Cyclical Population Dynamics of Automatic Versus Controlled Processing: 
An Evolutionary Pendulum

David G. Rand
Yale University

Damon Tomlin
University of Colorado, Colorado Springs

Adam Bear
Yale University

Elliot A. Ludvig
University of Warwick and Princeton University

Jonathan D. Cohen
Princeton University

Psychologists, neuroscientists, and economists often conceptualize decisions as arising from processes that lie along a continuum from automatic (i.e., “hardwired” or overlearned, but relatively inflexible) to controlled (less efficient and effortful, but more flexible). Control is central to human cognition, and plays a key role in our ability to modify the world to suit our needs. Given its advantages, reliance on controlled processing may seem predestined to increase within the population over time. Here, we examine whether this is so by introducing an evolutionary game theoretic model of agents that vary in their use of automatic versus controlled processes, and in which cognitive processing modifies the environment in which the agents interact. We find that, under a wide range of parameters and model assumptions, cycles emerge in which the prevalence of each type of processing in the population oscillates between 2 extremes. Rather than inexorably increasing, the emergence of control often creates conditions that lead to its own demise by allowing automaticity to also flourish, thereby undermining the progress made by the initial emergence of controlled processing. We speculate that this observation may have relevance for understanding similar cycles across human history, and may lend insight into some of the circumstances and challenges currently faced by our species.

Keywords: automaticity, cognitive control, cognitive processing, evolution, mathematical modeling

Cognitive processes have long been conceptualized as lying along a continuum from automatic to controlled (Allport, 1954; Cohen, Dunbar, & McClelland, 1990; Kahneman & Treisman, 1984; Posner & Snyder, 1975; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). More automatic processes are more “hard-wired” or overlearned, which leads to greater efficiency (e.g., greater speed and less effort) at the cost of reduced flexibility and less ability to adjust to the details of the current situation. More controlled processes, conversely, involve more deliberation and thought—requiring greater investment of time and effort, but allowing a greater degree of flexibility and sensitivity to specifics and/or circumstances of the particular decision. This distinction remains a fundamental tenet of cognitive psychology that has continued to be a focus of intensive research (Botvinick & Cohen, 2014; Evans & Stanovich, 2013), influencing thinking in the behavioral and neurobiological sciences more generally (e.g., Cushman (2013); Fudenberg and Levine (2006); Hare, Camerer, and Rangel (2009); Kahneman (2003, 2011); McClure, Laibson, Loewenstein, and Cohen (2004); Miller and Cohen (2001); Rand, Greene, and Nowak (2012); Stanovich and West (2000)).

The remarkable capacity for controlled processing is considered one of the distinguishing characteristics of human cognition. Often termed “cognitive control,” it is critical to every faculty that is considered to be distinctively human, including reasoning, problem solving, planning, and symbolic language, and the role that these play in the formation and function of societies. Given the virtues of controlled processing, and the externalities to which it gives rise (e.g., advanced technologies in virtually every domain of human function, including agriculture, housing, transportation, communication and large-scale economics), one might imagine
that the prevalence of cognitive control among agents in a population would be directly (positively) associated with the fitness of that population. If so, two corollaries would seem to follow: the spread of controlled processing within a population should be inexorable, and that spread should be associated with the inexorable success of the population. Here, we challenge these conclusions on theoretical grounds.

The work we present is inspired in part by observations of human history. Anthropological evidence suggests that many cultures that have developed advanced technologies—presumably evidence of the emergence of cognitive control within at least some proportion of the population—have ultimately met with demise (e.g., Diamond (2005); Richerson, Boyd, and Bettinger (2009); Schwindt et al. (2016); Turner and Sabloff (2012)). It is possible that this demise could have been induced by fully exogenous factors beyond the influence of the population (e.g., environmental shocks). It is also possible, however, that at least in some cases this demise could have arisen from a failure of the population as a whole to act with foresight (i.e., in a manner that controlled processing would seem to make possible)—for example, using new technologies to consume resources in an unsustainable way, leading to exhaustion of the environment or increased vulnerability to environmental shocks.

Although the empirical evidence is at best suggestive, we believe that this possibility of controlled-based success within a population leading to lack-of-control-based failure is intriguing, and warrants formal exploration. In part, this is because it raises a question that has potentially profound consequences for the present circumstances of our species: why do populations comprised of agents with the rational faculties necessary to produce sophisticated technologies sometimes fail to act as rational stewards of such technologies? There are of course many possible responses to this question. One is that modern circumstances themselves refute the assertion: we have not met with, nor are we at risk of, such failure—the risks are either contrived, or are ones that new technological innovations will resolve. This is a possibility. However, the cost of at least considering the risks of failure seems small, while the danger of failing to accurately identify and respond to that risk seems infinitely greater.

It is in this spirit that we consider potentially fundamental dynamics that may occur as the capacity for cognitive control emerges in a population, spreads, and affects the environment — dynamics that we show can lead to cycles of growth followed by dramatic collapse as easily as they can lead to stability. Describing these dynamics in formal terms may, at the least, lay the foundation for exploring their relevance to the complex circumstances in which the human species presently finds itself. At best, they may help identify factors that could be leveraged to mitigate potential risks, and increase the likelihood of a stable and promising future.

Specifically, we explore the possibility that the increasing prevalence of controlled processing in a population (within and/or across its individuals), and the impact this has on the environment, can lead to initial improvements in the fitness of the population; but that, under a range of circumstances, this growth can sow the seeds of its own demise. We examine several scenarios that can produce this effect—and the reasons for it—using a formal theoretical approach that applies mathematical methods from nonlinear dynamical systems analysis and population biology together with numerical methods and computational simulations.

The models we present here implement the distinction between controlled and automatic processing in simple, but principled forms. Although they certainly do not capture the full complexity of the cognitive processes of which humans are capable, nor the underlying continuity of the spectrum from controlled to automatic processing (e.g., Cohen et al. (1990); Kahneman and Treisman (1984)), the models we present do capture critical distinctions that underlie the dimension of controlled versus automatic processing. Furthermore, judicious simplification has allowed us to build population models of agents that incorporate this critical dimension of processing. This, in turn, has allowed us to pursue some of the first efforts to incorporate this fundamental construct of automatic versus controlled processing from cognitive psychology into population models, and to use these models to ask questions about the emergence, impact, and evolution of psychological processes at the population level.

Prior Theoretical Work

Although formal models have provided insight into the mechanisms underlying automatic and controlled processes and the impact of these processes on individual behavior (Cohen et al., 1990; Miller & Cohen, 2001), these models have not addressed their interaction at the population level over the course of evolutionary timescales (whether cultural or genetic). Conversely, population models have largely ignored the dimension of controlled versus automatic processing, instead just focusing on the evolution of agents’ behaviors rather than the underlying cognitive processes that drive that behavior.

Despite this overall lack of consideration of evolution along the dimension of automatic versus controlled cognition, some work has begun to explore the population dynamics of factors that have much in common with this dimension. For example, Wolf, Doorn, and Weissing (2008) consider the population dynamics that arise from the competition between agents with “unresponsive personalities” that are inflexible in the face of a fluctuating environment (akin to reliance on automatic processing), and agents with “responsive personalities” that are able to change flexibility but must pay a cost to do so (akin to reliance on controlled processing). Their simulated agents face an environment that alternates between two possible states and each agent must choose between two actions at any given time, with each action optimally matched to one of the two environmental states. Based on the assumption that the benefit of choosing the correct action is decreasing in the number of other agents who also choose that action (i.e., that the benefits of “responsiveness” are negatively frequency dependent), they show that stable coexistence of responsive and unresponsive agents is a robust feature of the resulting population dynamics. Notably, they do not observe any cyclical dynamics in the frequency of responsive versus unresponsive agents.

Another example is the work of Bear and Rand (2016) and Bear, Kagan, and Rand (2017), who examine automatic and controlled processing in the context of the evolution of cooperation. Their agents play Prisoner’s Dilemma games that sometimes are one-shot (such that defection is always payoff-maximizing) and at other times involve future consequences (such that cooperation is payoff-maximizing if the other player also cooperates). Thus, their agents face a varying social environment. Agents can either use
automatic processing, inflexibly choosing to cooperate or defect without conditioning on game type, or they can pay a cost to use controlled processing and base their action on game type. They find that evolution leads to a population in which automatic and controlled processing stably coexist within each individual if games with future consequences are sufficiently likely. That is, the equilibrium strategy is to (a) cooperate when using automatic processing, but (b) sometimes exercise control (in trials for which the cost of control is sufficiently small) and switch to defection if it turns out the game is 1-shot. Like Wolf et al. (2008), they do not observe cyclical dynamics in the extent of automatic versus controlled processing.

Although these models have begun to address the population dynamics of automatic versus controlled processing, they do not present a general framework for studying this issue. Rather, each considers one specific application of the distinction between these types of processing. More importantly, these models omit a key feature of the natural world suggested above: not only can the environment (physical and/or social) determine the adaptive advantage of a particular cognitive style, but the prevalence of that cognitive style within the population may, in turn, impact the environment; that is, there can be feedback between environment and cognition (Cohen, 2005).

The interaction between the behavior of agents in a population and the environment has been explored previously in evolutionary models (e.g., niche construction; Bergmüller and Taborsky, 2010; Kendal, Tehrani, and Odling-Smee, 2011; Laland, Odling-Smee, and Feldman, 1999)—but not, to our knowledge, the interaction between agents’ cognition and the environment. Our group has recently begun an examination of the influence that such cognition-environment feedback has on the evolutionary dynamics of the balance between controlled and automatic processing in the context of intertemporal choice (Tomlin, Rand, Ludvig, & Cohen, 2015; Toupo, Strogatz, Cohen, & Rand, 2015).

In this work, agents foraging for goods could engage in either automatic or controlled processing as they chose how much of those goods to consume and competed with their fellow agents for access to those goods. While automatic processing led to the immediate consumption of goods (and maximal instantaneous individual fitness), control—and the associated capacity for forethought and planning—allowed agents to make better use of the resources they acquired by consuming them in an optimal way (leading to higher long-term fitness). However, because control required time and effort, automatic processing led agents to be more likely to acquire goods during competitions. Furthermore, the intensity of competition and the abundance of resources (and therefore the importance of planning for the future) were allowed to vary based on the extent of controlled processing in the population. This feedback gave rise to cyclic dynamics under a robust set of parameters using both agent-based simulations (Tomlin et al., 2015) and differential equation modeling (Toupo et al., 2015), with populations alternating between high and low prevalence of controlled processing. However, like previous work, these models did not present a general analysis of the balance between automatic and controlled processing, but instead made a number of domain-specific assumptions tailored to the details of intertemporal decision-making in a particular context.

The Present Work

In this paper, we present a set of models that capture key features of automatic versus controlled processing, and their interactions with the environment, in a way that is fully general and not tied to any particular implementation. We begin with the simplest possible formulation (Minimal Model), in which controlled processing generates higher fitness than automatic processing, but also carries a cost (which can vary based on the fraction of the population engaging in automatic processing). The Minimal Model consists of two differential equations respectively characterizing the population (fraction of agents engaging in controlled vs. automatic processing) and the environment (extent to which controlling processing outperforms automatic processing) that are coupled with some lag. We examine the dynamics of this model in detail. We then demonstrate the robustness of the conclusions from this Minimal Model by considering a series of additional models that add complexity in varying ways, and show that all of these extensions also exhibit cyclical dynamics.

Minimal Model

Automatic Versus Controlled Processing

The minimal model of automatic versus controlled processing focuses on the trade-off between efficiency and flexibility of processing. Specifically, we assume that automatic processing supports efficient and typically effective behavior, achieved by encoding “pre-compiled” responses that are optimally adapted to a particular set of circumstances, but are slow to develop or change. In contrast, we assume that controlled processing supports a more flexible range of responses that can adjust more quickly to changes in contingencies and thereby generate advantageous responses under a wider range of conditions, but that this flexibility comes at a cost (as discussed further below). This distinction bears a close relationship to the distinction between compiled (efficient but rigid) and interpreted (slower, more demanding, but more flexible) procedures in computer science. In an evolutionary context, this distinction may align with different time scales of adaptation—automatic processing over longer (developmental, and/or traditional evolutionary) time scales, and controlled over much shorter (circumstance-by-circumstance).

We capture the advantage of flexibility conferred by control by stipulating that controlled processing results in a payoff from decision-making normalized to value 1, whereas automatic processing results in a discounted payoff of $1 - p$ (with $0 < p < 1$). The flexibility of controlled processing, however, comes at a cost (e.g., Daw, O’Doherty, Dayan, Seymour, & Dolan, 2006; Gershman, Horvitz, & Tenenbaum, 2015; Griffiths, Lieder, & Goodman, 2015; Keramati, Dezfoul, & Piray, 2011; Posner & Snyder, 1975; Shenhav, Botvinick, & Cohen, 2013): it requires time and effort to

---

1 We should emphasize that automatic and controlled processing are not necessarily always in conflict; both modes of processing can arrive at the same response. In the work reported here, however, we focus on competition between these two extremes of processing, as there is mounting evidence that they may indeed compete in determining responses (e.g., Evans and Stanovich (2013); Greene, Nystrom, Engell, Darley, and Cohen (2004); McClure et al. (2004)) and we seek to understand the influence that such a trade-off has at the population level.
attend to the relevant information and compute the optimal course of action that, at the least, imposes an opportunity cost with regard to other potentially advantageous behaviors (Kurzban, Duckworth, Kable, & Myers, 2013). To model this, we impose a fixed cost \( c \) upon the use of controlled processing.

For simplicity, in the Minimal Model we consider the evolutionary dynamics of a population of agents that act exclusively in either a controlled or automatic manner. Specifically, let \( x \) be the fraction of the population that is controlled (\( 0 \leq x \leq 1 \)), and therefore \( 1 - x \) be the fraction that is automatic. In the absence of cognition-environment feedback, the fitness of controlled agents is \( f_c = 1 - c \) (the decision-making payoff of 1 minus the cost of control, \( c \)), and the fitness of automatic agents is \( f_a = 1 - p \) (the inferior decision-making payoff of \( 1 - p \), but with no additional cost).

Thus, there are two environmental parameters that describe the nature of the world in which the agents operate: \( p \), capturing the factors that favor the value of flexible controlled processing relative to inflexible automatic processing (e.g., how unstable the environment is, how scarce resources are, etc.), and \( c \), capturing how costly it is to exert cognitive control.

### Evolutionary Dynamics

Within this simple framework, we allow the frequency of controlled agents \( x \) in the population to evolve according to the replicator equation (Hofbauer & Sigmund, 1998). The replicator equation implements a fairly general population dynamic, whereby the strategy with the higher payoff becomes more common in the population over time. This dynamic can equally well describe evolution that is genetic or cultural (e.g., in which social learning leads people to propagate successful behaviors observed in others).

For our system, the replicator equation is specified by

\[
\dot{x} = x(f_c - \phi),
\]

where \( \phi \) is the average fitness of the population,

\[
\phi = xf_c + (1 - x)f_a.
\]

Our subsequent analyses will use \( d \) as a proxy for population size, because the replicator equation does not directly describe the size of the population (only the fraction of the population that is automatic vs. controlled). Furthermore, in the Variable Population Size Model described below we consider an agent based model in which the population size does vary, and show equivalent results.

### Cognition-Environment Feedback

To incorporate cognition-environment feedback, we allow the prevalence of automatic versus controlled processing in the population \( x \) to influence both environmental parameters \( p \) and \( c \); that is, both \( p \) and \( c \) vary as a function of \( x \).

**Feedback of \( x \) on \( p \).** Our characterization of the effect of \( x \) on \( p \) is used to capture the influence that externalities of controlled processing can have on the environment (and therefore on the relative advantage of the flexibility allowed by control). For example, technological advances are an external consequence of the proliferation of controlled processing in the human population, and this has had clear consequences on our environment: abundance of food and shelter, fluidity and scope of social interaction, and so forth. Here, we begin by considering the simplest case in which such externalities close the gap between automatic and controlled processing: by making resources more plentiful and thus stabilizing the environment, the innovations created by controlled processing reduce the importance of being able to flexibly adapt and plan for the future. Automatic processing, conversely, can undermine many of these benefits (e.g., because of overconsumption, ill-advised, overuse, and/or inefficient use of resources, etc.).

To capture these influences, we link the value of \( p \) inversely to the fraction of controlled agents in the population \( x \): as control increases in the population, the advantage of being controlled decreases. Because it takes time both for the externalities associated with controlled agents to develop and for any deleterious effects of automatic processing to undermine the advantages of these externalities, we incorporate lag into the feedback between \( p \) and \( x \). Specifically, we implement the additional differential equation

\[
p = \frac{((1 - x) - p)}{\tau_p}
\]

such that \( p \) always moves in direct opposition to \( x \) (i.e., toward the current value of \( 1 - x \)), but with some time lag parameterized by \( \tau_p \).

**Feedback of \( x \) on \( c \).** Feedback of \( x \) on \( c \) is used to capture the influence that the prevalence of automatic processing can have on the cost of cognitive control. We focus primarily on the case in which the presence of more automatic agents increases the cost of cognitive control. For example, automatic agents may respond more quickly or efficiently (outcompeting controlled agents for access to resources) and/or consume resources without regard to future need, thereby diminishing resources upon which controlled agents had planned to rely and, as a result, increasing the relative cost of being controlled. At the same time, the costs of control might be decreased by a greater preponderance of controlled agents in the population. For example, this may give controlled agents a greater opportunity to form coalitions or design institutions that facilitate or reward the use of control, or selectively sanction automatic agents (which reduces the relative cost of being controlled).\(^2\)

The prevalence of control in the population \( x \) is likely to impact the cost of control \( c \) on a much faster time scale than the rate at which the externalities of control impact \( p \) (the relative advantage it has over automatic processing, as described above). This is because the former typically emerges from direct interactions between individuals, or potentially from influences on the environment that occur relatively quickly (e.g., depletion through consumption) compared to environmental influences that affect \( p \) (e.g., technological development and growth). Therefore, we implement the feedback of \( x \) on \( c \) as instantaneous (in the Two-lag Model we consider the more complex case where feedback of \( x \) on \( c \) is lagged).

\(^2\) Note that these effects all involve costs that weigh more heavily on controlled or automatic agents (i.e., affect the relative fitness of control). Effects that reduce the fitness of both types of agents equally do not alter the model’s dynamics, because the replicator equation is driven by how each strategy’s fitness compares with the average fitness (and so adding or subtracting a constant from all payoffs has no effect).
Specifically, we modify the fitness of controlled agents $f_c$ such that an additional cost $w(1-x)$ is added, where $w$ is the intensity of the impact of the population state on the cost of control, yielding $f_c = 1 - (c + w(1-x))$.

Although we focus on the case in which the presence of automatic agents increases the cost of control, this formulation can equally well describe the opposite case in which the presence of automatic agents decreases the cost of control. Such a situation might result from technologies or behaviors that, when employed by controlled agents, leverage the short-sighted behavior of automatic agents for the personal gain of the controlled agents (e.g., in the domain of intertemporal choice, the design and sale of products providing instant gratification, but long-term costs—products that would primarily appeal to agents engaging in automatic processing). These situations correspond to cases in which $w < 0$, which leads to the fitness of controls $f_c$ increasing with the frequency of automatic agents $1-x$.

**Results**

Putting these pieces together, the Minimal Model is specified by the following system of two ordinary differential equations (ODEs):

$$\dot{x} = x(f_c - f) = x((1 - (c + w(1-x))) \\
- (x(1 - (c + w(1-x))) + (1-x)(1-p)))$$

$$\rho = \frac{(1-x) - p}{\tau_p}$$

with $c > 0$ and $\tau_p > 0$.

Analyzing this system shows the existence of up to three fixed points. There are always fixed points at $[x = 0, p = 1]$ (exclusively automatic agents in an inhospitable world) and $[x = 1, p = 0]$ (exclusively controlled agents in a hospitable world).

When $c + w < 1$, there is a third (interior) fixed point at $[x = \frac{1-c-w}{1-w}, p = \frac{c}{1-w}]$ where automatic and controlled agents coexist.

These fixed points exhibit different stability characteristics. First, the fixed point $[x = 1, p = 0]$ is never stable given that $c > 0$. That is, in a maximally hospitable world, automatic processing is just as successful as controlled processing because when $p > 0$ there is no advantage of control. Thus, as long as there is any cost to control ($c > 0$), automatics will outperform controls.

When $c + w > 1$, the fixed point $[x = 0, p = 1]$ is stable, and the interior fixed point is not relevant (i.e., lies outside the interval $[0, 1]$). In this case, the cost of control in an entirely automatic population, $c + w$, is larger than the relative advantage of controlled over automatic processing in a maximally inhospitable world, 1. Under these conditions, controlled agents are at an insurmountable disadvantage and unable to proliferate.

When $c + w < 1$ (that is, when the cost of control in an entirely automatic population is less than control’s advantage in a maximally inhospitable world), the fixed point at $[x = 0, p = 1]$ becomes unstable and the interior fixed point becomes relevant (i.e., enters the interval $[0, 1]$). Everywhere in this region, we observe coexistence between automatic and controlled processing. Thus, coexistence is a robust feature of this model, as it has been in other models that did not involve feedback between agents’ cognition and the environment (Bear et al., 2017; Bear & Rand, 2016; Wolf et al., 2008). As long as the costs of control are not so large as to prevent controlled processing from emerging in the first place (i.e., to prevent controlled agents from invading the “state of nature” of automatic agents in an inhospitable world), both automatic and controlled processing will persist.

Interestingly, however, the dynamics of this coexistence depend on how quickly the prevalence of control in the population $x$ diminishes the relative advantage of controlled processing $p$ (as captured by the lag parameter $\tau_p$). There is a critical value of $\tau_p$:

$$\tau_p = \frac{(1-w)^2}{cw(1-c-w)}$$

around which the dynamics change.

When $\tau_p < \tau_p^*$ such that change occurs sufficiently quickly, the interior fixed point is stable and the population settles there. At $\tau_p = \tau_p^*$, however, the interior fixed point becomes unstable and a limit cycle is born (a Hopf bifurcation occurs). Thus, when $\tau_p > \tau_p^*$ (i.e., feedback from $x$ on $p$ is sufficiently lagged), the proliferation of control occurs more quickly than the rate at which this diminishes its advantage, and we observe persistent cycles in the relative balance of automatic and controlled processing—unlike prior models lacking cognition-environment feedback. A representative example of these cyclical dynamics is shown in Figure 1.

The cycles shown in Figure 1 exhibit several distinct phases:

1. The population begins with dominance by automatic agents (i.e., $x$ is small), in an inhospitable environment (i.e., $p$ is large). This population is small in size (i.e., has low average fitness).

2. Controlled agents outperform automatic agents because of the advantage controlled processing has over automatic processing in inhospitable environments. Thus, $x$ increases and correspondingly, subject to some lag, $p$ decreases. The population’s size increases as the controlled agents outperform the automatic agents and average fitness increases.

3. With time, the externalities of control associated with the prevalence of controlled agents in the population (i.e., large $x$) lead to a progressively more hospitable environment, and $p$ continues to decrease (with the associated increase in the fitness that would be achieved by an automatic agent).

4. Once $p$ becomes sufficiently small, automatic processing becomes successful enough that the cost of control outweighs the relative benefit of controlled processing. Thus automatic agents begin to outperform controlled agents and proliferate, and automatic agents come to dominate the population ($x$ decreases).

---

3 Fixed points are $[x, p]$ pairs at which $\dot{x} = \dot{p} = 0$, such that when at a fixed point the system will remain there. A fixed point is stable if the system returns to the fixed point when perturbed away, and unstable if even a tiny perturbation causes the system to leave the fixed point. Thus, it is the identification of stable fixed points that is our goal for understanding potential evolutionary outcomes.
creates inertia in cycles to emerge? The answer involves hysteresis: The lag Model’s alternative formulation of cognition-environment feedback is possible for cycles to arise with will see below, this particular result is not totally general: it is relative advantage of controlled processing remains relatively (i.e., reaches a state in which automatic and controlled processing consequently perform poorly), controlled agents begin to proliferate. If the relative advantage of control \( p \) diminishes rapidly enough, the population reaches equilibrium (i.e., reaches a state in which automatic and controlled processing have the same fitness). But if there is sufficient lag, the relative advantage of controlled processing remains relatively high as control proliferates, allowing the frequency of control to exceed the value it would occupy in the interior equilibrium. Once the advantage of controlled processing finally falls far enough, the system swings back in the opposite direction: automatics proliferate and enjoy a period of success, allowing the level of control to drop below that of the interior fixed point, thus reinitiating the cycle.

In sum, we find that not only is coexistence between automatic and controlled agents a robust feature of this model, but so is cyclicity. As long as there is cognition-environment feedback, with a sufficient lag in the impact of that effect on the relative advantage of controlled processing, and the cost of control is not too large, persistent cycles emerge: the population alternates between periods of dominance by automatic and controlled processing, and the population fitness (and thus size) fluctuates accordingly. It seems reasonable to imagine that such lags may be characteristic of real-world systems; that is, the pace at which the negative secondary consequence of newly developed technologies accrue (e.g., bacterial resistance, or environmental damage from resource use) typically lags behind the initial positive impact of those technologies (i.e., protection from infection, or energy availability).

**Robustness Across Model Specifications**

It is reasonable to ask whether the results observed for the Minimal Model are specific to the simplicity and/or specific assumptions of that model. To address this, we consider a series of related but more complex models that modify the Minimal Model in a number of different ways, and show that all of these produce results that are qualitatively equivalent to those of the Minimal Model.

The first of these is the Two-Lag Model, which addresses the possibility that lag exists not only in the effects of control on the environment, but also in the effects of automaticity on the cost of control. Although the impact of automatic agents on the cost of

---

4 It is important to note that when using the replicator equation, the fraction of controlled agents \( x \) can become arbitrarily small without actually reaching zero. Therefore, after the environment destabilizes and the population crashes, control is always able to reinvade. In reality, however, populations are finite and thus actual extinction may occur at the end of one of the downward spirals, or the complete elimination of control (although mutation and migration may also reintroduce control into an entirely automatic finite population, seeding a new cycle).
control is likely to occur relatively quickly, because it emerges from direct interactions between individuals (e.g., competition) or short-term influences on the environment (e.g., consumption), it cannot literally be instantaneous (as assumed in the Minimal Model). The Two-Lag Model assesses the impact of incorporating this extra lag by adding a third differential equation to the Minimal Model characterizing the extent to which automatic agents directly impact the fitness of controlled agents (also coupled to the population state with a lag). The second extension is the Consumption Model, which considers the possibility that automatic agents impact the cost of control via their consumption (rather than just their presence): Furthermore, this model considers another externality of control: in addition to stabilizing the environment (and therefore reducing the relative advantage of control’s flexibility), controlled processing and associated technological innovation increases the carrying capacity (i.e., maximum population size the environment can support). To examine the effect of these factors, the Variable Population Size Model modifies the agent-based simulations of the Multiprocess Agent Model to allow the population to vary in size, constrained by the frequency of controlled processing.

Two-lag Model

In the Minimal Model, we assumed that the prevalence of automaticity in the population \((1 - x)\) impacted the cost of control instantaneously by specifying the cost of control in \(f_c\) to be \(c + w(1 - x)\). Here we show that extending the model to the case in which this feedback, like the influence of \(x\) on the advantage of controlled processing \(p\), is also lagged yields similar results. To do so, we specify the cost of control in \(f_c\) to be \(c + w\), and add a differential equation for \(dw/dt\) whereby \(w\) changes to follow (1 - x) with lag (in the same way that \(p\) follows (1 - x) in the \(dp/dt\) equation of the Minimal Model). Furthermore, we specify \(dw/dt\) such that \(w\) need not vary fully between 0 and 1, but rather varies between 0 and some maximum value \(w_{\text{Max}}\) (with \(0 < w_{\text{Max}} \leq 1\)). For maximal generality, we also modify the \(dp/dt\) equation to have a maximum value \(p_{\text{Max}}\) (with \(0 < p_{\text{Max}} \leq 1\); the Minimal Model implicitly uses \(p_{\text{Max}} = 1\)).

This gives us the following three-dimensional system:

\[
\dot{x} = ax((1 - (c + w)) - (x(1 - (c + w)) + (1 - x)(1 - p)))
\]

\[
\dot{p} = \frac{((1 - x)p_{\text{Max}} - p)}{τ_p}
\]
Analyzing this system, we find a potential interior fixed point analogous to that in the Minimal Model at \( x = \frac{p_{\text{Max}} - w_{\text{Max}}}{p_{\text{Max}} - w_{\text{Max}}} \), \( p = \frac{p_{\text{Max}}}{p_{\text{Max}} - w_{\text{Max}}} \), \( w = \frac{w_{\text{Max}}}{p_{\text{Max}} - w_{\text{Max}}} \), which is relevant (i.e., on the interval \([0, 1]\)) when \( p_{\text{Max}} > c + w_{\text{Max}} \). The interpretation of this condition is straightforward: the maximum possible advantage of control must be larger than the maximum possible total cost of control (fixed cost + cost imposed by automatics).

As in the Minimal Model, a stability analysis finds that this interior fixed point can become unstable and give birth to a limit cycle via a Hopf bifurcation. Figure 3 shows a representative example of the dynamics of the 3-dimensional system that exhibits a series of phases similar to those in the simpler 2-dimensional Minimal Model.

Although we can analytically derive the exact condition required for the limit cycle to exist, this condition is too complex to be readily interpretable. To give a sense of its implications, Figure 4 shows where limit cycles occur in the \([\tau_p, \tau_w]\) plane for different values of \( c \) and \( w_{\text{Max}} \). (Fixing \( p_{\text{Max}} = 1 \), as in the Minimal Model).

The most salient feature of Figure 4 is that, in order for limit cycles to occur, \( \tau_p \gg \tau_w \) must be satisfied; that is, the prevalence of control over automaticity in the population must influence the relative advantage of controlled processing \( p \) substantially more slowly than the cost of control \( w \). We also see that the minimum lag required for cycling, as measured by the slope of the line in Figure 4, decreases as the maximum cost imposed by automatics \( w_{\text{Max}} \) increases.

In sum, the Two-lag Model demonstrates that the results of the Minimal Model are robust to accounting for the fact that, in reality, the processes through which automatic agents increase the cost of control need not involve instantaneous feedback. For example, if automatic agents consume resources that controlled agents had planned to rely on in the future, the consequences of the behavior of current automatic agents will not be felt by controlled agents until some time in the future (i.e., when they attempt to use the already-consumed resources).

### Consumption Model

In the Minimal Model and the Two-lag Model, automatic agents influenced the cost of control in direct proportion to their frequency in the population. Here we examine the consequence of having automatic agents instead impact the cost of control via their consumption rather than simply their prevalence. Specifically, we link the cost of control to the product of the proportion of the population that is automatic \( 1 - x \) and the average fitness (capturing consumption) of the automatic agents \( 1 - p \). Thus we replace the Minimal Model’s cost of control term in \( f_c \) of \( c + w(1 - x) \) with the alternative formulation \( c + w(1 - x) \). This yields the following system of two ODEs:

\[
x = x((1 - (c + w(1 - p)(1 - x)))
- (\alpha(1 - (c + w(1 - p)(1 - x))) + (1 - x)(1 - p))
\]

\[
\dot{p} = \left(\frac{(1 - x) - p}{\tau_p}\right)
\]

Analyzing this system shows that although the analytic expressions are now even more complex (and thus harder to directly interpret), we again observe regimes in which there is coexistence between automatic and controlled agents, as well as regimes in which there are limit cycles. Both coexistence and limit cycles are more robust than in the Minimal Model. Coexistence is more robust in that no homogeneous population make-up is ever stable in the Consumption Model: we only find coexistence or limit cycles. As in the Minimal Model, if automatic agents increase the cost of control \( (w > 0) \) then limit cycles can occur if cognition-environment feedback is sufficiently lagged (i.e., \( \tau_p \) exceeds a specified threshold). Figure 5 illustrates the minimum amount of feedback lag required to induce a limit cycle. Because coexistence is more robust in the Consumption Model than the Minimal Model, limit cycles are also more robust: they can occur no matter how large the magnitude of \( w \) (unlike in the Minimal Model, which requires \( w < 1 - c \) for either coexistence or limit cycles). Nonetheless, as in the Minimal Model, it is easiest to get limit cycles when the fixed cost of control \( c \) is small but the cost imposed by automaticity \( w \) is large.

![Figure 3](image-url)

Figure 3. Persistent cycles of automaticity and control also emerge when both forms of feedback are lagged in the Two-lag Model. Shown are the values of \( x \), \( 1 - x \), \( p \), \( w \), and \( \Phi \) as a function of time. Results were generated using numerical integration of the Two-lag Model ordinary differential equations (ODEs) using \( c = 0.5 \), \( w_{\text{Max}} = 0.3 \), \( p_{\text{Max}} = 1 \), \( \tau_p = 50 \), \( \tau_w = 10 \), and initial conditions \( x = 0.01 \), \( p = 0.9 \), \( w = 0.1 \). See the online article for the color version of this figure.
These observations show that the limit cycles observed in the Minimal Model when cognition-environment feedback is sufficiently lagged are robust to the alternative implementation of the cost of control in the Consumption Model, in which the consumption of automatic agents, rather than their prevalence per se, increases the cost of control.

Threshold Model

In all of the models presented thus far, cognition-environment feedback was implemented in a continuous form, such that the advantage of controlled processing always moved in opposition to the fraction of the population that was controlled: \( p \) followed \( x \), subject to lag. Here, we change the form of this feedback (i.e., the specification of \( dp/dt \)) to instead operate via a discrete threshold dynamic.

Specifically, the Threshold Model assumes that excess time/energy is required for controls to invest in technological innovation, and therefore that \( p \) decreases if the product of the fraction of controlled agents and the fitness of those controlled agents is greater than a threshold \( T \), and increases if not. To keep \( p \) bounded on the interval \([0, 1]\), we also add a factor of \( p(1/p) \) to \( dp/dt \). Combining this alternative formulation of \( dp/dt \) with the Consumption Model\(^5\) presented in the previous section yields the following set of two ODEs:

\[
\dot{x} = x(1 - (c + w(1 - p)(1 - x))) \\
- x((1 - (c + w(1 - p)(1 - x)) + (1 - x)(1 - p))) \\
\dot{p} = p(1 - p)T - x(1 - (c + w(1 - p)(1 - x)))
\]

Analyzing this model finds six possible fixed points. However, only three of these fixed points are ever stable. The resulting dynamics depend critically on \( w \), the impact that the consumption of automatic agents has on the cost of control.

\(^5\) The threshold implementation of the Minimal Model yields much more complex conditions which are intractable, so we focus on the Consumption Model when considering threshold updating of \( p \).
When the consumption of automatic agents increases the cost of control \((w > 0)\), the results are qualitatively similar to the Minimal Model and the Consumption Model. When \(T > 1 - c\), it is very difficult for controlled agents to make the environment more hospitable for automatics (and thereby reduce their own relative advantage \(p\)); Even when controls entirely dominate the population, \(x_f\) is not sufficiently large to exceed the threshold \(T\) and thereby decrease \(p\). As a result, the only stable fixed point involves the complete dominance of control, \(x = 1, p = 0\). However, so long as \(T < 1 - c\), there is an interior fixed point at \(x = \frac{\tau_1(w + 1)}{\tau_2 + c + 1}, p = \frac{c + w(1 - T)}{1 + w}\) which leads to coexistence when \(\tau_p < \frac{c + w(1 - T)}{\tau_1 + w}\), and limit cycles (via a Hopf bifurcation) when \(\tau_p > \frac{c + w(1 - T)}{\tau_1 + w}\). Thus, as in the other models, lag in the cognition-environment feedback can lead to cycling.\(^6\)

The foregoing analysis focused on the situation in which the consumption of automatic agents increased the cost of control \((w > 0)\). However, the behavior of the model is more complex and qualitatively distinct from the previous models when the consumption of automatic agents decreases the cost of control \((w < 0)\). Such a decrease could arise, for example, from, controlled agents profiting from the consumption of automatic agents by selling the automatic agents products that exploit their lack of control (e.g., unhealthy but delicious food). When \(w < 0\), it remains the case that \([x = 1, p = 1]\) is stable when \(T > 1 - c\). There is also the possibility of another stable fixed point involving the coexistence of automatic and controlled agents at \([x = (w + c)/w, p = 0]\) when \(w < -c\) and \(T < (w + c)/w\). Here, it is very easy for controlled agents to make life easier for automatics and as a result the advantage of controlled processing disappears \((p\) goes to \(0)\). However, because the consumption of automatic agents benefits controlled agents, some fraction of controls can still survive in the population even in the absence of a decision-making advantage. (Note that these two fixed points can co-occur, such that there is bistability between them and the initial conditions determine which fixed point evolution favors.)

Finally, when \((w + c)/w < T < 1 - c\), neither of these fixed points are stable and instead we again see the interior fixed point at \(x = \frac{\tau_1(w + 1)}{\tau_2 + c + 1}, p = \frac{c + w(1 - T)}{1 + w}\) as the only possibility. Unlike the case when \(w > 0\) (or the results from the previous models), this point is stable when \(\tau_p > \frac{c + w(1 - T)}{\tau_1 + w}\) and leads to limit cycles when \(\tau_p < \frac{c + w(1 - T)}{\tau_1 + w}\). In other words, when automatic consumption benefits control, limit cycles emerge when cognition-environment feedback occurs sufficiently quickly rather than sufficiently slowly. Figure 6 shows the critical lag required for cycles, and Figure 7 shows sample cycles arising when \(w > 0\) and \(w < 0\).

In sum, the Threshold Model provides further evidence that the occurrence of limit cycles is robust when automaticity increases the cost of control, and cognition-environment feedback is sufficiently lagged (although here it is easiest to get cycles when both the fixed and variable costs of control are small). Furthermore, the Threshold Model extends the conditions under which limit cycles can emerge, now including situations in which automaticity decreases the cost of control (i.e., when controls benefit from the presence of automatics), albeit through a different mechanism in which the cognition-environment feedback must occur quickly rather than slowly. Such a situation might result from technologies or behaviors that, when employed by controlled agents, leverage the short-sighted behavior of automatic agents (e.g., in the domain of intertemporal choice, the sale of products providing instant gratification, but long-term costs—products that would primarily appeal to agents engaging in automatic processing).

**Multiprocess Agent Model**

In all of the models described so far, agents were binary: they were either exclusively automatic or controlled. Here, following on previous work (Bear et al., 2017; Bear & Rand, 2016; Tomlin et al., 2015), we describe an agent-based model in which each agent exhibits a probabilistic mix of controlled and automatic processing. We conducted simulations of this model to examine how the probability of control within agents evolves over the course of generations in response to the same factors implemented in the ODEs described above.

In this model, each agent \(i\) implemented controlled processing with an agent-specific probability \(x_i\) (and correspondingly implemented automatic processing with an agent-specific probability \(1 - x_i\)). Adapting the Minimal Model formulation, the fitness in generation \(t\) of an agent \(i\) that exhibited control with probability \(x_i\) in a population for which the mean probability of control was \(<x>\) was given by the sum of the fitness of controlled processing and of automatic processing in the current population and environment, weighted by that agent’s probability of engaging in controlled and automatic processing, respectively:

\[
 f_i = x_i(1 - (c + w(1 - <x>))) + (1 - x_i)(1 - p_i)
\]

where \(c\) is the fixed cost of control, \(w\) is the impact of automatic processing on the cost of controlled processing, and \(p_i\) is the relative advantage of controlled processing in generation \(t\).

We examined the stochastic evolutionary dynamics of a population of \(N = 100\) such agents using the pairwise comparison process (Traulsen, Pacheco, & Nowak, 2007): In each generation, one agent was picked at random to potentially update its strategy, and another agent was picked at random to potentially reproduce. The updater was replaced by a copy of the reproducer with probability \(1/\tau_s\) where \(s\) is the intensity of selection, \(\tau_\Pi\) is the fitness of the potential reproducer (teacher), and \(\tau_\Sigma\) is the fitness of the potential updater (learner); otherwise, no change occurred in that generation. Alternatively, with probability \(w\) a mutation occurred; in that case, instead of adopting the other agent’s strategy, the updater adopted a new strategy sampled from a uniform distribution on the interval \([0, 1]\). (Simulations using local mutation produced equivalent results.)

In addition to this standard evolutionary dynamic, we also implemented cognition-environment feedback by updating the advantage of controlled processing in each generation, such that

\[
 p_t = p_{t-1} + \frac{(1 - <x>) - p_{t-1}}{\tau_p}
\]

Figure 8 shows that this agent-based model, in which agents implemented a probabilistic distribution of controlled versus automatic processing, displays qualitatively similar dynamics to the

---

\(^6\) Numerical simulations indicate the existence of additional limit cycles when \(w > 0\) not born out of a Hopf bifurcation, but we do not characterize the details of those limit cycles here.
analytic models described above, in which automatic versus controlled processing was a binary, deterministic variable. We see that when $\tau_p$ becomes sufficiently large, the dynamics of the system transition from coexistence to oscillations.

In this regime, once the agents (which were initialized to use exclusively automatic processing) developed sufficient control and began to improve the environment, environmental richness improved quickly and substantially. This improvement lessened agents’ dependence on controlled processing for survival, and the advantage of control was reduced therefore automatic processing became more prevalent. This increased prevalence of automatic processing exacerbated the competitive disadvantage of using cognitive control, thus further eroding the frequency of control. Eventually, the prevalence of control was not sufficient to maintain improvements to the environment, thereby returning the environment to its original state and reinitiating the cycle.

In sum, the Multiprocess Agent Model demonstrates that cycles observed in the previous models were not an artifact of the expression for updating of $\tau_p$ in the Multiprocess Agent Model (so that the rate of environmental updating would not vary with the population size).

Variable Population Size Model

In the service of tractability, the models described above did not consider changes in the size of the population, instead examining changes in the relative frequency of automatic versus controlled processing. Here, we examine the impact of allowing the size of the population to grow and shrink.

To do so, we define strategies, payoffs, and updating of the environmental parameter $p$ as in the Multiprocess Agent Model, and modify reproduction as follows. At the beginning of each simulation, the population is initialized at size $N_0$. Each generation, an agent is selected proportional to fitness to reproduce. When probability $u$, a mutation occurs and an agent with a random strategy is added to the population; with probability $1 - u$, a copy of the selected agent is added to the population. If the new population size $N$ exceeds the environment’s carrying capacity $K$, agents are selected at random to die until $N \leq K$.

Rather than fixing $K$ at some predetermined level, we allow $K$ to vary with the population’s make-up. Controlled processing’s ability to flexibly plan for the future, and to develop technological innovations, suggests that greater levels of control in the population should be associated with a larger carrying capacity: in the same way that control can make the environment more stable (as modeled by feedback on $p$), it can also make the environment richer and able to support a larger population. To implement this logic, we set $K = N_0 + \sum_{i=1}^{N_0} x_i$, such that the carrying capacity is increased above the baseline $N_0$ by the extent to which agents engage in controlled processing. Furthermore, because $N$ now varies over time, we removed the $N$ term from the denominator of the expression for updating of $P_i$ in the Multiprocess Agent Model (so that the rate of environmental updating would not vary with the population size).

Figure 9 shows that the Variable Population Size Model can generate similar cyclical dynamics to those of the previous models. However, the current simulations have the important added ability to directly demonstrate population booms and crashes associated with the rise and fall of controlled processing. As agents becomes more likely to exercise control, the environment’s carrying capacity $K$ increases, which in turn leads to an increase in population size $N$. As in the other models, the increase in control eventually leads to enough of a decrease in $p$ that automaticity can reinvade. As automaticity increases (i.e., control decreases), carrying capacity $K$ decreases, driving the population size back toward its initial baseline level of $N_0$.

In sum, the Varying Population Size model shows that the cycles observed in the previous models were not an artifact of
considering only relative frequency of automatic versus controlled processing, and provides a demonstration of the impact these oscillations can have on population size.

**Discussion**

We have described a series of models that examine the evolutionary dynamics of mixed populations of agents that implement different forms of cognitive processing along the dimension from automatic to controlled. Our implementation of cognitive processing was designed to capture, as simply as possible, the most fundamental and commonly assumed differences along this dimension: automatic processing that is efficient but rigid, and controlled processing that is costly but flexible (Kahneman, 2011; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). We implemented differences in efficiency by assigning a cost to the use of control, and differences in flexibility by assuming that controlled processing led to a fitness advantage relative to automatic processing in decision-making (e.g., because controls can plan for the future and adapt more quickly to the exigencies of the situations they encounter). We also implemented cognition-environment feedback by allowing increases in the frequency of controlled processing to make the environment more hospitable—and therefore to reduce the advantage of control.

We found that limit cycles arose in all of the models we considered, and across a wide range of parameters: recurrent boom-and-bust dynamics in which control flourished and the population grew, only to be undermined by an ensuing rise in automaticity, leading to a crash in control and population size. Consistently across models, it was the case that such cycles could arise when (a) the prevalence (or consumption) of automatic agents increased the cost of control, and (b) there was sufficient lag in the influence of controlled agents on the environment, relative to the rate at which automatic processing imposed its costs on control. This seems like a reasonable approximation of real world relationships: the stabilizing influence of control-based technologies on the environment takes longer to develop and have its impact than the direct costs that automatic agents impose on controlled agents via competition and/or short-sighted consumption. Moreover, we observed that these cyclical dynamics were typically most likely to arise in situations in which controlled processing entailed a relatively small fixed cost, but incurred a large loss of fitness from the presence of automatic agents, conditions that may also align well with some real-world circumstances (e.g., the overuse of antibiotics).

There was some divergence across models, however, regarding situations in which automaticity created a benefit for control (rather than imposing a cost). Although in most of the models cycles were not possible in this regime (only coexistence), the Threshold Model differed: in that case, cycles were possible provided cognition-environment feedback was sufficiently fast. Al-

---

**Figure 7.** Sample cyclical dynamics from the Threshold Model. Shown are the values of $x$, $1 - x$, $p$, and $\Phi$ as a function of time. Results generated using numerical integration of the Threshold Model ordinary differential equations (ODEs) using $c = 0.2$, $\tau_p = 0.05$, $T = 0.5$, and (a) $w = 0.2$ versus (b) $w = -0.2$. See the online article for the color version of this figure.
Figure 8. Agent-based simulations of the Multiprocess Agent Model using $N = 100$, $s = 10$, $c = 0.5$, $w = 0.15$, and (A) $\tau_p = 1$, (B) $\tau_p = 10$ or (C) $\tau_p = 100$. Shown are the population average value of $x$, the value of $p$, and the population average value of $\phi$ as a function of time. See the online article for the color version of this figure.
though this finding appears to be less general across the models we have considered,\(^7\) it is intriguing because it expands the space in which cycles can occur into a domain where controlled agents exploit the weakness of automatic agents. More generally, our results are interesting from a dynamical perspective: although cyclic behavior commonly emerges in evolutionary dynamics of three or more strategies (Szolnoki et al., 2014), here we observe that environmental feedback enables cycles with only two strategies.

It is important to note that the effects we report are independent of whether the underlying mechanisms of evolution are genetic or cultural (Richerson & Boyd, 2005). Thus, they may help explain observations of human societies in the past, and may have relevance to our own time. Anthropologists and archeologists have written about a repeated pattern in human history: the emergence of highly successful societies that expand greatly as a consequence of technological innovation, only to eventually collapse (Diamond, 2005; Richerson et al., 2009; Schwindt et al., 2016; Turner & Sabloff, 2012). Such collapses may have occurred for a number of reasons, including environmental shocks or other factors external to the population, or the overuse of technologies by those (presumably controlled) agents who created them. The models we introduce here suggest another possible route to collapse: even if controlled agents exercise forethought and use the technologies they generate wisely so as not—themselves—to overexploit the environment, the flourishing of control and its attendant technologies can invite the concomitant flourishing of automaticity, which in turn can increase the likelihood of collapse due to the shortsightedness and inability to adapt to changing environments (including those brought about by the new technologies) that are defining features of automatic processing. These findings illustrate a mechanism that may be responsible—at least in part—for cyclical dynamics: the pace at which controlled processing generates benefits, relative to the pace at which automatic “free-riders” impose costs on controlled processing (also see Cohen (2005)). In the present work, we leveraged formal models to demonstrate the plausibility of such cyclical dynamics, and to identify quantitative relationships and boundary conditions for these effects.

The models we describe may also be relevant to modern issues and concerns, providing a quantitative framework within which to consider and potentially address them. For example, the pace at which new antibiotics can be developed is slow relative to the pace at which their overuse can produce harm (particularly to those who exercise restraint). Similarly, the pace at which new forms of energy (and the technologies based on them) can be developed is slow relative to the pace at which their abuse can cause damage, and provide (short-term) relative advantage to those who overconsume. The emergence of these technologies is, without a doubt, a reflection of the uniquely human capacity for cognitive control. Similarly, the behaviors that afflict our society most (e.g., drug addiction and failures to save for retirement) are short-sighted forms of behaviors that are thought to reflect the influence of automatic processing (Angeletos, Laibson, Repetto, Tobacman, & Weinberg, 2001; O’Donoghue & Rabin, 1999; Wiers et al., 2007). It is likely that the same is true for subtler, but potentially just as damaging, behaviors (such as overuse of antibiotics, or overconsumption of nonrenewable resources). Considerable research has

\(^7\) We also found cyclical dynamics in a previously unreported agent-based intertemporal choice model where automatic agents benefited controlled agents (Tomlin et al., mimeo).
been devoted to understanding the dynamics of technological developments and their impact from historical, sociological, economic and environmental perspectives (e.g., Abernathy and Utterback 1978; Mokyr 1992; Perez 2003; Rogers 1962), but none of these studies appear to take account of the fundamental psychological processes that drive these dynamics. Conversely, considerable research in psychology and neuroscience has addressed the mechanisms underlying automaticity and control (Daw, Niv, & Dayan, 2005; Hare et al., 2009; McClure et al., 2004), but have not examined how these interact at the population level. The models we have described provide a first step toward bridging these levels of analysis, and suggest that doing so may reveal fundamental principles that yield consistent effects, as well as illuminating factors that may influence these effects.

In the tradition of theoretical work within evolutionary and population biology, the models we described here are as simple as possible. This simplicity naturally omits potentially important aspects of cognitive function. For example, although it is generally acknowledged that there is a continuum between automatic and controlled processing, and that the automaticity of many processes is dependent on the context in which they are executed (e.g., Cohen et al. 1990; Kahneman and Treisman 1984), automatic versus controlled processing was implemented in binary form in our models. Implementing automatic versus controlled processing in a more graded and context-sensitive manner, and more nuanced and realistic forms of controlled processing in population-scale models, is an important direction for future work. Nonetheless, the robustness of the effects we observed across a variety of model implementations considered here suggest the possibility that these are general properties of the evolution of populations comprised of agents with a heterogeneous mixture of proclivities for automatic versus controlled processing.

Future work should also investigate the effects of environments with nonuniform spatial structure, in which agents could flexibly adapt to localized distributions of resources to produce “cultural topologies” that may vary in their expression of automatic versus controlled processing, and cases in which the bias toward automatic versus controlled processing may anticipate (and attempt to counteract) the risks associated with automatic processing. It will also be informative to develop models tailored to specific domains in which the dimension of automaticity has been suggested to play an important role, such as we have begun to do for intertemporal choice (Tomlin et al., 2015; Toupo et al., 2015); this might include dietary and other health-related behaviors, savings, and behaviors that impact the environment. Cooperative social dilemmas are another important domain to explore using the current framework, as the individually optimal behavior may not be optimal at the population level and so controlled processing may itself lead to collapse (Rand, 2016)—although socially optimal cooperation can also be individually optimal if, for example, the interactions are stochastically repeated or institutions exist which sanction noncooperation (Jordan, Peysakhovich, & Rand, 2015), in which case control does support cooperation (Rand, 2016). Finally, an important direction for future work will be to examine domains of function in which the distinction between controlled and automatic processing is not as stark as we have treated it here. As noted in the introduction, it is widely recognized that processes lie along a continuum of automaticity, and that the degree to which processing relies on control depends in large measure on the context in which it occurs. Implementing this more realistic portrayal of control will add considerable complexity to any model, though it may be important for addressing some of the issues discussed above that may be sensitive to the context in which behavior occurs.

In sum, the models we described introduce a fundamental dimension of cognitive function into population-level models, and examine the consequences this has for evolutionary dynamics. Our findings suggest that a robust feature of these dynamics is a cyclic pattern, in which controlled processing initially flourishes, but then sets the stage for its own demise. This suggests that the advent of controlled processing in a population sets in motion a “treadmill,” in which the very advances that are afforded by controlled processing simultaneously set in motion regressive forces—engendered by the presence of automatic processing in the population—that must be outweighed and outpaced if the population as a whole is to progress and/or reach a stable state. It is our fervent hope that further analyses of the sort we have presented here may lead to strategies that will help promote such a positive outcome, and avert the fate that has befallen many previous advanced cultures.

References


Received March 21, 2017
Revision received May 18, 2017
Accepted May 19, 2017