17 The End of Human Behavioral Ecology

Richard McElreath and Jeremy Koster

17.1 Introduction

Edited volumes tend toward triumphalism. This one is no different. Human behavioral ecology has enjoyed a lot of success, both in terms of professional growth and scientific results. This success is more remarkable when we recognize that the field has no clear disciplinary home. Whereas many human behavioral ecologists are anthropologists, others are scattered across psychology and economics and biology departments. Some self-congratulation is deserved.

But at the time of this volume, the behavioral and biological sciences are in crisis. As a consequence of failures to replicate highly cited results (Camerer et al. 2018), the impossibility of verifying analyses due to lost data and materials (Minocher et al. 2021), and even widespread fraud (Bik et al. 2016), a variety of professional organizations, funding agencies, and governments have called for methodological reform (McNutt 2014; Munafò et al. 2017). A former editor-in-chief of *The Lancet* summarized the situation this way:

> The case against science is straightforward: much of the scientific literature, perhaps half, may simply be untrue. Afflicted by studies with small sample sizes, tiny effects, invalid exploratory analyses, and flagrant conflicts of interest, together with an obsession for pursuing fashionable trends of dubious importance, science has taken a turn towards darkness. [. . .] And individual scientists, including their most senior leaders, do little to alter a research culture that occasionally veers close to misconduct. (Horton 2015)

It is reasonable to disagree about the contributing causes of the crisis, but it is no longer reasonable to deny that we are in one. It would be hubris for human behavioral ecology to think it is immune to these systemic problems. Some introspection is merited. Accordingly, this essay outlines some short-term strategies that may help human behavioral ecology (HBE) and cognate fields to achieve their long-term scientific goals.

In order to decide what to do next, we should decide where we want the field to go in the long run. The current state of our field is partly explained by decisions a small number of researchers made in the early and middle twentieth century. That is not the whole story. But scientific fields are culturally evolved institutions. Their practice embodies their origins, and their origins strongly influence their practice. How we practice our field now influences where we go next. How can we deliberately steer the field in better directions, rather than waking up each day and shortsightedly chasing another grant or publication? (See Boxes 17.1 and 17.2.)

Box 17.1 | **Two Zombie Papers**

The purpose of this chapter is to examine the general conditions and development of the field of human behavioral ecology. We do not focus on particular papers. However, there are two papers which have been published often enough now that they deserve special consideration.[1] These papers contribute little to cumulative science but are difficult to eradicate.

The first paper is entitled something like "Social Networks, Reproductive Success, and the Evolution of Human Sociality." After an introduction that promotes the banal claim that social relationships and fertility may influence one another, the number of self-reported friends is shown to be associated with past fertility. The analysis adjusts for a basket of perfunctorily measured demographic variables, and every coefficient is interpreted as a total causal effect. The paper claims that this finding demonstrates the value of an evolutionary approach to human behavior, mainly because economists have ignored it. But since no evolutionary model predicted the result, the size and direction of the association have no impact on any evolutionary model. Any evolutionarily relevant variable like relatedness or age at first birth can be deployed in this paper to imitate, but not actually to perform, behavioral ecology. The paper is highly cited.

The second appears under the banner "The Evolution of Human Cooperation: A Bayesian Phylogenetic Analysis of Kinship and Pudding Recipes." A cross-cultural sample of stereotyped features of human societies is fed into software designed for the analysis of biological species. The procedure identifies an association between uxorilocal residence and the prominence of dessert. No model predicts this association nor its size. But rejecting a null model of no association somehow illustrates the value of an evolutionary approach. A few fatal confounds and methodological flaws flirt with the reader in the twilight paragraphs. The result has implications for the origins of human society and the possibility of world peace.

We have likewise written these papers, but we are trying to stop. If we could all avoid producing, reviewing, and reading these papers ever again, we would divert a substantial amount of talent toward productive research.

The authors of these papers could make a contribution by instead advancing an optimization model that resembles the models that are described in other chapters. Then the model's predictions could be made algorithmically and contrasted with similarly precise alternatives. Data collection and analytical procedures could be designed that have some hope of causal identification. But at least the assumptions that justify a causal interpretation would be clear. Making zombies does not train us to perform any of these tasks. Our students and their students deserve better.

[1] This box is an homage to Elias (1958) and can be read as a lament of the structural considerations that incentivize overstated inferences from limited data and rote modeling.

https://doi.org/10.1017/9781108377911.018 Published online by Cambridge University Press

Box 17.2 **The Human Behavioral Ecology of Academic Research: Status**

A key premise of human behavioral ecology is that individuals respond adaptively to their environments, including the cultural norms and institutions that shape the costs and benefits associated with different strategies. This theoretical assumption applies equally to all individuals and populations, regardless of nationality, ethnicity, gender, age, and other characteristics. With that in mind, it is helpful to consider how academics, including human behavioral ecologists, are incentivized to navigate the landscape of academic research.

Often working as members of academic departments, human behavioral ecologists are expected to conduct high-quality research that contributes to the cumulative advance of scientific knowledge. Assessing the quality of research at the creative edge of knowledge is difficult, however. It is a noisy process. Peer review has a subjective element; bias can intrude. Absent objective measures and pressed for time, evaluators sometimes fall back on expedient heuristics, status being an especially common one. That the process often works reasonably well does not absolve us from acknowledging and addressing its imperfections.

Status is a position in a social hierarchy that accumulates from social esteem, respect, and repeated acts of deference (Chapter 7; see also Sauder et al. 2012). In academic contexts, researchers can show deference via citations of other scholars' published work (Merton 1988). Accordingly, research is frequently assessed via citation-based metrics despite the broadly recognized limitations and ethical issues associated with these measures (Raff 2013; Chapman et al. 2019). Moreover, the peer review system is rife with deference to status, resulting in a Matthew Effect that inordinately rewards high-status scientists and their institutions (Merton 1968; Huber et al. 2022).[2]

This Matthew Effect is troubling given evidence that status and quality are imperfectly correlated (Sauder et al. 2012). If rewards disproportionately accrue to individuals with high-status mentors, institutional affiliations, and demographic characteristics, then high-quality research by individuals not sharing these profiles is disadvantaged (Lynn et al. 2009; Hofstra et al. 2020). Jockeying for status can quickly cease to be healthy competition and instead contribute to incentives for attention-grabbing rather than reliable and reproducible studies (Smaldino and McElreath 2016; Fraser et al. 2018). Ethical publishing choices can be overlooked or disregarded (Logan 2017). In short, the Matthew Effect can harm scientists and hinder scientific progress.

[2] The Matthew Effect derives its name from the biblical passage in the book of Matthew: "For unto every one that hath shall be given, and he shall have abundance: but from him that hath not shall be taken away even that which he hath" (Matthew 25:29; as cited in Merton 1968).

Behavioral ecologists can choose to guard against deference to status by reading carefully and widely, attending to sources outside the usual venues, and keeping in mind their own personal biases when making decisions that affect the fair appraisal of colleagues, students, and others in the academic community. More broadly, given the importance of institutional norms in academic research settings, we all have an obligation to support institutional reforms that incentivize high-quality, reproducible studies. We are optimistic that human behavioral ecologists will adjust to those incentives, improving the science.

Many behavioral ecologists will know an optimization technique called *stochastic dynamic programming*. The problem facing animals in their environments is not to optimize each decision, independent of all other decisions. Rather the problem is to optimize a sequence of decisions, usually under substantial uncertainty. Stochastic dynamic programming is a way of finding optimal conditional strategies in such circumstances (Houston et al. 1988). The key insight of the approach is that the optimal strategy is not found by thinking forwards about what to do next. Rather it is found by working backwards from the end goal.

Imagine the field many years in the future, when many of its original goals have been achieved. The field is not over. Instead, it enjoys successful integration with other basic and applied sciences, integration made possible through unique scientific contributions. What did researchers in the field do in the decade before reaching this end? And in the decade before that? And so on, all the way back to the present. Accordingly, the first step is not to decide what to do tomorrow. It is to decide what to do the day before the end. One must then ask: what is the end of human behavioral ecology?

In this rest of this essay, we develop a future-oriented, backward-looking view of the field of HBE. As an interdisciplinary field with an orientation toward formal theory and difficult, naturalistic investigation, HBE stands to make unique contributions to both basic and applied research. The best way to realize these contributions is to work backwards from one or more ends, developing pathways of investigation and training to guide us. We do not believe that this essay resolves these questions. Instead, it foregrounds them, presents credible reforms, and aims to defy the myopic path-dependency and substitution of scientific goals with individuals' professional rewards that have contributed to the contemporary crisis of the sciences.

17.2 The Ends

The goal of HBE is to understand the evolutionary origins, mechanisms, and dynamics of human behavioral adaptation, especially in the context of variable

https://doi.org/10.1017/9781108377911.018 Published online by Cambridge University Press

and culturally evolved environments. For the field to succeed, it must present a justified explanation for how, during the course of human evolution, human behavioral flexibility co-evolved with our life history traits, cognitive development, and capacity for culture. It must, together with other fields, contribute to the development of credible causal models of variation and change not only in individual behavior but also social institutions and technology over the course of human history.

Equally important is the future. Already, a majority of humans live in cities. The United Nations estimates that 70% of people will live in cities by the middle of the twenty-first century. An alien biologist observing us would logically conclude that we are evolved to build and live in cities. Therefore, another proper end of HBE is to contribute to an applied understanding of how our origins shape our present and future. This contribution might be accomplished through scientific insights. But it could be realized also through tools and approaches that applied research may adopt.

Many fields study human behavior, and many fields will make contributions to these goals. How can HBE make a unique and valuable contribution? Looking back from the ends of HBE, the research that maximizes its contribution is long-term, individual-based field research. Yet, before long-term studies can realize their potential, the field requires sustained investments in a culture of professional research design, data management, and data analysis. In the next sections, we outline the contributions of long-term research and the training necessary to make it sustainable and scientifically productive.

17.3 Long-Term, Individual-Based Research

In the early days of anthropological fieldwork, progress was made through discovery and systematic comparison. Fieldworkers encountered and documented substantial diversity in kinship and descent, political organization, religions, livelihood strategies, and so forth. The pace of discoveries eventually slowed, and HBE has largely transitioned out of seeking such discoveries and instead is busy trying to make sense of the diversity. Yet, much like anthropologists in the early twentieth century, the careers of human behavioral ecologists generally continue to feature a series of short-term studies on distinct topics.

Long-term individual-based field research provides advantages over both short-term and population-level research. By long-term research, we mean research that studies the same local population over many years. By individual-based research, we mean designs that collect data on discrete individuals and their social lives (Box 17.3).

There are accentuated structural and ethical obstacles to conducting long-term individual-based field studies on humans. From the auspicious ends of HBE, addressing these obstacles is just as important as the scientific advantages of the research. We discuss the advantages and obstacles in turn.

Box 17.3 **Longitudinal Research Designs and the Prospects for Long-Term, Individual-Based Studies**

Figure B17.3.1 depicts common research designs, including longitudinal designs that are commonly used in the social sciences (Gravlee et al. 2009). In general, cross-sectional studies have predominated in human behavioral ecology.

There are noteworthy exceptions to that generalization. Using a repeated cross-sectional design, for example, Urlacher and Kramer (2018) study separate cohorts of Yucatec Maya children to examine changes in physical activity level and anthropometric outcomes between 1992 and 2012. Among the Hadza of Tanzania, Pollom et al. (2020) use a similar research design to examine changes in juvenile foraging over time.

Meanwhile, because people can report on past events in interviews, it is often possible to compile datasets for retrospective panel studies. Respondents may be able to provide reliable data on significant events, such as the timing of their marriages and divorces, residence histories, and the births and deaths of their children and other family members. These data can then be organized into a unit-period format with entries for previous intervals of time that are suitable for an event history analysis (Sheppard et al. 2014; Blurton Jones 2016).

There are limitations to this approach, however, when unrecoverable time-varying variables are an important part of the data-generating process. As an

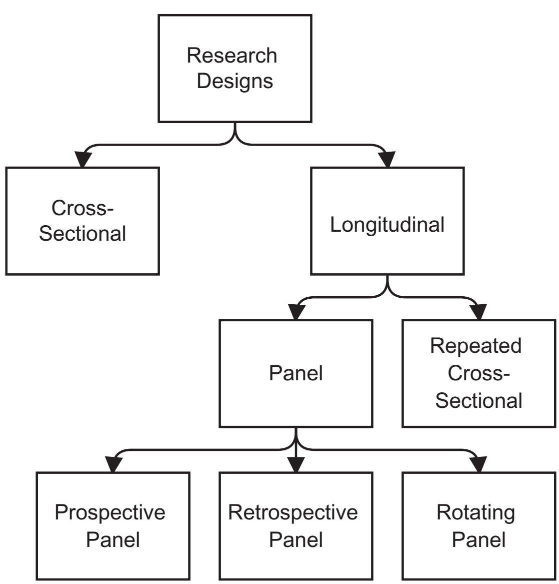

Figure B17.3.1 Common research designs in the social and ecological sciences. Adapted from Gravlee et al. (2009).

https://doi.org/10.1017/9781108377911.018 Published online by Cambridge University Press

example, consider the study of divorce by Winking and Koster (2021), which suggests that the risk of divorce decreases as married couples have more children together. This finding is consistent with behavioral ecology theories of serial mating and shared reproductive interests (Chapter 9).

However, the event history analysis does not include the effect of the couples' material wealth on divorce. This omission is problematic given the evidence from other studies that (1) financial problems increase the probability of divorces (Snopkowski 2016), and (2) fertility varies as a function of the wealth that is available to the married couple (Borgerhoff Mulder 1992a). If both of those relationships are evident in a particular study population, then an observed effect of couples' fertility on divorce might represent a spurious correlation. That is, if the statistical model were to include wealth as a predictor, then the effect of fertility might be negligible. In other words, the analysis of Winking and Koster (2021) potentially suffers from an omitted variable bias.

Although it is often possible to elicit measures of wealth at the time of data collection, in many study contexts it is unrealistic to expect respondents to report reliably on their annual wealth in the preceding years and decades (Bernard et al. 1984). In the absence of longitudinal data, researchers might consider using a measure of wealth collected at time t as a static predictor of the outcomes in preceding intervals $(t-1, t-2, \text{etc.})$. This approach, however, rests on the dubious assumption that wealth does not fluctuate over time. Overall, this example helps to illustrate the limitations of retrospective panel studies (see also Boxes 2.1 and 13.2).

There are variables that can be reliably collected only via prospective panel studies, which entail multiple waves of data collection on the same individuals or units. For human behavioral ecologists, key time-varying variables might include subsistence outcomes, short-term behavioral strategies, measures of knowledge and beliefs, reputations, emotional states, and biomarkers. For robust inferences about the dynamic processes that interest human behavioral ecologists, often there is no substitute for careful data collection, repeated over time. Accordingly, prospective panel designs are an ideal option for the long-term, individual-based studies that are needed to answer longstanding research questions.

17.3.1 Advantages

We focus here on five advantages of long-term individual-based field studies: (1) the possibility to study age-related changes, (2) the ability to model causes of events in different life stages, (3) the study of dyadic and higher-order relationships and their impacts, (4) the possibility of measuring life span variables like reproductive success and economic and cultural contributions, and (5) the measurement of behavioral change both within and between generations. These advantages are shared with

https://doi.org/10.1017/9781108377911.018 Published online by Cambridge University Press

long-term studies of nonhuman animals (Clutton-Brock and Sheldon 2010). However, the longer lives and complex, dynamic phenotypes of humans increase the marginal value of sustained, individual-level data collection. We will not address any of these advantages thoroughly. Each is worth an entire chapter on its own. Also, archival data can have the same advantages as field research, which is worth keeping in mind throughout.

Individual behavior is highly dynamic, and one important dimension of these dynamics is age. From an orthodox evolutionary ecology perspective, humans are age-structured organisms. Our long lives create overlapping generations, and individuals at different ages have different physiology and behavior. There are also distinct, culturally constructed stages of life in all human societies (Brown 1991). Studying the origins and evolution of human life history means studying both age-related changes and stage-related transformations. Short-term fieldwork is limited by its mostly cross-sectional nature. Many variables, potentially including the timing of key events, cannot be elicited through retrospective interviews. Finally, the study of stage-related changes will require longer studies simply because changes in stage happen more rarely than changes in age and because sustained cultural competence may be necessary to recognize and accurately describe local stages.

To study and explain human behavioral adaptation, we want to model causes of individual change. For humans, this often means modeling causal relationships between the events in individual lives in earlier ages and stages to events in later ages and stages. Only long-term individual-based research can capture the necessary evidence. It is always necessary to makes strong causal assumptions to get causal inference out of observational research. But the strength and plausibility of the assumptions are greatly reduced in long-term designs. A rigorous mathematical framework exists in evolutionary demography for structuring and modeling causes of phenotypic change (Ozgul et al. 2009). Our field should adopt it and modify it to include cultural processes (Beheim and Baldini 2012).

Much of the classic theory in evolutionary ecology uses a *mean-field approximation*. This means that social forces are averaged across individuals and their social relationships. For example, cooperation is studied by considering an average interaction partner or an average migration context. For short-lived organisms, this is a very powerful approach because relationships do not last long and take on their own causal histories. But it is an approximation, and for long-lived organisms like humans, it can be a poor one. When dyadic and higher-order relationships can last for decades, averaging across social contexts produces misleading estimates of costs and benefits (Hauert and Szabó 2005). Studying causal forces like these requires long-term data on specific individuals and their interaction partners.

Many of the theoretically important quantities in HBE are life span variables or lineage variables that require generational and even intergenerational observations for estimation. Life span variables include reproductive success, number of spouses, income, and capital accumulation. Cultural analogs of these traditional variables include number of innovations, number of languages learned, number of artworks produced, as examples. Intergenerational processes such as lineage growth, change in any of life span quantities, and transmission of wealth and knowledge between generations are similarly central to HBE. With strong assumptions about the absence of cohort effects, these quantities can be estimated from short-term and cross-sectional data. However, because the necessary strong assumptions cannot be tested, long-term research is really the only way to observe the events that allow for proper description of life span and intergenerational processes.

17.3.2 Obstacles

There is a sense in which all of these assertion are obvious. Of course, long-term, individual-based research is essential for studying the processes of human behavioral adaptation.

Despite obvious benefits, long-term studies are underused because there are substantial obstacles to planning and maintaining them. Some of the most substantial difficulties are (1) the short-term focus of funding for basic research, (2) the stochasticity of research funding, (3) the ethical and practical obligations to cooperate with and transfer benefits to research communities, (4) a lack of expertise in and training for research design and data management for long-term, individual-based designs. We address the last of these in detail in the next major section. We briefly address the others here.

Even at long-term field sites, project designs do not typically center long-term research. Instead, many research projects promise to be able to shed new light on important questions using a few months of data collection. And some important work can be done that way. Yet, contemporary projects tend to promise that short-term data can address long-term questions about adaptation in part because they must advance that perspective in order to win competitive grants. And competitive grants rarely cover more than a few years of support.

In principle, iterated grant proposals could support longer-term research designs. Some successful long-term projects have done just that. But proposals that promise to continue doing what was done in the previous funding period compete poorly in a research culture that values innovation and rapid results. This results in shifting priorities, dropped measurements, loss of continuity, and poor data management.

For the ends of HBE, these structural obstacles must be overcome. Funding agencies and research institutions could choose to support and reward long-term designs and evaluate progress toward long-term scientific goals rather than short-term publication and citation impact. Some projects can support themselves with crowd-funding, especially since necessary maintenance funding for HBE is often modest. All of these problems are shared with other fields (Clutton-Brock and Sheldon 2010). Accordingly, solutions do not have to come only from HBE initiatives, but energy must be allocated toward reforming the support and incentive environment.

Perhaps the most important component of successful long-term field research is cooperation with and ethical involvement of the communities that provide data (Broesch et al. 2020; Urassa et al. 2021). There are ethical obligations but also potential advantages of increased community involvement in research. Long-term projects cannot succeed without the long-term consent of the participating community. And in most cases, this requires some sharing of rewards, either indirectly or directly through meaningful involvement of local communities in research design. Trained community participation can itself be an asset to data collection and the cumulative improvement in research designs.

More substantial are problems in the integration of research design with data management and analysis.

17.3.3 Training

As seen throughout this book, classic papers in human behavioral ecology and cultural evolution used mathematical models, solved through optimization or game theoretic dynamics, to make non-null predictions about behavior.

HBE is a highly quantitative field, but it has not succeeded yet in developing professional norms of data analysis or data management. In this regard, it is like most quantitative sciences. Yet, that should not stop us from acknowledging the field's limitations and charting a path toward improvements.

In this section we briefly outline two major projects in reforming our field's approach to quantitative inference. The first is the development of professional norms of data and code management. The second is the importance of returning to the field's roots and placing formal scientific causal models in charge of data analysis.

To some extent, it may seem unrealistic or onerous to add more items to the already long list of skills that human behavioral ecologists are expected to master. For instance, consider a graduate student who is conducting dissertation research on an HBE topic in an international setting. Typically, the student is expected to develop a theoretically novel research question, independently secure grant funding for the fieldwork, navigate potentially complex bureaucracies to obtain permits for research, gain proficiency in another language, adapt to residing and working in a different cultural context, collect and organize high-quality data, analyze the data with complex statistical methods, and exercise superb technical writing skills while publishing the findings in peer-reviewed journals. This diverse skillset has few parallels in other academic disciplines, and it is a genuinely masterful scholar who meets all of these expectations.[3]

Given the benefits of improved data management and analysis, however, there are compelling motivations to support the dissemination of these skills. Fortunately, the broader scientific community develops tools and innovations that can be adapted to the HBE workflow relatively easily. Mentors of early-career researchers can facilitate the acquisition of valuable skills by researching and supporting relevant opportunities for extracurricular training. The pace of scientific progress is swift. Few graduate programs can maintain the faculty expertise to provide the needed skills via the traditional reliance on coursework and mentorship. The following overview mentions

[3] There is a case to be made, of course, that expecting individual researchers to embody all of these skills is detrimental to scientific progress. Diverse teams of specialized researchers who can leverage their comparative advantages often produce better science than individuals who attempt to implement the necessary skills independently (as in the subsistence contexts described in Chapter 6). Human behavioral ecologists are often housed in academic departments and disciplines, however, that have been slow to recognize the benefits of collaboration and to rethink its evaluations of scholars' contributions to scientific teams.

tools and methods to consider, but better alternatives may soon appear. The broader point is that progress in HBE will accelerate when the quality of data management and analysis rivals the quality of the theorizing and data collection.

17.3.4 Data Management

Especially for long-term research projects, HBE will benefit from greater emphasis on data management and design. Too many projects begin without any formal training in longitudinal data management. This results in awkward and error-prone databases. Problems in data management are also problems in data collection. Seemingly simple tasks like maintaining persistent identities of individuals can be deceptively challenging. Integrated solutions that design for these problems at both time of data collection and database integration exist, but they are not well-known or commonly utilized. Additional problems arise in long-term research because eventually key personnel retire, but the data must remain intelligible to new personnel. Too many HBE projects have grown with databases designed for and understood by only one or two people, which places all of the data at risk. HBE can quickly improve by modifying solutions from allied fields like development economics and global health, in which longitudinal databases with distributed users are more common. A growing professional culture of data "carpentry" is also helping, with many students acquiring relevant data-science skills and data-design principles through professional workshop curricula (Teal et al. 2015). The problems of data management remain substantial, but HBE does not have to solve them on its own. It merely has to participate.

Transparent sharing of data and analytical code is a core principle of open science (Munafò et al. 2017). Without that transparency, skepticism of published findings may be warranted given the potential for questionable research practices, which can range from fraudulent manipulations of data to seemingly minor undisclosed decisions about outlying data points and statistical modeling. For individual researchers, however, sharing data and code might seem perilous. When analysis and data can be scrutinized by others, mistakes may be revealed, potentially leading to retractions (e.g., Whitehouse et al. 2019, 2021; Beheim et al. 2021). In terms of the collective advancement of science, though, such scrutiny is unambiguously helpful. A retracted finding is preferable to unnoticed errors in the literature. Institutionalized standards are needed that reward researchers who make their data and code available.

It is important, however, to consider the ethical dimensions of data sharing. Whereas research with human subjects always requires thoughtfulness, the work done by human behavioral ecologists may need extra consideration (Kraft et al. 2020). Fieldwork is often conducted among Indigenous, marginalized, or socioeconomically disadvantaged populations. These societies may exhibit familial, marital, or reproductive norms that depart from the conventions and laws of countries in which they reside. In some cases, there is potential for published data to enable discriminatory treatment by authorities. Amid ongoing conversations about best practices for data management and sharing, it is important for human behavioral ecologists to weigh local contexts alongside general principles (e.g., Carroll et al. 2021). For particularly sensitive data, it may be necessary to consider repositories that curate access to the data, limiting access only to researchers who pledge to maintain the anonymity of the data. Reasonable embargo periods that limit access until time has elapsed are another method to mitigate unethical repurposing of data. With these alternatives, there are few scenarios in which HBE datasets cannot be made available to others.

17.3.5 Causal Inference

The study of behavioral adaptation is not just the description of behavioral and ecological change. It also requires some way of inferring causes of behavioral change.

The preferred method of causal inference in many sciences is the controlled, randomized experiment. However randomized experiments are not, and never have been, the dominant style of causal inference in the human sciences. One reason is that many of the most important questions about human behavior cannot be studied experimentally. Ultimately, behavior must be studied in the natural environments that people construct because those environments are both causes and consequences of behavior. Without studying people in their own communities, we would not even know what we are trying to explain. But equally important are the ethical limitations on experimental intervention that would be required to study, for example, the influence of family structure on behavior.

Causal inference is clearly possible in observational settings, and statisticians have developed specialized methods for designing analyses that address stated causal estimands, the targets of inference (Pearl et al. 2016). Conventionally, few HBE studies have made use of these methods, often relying instead on a mixture of predictive model selection and interpretation of every coefficient as an unmediated total effect (McElreath 2020).

It has become clear, however, that statistical models are not enough. What is required is a conceptual model, distinct from any statistical models, that embeds the assumptions about relationships between variables in ways that enable predictions about the consequences of interventions (Pearl et al. 2016).

When the only model a researcher uses is a statistical model, there is no principled way to justify the structure of the model and no logical way to interpret its output. Not only will the estimate of interest be uninterpretable, but so will all of the other coefficients from the model. It is never justified to simply include control variables in a multiple regression and interpret every coefficient equivalently as a direct causal effect. Ignoring the overall causal structure when interpreting coefficients is known as the "Table 2 Fallacy" (Westreich and Greenland 2013). The fallacy is unfortunately very common, even in fields important to public welfare (Westreich et al. 2021).

A core challenge for HBE, therefore, is to make scientific model construction an integral part of student training, research design, and peer criticism. Of course, more detailed scientific models are needed as well. But models at different levels of abstraction bring different benefits. Abstract models that can be represented as simple diagrams excel at communication while still having clear implications for measurement. More detailed causal models, based for example on dynamical systems models or agent-based simulations, are harder to communicate but make more detailed predictions and are therefore more powerful in measurement.

To understand why, consider the example of the role of ecological knowledge in foraging production. To prosecute this example, we will use a type of causal diagram known as a directed acyclic graph (DAG). DAGs are graphical causal models that communicate abstract causal assumptions and allow us to logically deduce estimation strategies. See Box 17.4 for general information about DAGs if you are unfamiliar with them.

Box 17.4 **Causal Inference with Directed Acyclic Graphs**

The basis of formal causal inference is to specify one or more generative models of the phenomenon and then to deduce, or to deduce the impossibility of deducing, the consequences of intervening (altering) one or more variables in the model. One problem is that most working scientists do not have complete generative models of the systems they study. But even when they do not have a complete generative model, a lot of logical work can be accomplished. Indeed, it is dangerous to attempt statistical inference without thinking about causality, even if the goal is mere description.

A popular and axiomatic approach is the use of directed acyclic graphs (DAGs) to represent and communicate assumptions about causal relationships.

As an illustration, consider these three schematic generative models, drawn as DAGs:

$$Z \leftarrow X \rightarrow Y \qquad X \rightarrow Y \rightarrow Z \qquad X \rightarrow Z \leftarrow Y$$

In each DAG, the letters are measured variables and the arrows represent causal influence. The scientific goal is to estimate the influence of X on Y, indicated by $X \rightarrow Y$. In each case, a third variable Z is also available. The question is, should you include Z in your analysis? A contemporary norm in HBE is to include Z in a regression analysis and either interpret it as a control variable or use a criterion like AIC to justify its inclusion. However, in each of the three models, including Z has very different consequences.

In the first example, on the left, X influences both Z and Y. If we include Z in a regression of Y on X, we lose precision on the estimate of the influence of X on Y. But it does not systematically distort the estimate. In the second and third examples, however, including Z is a mistake. In the middle model, including Z creates a form of selection bias and can induce large systematic error in the inference of the influence of X on Y. In the third model, X does not influence Y at all. But when we include Z in the analysis, it can lead to a very strong association between X and Y via a collider bias. Furthermore, although HBE rarely obtains experimental control of a variable like X, even when X is randomly assigned, including Z in a regression could ruin inference. Even experiments are not safe. These claims can be proved algebraically (Pearl 1995; Pearl et al. 2016; Cinelli et al. 2022) or demonstrated using simulation (McElreath 2020).

https://doi.org/10.1017/9781108377911.018 Published online by Cambridge University Press

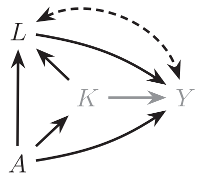

Figure 17.1 This Directed Acyclic Graph (DAG) depicts the hypothesized causal relationships among foraging returns Y, ecological knowledge K, age A, and labor allocations to foraging L. In this example, the dashed path between labor allocations and foraging returns is representative of unmeasured confounds. In contrast to the black arrows, the gray arrow indicates that the target of inference, also known as the estimand, is the effect of knowledge on foraging returns. Generally, for a given estimand, the structure of a DAG allows researchers to determine which variables should be added or omitted in a statistical model and how coefficients should be interpreted. Also, a DAG can constructively inform research designs by illustrating which variables should be measured during data collection.

We adopt for the sake of example the causal DAG in Figure 17.1. In this example, the estimand of interest is the causal influence of knowledge K on foraging returns Y, highlighted in gray. However, there are other variables that influence both K and Y and make the task of estimation more difficult than simply regressing returns on knowledge, even with controls. In this example, there are two competing causes of foraging returns, age A and labor L allocated to foraging. Age A influences foraging returns Y through unmeasured mediating causes like physical condition, represented by the path $A \rightarrow Y$. But age A also influences knowledge K, since older individuals have had more time to learn. And age A also influences the labor L allocated to foraging. Knowledge K itself influences labor L, since more knowledgeable individuals allocate labor differently.[4] Finally, the dashed path on top between L and Y represents unmeasured confounds, such as illness, that influence both labor and foraging success. For example, when people are ill, they will forage less and be less successful when they do forage.

It is easy to imagine other paths in this diagram, or alternative diagrams. But the point is not to argue that this represents the true relationships among the variables. Rather the point is to argue that meaningful research in the evolutionary behavioral sciences requires some causal model and must be able to deduce from such a model how to use data to estimate an estimand of interest. How would you use observational data on Y, K, A, and L to estimate $K \rightarrow Y$? Should you control for any of the other variables? If so, how?

The value of an explicit causal model is that it provides a logical, objective way to answer these questions. In fact, a computer can do it upon being taught the causal model. But for simple models like this one, it is simple enough that a computer is not needed. In this example, in order to estimate the direct causal effect of K on Y, the

[4] For this simple illustrative example, the model assumes relationships only at a particular moment in time. It is possible to extend DAGs to a longitudinal framework. For example, knowledge K could influence labor L at time t, and then the experience gained at time t could enhance knowledge at time $t + 1$. These dynamics can be incorporated into a DAG with additional time-varying nodes, such as future knowledge K_{t+1}. When long-term, individual-based datasets permit longitudinal analyses, expanded DAGs with time-varying nodes can help researchers to fit sensible and robust statistical models.

https://doi.org/10.1017/9781108377911.018 Published online by Cambridge University Press

Figure 17.2 Replicating the causal structure of the DAG in Figure 17.1, the variations in the gray arrows indicate two important inferential considerations, the treatment of confounders (left) and distinctions between total effects and direct effects (right). In the left-hand panel, the inclusion of age A in the statistical model remedies the confounding effect of age on the causal relationship between knowledge K and foraging returns Y. In the right-hand panel, the total effect of knowledge on foraging returns includes the direct effect $K \to Y$ and the indirect effect that is mediated by labor allocations $K \to L \to Y$. A statistical model that does not include labor L as a predictor variable permits the estimation of the total effect of K. As noted in the text, however, the unmeasured confounds of L and Y preclude unbiased estimates of the direct effect of K on Y.

gray arrow, we need our statistical machinery to do two things. First, we need to remove any noncausal sources of association between K and Y. In this example, age A is a common cause of both K and Y. The observed association between K and Y is partly a result of shared variation stemming from A. That is, age A is a confound. Second, we need to separate the direct effect of K from its indirect effects, if any. Any observed association between K and Y is a combination of its direct and indirect effects. In this example, there is an indirect effect of K on Y that is mediated by L.

Let us redraw the diagram in Figure 17.2, highlighting each of these problems. Then let us consider, in nontechnical terms, how to accomplish our objectives.

The first problem is to remove the confounding influence of A. This is accomplished by stratifying by A, estimating the association between K and Y for each level of A. Once variation in A is statistically "controlled" this way, the average association between K and Y is free of the confounding influence of A. In ordinary linear models, the way to accomplish this control is by including A as a predictor variable. But it is important to justify its inclusion in the model, as well as how it is included. It is dangerous to simply add variables and watch how coefficients change.

The second problem is separating the direct and indirect effects of K. After stratifying by A, according to the causal diagram, the association between K and Y measures the total causal influence of K on Y through both direct and indirect paths. How can we isolate the direct effect? It turns out that we cannot. Similar to removing the association due to A, we might also stratify by labor L, examining the association between K and Y for each level of L. This statistically "controls" for the influence of labor. However, in this example, L and Y are confounded by unobserved influences, as indicated by the dashed path connecting them. This means we cannot estimate the causal influence of L on Y. This is no problem for estimating the total causal effect of K on Y, because we do not need to measure $L \to Y$ to estimate the total effect of K. But in order to isolate the direct effect of K, we do need to estimate $L \to Y$. So according to this causal model, there is no way to estimate that direct effect. By contrast, if we were willing to assume that there are no unobserved confounds of L and Y, then stratifying by both L and A would be sufficient to estimate the direct effect of K. In this respect, devices like DAGs are valuable for communicating assumptions, which can then be considered and critiqued by others.

In general, the language of causal inference has not characterized the initial decades of HBE research. Until recently, distinctions between "total effects" and "direct effects" and "indirect effects" and "mediators" have not been common in the literature, nor have biases such as collider biases been considered regularly. To the contrary, the overarching concern has been the possibility of an omitted variable bias, a preoccupation that can lead to unwieldy statistical models when researchers consider nearly any available variable to be a possible confounder. It is now clear, however, that adding control variables in this way can lead to spurious or misguided inferences. It is important that we teach and learn to recognize that there are both "good controls" and "bad controls" and that distinguishing between the two requires a full causal model (Cinelli et al. 2022). There is no substitute for theory when fitting and interpreting statistical models.

None of these points are novel nor controversial in the statistical community. But only recently have relevant methods started to become part of HBE's standard training. Formerly, standard statistical practice in HBE, as reflected in many of the field's most highly cited papers, was based on mistaken understandings of the relationship between scientific and statistical models. As a result, much of the research published in the last few decades contains a tangle of estimates of debatable value. We must tackle this both by stopping to analyze data in similar ways and by revisiting older publications and critically evaluating their results in light of explicit, theoretically informed scientific models.

Fortunately, to demystify the logic of DAGs and causal modeling, software tools are available to facilitate analyses (e.g., Textor et al. 2016). To use these tools, researchers supply the DAG that represents their view of the scientific model. The software can then generate the recommended statistical model for estimating the effect of interest. With moderate effort, therefore, researchers can leverage their theoretical expertise to advance statistical models that align convincingly with key research questions. Overall, these methods enable an approach in which a scientific model asserts a data-generating process, and then statistical models are used to examine the extent to which the asserted scientific model is evident in the empirical data.[5]

17.3.5 Methods Inform Theory and Research Design

Our overview of methodological training has been deliberately pointed, focusing on areas for improvement. Still, there is also progress to be celebrated, as multilevel

[5] It is important to reiterate that multiple statistical models with heterogeneous combinations of variables may be needed when researchers aim to make inferences about different aspects of the causal structure for a particular data-generating process (Westreich and Greenland 2013). One variable's confounder is the other variable's mediator. Not all regression coefficients can be interpreted equivalently.

https://doi.org/10.1017/9781108377911.018 Published online by Cambridge University Press

modeling and DAGs and other methodological advances have been enthusiastically adopted by rising cohorts of human behavioral ecologists (e.g., Hubbard et al. 2022; Pretelli et al. 2022; Hillemann et al. 2023). Future advances will be embraced, too. Relative to other fields, HBE is generally welcoming of methodological advances, and there are abundant opportunities for early-career researchers to pioneer new and better methods. That is fortunate.

On the other hand, substantial time has elapsed since Winterhalder and Smith (2000) observed that advances in HBE theory had outpaced the quality and rigor of empirical analyses. Their observation remains salient, particularly in cases when causal reasoning and statistical modeling are post hoc considerations after data have already been collected. The ideal scenario is for researchers to anticipate the causal structure and statistical models that will be needed for convincing inferences so that the research project can be designed from the outset to collect the necessary data.

New methodological tools can spur enhanced research designs and theorizing. Human behavioral ecologists who invest in learning analytical methods for longitudinal data can deduce how these approaches will be pivotal for answering fundamental research questions. This foresight, in turn, provides indelible motivation for surmounting the structural obstacles to long-term, individual-based studies.

17.4 Conclusion

Compared to many scientific fields, human behavioral ecology has a strong theoretical foundation. As seen throughout this volume, evolutionary principles and the logic of optimization provide a coherent framework for explaining phenotypic variation. This framework is intrinsically causal in its orientation. That is, the variation in individual-level outcomes is posited to arise from variation in key predictor variables.

Causal reasoning is necessarily chronological. In other words, a predictor X that is observed at time t leads to an outcome Y at time $t + 1$. Among relatively long-lived humans, these causal processes plausibly unfold over long intervals, potentially decades in some cases. Cross-sectional studies can provide valuable glimpses into these processes, but they are often limited by possible endogeneity biases, such as omitted variable biases and concerns about reverse causality. Long-term, individual-based studies are needed for stronger tests of HBE theories.

Institutions and peers can help to minimize the barriers to long-term, individual-based studies. Employers and funding agencies can provide the time and resources for researchers to conduct fieldwork at regular intervals. Assessments of early-career scholars can encompass not only published works but also their progress toward the compilation of compelling long-term datasets. Mentors can encourage training in data carpentry and longitudinal statistical analysis that enable advances, as seen in other scientific disciplines. Commitments to open science and ethical research and publishing choices can be rewarded.

Human behavioral ecologists are uniquely positioned to contribute holistic and incredibly detailed longitudinal studies of individuals in diverse societies. Those studies will contribute to the end of human behavioral ecology, which is attainable with sustained investments.