10

Toward a Science of Augmented Agency

Previous chapters identify major dilemmas for digitally augmented humanity, which this book defines in terms of close collaboration between human and artificial agents. These dilemmas are largely owing to differences between human and artificial capabilities and potentialities, and the resulting tensions in their collaboration. Working together, human and artificial agents must learn to manage these challenges. Joint supervision will be critical. Otherwise, augmented agents will tend toward divergent or convergent, dysfunctional form and function. As preceding chapters also explain, these dilemmas give rise to the following novel problematics: how can human beings collaborate closely with artificial agents, while remaining genuinely autonomous in reasoning, belief, and choice; relatedly, how can humans integrate digital augmentation into their subjective and intersubjective lives, while preserving their identities, commitments, and psychosocial coherence; how can digitally augmented institutions and organizations, conceived as collective agents, fully exploit artificial capabilities, while avoiding extremes of digitalized docility, dependence, and determinism; how can humanity ensure fair access to the benefits of digital augmentation, and not allow them to perpetuate systemic discrimination, deprivation, and injustice; and finally, the most

novel and controversial challenge, which is how will human and artificial agents learn to understand, trust, and respect each other, despite their different capabilities and potentialities?

In fact, comparable challenges have arisen before, albeit in less advanced technological contexts. During each historical period of civilized humanity, there have been new agentic forms, functions, and associated challenges of supervision. In parallel, procedures and institutions have evolved to exploit and govern these transitions. For example, the premodern world produced artisan guilds and councils, while modernity created institutions to regulate markets, industries, professions and more. However, such historic transitions are problematic too, because major socioeconomic change disrupts established orders and exposes the limitations of existing institutions. Once again, modernity is illustrative. Throughout the modern industrial period, different stakeholder groups have struggled over issues of governance, the distribution of resources, access to opportunities, and the rights and duties of employees versus owners. This has often led to major social and political disruption, and sometimes revolution.

Mass digitalization continues this historic trend, at unprecedented speed and scale. In fact, as earlier chapters explain, digitalization signals a major shift in human experience and organization. Digital technologies reach far more deeply into all aspects of agentic form and function, compared to earlier periods. Prompting some to predict a type of singularity, in which artificial intelligence equals and perhaps surpasses the human, and both then fuse to become effectively one (Eden et al., 2015). Preceding chapters list some of the enabling technologies, including artificial empathy and personality, and brain-machine engineering. In any case, whether singularity happens or not, human and artificial agents are ready subjects for a science of digitally augmented agency.

10.1 Science of Augmented Agency

This science is clearly needed. Augmented agents must know how to supervise their increasingly close collaboration, maintaining appropriate levels of convergence and divergence, and thus maximizing metamodel

fit. A science of augmented agency will support many of the required tools and techniques. Without such capabilities, however, poor supervision will result in dysfunctional patterns of ambimodality, ambiopia, ambiactivity, and ambidexterity. Moreover, as this terminology clearly demonstrates, the conceptual architecture of modern human science fails to capture important features of digital augmentation. I therefore import a few concepts from other fields. Table 10.1 lists these conceptual innovations.

The first new concept is ambimodal, which comes from chemistry and refers to transition or transformation processes which lead to multiple outcome states. In this book, the term refers to single processes which generate different modal characteristics, and more specifically, to agentic forms and functions which combine artificial compression with human layering. The second conceptual innovation is hyperopia, which is borrowed from ophthalmology, and refers to farsighted vision, the opposite of myopia. In this book, hyperopia refers to farsighted problem sampling

Table 10.1 New concepts and terms

Term	Conceptual meaning	Sample uses
Ambimodal	Processes and systems which produce multiple outcome states, for example, both compressed and layered agentic modalities	Yang et al. (2018), Chen et al. (2018)
Hyperopic	Farsighted expansive processes, especially in sampling and search, the opposite of myopia	Remeseiro et al. (2018), Tunyi et al. (2019)
Ambiopic	Processes and systems which combine nearsighted myopia and farsighted hyperopia	Smolarz-Dudarewicz et al. (1980), Buetow (2020)
Empathice	To satisfice in solving problems of other minds, rather than seeking to optimize in cognitive empathizing	This term is original and new to the literature
Ambiactive	Processes and systems which simultaneously stimulate and suppress complexity, sensitivity, and/or process cycle rates	Zukowski (2012), Carceroni et al. (2017)
Entrogenous	Systematic mediators of in-betweenness among the different phases and modalities of digitally augmented agency	This term is original and new to the literature

and solution search. In fact, the concept is already applied in some social and behavioral sciences. Third is the concept of ambiopia, which refers to double vision in ophthalmology, when one eye is myopic and the other is hyperopic. I use the term to describe processes which combine myopic and hyperopic, sampling and search, especially in problem-solving and cognitive empathizing. Fourth, the concept of empathicing is original and refers to satisficing in solving problems of other minds, rather than seeking to optimize in cognitive empathizing. Fifth, the concept ambiactive is borrowed from biology and refers to processes which simultaneously suppress and stimulate the same type of effect. Here the term refers to processes which both suppress and stimulate levels of complexity, sensitivity to variance, and processing rates. For example, an ambiactive system of augmented agency could suppress human sensitivity and processing rates, while also stimulating artificial hypersensitivity and hyperactive rates. This book also employs the established concept of ambidexterity, to describe the combination of different modes of human and artificial self-generation.

As noted previously, the prefix "ambi" is consistent, meaning "both" in Latin. It captures the fundamental combinatorics of augmented humanity, which integrates human and artificial agents. In fact, comparable concerns occur throughout Western thought. During the premodern period, for instance, agentic combinatorics focused on the relationship between human and divine beings. Whereas in modernity, scholars investigate the combination of autonomous, reasoning persons within social collectives. Both periods emphasize different combinatorics, reflecting the stage of social and technological development at the time. In the period of digitalization, greater focus will be on human-machine interaction. Granted, such combinatorics are observed in every period of civilization, albeit involving lower levels of technological sophistication and capability. Human-machine processing has always exhibited divergent rates, ranges, and levels of complexity, combining fast and slow, near and far, simplification and complexity. However, contemporary digitalization massively expands such effects. The scale, scope, and speed of digital augmentation are transformative. In consequence, augmented humanity will be characterized by dynamic agentic combinatorics. That said, human spirituality and autonomous reason will continue to matter greatly, but in

the context of increasingly augmented realities. The major risk will be that combining different human and artificial capabilities could result in distorted agentic forms and functions, which are either too divergent or convergent for a particular context.

Table 10.1 also includes another concept which is original and captures important novelties of the digitally augmented world, namely the concept of entrogenous, which refers to the systematic in-betweenness of digitalized mediators. Chapter 2 identifies three such mediators, which are central to augmented agency: intelligent sensory perception, performative action generation, and contextual learning. Together, they allow augmented agents to learn, compose, and recompose, in a dynamic fashion, updating form and function in real time. Importantly, these mediators are neither endogenous nor exogenous, relative to the boundaries they help to define. Rather, they are consistently in-between, processing potential form and function, and hence entrogenous. Recall that Fig. 2.4 illustrates this type of mediation. It depicts three levels and rates of processing and highlights the way in which human and artificial processes might diverge. The major driver of this effect is that artificial agents are inherently hyperopic, hyperactive, and hypersensitive, while humans are naturally myopic, relative sluggish, and insensitive. Hence, artificial and human processes could easily diverge in terms of their ranges, rates, and levels of precision and complexity.

Dilemmas of Digital Augmentation

Extreme divergence will manifest in numerous ways. This book exposes a number of critical manifestations. First, ambimodal distortion will create poorly integrated agentic forms and functions, which are overly compressed and layered at the same time. Second, ambiopic distortion will lead to problem-solving which is overly complex and simplified on different dimensions of problem representation and solution. Cognitive empathizing will be equally affected, when viewed as a type of complex problem-solving. Third, ambiactive distortion will produce dysfunctional self-regulation, evaluation of performance, and learning, in which relative human simplicity, sluggishness, and insensitivity diverge from artificial complexity,

hypersensitivity, and hyperactivity. As a further consequence, ambiactive distortion heightens the risk of cognitive dissonance, extreme ambiguity, and ambivalence, especially regarding core beliefs and commitments, where commitment in this context is defined as being dedicated, feeling obligated and bound, to some value, belief, or pattern of action (Sen, 1985). Fourth, these distortions compound to produce divergent ambidextrous self-generation, in which augmented agents adopt poorly synchronized, conflicting modes of human and artificial self-generation. In summary, digital augmentation could either enhance or diminish agentic form and function. Table 10.2 summarizes the resulting dilemmas of ambimodality, ambiopia, ambiactivity, and ambidexterity, plus the human and artificial tendencies for each, their potential risks and impact.

To mitigate these risks and maximize the opportunities of digitalization, human and artificial agents must therefore strengthen the supervision of their combinatorics. More specifically, when joined in augmented agency, human and artificial agents must be sensitive to contextual variance and ecological dynamics, while managing their complementary strengths and weaknesses. In doing so, they will regularly compose and recompose metamodels and methods. The primary goal will be to achieve and maintain maximal fit on every dimension. But to achieve this type of supervision, we need to develop the science of augmented agency.

10.2   Hyperparameters of Future Science

Humans are quintessentially agentic when they seek scientific understanding: purposive, forward looking, reflective, and self-directed. This includes the effort to interpret and explain their own patterns of thought and action. In this respect, civilized humans have always been their own object of interpretation and study. The agentic self has been a problem for the self, even in premodern worlds of narrative myth. In like fashion, the scientific study of augmented agency will be a major domain of augmented, agentic activity, which leads to an important insight: the science of augmented agency will exemplify the phenomena examined in earlier chapters. This science will be, itself, an expression of digitally augmented agency and its dilemmas, just as scientific thought and method are subjects of study in modern human science.

Table 10.2 Risks for digitally augmented agency

	Ambimodality	Ambiopia	Ambiactivity	Ambidexterity
Human tendency	Low modal compression, and hence layered agentic form and function	Myopic sampling and search (nearsighted, simplified processing)	Sluggish and insensitive (relatively low levels of activation)	Singular, mimetic modes of self-generation
Artificial tendency	High modal compression, and hence flattened agentic form and function	Hyperopic sampling and search (farsighted, complex processing)	Hyperactive and hypersensitive (relatively high levels of activation)	Flexible, original modes of self-generation
Combined tendency	Ambimodality produces mixed agentic forms and functions, which are compressed and layered	Ambiopic problem-solving which blends divergent myopia and hyperopia	Ambiactivity which both suppresses and stimulates activation mechanisms	Ambidextrous blend of singular human and flexible artificial modes of self-generation
Major risks	Incoherent, fragile, and ineffective, augmented agentic ambimodality	Highly divergent and ambiopic problem-solving, including cognitive empathizing	Highly ambiactive self-regulation, evaluation of performance, and learning	Highly divergent and incoherent patterns of ambidextrous self-generation

The science of augmented agency therefore faces the same challenges as other expressions of augmented agency. That is, issues arise regarding the specification of hyperparameters and metamodeling, plus the activation and upregulation, or deactivation and downregulation, of artificial and human processes. To begin with, metamodeling entails ontological hyperparameters, or the specification of fundamental categories of reality, both visible and hidden. Additional hyperparameters relate to epistemological properties, which specify logics and models of reasoning. Next, there are hyperparameters which define core activation and change

mechanisms, including potential sensitivity to variance and cycle rates. Following sections discuss each type of hyperparameter in relation to the science of augmented agency.

Ontological Principles

In the science of augmented agency, the fundamental categories of reality will transcend traditional conceptions of material nature and conscious mind. However, this does not imply the reduction of mind and consciousness to purely material cause. Rather, these categories are reconceived as higher order expressions of generative, augmented systems, which in turn result from complex neurophysiological, symbolic, and digital interactions. In this science, therefore, ontological commitments will be contextual, systematic, and rigorous. The resulting shift is comparable to earlier historical transitions. Just as the ancient concept of soul was demystified and naturalized by modernity, and human psyche then became a topic of science, so conscious mind will be digitally naturalized within the science of augmented agency (see Quine, 1995). In both cases, the shift is from anthropomorphic conceptions based on ordinary experience to a deeper understanding of reality which requires specialized techniques of observation and analysis. This also suggests that a new domain of enquiry may be required, focusing on the study of digitally augmented, agentic combinatorics (see Bandura, 2012; Latour, 2013). Neither the existing human sciences nor computer sciences adequately capture the forms and functions of augmented agency and mind. The recombination of prior fields is not enough. Radically new phenomena of this kind will require fresh concepts and frameworks.

In addition, the science of augmented agency will investigate novel forms of entrogenous mediation—previously defined as digitalized mediators of in-betweenness—which facilitate the dynamic composition and functioning of augmented agents. As noted earlier, this book has identified three such mechanisms: intelligent sensory perception, performative action generation, and contextual learning. It is important to stress, once again, that entrogeneity does not entail unfettered relativism or irregularity. Rather, the science of augmented agency will accommodate the dynamic generation of alternative categories and their boundaries. In this fashion,

entrogenous mechanisms will set and reset system boundaries, but they are neither endogenous nor exogenous with respect to these boundaries. Notably, this mirrors the approach to agentic hybridity proposed in numerous behavioral and social sciences (e.g., Battilana et al., 2015; Seibel, 2015). And not by coincidence, hybridity often emerges in digitalized contexts.

Epistemological Principles

Digital technologies also massively enhance intelligent processing capabilities. By leveraging these capabilities, augmented agents will gather and process information with far greater precision and speed. At least, expansion is feasible, notwithstanding persistent human limitations. In these respects, augmented agents will be bounded and unbounded, at the same time. This will occur, because human agents retain significant degrees of boundedness, especially in everyday cognitive functioning. Yet at the same time, artificial agents are increasingly unbounded. In effect, augmented agents will exhibit functional ambimodality with respect to rationality, as distinct from the organizational ambimodality discussed in Chap. 3. That is, digital augmentation will combine two different modes of reasoning, thinking far and fast, as well as near and slow. The supervisory challenge, therefore, is to manage the potential divergence or convergence of simultaneously bounded and unbounded, ambimodal patterns of reasoning.

Satisficing then becomes more dynamic and complex, and arguably more important. Most notably, because satisficing both simplifies and maximizes, it helps to mitigate the risks of overprocessing. Satisficing will constrain overly hyperopic sampling and search, and overly hypersensitive and hyperactive responses to variance. Hence, in addition to satisficing because of limited capabilities, as Simon (1955) originally argued, augmented agents will also satisfice to restrain excessive capabilities. Put another way, digitally augmented agents will satisfice, not only because of limits, but to impose limits. They will choose to satisfice, even when ideal optimization is feasible, to avoid unnecessary processing. In fact, artificial systems do this already, when they limit their own processes to improve speed and efficiency. Augmented agents will do the same, choosing to forgo optimization for good reasons, just as humans already do (Gigerenzer

& Gaissmaier, 2011). Hyperparameters will specify these epistemic features in metamodels of augmented science.

Another major epistemic shift is the extension of systematic reasoning to problem sampling and representation. This occurs because intelligent sensory perception will allow augmented agents to apply systematic reasoning to problem sampling and representation. In such a world, problems will emerge in an intelligent fashion, similar to the sampling and representation of problems in empirical science. By contrast, even in the recent past, the ordinary sampling and representation of problems are not viewed as reasoned activities. At most, they involve selective attention and observation (Ocasio, 2012). Intelligent sampling and representation only consistently occur in expert domains, such as experimental science. Even behavioral research rarely focuses on the cognitive-affective mechanisms of sampling and problem representation (Fiedler & Juslin, 2006). Similarly, behavioral research largely neglects normative satisficing. Rationality is applied to solution search, not to problem sampling and representation. Digital augmentation upends these assumptions and suggests a fusion of ecological realism and rationality.

Furthermore, digitalization supports the dynamic composition of metamodels of reasoning, using methods which can be described as "compositive" (see Latour, 2010). Such methods do not rely on predetermined models or axioms, nor do they rely on traditional descriptive and normative templates. Rather, compositive methods employ digitalized processes to develop customized metamodels which best fit the problem context. At the same time, compositive methods are systematic and rigorous, neither ad hoc nor idiosyncratic (e.g., Pappa et al., 2014; Wang et al., 2015). In fact, many digital systems already exhibit these capabilities. As noted in earlier chapters, advanced artificial agents are already compositive in this sense, and require minimal or no supervision. Evolutionary deep learning systems and GANs function in exactly this way (Shwartz-Ziv & Tishby, 2017). Via rapid inductive, abductive, and reinforcement learning, they process massive volumes of information, identifying hitherto undetectable patterns, to compose new explanatory methods and models without external guidance. In this fashion, generative metamodeling will translate the techniques of experimental computer science into all domains of augmented agency.

Mechanisms of Adaptation and Change

Hyperparameters also specify core mechanisms and processing rates. During modernity, the human and natural sciences bifurcated, in these respects. Within the biological sciences, the major change mechanisms are organic processes of variation and selection. In contrast, within many human sciences, conscious thought, will, and intention are seen as primary drivers of change. In consequence, modern scholarship often struggles to integrate bifurcated science. Scholars are unsure how to integrate biological processes of random variation, natural or ecological selection, and material cause, with conscious processes of intentional variation, preferential choice, and intelligent cause. Polarizing debates therefore persist about materialism versus idealism, the distinction of mind from body, reductionism versus holism, and positivist versus interpretive explanation. These bifurcations also partly explain the poor integration of ecological and behavioral mechanisms, especially during modern industrialization (Latour, 2017).

Herbert Simon (2000) had foresight on these issues as well. At the dawn of the twenty-first century, in the last year of his life, he proposed three priorities for digitally augmenting humanity. They were his minimal requirements for "designing a sustainable acceptable world." In effect, he described a program of global recomposition or digitally augmented worldmaking. First, he argued that humanity must learn to live at peace with all of nature, in a sustainable collaborative way, and overcome the "false pride" of being separate from, and superior to, the rest of the natural world. Second, he argued that humanity must share goods and wealth fairly and productively, so that all persons will enjoy comparable benefits and opportunities. Third, to achieve such fairness, he said humanity must eliminate the divisions which arise from cultural and social antipathy and stop viewing the world in terms of "we versus them." In fact, Simon was rejecting the classic bifurcations of modernity, that mind and consciousness are distinct from nature, that the autonomous self stands apart from the other, and that empathy is inevitably limited and local.

Simon was correct, then and now. Just as he predicted, digitalization problematizes the conceptual architecture of modernity. Augmented agents will better connect material nature and conscious mind. Likewise, the science of augmented agency will synthesize the study of human

agency with the natural and computer sciences. Research methods will be contextual and compositive, adapting to maximize and maintain metamodel fit. Entrogenous mediation will be critical, and many polarities will thus resolve, as we incorporate intelligent sensory perception, performative action generation, and contextual learning. In a digitally augmented world, moreover, change will occur through generative variation and intelligent adaptation, not merely through random mutation and natural or ecological selection. Agentic evolution will be experimental and intelligent, similar to Gregor Mendel's cultivation of new plant varieties through guided variation and selection (Levinthal, 2021). It is also likely that in future, advanced digital systems will fully integrate with the biological, geophysical world. When this occurs, digitalization will augment organic variation and selection as well. Augmented agency could become a truly positive force in the natural world, enabling self-generation and renewal, rather than destruction and exploitation. All this is possible, assuming a future science of augmented agency and appropriate supervision of its application. The overall effect would be transformative.

Table 10.3 summarizes the paradigmatic shift just described. It shows three historical periods—premodern, modern, and digitalization—their major ontological and epistemological commitments, plus the dominant mechanisms of change and scientific methods. Most notably, the table summarizes the emerging shift toward generative, augmented pluralism and compositive methods. It is also important to note that all three systems may continue adding value to agentic experience and understanding, assuming appropriate supervision and application.

10.3   Domains of Augmented Science

While the future unfolds, contemporary human science still grapples with the dilemmas of modernity. Numerous dialectics accompany these concerns: explaining the interaction of nature and nurture; how material cause relates to meaning and intention; developing autonomous personality as well as sociable collectivity; seeking order and continuity while embracing change (Giddens, 2013). Reflecting these dialectics, the human sciences divide into separate disciplines, most of which focus on different agentic modalities and functional domains. For example,

Table 10.3 Summary of scientific metamodels

Metamodel	Ontology	Epistemology	Mechanisms	Methods
Replicative, premodern, monism	Essential categories, forms, and states	Revealed truth and narrative order	Teleological final cause and imitation	Narrative and discursive
Adaptive, modern, dualism	Material nature, distinct from mind and consciousness	Axiomatic reasoning and interpretative sense-making	Organic variation and selection, and iterative adaptive learning	Quantitative, qualitative, multiple, and mixed
Generative, digitally augmented, pluralism	Augmented fusion of material, artificial, and human forms	Generative, composed models of reasoning and judgment	Intelligent variation and selection, and continual real-time learning	Compositive, contextual, and blended

psychology focuses on the study of mind and behavior within individuals, groups, and collectives. In contrast, sociology focuses on social life and collectivity and then examines the role of individuals in these contexts. Other human sciences, such as education and management studies, combine modalities in particular activity domains. Contemporary research also organizes around multidisciplinary, hybrid approaches to complex problems (Seibel, 2015; Skelcher & Smith, 2015). In this respect, contemporary human science recognizes the increasing integration and interdependency of agentic modalities and contexts. Ecological and environmental factors receive increasing attention as well.

Digitalization accelerates these trends. It also generates new questions for human science, especially regarding the dilemmas of augmented combinatorics, or how to combine human and artificial agents. To investigate these questions, the future science of augmented agency will organize around complex problems too. It will be less divided into siloed disciplines, and less oriented toward different modalities (Latour, 2011). Disciplinary categories and boundaries will be more flexible and fluid. In fact, recent scholarship is moving in this direction already, illustrated by ecological theories of social organization, and neurocognitive models of personality and culture (Chimirri & Schraube, 2019; Kitayama & Salvador, 2017). Through this type of research, scholars develop

multidisciplinary theories of agentic form and function (Fiedler, 2017). This will be the norm in the science of digitally augmented agency.

Furthermore, the science of augmented agency will treat modality itself as generative and contextual. In fact, some scholars already view agentic modality as epiphenomenal to performance, meaning it is mediated by action in context, rather than expressing autonomous form (Hwang & Colyvas, 2021; Pentland et al., 2012). Collective hybridity emerges in this fashion too. Postmodern thinkers go even further. They view autonomous agency as chimerical, a device which dissolves in the deconstruction of text and context. Many of these thinkers take inspiration from Freud's argument that conscious ego reflects the hidden subconscious (Tauber, 2013). However, my proposals take a markedly different approach. They anticipate a systematic, empirical science of emergent phenomena, with clearly defined metamodels and mechanisms.

To illustrate such a science, consider the needs of autonomous mobility systems, in which human and artificial agents collaborate as augmented agents. These systems will digitalize and integrate every level and modality of agency, both organizational and functional. Smart cities will digitalize the transport infrastructure, to create the necessary environment for immediate contextual learning. Vehicle manufacturers will incorporate intelligent sensory perception which supports fully augmented problem representation and feedforward response. Advanced, empathic artificial agents will be embedded throughout, enabling performative action generation in real time. And network management agents will supervise and govern the entire system, to ensure efficiency, safety, sustainability, and social inclusion. In summary, the augmented science of autonomous mobility systems will be generative, compositive, and integrate multiple disciplines, technologies, and human factors.

Science of Consciousness

These developments impact the role of ordinary consciousness in the science of augmented agency. Many disciplines research such questions, including the philosophy of science, cognitive psychology, and the human sciences more broadly (Metcalfe & Schwartz, 2016). However, the nature

and role of consciousness are not yet fully understood. That said, ample evidence shows that ordinary consciousness is an imperfect means of observation in rigorous, scientific pursuits. Unassisted, it often leads to anthropomorphic assumptions which are inherently myopic and misleading. For this reason, ordinary consciousness will have a different role in the science of augmented agency. It will be less a means of access to fundamental reality and truth, and more a source of humanistic reference for augmented agency, which is an equally vital role. Ordinary consciousness will remain important, therefore, but for different reasons, compared to the past.

However, as the history of science shows, humanity always struggles to reset the role of consciousness as a source of reality and truth. Over time, the trend is to expose anthropomorphic assumptions and demote the status of consciousness as such. For example, as noted earlier in this chapter, the ancient soul was naturalized to become a topic of study for modern psychological science. Such shifts often incite trouble, because they threaten embedded narratives and identities. Similar shifts were primary sources of opposition to Copernican cosmology, Galilean mechanics, and Darwin's theory of evolution. Intuitions of the world run deep and are resilient. Natural scientists acknowledged this problem long ago and worked hard to liberate their thinking. They largely succeeded. Every educated person now knows that it takes sophisticated technological means to observe the deeper realities of the physical world.

In the human sciences, by contrast, there are ongoing debates. Some maintain that ordinary consciousness does provide access to the fundamental realities of human form and function. In branches of linguistics and psychology, for example, some rely on self-reports to illuminate core processes of language acquisition and reasoning. By implication, they believe that subjective mental states can be treated as primitive and are not decomposable. Others disagree and argue that ordinary consciousness is not adequate for such purposes (e.g., Wilson & Dunn, 2004). They contend that, just as natural science demoted ordinary consciousness and turned to technological tools and formal methods, the human sciences must do the same. Granted, the humanities will continue to treat mind and consciousness as fundamental. These disciplines are concerned with the interpretation of hermeneutic and cultural phenomena. But any attempt at an empirical science of human agency—and

especially digitally augmented agency—must adopt the tools and techniques of neurophysiology, cognitive psychology, computer science, and the like. How phenomena transform into consciousness and subjective mental states are then questions for ongoing research (Sohn, 2019). In like fashion, the future science of augmented agency will investigate the role of consciousness in humanistic supervision, and how best to regulate its influence (see Lovelock, 2019).

Generative Commitments

This also points toward a science of generative commitments. That is, augmented agents will have the capability to relax, update, and recompose their commitments, which constitutes another significant departure from traditional assumptions. Throughout most of history, cultures have assumed that core commitments and reference criteria are fixed, often inviolable. Deviation has prompted sanction and conflict. It still does, in many places. More recently, however, as humanity becomes globally connected and mobile, commitments are more pluralistic and embracing, even if such pluralism sometimes triggers anxiety and antipathy, which is not surprising, given the deep role of shared commitments in culture and identity (Appiah, 2010). To be sure, many traditional commitments warrant preservation. If supervision is flawed, important aspects of human experience could erode, including shared commitments about reality, truth, beauty, and justice. Noting these risks, the science of augmented agency will need to resolve how to supervise generative commitments.

Some already research the closely related topic of holistic value. New theories of economics and management, for example, incorporate diverse concepts of human welfare and socioeconomic value creation (e.g., Raworth, 2017; Sachs et al., 2019; Stiglitz et al., 2009). Similarly, in contemporary theories of agency itself, scholars are expanding their conception of human flourishing and psychosocial well-being to accommodate richer, alternative commitments (Seligman et al., 2013). Some psychologists are exploring new theories of virtue, referencing classic thinking about holistic well-being (Fowers et al., 2021). In addition, as noted earlier, agentic hybridity is increasingly recognized in many fields. The future

science of generative commitments can build on these contributions. In fact, this prospective enquiry harks back to Aristotle's (1980) concept of eudaimonia, about living a good life with practical wisdom. From this perspective, a science of generative commitments will be a science of eudaimonics. It will study how to compose and live a complete, flourishing life in a digitalized world. The inquiry would encompass all value commitments, complementing the existing study of specific types of value in economics, ethics, and aesthetics (see Di Fabio & Palazzeschi, 2015).

For example, consider digitally augmented health care. In these contexts, emerging priorities are overall health, well-being, and quality of life, or in other words, eudaimonic concerns. Systems will be designed and evaluated based on holistic human outcomes, and not simply on crude metrics of service delivery. Once maximized in this way, health care will be value based, personal, precise, and fully integrated into social life. Relevant technologies will include wearable and implantable devices. Importantly, this kind of system will require generative commitments, developing and adapting values and goals for both individuals and collectives. The responsible, augmented agents will recognize and/or generate a range of cultural, social, and personal commitments and preferences. There will also be empathic artificial agents enabling performative action generation in real time. Generative commitments will thus guide the design and delivery of services, while network management agents will supervise and govern the entire system. In this fashion, augmented agents in health care will generate commitments.

As Aristotle further understood, shared commitments underpin the good governance of communal life (Nussbaum, 2000). In ancient Athens, this was centered in the polis, its rituals, and celebrated by dramatic chorus. Whereas, in the modern period, good governance calls for reasoned public debate, the fair determination of collective choice, and participatory decision-making. In the period of digitalization, the governance of collective agency will be transformed as well. For example, civic participation could become more inclusive and globally integrated. As in the past, therefore, a new period of agentic experience will require a fresh approach toward collective governance and politics. And if history is any guide, we should expect to see more strife and struggle in this regard, as the impact of digitalization continues to grow. Institutional systems of power and

influence are never easy to change, and digitalized societies will be no different. We see evidence of such conflict already, as opposing political and cultural groups struggle to control online social networks.

Science and History

Throughout this book, history is a guide. The argument consistently refers to three major periods of civilized humanity and agency: premodernity, modernity, and the contemporary period of digitalization. Among other key features, each period is characterized by stages of technological assistance: from the primitive technologies of premodernity, to the mechanical and analogue technologies of modernity, to the digital and neural technologies of the contemporary period. Hence, my argument also speaks to the history of human science, conceived as the study of human self-understanding over time. In fact, the agentic metamodels presented in this book constitute a broad framework for reconceiving the history of human science.

Some historians take an equally broad perspective on the past. This is true of the Annals School, founded in the mid-twentieth century by Lucien Febvre and Marc Bloch (2014). For them, history is a long narrative of unfolding worlds of lived experience and mentality. Politics and princes are then expressions of their periods, not the primary forces of history. Thomas Kuhn's (1970) work on paradigms of scientific investigation and knowledge is equally broad and long term. In fact, his analysis of paradigms could be restated in terms of metamodels and their hyperparameters. Each successive paradigm exhibits major shifts in core ontology, epistemology, and mechanisms of knowledge generation and diffusion. In many ways, Kuhn's view of the past aligns with the historical perspective of this book. Both identify long periods and general frameworks, although my argument articulates alternative processes, mechanisms, and metamodels and tries to bring fresh clarity and organization to this narrative.

More practically, digitalization accelerates historical time. During premodernity, rates of change were slow and often imperceptible. Societies were relatively stable and evolved slowly. For this reason, the premodern concept of historical time was expressed in legend and myth, rather than narratives of social and political change. Then during modernity, history

accelerated. Indeed, for modern societies and persons, historical change is a central feature of communal and autobiographical narrative, not some distant horizon or myth, although for many indigenous cultures, the modern compression of time has been a source of cultural distress and alienation. They struggle to maintain traditional narratives in the face of imperialistic and industrial forces. In a digitalized world, history accelerates yet again. Now all humanity will share the indigenous struggle to maintain cultural narratives. In these respects, augmented humanity can look to indigenous peoples for lessons about cultural survival in the face of overwhelming social and technological change (Hogan & Singh, 2018).

For without doubt, in a digitalized world, dynamic change will be constant and ubiquitous. This will not be the end of history, by any means, but it does imply significant acceleration. Historical transformation will occur within generations and seasons, not only across the life span. Viewed positively, this will enable a new type of self-generativity, empowering augmented humanity to make and remake the world, while living within it (Latour, 2013). Augmented humanity will move "off the edge of history," as Anthony Giddens (2015) puts it, by compressing and transcending the classic parameters of historical time. Change will be discontinuous and the past will explain less and less about the future. But what comes next is not yet assured. Moving off the edge of history can be perilous or liberating. Regarding peril, some people and communities might lose their bearings or surrender to artificial control. In terms of liberation, a new period of self-generative freedom and flourishing is possible, assuming humanity meets the supervisory challenge of digital augmentation.

Research Methods

Additional consequences follow for research methods. In standard approaches, researchers in the human sciences gather qualitative data to support descriptive, interpretive models of human experience and behavior, and quantitative data to support calculative, causal models of such phenomena. The former methods focus on rich, holistic description, narratives, and sense-making, hoping to interpret meaning and intention, while the latter methods seek measurable data and discrete mechanisms, to explain causation. Multiple and mixed methods blend these

approaches. Scholars debate which approach is more reliable and enlightening: rich, holistic descriptions of experience and the interpretation of meaning, or measurable mechanisms of variation in causal explanation; or some combination of these approaches (Creswell, 2003). Of course, this takes us back to Herbert Simon (1979) again, and the dilemmas of simplification in the modeling of human thought and behavior.

In parallel, scholars debate ontological and epistemological priorities. On the one hand, those who privilege qualitative methods and interpretation, typically argue that holistic description, consciousness, and meaning take priority and cannot be reduced to mechanistic cause, while on the other hand those who privilege quantitative methods and causation argue that functional mechanisms and assisted observation take priority and reject any reliance on subjective meaning and interpretation. Not surprisingly, many regard qualitative and quantitative methods as deeply incommensurable. That said, a significant research community now advocates for blended, mixed, and multiple methods (Denzin, 2010).

In a period of digitalization, these distinctions will blur even further and faster. For example, it is already possible to achieve machine-based outcomes which were previously deemed impossible, such as automated pattern recognition, associative computation, artificial empathy, intuition, and creativity (Choudhury et al., 2020; Varshney et al., 2015). Quite simply, digital systems are replicating many of the more complex, holistic functions of human cognition. As noted in earlier chapters, within the foreseeable future, there will be no detectable difference between human and artificial agents in these domains, although whether artificial agents should be classified as truly sentient and conscious is another question. Nevertheless, in consequence of these developments, it is feasible to compose blended research methods at massive scale. Different tools and techniques will be combined and recombined to match problem contexts. Studies will apply quantitative techniques to the interpretation of meaning, including self-narratives and sense-making, while also scaling qualitative techniques to predict complex patterns of thought and behavior.

Using such compositive methods, augmented science will customize different techniques of sampling and search to the phenomena and questions of interest. Here again, entrogenous mediators will play a central role, updating metamodels and methods in real time. Therefore, just as the metamodels of augmented science will be contextual and generative,

so will be the methods used to gather, interpret, and analyze information. Methods will be composed to fit the problem space. They will be compositive, as Hayek (1952) originally proposed, not simply qualitative, quantitative, or mixed, in the traditional sense. In fact, advanced forms of artificial intelligence and machine learning already do this (e.g., Mehta et al., 2019). Some social scientists do as well (e.g., Latour, 2011).

Future Prospect

James March (2006), the great scholar of organizations, argues that it requires courage and positive deviance to embrace the ambiguity and ambivalence of exploratory thought. It also requires patience and persistence, to see whether fruits ripen or not. And it should, given the need for rigor and replication. However, the science of augmented agency calls for extra effort and speed, in these respects. Digitalization is rapidly infusing agentic domains, bringing unprecedented gains in capability and potentiality. In consequence, it problematizes the traditional assumptions of modernity, and presents new and urgent challenges. Most particularly, the combinatorics of digitally augmented humanity are transforming and confronting. Human agents will likely remain relatively myopic, sluggish, layered, and insensitive to variance, while artificial agents will be increasingly farsighted, fast, compressed, and hypersensitive. As both collaborate more closely, they risk amplifying the tendencies of the other, leading to internal divergence or convergence and dysfunction.

   Novel problematics and dilemmas emerge. Inadequate supervision of these could produce the following dysfunctions: highly ambimodal systems, resulting in incoherent agentic form and function; highly ambiopic problem-solving and cognitive empathicing will skew judgments of the world and other minds; highly ambiactive self-regulation, evaluation of performance, and learning, would risk incoherence, extreme ambiguity and ambivalence; all contributing to dysfunctional patterns of ambidextrous, human and artificial self-generation. To mitigate these risks, human and artificial agents must develop the capability for collaborative supervision grounded in mutual understanding, trust, and respect. Achieving all this will be contingent on the development of a science of augmented agency. The core features of this science will include the following:

digitally augmented mind will be treated as a fundamental category of reality; the science of augmented agency will employ contextually sensitive, compositive methods; its metamodels will be highly generative, rather than replicative or slowly adaptive; problem sampling and representation will be intelligent and reasoned, complementing ecological rationality; augmented agency will rely deeply on the entrogenous mediation of intelligent sensory perception, performative action generation, and contextual learning; ordinary consciousness and commitments will play important roles in humanizing the science of augmented agency, rather than being sources of fundamental insight about the world itself.

Granted, the exact shape of this future science is not yet clear. Much of the current chapter—indeed, this book as a whole—is therefore prospective. It anticipates the future, grounded in the best knowledge currently available, while acknowledging that its proposals will require further elaboration and testing. Nor is this book a comprehensive treatment of the phenomena. Rather, it takes steps toward a science of augmented agency. But the process remains emergent. The trajectory of digital augmentation could change, as the natural, human, and virtual worlds continue evolving, interacting, and often conflicting. That said, we need to move forward. Prospective theorizing helps, by shedding light on unfamiliar territory. The history of science also teaches that radically new phenomena often require fresh conceptual architecture. Existing frameworks rarely suffice and waiting for certainty and normality is unlikely to succeed. Digitalization is too novel and dynamic. We could wait in vain, while the world moves on. This would be unproductive and arguably negligent, given the accelerating impact of digitalization. The augmentation of humanity has clearly begun. Its dilemmas are present and increasingly urgent. Science must respond with matching speed and purpose.

References

Appiah, K. A. (2010). *Cosmopolitanism: Ethics in a world of strangers*. WW Norton & Company.

Aristotle. (1980). *The Nicomachean Ethics* (D. Ross, Trans.). Oxford University Press.

Bandura, A. (2012). On the functional properties of perceived self-efficacy revisited. *Journal of Management, 38*(1), 9–44.

Battilana, J., Sengul, M., Pache, A.-C., & Model, J. (2015). Harnessing productive tensions in hybrid organizations: The case of work integration social enterprises. *Academy of Management Journal, 58*(6), 1658–1685.

Bloch, M. (2014). *Feudal society*. Routledge.

Buetow, S. A. (2020). Psychological preconditions for flourishing through ultrabilitation: a descriptive framework. *Disability and rehabilitation, 42*(11), 1503–1510.

Carceroni, R. L., Sanketi, P. R., Shah, S., Ozkan, D., Mariooryad, S., Tarzjani, S. M. S., Lider, B., & Ludwig, P. W. (2017). *Touchless user interface navigation using gestures*. U.S. PTO.

Chen, S., Yu, P., & Houk, K. N. (2018). Ambimodal dipolar/diels-aldercycloaddition transition states involving proton transfers. *Journal of the American Chemical Society, 140*(51), 18124–18131.

Chimirri, N. A., & Schraube, E. (2019). Rethinking psychology of technology for future society: Exploring subjectivity from within more-than-human everyday life. In *Psychological studies of science and technology* (pp. 49–76). Springer International Publishing.

Choudhury, P., Allen, R. T., & Endres, M. G. (2020). Machine learning for pattern discovery in management research. *Strategic Management Journal, 42*(1), 30–57.

Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches* (2nd ed.). Sage.

Denzin, N. K. (2010). Moments, mixed methods, and paradigm dialogs. *Qualitative Inquiry, 16*(6), 419–427.

Di Fabio, A., & Palazzeschi, L. (2015). Hedonic and eudaimonic well-being: The role of resilience beyond fluid intelligence and personality traits. *Frontiers in Psychology, 6*, 1367.

Eden, A. H., Moor, J. H., Søraker, J. H., & Steinhart, E. (2015). *Singularity hypotheses*. Springer.

Fiedler, K. (2017). What constitutes strong psychological science? The (neglected) role of diagnosticity and a priori theorizing. *Perspectives on Psychological Science, 12*(1), 46–61.

Fiedler, K., & Juslin, P. (Eds.). (2006). *Information sampling and adaptive cognition*. Cambridge University Press.

Fowers, B. J., Carroll, J. S., Leonhardt, N. D., & Cokelet, B. (2021). The emerging science of virtue. *Perspectives on Psychological Science, 16*(1), 118–147.

Giddens, A. (2013). *The consequences of modernity*. Wiley.

Giddens, A. (2015). Off the edge of history. *The world in the 21st century. Teacher Training Agency Annual Lecture*, London.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology, 62*, 451–482.

Hayek, F. A. (1952). *The counter-revolution of science: Studies on the abuse of reason*. The Free Press.

Hogan, T., & Singh, P. (2018). Modes of indigenous modernity: Identities, stories, pathways. *Thesis Eleven, 145*(1), 3–9.

Hwang, H., & Colyvas, J. A. (2021). Constructed actors and constitutive institutions for a contemporary world. *Academy of Management Review, 46*(1), 214–219.

Kitayama, S., & Salvador, C. E. (2017). Culture embrained: Going beyond the nature-nurture dichotomy. *Perspectives on Psychological Science, 12*(5), 841–854.

Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd ed.). University of Chicago Press.

Latour, B. (2010). An attempt at a "compositionist manifesto". *New Literary History, 41*(3), 471–490.

Latour, B. (2011). From multiculturalism to multinaturalism: What rules of method for the new socio-scientific experiments? *Nature and Culture, 6*(1), 1–17.

Latour, B. (2013). *An inquiry into modes of existence* (C. Porter, Trans.). Harvard University Press.

Latour, B. (2017). Anthropology at the time of the anthropocene: A personal view of what is to be studied. In *The anthropology of sustainability* (pp. 35–49). Springer.

Levinthal, D. A. (2021). *Evolutionary processes and organizational adaptation: A mendelian perspective on strategic management.* Oxford University Press.

Lovelock, J. (2019). *Novacene: The coming age of hyperintelligence*. MIT Press.

March, J. G. (2006). Rationality, foolishness, and adaptive intelligence. *Strategic Management Journal, 27*(3), 201–214.

Mehta, Y., Majumder, N., Gelbukh, A., & Cambria, E. (2019). Recent trends in deep learning based personality detection. *Artificial Intelligence Review*, 1–27.

Metcalfe, J., & Schwartz, B. L. (2016). *The ghost in the machine: Self-reflective consciousness and the neuroscience of metacognition*. Oxford University Press.

Nussbaum, M. C. (2000). Aristotle, politics, and human capabilities: A response to Antony, Arneson, Charlesworth, and Mulgan. *Ethics, 111*(1), 102–140.

Ocasio, W. (2012). Attention to attention. *Organization Science, 22*(5), 1286–1296.

Pappa, G. L., Ochoa, G., Hyde, M. R., Freitas, A. A., Woodward, J., & Swan, J. (2014). Contrasting meta-learning and hyper-heuristic research: The role of evolutionary algorithms. *Genetic Programming and Evolvable Machines, 15*(1), 3–35.

Pentland, B. T., Feldman, M. S., Becker, M. C., & Liu, P. (2012). Dynamics of organizational routines: A generative model. *Journal of Management Studies, 49*(8), 1484–1508.

Quine, W. V. (1995). Naturalism; or, living within one's means. *Dialectica, 49*(2–4), 251–263.

Raworth, K. (2017). *Doughnut economics: Seven ways to think like a 21st-century economist.* Chelsea Green Publishing.

Remeseiro, B., Barreira, N., Sanchez-Brea, L., Ramos, L., & Mosquera, A. (2018). Machine learning applied to optometry data. In *Advances in biomedical informatics* (pp. 123–160). Springer.

Sachs, J. D., Schmidt-Traub, G., Mazzucato, M., Messner, D., Nakicenovic, N., & Rockström, J. (2019). Six transformations to achieve the sustainable development goals. *Nature Sustainability, 2*(9), 805–814.

Seibel, W. (2015). Studying hybrids: Sectors and mechanisms. *Organization Studies, 36*(6), 697–712.

Seligman, M. E. P., Railton, P., Baumeister, R. F., & Sripada, C. (2013). Navigating into the future or driven by the past. *Perspectives on Psychological Science, 8*(2), 119–141.

Sen, A. (1985). Goals, commitment, and identity. *Journal of Law, Economics & Organization, 1*(2), 341–355.

Shwartz-Ziv, R., & Tishby, N. (2017). Opening the black box of deep neural networks via information. *arXiv preprint arXiv:1703.00810.*

Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics, 69*(1), 99–118.

Simon, H. A. (1979). *Models of thought* (Vol. 352). Yale University Press.

Simon, H. (2000). Keynote address. Earthware Symposium, Carnegie Mellon University.

Skelcher, C., & Smith, S. R. (2015). Theorizing hybridity: Institutional logics, complex organizations, and actor identities: The case of nonprofits. *Public Administration, 93*(2), 433–448.

Smolarz-Dudarewicz, J., Poborc-Godlewska, J., & Lesnik, H. (1980). Comparative evaluation of the usefulness of the methods of studying binocular vision for purposes of vocational guidance. *Medycyna Pracy, 31*(2), 109–114.

Sohn, E. (2019). Decoding the neuroscience of consciousness. *Nature, 571*(7766), S2–S5.

Stiglitz, J. E., Sen, A., & Fitoussi, J.-P. (2009). *Report of the commission on the measurement of economic performance and social progress.*

Tauber, A. I. (2013). *Requiem for the ego: Freud and the origins of postmodernism.* Stanford University Press.

Tunyi, A. A., Ntim, C. G., & Danbolt, J. (2019). Decoupling managementinefficiency: Myopia, hyperopia and takeover likelihood. *International Review of Financial Analysis, 62*, 1–20.

Varshney, L. R., Wang, J., & Varshney, K. R. (2015). Associative algorithms for computational creativity. *The Journal of Creative Behavior, 50*(3), 211–223.

Wang, B., Xu, S., Yu, X., & Li, P. (2015). Time series forecasting based on cloud process neural network. *International Journal of Computational Intelligence Systems, 8*(5), 992–1003.

Wilson, T. D., & Dunn, E. W. (2004). Self-knowledge: Its limits, value, and potential for improvement. *Annual Review of Psychology, 55*, 493–518.

Yang, Z., Dong, X., Yu, Y., Yu, P., Li, Y., Jamieson, C., & Houk, K. N. (2018). Relationships between Product Ratios in Ambimodal Pericyclic Reactions and Bond Lengths in Transition Structures. *Journal of the American Chemical Society, 140*(8), 3061–3067.

Zukowski, M. M. (2012). *Genetics and biotechnology of bacilli.* Elsevier.