could ask: *Why is this stone this size?* (**Figure 3**). Just as with all 'why'
191 questions about animal behavior, this can be answered in two logically distinct ways; but one is
192 a 'how' answer and one is a 'why' answer. The 'how' answer satisfies the question: *How did this*
193 *stone come to have this particular size?* (**Figure 3A**). For this answer, we need to know the
194 history of the stone and the breakage and erosion events that caused it to arrive at the size it
195 has. This is a causal question about each individual stone that asks what events directly led to
196 the stone having its current form. This is equivalent to Tinbergen's how-animals-got-their-form
197 side of the evolution level. In this case, the filtering process is irrelevant. Alternatively, the 'why'
198 answer satisfies the question: *Why did I grab a stone of this size?* (**Figure 3B**). For this answer,
199 we need to know the process of stone-sorting that happened in the past. This is a population-
200 level sampling question that depends entirely on the size of the holes in the sieve and the
201 sorting process. The output of the sorting process is a consequence of each stone's size. This is

equivalent to Tinbergen's why-evolution-proceeded-the-way-it-did side of the evolution level. Keep in mind, this process tells us nothing about how each stone came to have its size. Natural selection (or drift, or any evolutionary filter) does not cause individual traits any more than the size of the holes in a sieve create the sizes of the stones that pass through them. It is possible for "filters" to change the things that pass through them (like potato ricers) where the shape of the trait is caused by the filter, but we know that is not how natural selection operates. In biology, variation is caused by only two generative processes - *mutation* and *recombination* - which can variably be expressed through epigenetic processes that we simplify here as *gene expression*. Importantly, the *consequence* of the sorting (the traits that remain after population-level sorting) are not the *cause* for the trait. The sieve as an analogy helps us understand the process of selection, but it is imperfect in that it suggests a teleological process with a particular goal (size of stone) as the outcome. Natural selection is a filter but not a goal-directed one; it emerges from an ever-changing environment.

INSERT FIGURE 3 HERE

Evolution by natural selection is an iterative process. The variation that is present in one generation results, in part, from natural selection in the previous generation. However, as soon as we start describing processes in this way, we have moved from individual-level thinking to population-level thinking [10]. We have started to answer the *why-did-I-grab-a-stone-of-this-size* question, which involves looking at the prior consequences of the behavior (i.e., the *historical consequences,* **Figure 3B)**. Although current consequences of behaviors are the stuff of most

224 animal behavior manuscripts, historical consequences can also be studied but only *indirectly*
225 using comparative methods to identify the selective pressures that shaped the trait in the past
226 [25]. Historical consequences of a behavior *could* be the same as current consequences (e.g., a
227 history of directional selection that continues in the present), but they may also be entirely
228 different (e.g., a history of directional selection followed by stabilizing selection). Historical and
229 current consequences are two separate, independent questions. The comparative method is a
230 powerful tool, but it does not identify the 'ultimate causation' of a behavior, it identifies the
231 historical consequences.

232 Finally, it is critical to keep in mind that the fitness consequences of a behavior are not
233 always beneficial. Traits can be neutral - or even detrimental - to fitness. Another benefit of
234 using the term 'consequences' (rather than the weighted terms 'survival value', 'function', or
235 'current utility') is that the valence of this term leaves open the possibility for positive, neutral,
236 or negative effects.

237

238 **3.3 Consequence ripples.** Generally, the environment (context, social, ecological, physical) is
239 what determines the fitness consequences of a behavior. But, behaviors can also alter the
240 environment in ways that change future fitness consequences. This bi-directional relationship
241 was termed 'reciprocal causation' and can be seen in processes such as niche construction,
242 coevolution, habitat selection, and cultural evolution [9]. Certainly, the effect of behavior on the
243 environment is an important part of the story. However, the term 'reciprocal causation' has the
244 same problems as the term 'ultimate causation' - mainly, it is not causal at all [24]. Phenotypes
245 can and do cause changes in the environment, which then alter the selective pressures on

themselves (and other phenotypes); but this is very different from selection *causing* a phenotype. Rather, consequences that change the environment for future behaviors may be thought of as *consequence ripples* that extend forward in time. Consequences are often ephemeral and are essentially reset with each instance of a behavior. For example, having escaped a lion yesterday has little bearing on your chances of escaping a different lion today. However, consequences can persist and alter future consequences. For example, the nest you build today continues to provide nest-related benefits in the future, even for other individuals and future generations. In sum, current behaviors change the future consequences of other behaviors (they do not cause those behaviors).

3.4 Cause-driven and consequence-driven traits. To further understand the difference between cause and consequence in the evolution of a trait, it helps to understand the two extremes - in what we will call *cause-driven traits* versus *consequence-driven traits*. Consider an evolutionarily stable trait. This trait does not change across large fluctuations in the environment, and it does not respond to selection. In such cases, causal processes either (1) constrain variation in a trait, so there is no raw material for selection to act on [26], or (2) yoke variation in a trait to negative consequences of other traits, so any potential benefits of variations in the trait are invisible to selection [27]. In short, this trait has low evolvability [26]. For example, all primates have four limbs. There is no variation in this trait (four limbs), and it appears to be highly constrained across primates. Such traits are *cause-driven* because the trait we see today is driven by the process of producing the trait (**Figure 3A**) rather than the process of sorting it (selection). Note that cause-driven traits still have consequences (just as

consequence-driven traits still have causes, see below). To measure these consequences, however, we would need to be able to isolate and compare variants of the trait (which is often impossible for the same reasons that natural selection cannot act on the trait). Five limbs might be better than four, but we have no way of studying this.

By contrast, a trait with high-evolvability - heavily shaped by natural selection with a high degree of variation that maps onto the environment - is a *consequence-driven trait*. The process that shaped the trait (the cause) takes a backseat to understanding the exact form of the trait we see today. Instead, because there are so many variants to sort (e.g., **Figure 3B)**, it is the sorting process (selection) that is largely responsible for the traits that persist. For example (returning to the limbs of primates), while the *number* of limbs is invariable, the *proportions* of limb lengths across primates varies considerably. These limb proportions covary with the locomotor style and habitat use of different species in ways that are adaptive [28]. The consequences from having different limb proportions in the past contributed to which limb proportions we see today. For consequence-driven traits with a high degree of variability, this is where we need to address our adaptive hypotheses. The extent to which phenotypes are shaped by selection (consequence-driven) remains an open question, and this is an active area of debate in evolutionary thinking [29].

4 Human sex differences example

We end with an example of how this framework can clarify confusion about answers from different levels of analysis. A recent analysis, from this journal, addressed the question, *why are there sex differences in human stature?* [30]. The author, Dunsworth, focused analyses on

both stature and pelvic shape, but for simplicity, we focus on the stature question because the logic is the same. Dunsworth states that sexual selection explanations for sex differences in human stature (e.g., that male competition favors the evolution of larger men, e.g., Puts [31]) have been over-emphasized in the story of human evolution. Instead, they propose that sex differences in human stature are due to differential estrogen secretion (because estrogens fuse the epiphyses of long bones):

> *For humans and likely other hominids, male skeletons continue to grow after females' stop because their bodies take longer to produce enough estradiol to surpass the amount that stimulates continued growth and to achieve a level that closes long bone epiphyses* [30], p. 111).

They additionally state that the estrogens explanation means that "the sexual selection perspective on male height seems unnecessary" [30] (p.110). Two published responses have already disputed the logic of this approach, [32,33] saying that explanations for sex differences in human stature in terms of estrogens and sexual selection are not mutually exclusive but are answers to different questions [33]. In short, support for a mechanism explanation cannot reject a functional one.

The cause-consequence framework can help clarify this debate in two ways. First, the framework highlights the temporal relationship between a cause and a consequence making it clear why one can never be substituted for the other. The pattern of estrogens secretion *precedes*, and is therefore a potential cause of, adult stature (i.e., it stops further growth). By contrast, sexual selection *follows*, and is therefore a potential consequence of, the preexisting

stature. Once the adult stature is achieved and the phenotype is active in the environment, the process of how that phenotype came about (the cause) is largely invisible to the selection process (that yields the consequences). Any selection acting on the length of a giraffe's neck does not 'care' if the neck is long because it has *extra* vertebrae or because it has *longer* vertebrae. Any selection acting on sex differences in body size does not 'care' if men are taller than women because of differential estrogen secretion or (hypothetically) growth hormone secretion. Even if we are able to reject the growth hormone hypothesis for why men are taller than women, this does not make the sexual selection hypothesis any more (or less) likely. The hormonal causes are entirely orthogonal to testing the consequences of differential growth. Dunsworth does make the important point that phenotypes do not always have an adaptive explanation [30]. Certainly, traits are not always adaptive. They could emerge simply as a byproduct of another trait [34] or by chance [35]. But, such traits still produce consequences - adaptive, neutral, or detrimental to fitness. This then raises the question, if traits are not driven or maintained by natural selection, how do they persist over evolutionary time?

In cases where selection is unable to act on a trait, we consider these *cause-driven* traits (e.g., the four limbs present in all primates). A cause-driven phenotype is likely what Dunsworth [30] is arguing for the estrogens explanation for sex differences in human stature. This would mean that human stature is largely a product of constraints (e.g., relating to reproduction and estrogen secretion) rather than selective consequences (e.g., relating to sexual selection). Like the four limbs in primates, a cause-driven hypothesis predicts that stature dimorphism will show little variation from humans to apes to monkeys. Comparative data do not, however, support this prediction. Size dimorphism is immensely variable both within humans and across

primates. Across primates, females are larger than males in some species and males more than three times the size of females in others, and these differences closely map onto different social and mating systems [33,36]. Contrary to a cause-driven hypothesis, these data suggest that differences in body size (across primates, and even across vertebrates) are enormously plastic and what we would consider to be *consequence-driven,* with very high evolvability.

Although support for the estrogens hypothesis explaining differences in human stature cannot be used to reject the sexual selection hypothesis, the high evolvability in primate body size dimorphism *actually supports* Dunsworth's primary claim that sexual selection plays a reduced role in the recent history of humans. Indeed, other authors have successfully argued using comparative datasets that sexual selection, if anything, is very much relaxed in humans compared to other closely related primates [36].

5 Conclusion

Sapolsky warns us against categorical thinking, what he calls 'thinking in bins'. We wholeheartedly agree - not just because thinking across bin boundaries is necessary for integrative science; but also because now that we understand so much more about the processes that contribute to behavioral outcomes, the bins themselves are confusing. Moving away from conceptual bins to a temporal continuum is more compatible with our current understanding of integrative biology. A temporal continuum makes a fundamental distinction between processes that precede (and could cause) a phenotype and processes that come after (and could be consequences of) a phenotype. Given the iterative nature of natural selection, this distinction is particularly important. Natural selection links cause and consequence because

current fitness consequences determine which 'causes' persist into the future. Despite this link, population-level sorting processes (consequences) remain logically distinct from individual-level determinants (causes). Tinbergen himself said that there are only two problems in biology - that of causation and that of function [16]. Therefore, we recognize that a more appropriate title for our manuscript might have been "Leveling *along with* Tinbergen…", since we are simply advocating what he first championed almost 50 years ago.

Acknowledgements

We thank our graduate school professors (Garland Allen, Jim Cheverud, Cliff Jolly, Allan Larson, Jonathan Losos, Jane Phillips-Conroy, Alan Templeton) for giving us a strong foundation in evolutionary biology; and we thank our postdoctoral mentors (Jeanne Altmann, Dorothy Cheney, and Robert Seyfarth) for giving us a strong foundation in animal behavior.

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Author Biographies

THORE J BERGMAN is a Professor in the Department of Psychology and in the Department of Ecology and Evolutionary Biology at the University of Michigan. He is broadly interested the evolution of cognition and communication, specifically as they relate to social knowledge and evolutionary fitness. He founded and currently directs two field sites focused on wild primates:

the Simien Mountains Gelada Research Project in Ethiopia (studying geladas) and the Capuchins at Taboga Research Project in Costa Rica (studying white-faced capuchins).

JACINTA C BEEHNER is a Professor in the Department of Psychology and in the Department of Anthropology at the University of Michigan. She is broadly interested in hormones and behavior, specifically as they relate to reproductive success. She founded and currently directs two field sites focused on wild primates: the Simien Mountains Gelada Research Project in Ethiopia (studying geladas) and the Capuchins at Taboga Research Project in Costa Rica (studying white-faced capuchins). She also directs two hormone laboratories – one at the University of Michigan (Beehner Endocrine Laboratory) and one at the Capuchins at Taboga field site (TREX Endocrine Laboratory).

References

1 Tinbergen N. 1963. On aims and methods of Ethology. Z Tierpsychol Wiley. 20:410–433.

2 Mayr E. 1961. Cause and effect in biology. Science 134:1501–1506.

3 Rubenstein DR, Alcock J. 2018. Animal Behavior: An Evolutionary Approach. Sinauer Inc., Sunderland, Massachusetts.

4 Dugatkin LA. 2020. Principles of Animal Behavior, 4th Edition. University of Chicago Press.

5 Klopfer PH et al. 1972. Function and Evolution of Behavior. Mass., Addison-Wesley.

6 Alcock J. 1975. Animal behavior: An Evolutionary Approach. Sinauer Associates Sunderland.

7 Nesse RM. 2013. Tinbergen's four questions, organized: a response to Bateson and Laland. Trends Ecol. Evol. p 681–682.

8 Sapolsky RM. 2017. Behave: The Biology of Humans at Our Best and Worst. Penguin.

9 Laland KN et al. 2011. Cause and effect in biology revisited: is Mayr's proximate-ultimate dichotomy still useful? Science 334:1512–1516.

10 Ariew A. 2003. Ernst Mayr's "ultimate/proximate" distinction reconsidered and reconstructed. Biol Philos Springer Nature. 18:553–565.

11 Arnold AP, Chen X. 2009. What does the "four core genotypes" mouse model tell us about sex differences in the brain and other tissues? Front Neuroendocrinol 30:1–9.

12 Nettle D, Bateson M. 2015. Adaptive developmental plasticity: what is it, how can we recognize it and when can it evolve? Proc R Soc Lond B Biol Sci 282:20151005.

13 Champagne FA. 2012. Interplay between social experiences and the genome: epigenetic consequences for behavior. Adv Genet 77:33–57.

14 Phoenix CH et al. 1959. Organizing action of prenatally administered testosterone propionate on the tissues mediating mating behavior in the female guinea pig. Endocrinology 65:369–382.

15 Goldman SA, Nottebohm F. 1983. Neuronal production, migration, and differentiation in a vocal control nucleus of the adult female canary brain. Proc Natl Acad Sci U S A 80:2390–2394.

16 Tinbergen N. 1972. The Animal in Its World: Explorations of an Ethologist, 1932-1972.

418 Harvard University Press.

419 **17** Shettleworth SJ. 1974. Function, causation, evolution, and development of behavior: A
420 review of The Animal in Its World, by N. Tinbergen. Journal of the Experimental Analysis of
421 Behavior. p 581–590.

422 **18** Sherman PW. 1988. The levels of analysis. Anim Behav 36:616–619.

423 **19** D'Urso A, Brickner JH. 2014. Mechanisms of epigenetic memory. Trends Genet 30:230–236.

424 **20** Zambrano E et al. 2005. Sex differences in transgenerational alterations of growth and
425 metabolism in progeny (F2) of female offspring (F1) of rats fed a low protein diet during
426 pregnancy and lactation. J Physiol Wiley. 566:225–236.

427 **21** Ader R, Grota LJ. 1969. Effects of early experience on adrenocortical reactivity. Physiol Behav
428 4:303–305.

429 **22** Oyegbile TO, Marler CA. 2005. Winning fights elevates testosterone levels in California mice
430 and enhances future ability to win fights. Horm Behav 48:259–267.

431 **23** Joëls M, de Kloet ER. 1992. Control of neuronal excitability by corticosteroid hormones.
432 Trends Neurosci 15:25–30.

433 **24** Dewsbury DA. 1999. The proximate and the ultimate: past, present, and future. Behav
434 Processes 46:189–199.

435 **25** Martins EP. 2000. Adaptation and the comparative method. Tren Ecol Evol. p 296–299.

436 **26** Kirschner M, Gerhart J. 1998. Evolvability. Proc Natl Acad Sci U S A 95:8420–8427.

437 **27** Futuyma DJ. 2010. Evolutionary constraint and ecological consequences. Evolution 64:1865–
438 1884.

439 **28** Jungers WL. 1985. Body size and scaling of limb proportions in primates. In: Jungers WL,
440 editor. Size and Scaling in Primate Biology. Boston, MA: Springer US. p 345–381.

441 **29** Shanahan T. 2017. Selfish genes and lucky breaks: Richard Dawkins' and Stephen Jay Gould's
442 divergent Darwinian agendas. The Darwinian Tradition in Context. Cham: Springer International
443 Publishing. p 11–36.

444 **30** Dunsworth HM. 2020. Expanding the evolutionary explanations for sex differences in the
445 human skeleton. Evol Anthropol 29:108–116.

446 **31** Puts D. 2016. Human sexual selection. Curr Opin Psychol 7:28–32.

447 **32** Galipaud M, Kokko H. 2020. Adaptation and plasticity in life-history theory: How to derive
448 predictions. Evol Hum Behav 41:493–501.

449 **33** Font E, Carazo P. 2021. False dichotomies and human sexual size dimorphism: A comment of
450 Dunsworth (2020). Evol Hum Behav. p 176–178.

451 **34** Paaby AB, Rockman MV. 2013. The many faces of pleiotropy. Trends Genet 29:66–73.

452 **35** Gould SJ, Lewontin RC. 1979. The spandrels of San Marco and the Panglossian paradigm: a
453 critique of the adaptationist programme. Proc R Soc Lond B Biol Sci 205:581–598.

454 **36** Lindenfors P, Tullberg BS. 1998. Phylogenetic analyses of primate size evolution: the
455 consequences of sexual selection. Biol J Linn Soc Lond Oxford University Press (OUP). 64:413–
456 447.

457

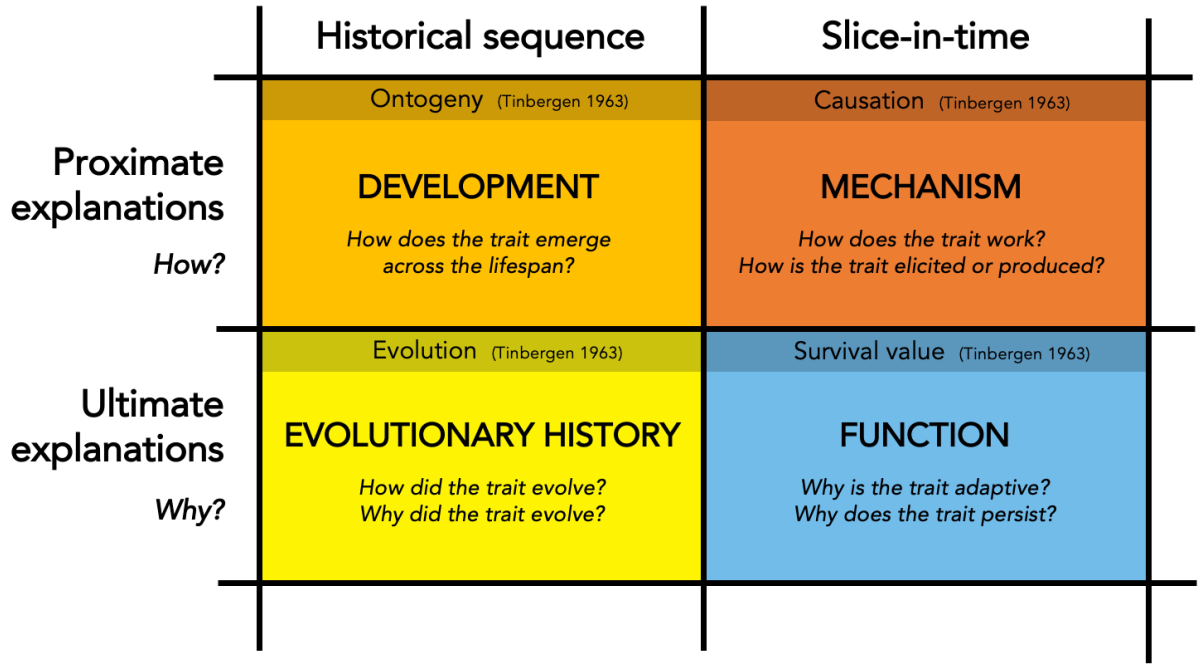

Figure 1. Tinbergen's four levels of analysis, modified from Nesse (2013).

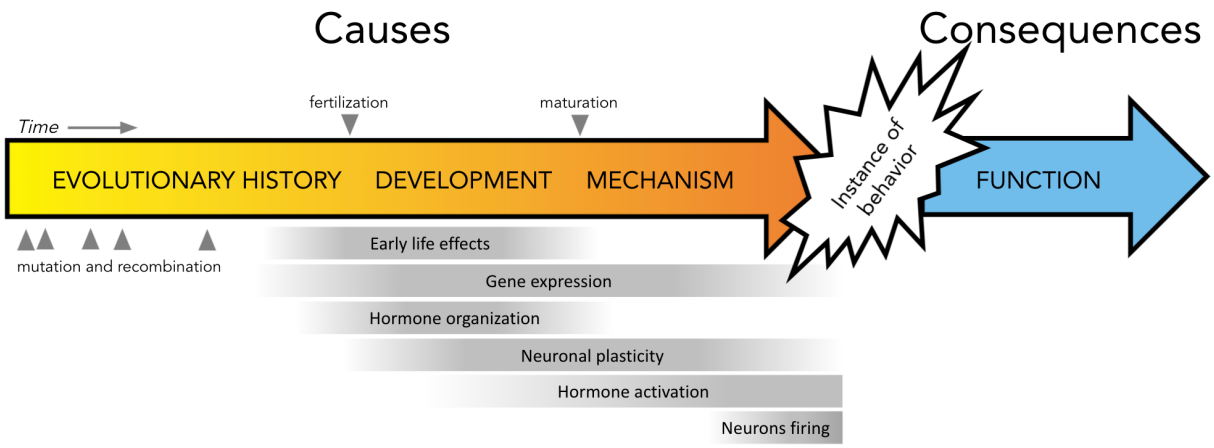

Figure 2. Tinbergen's four levels, reduced to a single temporal continuum separated into the processes that precede (and, therefore, can lead to) a single instance of a behavior, and those that come after (and, could possibly, result from) this particular behavior. Examples of causal processes and their approximate timescales are in grey below the arrows.

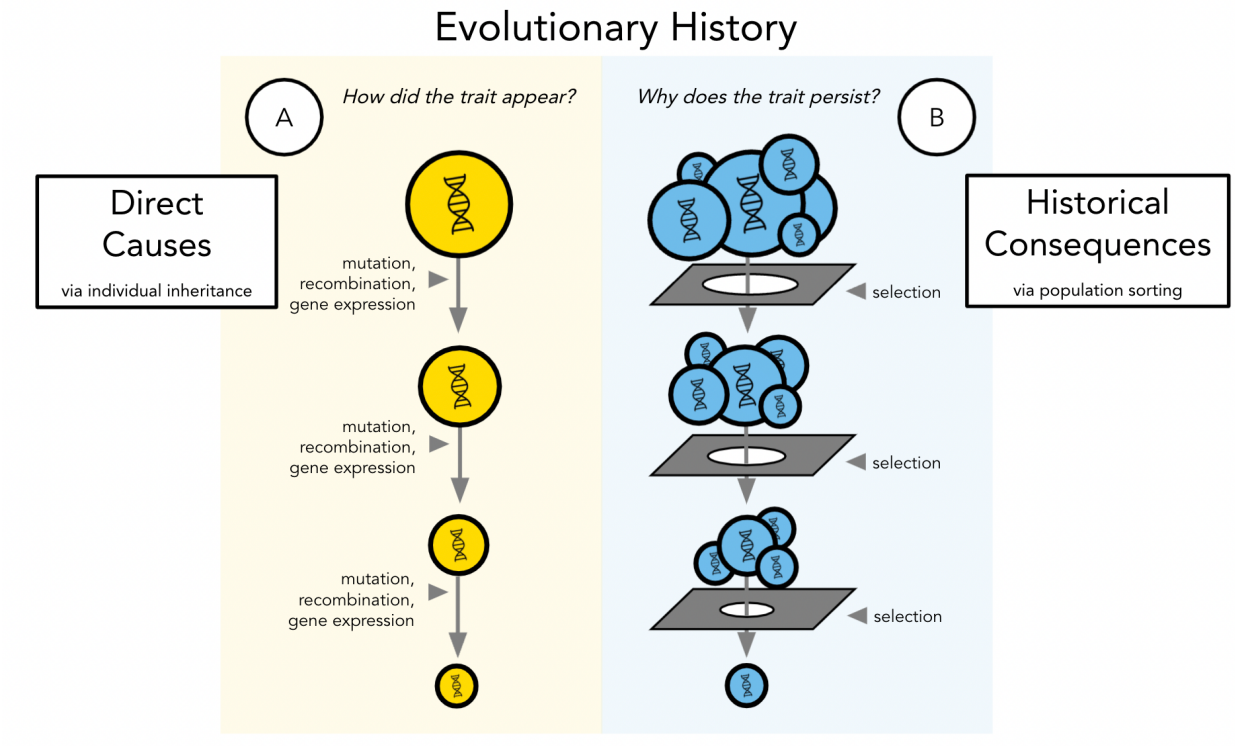

Figure 3. Tinbergen did not distinguish between cause and consequence processes with his *evolution* level. The left side indicates a change in phenotype due to the causal mechanisms of mutation, recombination, differential gene expression, and inheritance. The right side indicates a change in phenotypic frequency due to the consequences of selection.