AFTERWORD: EVOLUTIONARY EMOTION RESEARCH AT THE CROSSROADS

Marco Del Giudice

This is an exciting book. Not just in the sense that it is a fine volume, full of sophisticated and up-to-date contributions. (Compared with my memories when I first got interested in emotion research as a student, the field has been transformed almost beyond recognition; it is amazing how much progress has been made in this relatively short time.) Not just because the coverage of specific emotions and domains is exceptionally broad and fine-grained (from love and grief to *kama muta*, from conflict and reconciliation to care for the sick), or because so many chapters display that unique combination of explanatory power and trueness to life that is the hallmark of the evolutionary behavioral sciences (Pinker, 2015). To me, the book as a whole transmits a palpable excitement, not only for what it already contains, but also for what it foreshadows—a truly integrated theory of emotions, in which a set of shared functional principles link together the myriad particulars of our affective lives, with their phenomenal richness and subtlety.

Although this vision is still unrealized, and there remain many loose threads and gaps to fill, I believe that this volume marks an important step in the right direction. A clear tension toward integration runs through the book—as if the chapters were magnetic pieces, ready to snap into place and self-assemble into a unified construction. But before this can happen, evolutionary scholars will have to do some important conceptual groundwork to clarify the terms on which integration will take place. Evolutionary emotion research is at a crossroads—and we are lucky to be here in what promise to be very exciting times. Having been given the honor of writing this concluding chapter, I want to use the opportunity to review and critically discuss what I see as the most pressing theoretical issues in the field, then propose a research agenda for addressing them. While the items in the agenda lend themselves to a variety of approaches, I will argue that they can be tackled most effectively from an explicitly motivational perspective—a perspective already adopted by several authors of this *Handbook*, and presented somewhat more formally in my earlier chapter (Del Giudice, Chapter 5 in this volume).

Key Theoretical Issues

Specialized Mechanisms and the Great Bifurcation

The differences and similarities between emotion theories can be parsed using a number of different criteria (see, e.g., Scarantino, 2018). Without doubt, the most significant contrast in the evolutionary study of emotions is between theories that postulate the existence of multiple *specialized mechanisms* (or *programs*) dedicated to emotions and theories that deny it. The mechanism-rich camp includes the various incarnations of Basic Emotions Theory (BET; see Cordaro, Chapter 1 in this volume), the coordination approach synthesized by Tooby, Cosmides, and others (see Tooby & Cosmides, Chapter 2 in this volume; Al-Shawaf et al., 2016; Nesse, 1990), and LeDoux's theory of "survival circuits" (e.g., LeDoux & Brown, 2017). The main representatives of the mechanism-free camp are the constructionist theories advanced by L. F. Barrett (2006, 2017) and Russell (2003); Fridlund and Russell (Chapter 3 in this volume) also argue that the Behavioral Ecology View (BECV) of facial displays aligns with a constructivist, mechanism-free stance that denies the existence of central coordination machinery.

With a few notable exceptions (e.g., Lin & Westgate, Chapter 17 in this volume), the authors of this *Handbook* adopt one version or another of the mechanism-rich perspective. On the one hand, this probably has something to do with the theoretical preferences of the editors; on the other hand, I suspect that the assumption of specialized mechanisms is more empirically generative, as it encourages researchers to focus on particular emotions and their possible adaptive functions. Be as it may, it is important to acknowledge that constructionist, mechanism-free theories are becoming increasingly popular in emotion research, especially outside the circle of scholars with an explicitly evolutionary perspective.

While I believe that constructionists have made hugely important contributions—not least by sharpening the methodological tools of emotion research, and successfully challenging some untenable assumptions of classical BET—I also believe that their arguments against specialized mechanisms are much weaker than they look on the surface. As others have noted, the evidence offered in support is perfectly compatible with the (mechanism-rich) coordination approach, and with revisions of BET that integrate the same functional principles (e.g., Faucher, 2013; Nesse, 2020; Scarantino, 2012a, 2015; Sznycer & Cohen, 2021).

Take the finding that episodes of "the same emotion" (e.g., anger) within and between individuals are heterogeneous, with relatively low coherence among different components of the response (subjective feelings, facial expressions, autonomic patterns, etc.). In the constructivist literature, this is regarded as evidence that emotions arise from the action of multiple independent processes without a central program, even if they are subjectively *categorized* as unitary phenomena. There are several problems with this inference. To begin, folk emotion labels like "anger" likely refer to the outputs of multiple mechanisms, each with its own specific rules of operation; thus, relying on folk categories is going to inflate the apparent variability of emotional responses (Scarantino, 2012a, 2015; Sznycer et al., 2017). Even more importantly, dealing effectively with fitness-critical situations requires emotion programs to be sensitive to context and produce open-ended response tendencies, rather than rigid and reflex-like behaviors (Al-Shawaf & Lewis, 2017; Al-Shawaf et al., 2016; Scarantino, 2015). Invariance is expected at the level of the function and information-processing structure of the mechanism, but not at that of the mechanism's output, which should display systematic variability across persons and situations. As noted by Sznycer and Cohen (2021), "adaptive orchestration may call for response coherence in some contexts and for low coherence or discoherence in other contexts" (p. 3).

Finally, one has to factor in the existence of pervasive individual differences (partly caused by genetic variation), both in the emotion mechanisms themselves and in the cognitive and physiological mechanisms they coordinate (Nesse, 2020). Clearly, response variability does not militate against the existence of (context-sensitive) emotion mechanisms; on the contrary, it may signal the need to postulate *more* mechanisms than would be suggested by folk emotion categories like "anger," "sadness," and "fear." The inference from response variability to the absence of specialized mechanisms is an unwarranted move, buttressed by rhetorical devices such as charges of "essentialism"—a term that by now has very little meaning, except as name-calling—and invocations of Mayr's authoritative-sounding but nebulous concept of "population thinking" (e.g., L. F. Barrett, 2017; on population thinking, see Hey, 2011; Witteveen, 2015, 2016, 2018).

In the constructivist literature, neurobiological evidence of "degeneracy" (i.e., many brain regions or networks can be mapped to one emotion category) and "neural reuse" (i.e., a single region or network can be mapped to many emotion categories) are used to argue that there are no specialized brain centers that coordinate emotional responses (see L. F. Barrett, 2017). But in fact, degeneracy and reuse are exactly what one expects if the overarching function of emotions is to orchestrate the activity of multiple mechanisms across the entire brain. If emotions work as coordination programs, the corresponding patterns of brain activity are inevitably going to look highly distributed and to overlap substantially with one another, as different emotions recruit many if not most of the same mechanisms (e.g., De Gelder & Vandenbulcke, 2012; Scarantino, 2012b). Likewise, findings of "reuse" are expected if emotions are not neatly separated neural mechanisms that evolved independently from one another, but partially differentiated adaptations that evolved from common precursors and may share substantial portions of their neural machinery (Nesse, 2004, 2020).[1]

The constructivist interpretation of the neuroscientific data is further vitiated by what I have called the *prediction-explanation fallacy*, or the uncritical use of prediction-optimized models for explanatory purposes (Del Giudice, 2021). For example, Barrett (2017) stated:

> Ironically, perhaps the strongest evidence to date for the theory [of constructed emotions] comes from studies that use pattern classification to distinguish categories of emotion—a finding that is routinely construed as providing the long awaited support for the classical view [. . .]. However, patterns that distinguish among the categories in one study do not replicate in the other studies. (p. 15)

The problem is that, in many applications of machine-learning techniques, prediction-optimized models deliberately introduce biases to improve predictive performance, at the expense of accurate explanation (see James et al., 2021; Shmueli, 2010; Yarkoni & Westfall, 2017). And because different models tend to be biased in different ways, they may suggest conflicting explanations of the same phenomenon—even when they are trained on the same or very similar data, and achieve similar levels of predictive accuracy (Breiman, 2001). Hence, it is a fallacy to interpret discordances between prediction-optimized models as straightforward evidence that the underlying phenomenon lacks consistency (for details see Del Giudice, 2021).

Fridlund and Russell (Chapter 3 in this volume) argue that facial behaviors (and presumably other behavioral and physiological phenomena that are usually regarded as components of emotional responses) evolved "in granular fashion" with no central coordination machinery, so that there are no "emotion centers in the brain." But this is a non sequitur, based on the unsubstantiated assertion that specialized computational mechanisms are somehow problematic from an evolutionary standpoint.

While adaptations *can* evolve as mosaics (and bipedalism seems a fitting example), this does not mean that all or even most of them do; in reality, organisms are packed with regulation and coordination mechanisms that control functionally related suites of traits and behaviors, and can be targets of selection in their own respect. Hormonal systems (and their regulatory centers in the hypothalamus) are obvious examples (Cox et al., 2016). Another striking illustration comes from the medial preoptic area (MPOA) in rodents, whose function is increasingly well understood thanks to progress in cellular and molecular neuroscience. This region of the hypothalamus collects and integrates parenting-relevant information from the environment (tactile, auditory, and chemical signals from pups, cues of nearby predators, etc.) and coordinates parental behaviors in a modular fashion by projecting to multiple regions of the brain (Kohl & Dulac, 2018; Kohl et al., 2018). This is as clear a counterexample as any to the idea that complex behavioral adaptations must evolve as mosaics, without specialized coordination centers. Even highly composite, multifaceted, delocalized adaptations like the immune system have shared regulatory nodes that coordinate their activity on a broad scale (Kitano & Oda, 2006). The ubiquitous logic of bow-tie architectures (Csete & Doyle, 2004; Doyle & Csete, 2011; see Del Giudice, Chapter 5 in this volume) shows that "outside-in" granularity and "inside-out" centralized coordination are not mutually exclusive, and instead can effectively complement each other.

The Readout Problem

Another important contrast that runs through the pages of this *Handbook* concerns the evolutionary plausibility of "readouts"—nonverbal displays that reliably signal the internal state of the individual in an involuntary fashion. Both BET and the coordination approach assume that displays *can* function as involuntary expressions of emotional states, at least for certain emotions and in certain contexts (Al-Shawaf et al., 2016; Tooby & Cosmides, Chapter 2 in this volume). BECV decisively rejects this assumption, in favor of the idea that nonverbal displays evolve as "social tools" designed to influence the behavior of other individuals (e.g., eliciting affiliation or prompting retreat), without expressing anything about the signaler's internal states (Fridlund, 1994, 2017; Fridlund & Russell, Chapter 3 in this volume).

Without question, BECV has had a salutary effect on the field by pushing scholars beyond naïve conceptions of uniformity and automaticity in the expression of emotions, thanks to its emphasis on the fitness costs and benefits of signals and on the role of contextual factors (e.g., audience effects; see Keltner et al., 2019; Scarantino, 2017a). That said, the wholesale rejection of "readouts" (and, by extension, expressions of internal states) turns out to be another non sequitur: BECV begins an important evolutionary argument about the costs and benefits of communicative signals, but stops before considering its implications in full. The argument starts from the assumptions that (a) signals evolve to serve the biological interest of signalers; and that (b) involuntary readouts of internal states would benefit receivers, thus imposing a potential fitness cost on the signaler. While these premises are both true, it is wrong to conclude that readouts are inevitably selected against "in the service of deception, economy, and privacy" (Fridlund, 1994, p. 109).

The first corrective is that, *pace* Fridlund, expressions need not be *entirely* automatic and reflex-like; instead, they can lie on continua of flexibility–inflexibility and voluntariness–involuntariness, thus allowing a degree of strategic control and cost mitigation (e.g., Scarantino, 2017a). In the BECV literature, this is ruled out by fiat because "expressions" are defined from the start as fully involuntary, reflex-like displays with no sensitivity to context; however, this is a way to avoid the issue rather than confronting it. Displays can be statistically reliable without being always and perfectly truthful, and can have an involuntary quality without being a hundred percent automatic and inflexible (Scarantino, 2017a, 2017b).

Even more importantly, honest signaling of one's state *can* be fitness-enhancing when there is common interest between signalers and receivers: this may happen when the receiver is genetically related to the signaler, or when transmitting reliable information about internal states permits coordination and reciprocity with one's cooperation partners (see Svetieva, Chapter 35 in this volume; Al-Shawaf et al., 2016; Bergstrom & Lachmann, 1997, 1998; Keltner et al., 2019; Krebs & Dawkins, 1984; Planer & Godfrey-Smith, 2021). More generally, it is crucial to realize that receivers can drive the evolution of reliable signals—including signals that carry fitness costs—as long as they are in a position to impose even larger costs on the signalers (e.g., by ostracizing or otherwise punishing them), and/or provide them with benefits that outweigh the costs (e.g., by choosing them as social or sexual partners). The fact that readouts have disadvantages is not a particular obstacle to their evolution, as long as their *net* effect on the signaler's fitness is positive. Indeed, social and sexual selection can favor the evolution of high levels of cooperativeness and altruism that benefit other group members at a significant cost to the individual (Aktipis, 2011; Nesse, 2007), and can lead individuals to produce credible, often expensive displays that reveal their internal condition and genetic quality (Biernaskie et al., 2018; Clutton-Brock & Huchard, 2013). Why should emotion expressions be an exception? By rejecting social partners who fail to provide legible information about their emotions, or who provide it in ways that turn out to be too confusing or misleading (and hence not predictive), receivers can select for the evolution of emotional readouts despite their potential costs.

On this count, there is intriguing evidence that emotional expressivity—not just for positive, but also for negative emotions—increases affiliation, cooperation, and trust (Boone & Buck, 2003; Feinberg et al., 2012; Schug et al., 2010; Wubben et al., 2011). Trust and likeability also increase when a person's emotional expressions are predictable (i.e., they match the observer's expectations based on contextual information; Chanes et al., 2018). The flip side is that we tend to avoid and mistrust people who come across as emotionally opaque or whose emotional responses are hard to predict. When receivers equipped with mind-reading adaptations can respond to differences in signaling behavior with affiliation or ostracism, excessive restraint in revealing one's inner states can be as damaging as excessive transparency; this demands a careful consideration of the payoff structures associated with social interactions, rather than a blanket rejection of the idea of readouts/expressions.

On a related point, Dezecache and Mercier (Chapter 44 in this volume) suggest that mechanisms of *emotional vigilance* in receivers may play a critical role in maintaining the (average) honesty of emotional displays, which is not guaranteed by their intrinsic features (e.g., metabolic costs). At various points in their fascinating chapter, the authors seem to embrace the BECV postulate that displays cannot function as readouts of emotional states; I find this somewhat ironic, because emotional vigilance is precisely the kind of receiver adaptation that *could* drive the evolution of credible readouts (see also McCullough & Reed, 2016).

Another questionable assumption of BECV is that displays benefit receivers mainly by allowing them to predict the signaler's actions in the immediate future; for example, a scowling, bared-teeth display might indicate that a person is more likely to start behaving aggressively. From this perspective, emotional states (to the extent that they can be said to actually exist; see Fridlund & Russell, Chapter 3 in this volume) would add nothing of value to the receiver, because they would fail to predict future behavior with any precision (for extended discussion of this point, see Scarantino, 2017a). This strikes me as a surprisingly narrow view of prediction: we build models of other people to anticipate their behavior on timescales that extend far beyond the immediate future—and not just in response to the present state of reality, but also in hypothetical future scenarios. Granted, knowing that someone is angry says little about what *exactly* they are going to do the next minute (though it remains predictive in a statistical sense); but, combined with contextual knowledge about the cause of anger, it can provide many other kinds of information that are invaluable in the long run—information about their goals and priorities, perception of that situation (and possibly similar ones), self-perceptions and self-evaluations, expectations from others, sensitivity to specific "triggers," regulation abilities, broader personality, and so forth. Indeed, people routinely use others' emotional reactions to "reverse-engineer" the appraisal process and draw inferences about their goals, values, motivations, and other traits of interest, and hence predict their future decisions more accurately (Gratch & de Melo, 2019: Hareli & Hess, 2010, 2019; Keltner et al., 2019). Emotional readouts can be hugely valuable to mind-reading receivers, setting the stage for social selection favoring displays that signal emotions with some degree of reliability.

Whither Basic Emotions?

It is hard to overstate the importance of BET in the history of emotion research. Over the years, the initial list of six basic emotions has been expanded to include more than twenty; Cordaro (Chapter 1 in this volume) lists 26 candidates, ranked at various levels of evidentiary support. In recent years, researchers working in the BET tradition have started to collect very large data sets of expressions and self-reports (Cowen & Keltner, 2020; Cowen et al., 2019); analyses usually yield about 25–30 dimensions of variation, which overlap largely—but not completely—across different studies and with the criteria-based lists offered by Cordaro (Chapter 1 in this volume), Keltner et al. (2019), and others. These and similar findings are quite impressive (see Keltner et al., 2019), but the edifice of BET is increasingly showing its cracks, which are largely conceptual rather than empirical in nature.

The first and most pressing issue with BET is that the criteria used to classify emotions as basic are not grounded in a functional and computational analysis of each emotion, but reflect the researchers' somewhat arbitrary intuitions of what an evolved emotion should look like. The three criteria of *brief duration* (usually specified as "less than a minute"), *universal signals*, and *primate homologies* (sometimes extended more broadly to mammals; e.g., Keltner et al., 2019) are particularly problematic and lack any coherent functional rationale. To give just an example, jealousy fails the first two criteria and possibly all three,[2] despite very strong evidence of adaptive design (Al-Shawaf et al., 2016; Buss, 2014). More generally, two evolved emotions may easily match different subsets of the "basic" criteria because of their specific functions and phylogenetic histories. Consider two hypothetical emotions A and B. Emotion A has clear homologues in other primates and occurs in brief episodes; whereas B has evolved more recently, and tends to manifest itself in longer-lasting episodes because its function unfolds on a longer timescale (e.g., by promoting certain forms of sustained cognitive processing). In the framework of BET, functional differences of this kind would be misinterpreted as mere differences in evidentiary support, so that emotion A would be classified as a "strong" candidate while B would be regarded as a "weak" contender for the title of basic emotion (see Table 1.2 in Cordaro, Chapter 1 in this volume). The deeper problem is that BET relies on a fixed set of criteria that are supposed to apply across the board; this use of checklists is arguably the original sin of BET, and at this point is holding back the theory rather than providing useful heuristic guidance.

A related problem—which also highlights the need for deeper functional thinking—is that BET tends to take folk emotional categories at face value, raising doubts about whether the "basic" emotions identified by researchers correspond to specific evolved mechanisms (see also Ortony, 2022). For example, different emotion programs may well elicit similar nonverbal displays—there is no reason to expect a strict one-to-one correspondence between emotions and expressions. As I noted in my earlier chapter, the reduction methods used to extract emotional dimensions from empirical data sets (PCA and such) do not help in this regard, because they are simply not capable of identifying the underlying mechanisms, except in unrealistic special cases (Del Giudice, Chapter 5 in this volume). The only way out of this quandary is through functional and computational analysis. This is the focus of the "New BET" (NBET) formulated by Scarantino (2015); in this important revision of the theory, folk categories are regarded as potentially heterogeneous, so that a single category like "fear" may comprise more than one basic emotion; an emotion is not regarded as basic because it fulfills a list of preset criteria, but because it can be shown to correspond to a specific evolved mechanism. Also, NBET answers the question of response variability by postulating that at least some emotional programs are "output-flexible," in the sense that they produce context-sensitive response *tendencies* instead of the mandatory, "output-rigid" responses of classical BET. Clearly, NBET goes a long way toward bridging the gap between the basic emotions tradition and the coordination model (e.g., Sznycer et al. 2017). In my view, the confluence of these approaches into a unified mechanism-rich theory is inevitable in the long run—and the sooner it happens, the better.

A Big-Picture Agenda

The theoretical issues I just surveyed have implications for all the major debates in emotion research; they need to be addressed and resolved, before an integrated evolutionary theory of emotions can emerge from the competing approaches that dominate the scene today. To this end, I would like to propose an agenda for evolutionary scholars who wish to tackle the big questions in the field. I formulate my suggestions as a member of the mechanism-rich team; but of course, many of the same points are just as relevant when viewed from the perspective of mechanism-free approaches.

Mechanisms and Variability

One of the main obstacles to the functional analysis of emotions is the fact that folk emotion categories may not identify specific adaptations, but rather collections of superficially similar mechanisms, with different purposes and rules of operation. While this is not a new insight (e.g., Cosmides & Tooby, 2000; Scarantino, 2012a), I note that many chapters in this *Handbook* still take folk categories at face value, and do not probe them for the possible existence of multiple underlying mechanisms. This should be a top research priority for mechanism-rich approaches, including BET. In the crucial task of "carving nature at its joints," functional and computational analyses can be usefully complemented by the fine-grained neurobiological dissection of brain regions and pathways (e.g., Panksepp, 1998, 2011).

Other important contributions in this sense can come from comparative and phylogenetic data, which are surprisingly under-utilized in the evolutionary study of emotion (but see Clay & Austry, Chapter 41 in this volume; Habecker & Flinn, Chapter 12 in this volume; Kujala & Braüer, Chapter 42 in this volume; O'Connell et al., Chapter 26 in this volume; Torres & Papini, Chapter 43 in this volume; Vonk et al., Chapter 40 in this volume). The view of emotions as partially differentiated mechanisms that evolve via reuse, duplication, and modification (Nesse, 2004, 2020; see also H. C. Barrett, 2012; Holbrook & Fessler, 2015; West-Eberhard, 2003) opens up many interesting questions about the phylogeny of emotions and corresponding neural mechanisms, from early animals to humans via the common ancestors of mammals and primates. Crucially, scholars need to abandon the preconception that evolved emotion mechanisms must be highly conserved, and instead expect a mixture of stability, novelty, and species-specific variation. The interplay between "granular" and "centralized" modes of evolution during phylogeny could be usefully tackled with simulations; for example, agent-based models could be used to explore the evolution of emotional mechanisms and their components under different conditions. The concept of bow-tie architectures (Csete & Doyle, 2004) could inform the design of these simulations and help interpret their results.

The problem of response coherence vs. variability is particularly tricky, because even if virtually all scholars agree that perfect coherence is not a reasonable expectation, it remains unclear how much variability is predicted by different theoretical models (and hence what amount of variability should count as supporting vs. disconfirming evidence). On this turf, constructivist models play with a tactical advantage: since they deny the existence of specialized mechanism, they have no grounds to expect a particular degree of coherence, and any evidence of variability can be loosely counted in their favor. From a mechanism-rich perspective, the solution is to start building explicit computational models of emotional responses, and—once the models become sufficiently realistic—to use their predictions as benchmarks to evaluate the empirical data. Models of this kind would also allow researchers to formulate predictions about *qualitative* patterns of variability (e.g., which components of a response are going to be the most/least variable), and about the individual and contextual factors that influence the coherence of emotional responses. Given the current lack of information and the fact that both sides of the debate rely on intuitions, even drastically simplified models are likely to provide useful insights. As far as I know, this research direction has remained virtually unexplored in the literature; the recent paper by Sznycer and Cohen (2021) is a step in the right direction, but still a long way from the explicit formal models that would be needed to adequately meet the constructivist challenge.

Eventually, computational models of emotions will need to explicitly integrate the physiological dimensions of the response (including, but not limited to, patterns of autonomic activation; see Tooby & Cosmides, Chapter 2 in this volume). Physiological changes modulate the internal state of the organism, but also have a range of perceptible effects that may function as signals (e.g., eyeblinks, blushing, changes in pupil size; Kret, 2015); also, physiological measures are often used in empirical studies of response coherence (e.g., Mauss et al., 2005; Rattel et al., 2020). In principle, then, physiology plays an important role in the coordination approach. However, current computational models of anger (Sell & Sznycer, Chapter 6 in this volume), shame (Landers et al., Chapter 7 in this volume), and many other emotions are still largely silent about the interplay between physiology and the cognitive/behavioral components on which they mainly focus. To include physiology in their models, evolutionary psychologists will have to get out of their "comfort zone" and cross the borders of evolutionary neurobiology and endocrinology; this is a vital but challenging task, which is made even more challenging by our very limited understanding of individual differences in physiological parameters and their functional meaning (see, e.g., Ellis & Del Giudice, 2019; Ellis et al., 2021; Gangestad & Grebe, 2017). At the same time, a computational mindset can greatly enrich the study of physiology and hormones—for example by underscoring the information-processing aspects of neurobiological mechanisms, framing individual differences in terms of internal regulatory variables (Tooby & Cosmides, 2008), and generally helping researchers ask important "why" questions in addition to the "how" questions of the standard mechanistic approach.

Readouts and Inferences

As I discussed earlier, BECV's negative argument against readouts does not follow from its premises, and readouts—in the sense of statistically reliable, at least partially involuntary signals of internal states—can certainly evolve, even if they have some costs and drawbacks. But this does not mean that emotional readouts *will* evolve in practice. To properly answer the critique from BECV, proponents of mechanism-rich models should clarify the conditions in which readouts are more or less likely to evolve, and identify the factors that tend to increase or decrease their "honesty" and credibility. These and similar questions could be usefully tackled with evolutionary simulations in which signaler and receiver strategies are free to coevolve (for an interesting but still rudimentary example, see Lowe et al., 2004); and when the models become sufficiently realistic, they could be tested against comparative data from animals with different ecologies and social structures.

Simulations are also an ideal way to formalize existing theories of emotional inference (Hareli & Hess, 2019), and refine them with the addition of mechanisms of emotional vigilance (Dezecache & Mercier, Chapter 44 in this volume). More generally, an evolutionary perspective would add specificity and depth to our understanding of inferential processes; the range of potential inferences that an observer could draw is almost limitless, but some aspects of the signaler's internal state are likely to be much more informative and/or fitness-relevant than others. The focus of emotional inferences may also shift depending on the specific social relationship between signaler and receiver—for example, the kind of information we seek can be quite different if we are dealing with a friend, an enemy, a relative, an attractive stranger, or a romantic partner.

Integration between Basic Emotions Theory and the Coordination Approach

I believe that the future will see a convergence between BET and the coordination approach, and would like to propose some ways to help the process along. First, coordination theorists should pay more attention to the large-scale empirical findings by researchers such as Cowen, Keltner, and others (e.g., Cowen & Keltner, 2020; Cowen et al., 2019). Patterns of perceived similarity can suggest ways to refine existing hypotheses; for example, emotion concepts form a continuum from anger and hate to sadness and despair, and while "righteous indignation" and "feeling oppressed" lie close to anger, "feeling let down" and "feeling betrayed" are much closer to sadness (Fig. 2 in Cowen et al., 2019). This seems potentially relevant to the recalibration theory of anger (Sell & Sznycer, Chapter 6 in this volume), and may point to a deep functional link between anger and sadness that has not been discussed in the recalibration literature (e.g., certain forms of sadness may reflect the recalibration of one's overly optimistic estimates of other people's welfare trade-off ratios).

From a different perspective, a crucial test of the coordination approach will be its ability to reproduce the observed patterns of correlation among emotional concepts, experiences, and/or expressions—not inductively, but starting from well-defined mechanisms and their functional relations with one another. The recent work on valuation as the common grammar of social emotions (Scrivner et al., Chapter 4 in this volume) is the beginning of a much-needed shift from a narrow, analytical focus on single emotions to a more synthetic perspective. An important next step would be to reproduce the broad structure of people's emotion-related concepts and perceptions, starting from a model of the underlying mechanisms. Of note, any discrepancies between the actual and predicted structure would provide invaluable information about gaps and inconsistencies in the model.

Finally, even though the BET criteria for basic emotions are blunt tools when it comes to identifying adaptations, they can suggest interesting questions about the reasons why different emotions show different patterns of correlates. For example, what are the main exceptions to the criterion of "brief duration"? What do those emotions have in common? What are the cognitive and neurobiological processes that sustain emotions for longer periods of time, and what is their design logic? Other potentially fruitful questions concern the degree of universality vs. cross-cultural variability of different emotions, and the functional meaning of differences in expression modalities: What are the relative strengths and weaknesses of visual, auditory, tactile, and olfactory cues (e.g., Bradbury & Vehrencamp, 2011)? Are different nonverbal modalities better suited to different purposes and social relationships? So far, these questions have been addressed only piecemeal in the BET literature; an integrated approach would have the tools to formulate and answer them in a systematic fashion.

The Advantages of a Motivational Perspective

The motivational perspective I synthesized in my earlier chapter (Del Giudice, Chapter 5 in this volume) is the antithesis of the mechanism-free approach—it takes an already rich model of emotions and extends it by adding *more* layers of specialized mechanisms (second-order motivational systems and third-order mood programs). Here I wish to briefly make the case that the resulting increase in complexity is worth the price, and that a focus on motivational systems can help with many of the items in the research agenda I just sketched—starting with the task of looking for multiple mechanisms within folk categories such as "anger" or "sadness." Carving nature at its joints gets much easier when the analysis can be guided by well-grounded functional distinctions; the distinctions between motivational domains can be used to generate plausible hypotheses about, say, various kinds of "anger" and their adaptive purposes.

From a neurobiological standpoint, the neural substrates of motivational systems may be easier to identify and study than those of specific emotions; for instance, the MPOA is a very promising candidate as the control hub of the parenting/caregiving motivational system. Also, the logic of bow-tie architectures suggests that motivational systems should be phylogenetically more conserved than emotions, and hence easier to study and compare across species. When it comes to integrating physiology in current models of emotions, the literature on motivational systems provides some useful computational principles, including control-theoretic ideas about feedback and feedforward regulation. The concept of internal regulatory variables has parallels and precursors in attachment theory ("internal working models"; Bowlby, 1969) and other classical theories of motivation, which I believe can still be mined for fresh insights.

A motivational perspective also helps frame the issue of readouts and emotional inferences: the functional parameters of motivational systems (e.g., their relative priorities and activation sensitivities) bridge emotions with personality, and can be extremely useful for predicting behavior in the long run. Because of this, people should often use emotions as indicators of the underlying motivational processes, which I suggested work as a privileged level of analysis for inference and prediction. Not least, motivational systems may offer a parsimonious way to organize the high-dimensional maps of "emotion spaces" that are emerging from the large descriptive studies carried out by BET scholars. Indeed, motivational states can be thought of as attractors of the more complex dynamics of emotions and moods; as such, they should help make sense of the increasingly rich, detailed, but also potentially overwhelming data on specific emotions by revealing their underlying functional structure.

Conclusion

I am grateful to the authors and editors of this *Handbook* for their contribution to what I think will become a landmark for the discipline. Researchers are already busy extending the reach of evolutionary approaches to more and more emotions, domains, and areas of application. To complement this enterprise and bring it to full fruition, they will have to address the most pressing conceptual issues head-on, and find convincing answers to some long-standing questions in the study of emotion. The call is not just for bigger and better data, but for a major investment in good theory, which—as per Kurt Lewin's maxim—is really the most practical of things. The day is young; let's get to work.

Acknowledgments

I am indebted to Laith Al-Shawaf, Andrea Scarantino, and Daniel Sznycer for their many insightful comments and suggestions.

Notes

1. On a side note, the predictive Bayesian models of the brain that inform Barrett's current *theory of constructed emotions* (L. F. Barrett, 2017) offer powerful explanations of perception and (perhaps) action control, but struggle to provide a realistic, tractable account of motivation (Colombo, 2016; Klein, 2018; Sun & Firestone, 2020a, 2020b). In light of the deep connections between motivation and emotion, I see this as more reason to be skeptical about the conceptual adequacy of this framework.
2. The jury is still out on the existence of proto-jealousy in other primates; see, e.g., Maninger et al. (2017).

References

Aktipis, C. A. (2011). Is cooperation viable in mobile organisms? Simple Walk Away rule favors the evolution of cooperation in groups. *Evolution and Human Behavior*, *32*, 263–276.

Al-Shawaf, L., Conroy-Beam, D., Asao, K., & Buss, D. M. (2016). Human emotions: An evolutionary psychological perspective. *Emotion Review*, *8*, 173–186.

Al-Shawaf, L., & Lewis, D. M. G. (2017). Evolutionary psychology and the emotions. In V. Zeigler-Hill & T.K. Shackelford (Eds.), *Encyclopedia of personality and individual differences* (pp. 1452–1461). Springer.

Barrett, H. C. (2012). A hierarchical model of the evolution of human brain specializations. *Proceedings of the National Academy of Sciences*, *109*, 10733–10740.

Barrett, L. F. (2006). Are emotions natural kinds? *Perspectives on Psychological Science*, *1*, 28–58.

Barrett, L. F. (2017). The theory of constructed emotion: An active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience*, *12*, 1–23.

Bergstrom, C. T., & Lachmann, M. (1997). Signalling among relatives. I. Is costly signalling too costly? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *352*, 609–617.

Bergstrom, C. T., & Lachmann, M. (1998). Signaling among relatives. III. Talk is cheap. *Proceedings of the National Academy of Sciences*, *95*, 5100–5105.

Biernaskie, J. M., Perry, J. C., & Grafen, A. (2018). A general model of biological signals, from cues to handicaps. *Evolution Letters*, *2*, 201–209.

Boone, R. T., & Buck, R. (2003). Emotional expressivity and trustworthiness: The role of nonverbal behavior in the evolution of cooperation. *Journal of Nonverbal Behavior*, *27*, 163–182.

Bowlby, J. (1969). *Attachment and loss*, Vol. I: *Attachment*. Basic Books.

Bradbury, J. W., & Vehrencamp, S. L. (2011). *Principles of animal communication* (2nd ed.). Sinauer.

Breiman, L. (2001). Statistical modeling: The two cultures. *Statistical Science*, *16*, 199–231.

Buss, D. M. (2014). Comment: Evolutionary criteria for considering an emotion "basic": Jealousy as an illustration. *Emotion Review*, *6*, 313–315.

Chanes, L., Wormwood, J. B., Betz, N., & Barrett, L. F. (2018). Facial expression predictions as drivers of social perception. *Journal of Personality and Social Psychology*, *114*, 380–396.

Clutton-Brock, T. H., & Huchard, E. (2013). Social competition and selection in males and females. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*, 20130074.

Colombo, M. (2016). Social motivation in computational neuroscience (Or, if brains are prediction machines, then the Humean theory of motivation is false). In J. Kiverstein (Ed.), *The Routledge handbook of philosophy of the social mind* (pp. 336–356). Routledge.

Cosmides, L., & Tooby, J. (2000). Evolutionary psychology and the emotions. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (2nd ed., pp. 91–115). Guilford Press.

Cowen, A. S., & Keltner, D. (2020). What the face displays: Mapping 28 emotions conveyed by naturalistic expression. *American Psychologist*, *75*, 349–364.

Cowen, A., Sauter, D., Tracy, J. L., & Keltner, D. (2019). Mapping the passions: Toward a high-dimensional taxonomy of emotional experience and expression. *Psychological Science in the Public Interest*, *20*, 69–90.

Cox, R. M., McGlothlin, J. W., & Bonier, F. (2016). Hormones as mediators of phenotypic and genetic integration: An evolutionary genetics approach. *Integrative and Comparative Biology*, *56*, 126–137.

Csete, M., & Doyle, J. (2004). Bow ties, metabolism and disease. *Trends in Biotechnology*, *22*, 446–450.

De Gelder, B., & Vandenbulcke, M. (2012). Emotions as mind organs. *Behavioral and Brain Sciences*, *35*, 147–148.

Del Giudice, M. (2021). The prediction-explanation fallacy: A pervasive problem in scientific applications of machine learning. *PsyArXiv*. https://doi.org/10.31234/osf.io/4vq8f

Doyle, J. C., & Csete, M. (2011). Architecture, constraints, and behavior. *Proceedings of the National Academy of Sciences*, *108*, 15624–15630.

Ellis, B. J., & Del Giudice, M. (2019). Developmental adaptation to stress: An evolutionary perspective. *Annual Review of Psychology*, *70*, 111–139.

Ellis, B. J., Horn, A. J., Carter, C. S., van IJzendoorn, M. H., & Bakermans-Kranenburg, M. J. (2021). Developmental programming of oxytocin through variation in early-life stress: Four meta-analyses and a theoretical reinterpretation. *Clinical Psychology Review*, *86*, 101985.

Faucher, L. (2013). Comment: Constructionisms? *Emotion Review*, *5*, 374–378.

Feinberg, M., Willer, R., & Keltner, D. (2012). Flustered and faithful: Embarrassment as a signal of prosociality. *Journal of Personality and Social Psychology*, *102*, 81–97.

Fridlund, A. J. (1994). *Human facial expression: An evolutionary view*. Academic Press.

Fridlund, A. J. (2017). The behavioral ecology view of facial displays, 25 years later. In J. M. Fernández-Dols & J. A. Russell (Eds.), *The science of facial expression* (pp. 77–92). Oxford University Press.

Gangestad, S. W., & Grebe, N. M. (2017). Hormonal systems, human social bonding, and affiliation. *Hormones and Behavior*, *91*, 122–135.

Gratch, J., & de Melo, C. (2019). Inferring intentions from emotion expressions in social decision making. In U. Hess & S. Hareli (Eds.), *The social nature of emotion expression: What emotions can tell us about the world* (pp. 141–160). Springer.

Hareli, S., & Hess, U. (2010). What emotional reactions can tell us about the nature of others: An appraisal perspective on person perception. *Cognition and Emotion*, *24*, 128–140.

Hareli, S., & Hess, U. (2019). The reverse engineering of emotions: Observers of others' emotions as naïve personality psychologists. In U. Hess & S. Hareli (Eds.), *The social nature of emotion expression: What emotions can tell us about the world* (pp. 103–118). Springer.

Hey, J. (2011). Regarding the confusion between the population concept and Mayr's "population thinking." *Quarterly Review of Biology*, *86*, 253–264.

Holbrook, C., & Fessler, D. M. (2015). The same, only different: Threat management systems as homologues in the tree of life. In P. J. Carroll, R. M. Arkin, & A L. Wichman (Eds.), *Handbook of personal security* (pp. 95–110). Psychology Press.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An introduction to statistical learning with applications in R* (2nd ed.). Springer.

Keltner, D., Sauter, D., Tracy, J., & Cowen, A. (2019). Emotional expression: Advances in basic emotion theory. *Journal of Nonverbal Behavior*, *43*, 133–160.

Kitano, H., & Oda, K. (2006). Robustness trade-offs and host–microbial symbiosis in the immune system. *Molecular Systems Biology*, *2*, 2006.0022.

Klein, C. (2018). What do predictive coders want? *Synthese*, *195*, 2541–2557.

Kohl, J., Babayan, B. M., Rubinstein, N. D., Autry, A. E., Marin-Rodriguez, B., Kapoor, V., Miyamishi, K., Zweifel, L. S., Luo, L., Uchida, N., & Dulac, C. (2018). Functional circuit architecture underlying parental behaviour. *Nature*, *556*, 326–331.

Kohl, J., & Dulac, C. (2018). Neural control of parental behaviors. *Current Opinion in Neurobiology*, *49*, 116–122.

Krebs, J. R., & Dawkins, R. (1984). Animal signals: Mind-reading and manipulation. In J. R. Krebs & N. B. Davies (Eds.), *Behavioural ecology: An evolutionary approach* (2nd ed., pp. 380–402). Blackwell.

Kret, M. E. (2015). Emotional expressions beyond facial muscle actions. A call for studying autonomic signals and their impact on social perception. *Frontiers in Psychology*, *6*, 711.

LeDoux, J. E., & Brown, R. (2017). A higher-order theory of emotional consciousness. *Proceedings of the National Academy of Sciences*, *114*, E2016–E2025.

Lowe, R., Cañamero, L., Nehaniv, C. L., & Polani, D. (2004). The evolution of affect-related displays, recognition and related strategies. In J Pollack, M. Bedau, P. Husbands, T. Ikegami, & R. A. Watson (Eds.), *Artificial life IX: Proceedings of the Ninth International Conference on the Simulation and Synthesis of Living Systems* (pp. 176–181). MIT Press.

Maninger, N., Mendoza, S. P., Williams, D. R., Mason, W. A., Cherry, S. R., Rowland, D. J., Schaefer, T., & Bales, K. L. (2017). Imaging, behavior and endocrine analysis of "jealousy" in a monogamous primate. *Frontiers in Ecology and Evolution*, *5*, 119.

Mauss, I. B., Levenson, R. W., McCarter, L., Wilhelm, F. H., & Gross, J. J. (2005). The tie that binds? Coherence among emotion experience, behavior, and physiology. *Emotion*, *5*, 175–190.

McCullough, M. E., & Reed, L. I. (2016). What the face communicates: Clearing the conceptual ground. *Current Opinion in Psychology*, *7*, 110–114.

Nesse, R. M. (1990). Evolutionary explanations of emotions. *Human Nature*, *1*, 261–289.

Nesse, R. M. (2004). Natural selection and the elusiveness of happiness. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *359*, 1333–1347.

Nesse, R. M. (2007). Runaway social selection for displays of partner value and altruism. *Biological Theory*, *2*, 143–155.

Nesse, R. M. (2020). Tacit creationism in emotion research. In C. Price & E. Walle (Eds.), *Emotion researcher: ISRE's sourcebook for research on emotion and affect*. International Society for Research on Emotion. https://emotionresearcher.com/tacit-creationism-in-emotion-research

Ortony, A. (2022). Are all "basic emotions" emotions? A problem for the (basic) emotions construct. *Perspectives on Psychological Science*, *17,* 41–61.

Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions*. Oxford University Press.

Panksepp, J. (2011). Cross-species affective neuroscience decoding of the primal affective experiences of humans and related animals. *PLoS ONE*, *6*, e21236.

Pinker, S. (2015). Foreword. In D. M. Buss (Ed.), *The handbook of evolutionary psychology* (2nd ed., pp. ix–xvi). Wiley.

Planer, R. J., & Godfrey-Smith, P. (2021). Communication and representation understood as sender–receiver coordination. *Mind & Language*, *36,* 750–770.

Rattel, J. A., Mauss, I. B., Liedlgruber, M., & Wilhelm, F. H. (2020). Sex differences in emotional concordance. *Biological Psychology*, *151*, 107845.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*, 145–172.

Scarantino, A. (2012a). How to define emotions scientifically. *Emotion Review*, 4, 358–368.

Scarantino, A. (2012b). Functional specialization does not require a one-to-one mapping between brain regions and emotions. *Behavioral and Brain Sciences*, *35*, 161–162.

Scarantino, A. (2015). Basic emotions, psychological construction, and the problem of variability. In L. F. Barrett & J. A. Russell (Eds.), *The psychological construction of emotion* (pp. 334–376). Guilford Press.

Scarantino, A. (2017a). How to do things with emotional expressions: The theory of affective pragmatics. *Psychological Inquiry*, *28*, 165–185.

Scarantino, A. (2017b). Twelve questions for the theory of affective pragmatics. *Psychological Inquiry*, *28*, 217–232.

Scarantino, A. (2018). The philosophy of emotions and its impact on affective science. In L. F. Barrett, M. Lewis, & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (4th ed., pp. 3–48). Guilford Press.

Schug, J., Matsumoto, D., Horita, Y., Yamagishi, T., & Bonnet, K. (2010). Emotional expressivity as a signal of cooperation. *Evolution and Human Behavior*, *31*, 87–94.

Shmueli, G. (2010). To explain or to predict? *Statistical Science*, *25*, 289–310.

Sun, Z., & Firestone, C. (2020a). The dark room problem. *Trends in Cognitive Sciences*, *24*, 346–348.

Sun, Z., & Firestone, C. (2020b). Optimism and pessimism in the predictive brain. *Trends in Cognitive Sciences*, *24*, 683–685.

Sznycer, D., & Cohen, A. S. (2021). Are emotions natural kinds after all? Rethinking the issue of response coherence. *Evolutionary Psychology*, *19*, 14747049211016009.

Sznycer, D., Cosmides, L., & Tooby, J. (2017). Adaptationism carves emotions at their functional joints. *Psychological Inquiry*, *28*, 56–62.

Tooby, J., & Cosmides, L. (2008). The evolutionary psychology of the emotions and their relationship to internal regulatory variables. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 114–137). Guilford Press.

West-Eberhard, M. J. (2003). *Developmental plasticity and evolution*. Oxford University Press.

Witteveen, J. (2015). "A temporary oversimplification": Mayr, Simpson, Dobzhansky, and the origins of the typology/population dichotomy (part 1 of 2). *Studies in History and Philosophy of Science C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *54*, 20–33.

Witteveen, J. (2016). "A temporary oversimplification": Mayr, Simpson, Dobzhansky, and the origins of the typology/population dichotomy (part 2 of 2). *Studies in History and Philosophy of Science C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *57*, 96–105.

Witteveen, J. (2018). Typological thinking: Then and now. *Journal of Experimental Zoology B: Molecular and Developmental Evolution*, *330*, 123–131.

Wubben, M. J., De Cremer, D., & van Dijk, E. (2011). The communication of anger and disappointment helps to establish cooperation through indirect reciprocity. *Journal of Economic Psychology*, *32*, 489–501.

Yarkoni, T., & Westfall, J. (2017). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on Psychological Science*, *12*, 1100–1122.