Studying Performance and Learning With ABIR

The Effects of Knowledge, Mobilizing Agents, and Predictability

A. MAURITS VAN DER VEEN
IAN S. LUSTICK
DAN MIODOWNIK
University of Pennsylvania

This study uses the Agent-Based Identity Repertoire model to investigate the ability of populations to adapt and learn in an unpredictable environment. The authors' findings highlight the trade-off between adaptation and diversity in the pursuit of performance but also show that this trade-off is far from straightforward. Increasing sophistication improves the ability to adapt but reduces diversity, imposing high costs down the line. However, high levels of sophistication also produce small, stable homogeneous clusters of agents, which slow down declines in diversity. Innovative or entrepreneurial agents reacting more rapidly to environmental signals increase the prevalence of such clusters, helping diversity but hampering adaptability. The authors also show that more predictable environments facilitate successful adaptation, especially for populations of intermediate sophistication. Finally, the authors conclude that the trade-off between adaptation and diversity is such that in the present model, long-term learning is difficult to achieve.

Keywords: agent-based modeling, learning, adaptation, diversity, knowledge, predictability, entrepreneurs, innovators

> One cannot infer from the fact that an organization functions smoothly that it is a rational and "intelligent" organism that will cope successfully with novel challenges. If anything, one should expect environmental change to make manifest the sacrifice of flexibility that is the price paid for highly effective capabilities of limited scope.
>
> —Nelson and Winter (1982, p. 126)

A fundamental characteristic of many social contexts is the presence (or absence) of learning. Indeed, one often hears the claim that our ability to learn sets us apart, as humans, from the rest of the animal kingdom. In this article, we investigate learning—or perhaps more appropriately, of adaptation and performance—at the level of organizations or societies. Our approach is somewhat unusual in that the amount of knowledge at the disposal of every individual remains constant over time, and the population must face an unpredictable environment. In this context, we examine the impact on performance and learning of variation in four different factors: the sophistication (or amount of knowledge) of each agent, the

AUTHORS' NOTE: A preliminary version of this material was presented at Agent 2000—A Workshop on the Simulation of Social Agents: Architectures and Institutions, University of Chicago/Argonne National Laboratory, October 6-7, 2000.

Social Science Computer Review, Vol. 19 No. 3, Fall 2001 263-280

readiness with which agents adapt, their influence on other agents, and the relative predictability of the environment.

Using a very simple model, our findings highlight the fundamental trade-off between adaptation and diversity in social learning. Moreover, they suggest that this trade-off is more complex than one might expect: The effects of increasing knowledge or changing the composition of a population are neither linear nor monotonous. Thus, certain intermediate conditions often produce either better or worse results than the extremes. Specifically, we find that (a) on balance, greater knowledge increases performance, especially in the long run; (b) it is difficult to improve performance by introducing agents that adapt more readily to their local and global environment; and (c) more predictable environments facilitate performance for intermediate knowledge conditions, whereas they tend to eliminate potential benefits provided by changing the population's composition. Finally, we conclude that the trade-off between adaptation and diversity in our model is such that it is virtually impossible to generate populations whose adaptations allow them to outperform corresponding nonadapting populations at all future points in time.

LEARNING IN POPULATIONS

How do groups, organizations, or populations learn? To answer this question, we first need to get a better sense of what learning entails. No simple definition can capture all the different meanings people ascribe to the concept of learning. Central to most characterizations of learning are two different notions: First, the accumulation of knowledge and second, adaptation to an external environment. In addition, it is usually expected that both processes will improve some measure of performance.

Groups or organizations can learn in several different ways. First, they might learn by replacing individuals within the group by others who have more knowledge or are better adapted to the current environment. Hiring and firing in organizations, or death and birth in populations are two examples. A central feature of this kind of learning is that individuals themselves do not learn.[1] Another type of social learning without individuals accumulating knowledge or adapting is the reorganization of connections within a group. This may improve the flow of information and allow a group to exploit better the distributed knowledge it contains. In the psychological literature, this is sometimes called strategic learning (see Carley, 2000).

When individuals do accumulate knowledge or adapt, performance of the group as a whole may (but need not) increase as well, giving us a third kind of group learning, in which both the composition of the group and its interconnectedness remain unchanged. Carley (2000) referred to this as experiential learning. In the present article, we are concerned with a particular instance of this form of learning. To be specific, in our model, individuals neither have memory (which would allow them to store past experiences) nor can they expand the overall volume of their knowledge (which would allow them to abstract new lessons from past experiences). Instead, they may only change what they know, which in turn may have implications for their behavior.

Some would argue that in such a model, individuals are not learning. After all, the amount of knowledge at each individual's disposal does not increase—each piece of information acquired (learned) requires jettisoning (unlearning) a comparable piece of information. In a static environment, such adaptation may dramatically increase performance, as only a few pieces of information may be required to perform optimally in that environment. In a dynamic environment, however, this becomes much less likely (cf. Carley, 2000, p. 247), especially when the number of possible environmental contexts is far greater than can be implicitly represented by an individual's knowledge store.

If an organization or population is sufficiently large, one might expect to observe what might be called "emergent" learning, where aggregate group performance shows a clear increase over time, even if the performance of most individuals shows no persistent trend. Furthermore, it seems plausible that the ability of a group to learn in this fashion will vary depending on the nature and variability of the external environment and on two fundamental characteristics of the population: (a) the size of the knowledge base of individual agents and (b) the presence within a population of individuals with varying levels of influence on others and/or with varying predilections for changing their repertoire of known (or learned) pieces of information.

Groups, organizations, and populations may be modeled in many different ways. In this article, we use the Agent-Based Identity Repertoire (ABIR) model developed by Ian Lustick.[2] It models a population as a two-dimensional rectangular grid, with a single agent at each location. Agents interact only with those agents that are directly contiguous (i.e., the eight-member Moore neighborhood). In other words, the system provides only a minimal amount of overall structure, making it possible to focus quite explicitly on the implications of varying the three factors mentioned in the preceding paragraph. We examine the performance of the system under two distinct models of change in the environment, for a broad range of different knowledge repertoires, and with five different types of population composition. In the process of examining the implications for performance of the various specifications tested, we will also be able to assess whether any of these populations can be said to be learning (and if so, in what fashion).

ABIR—MODEL AND OPERATIONALIZATION

Before we present our findings, it will be helpful to describe the ABIR model in some more detail. The model was developed by Ian Lustick (2000). It was originally designed, and is still being used, to refine, to explore, and to test propositions associated with constructivist theories of political identity (Lustick 2000; Lustick, Miodownik, & Philbrick 2000), but it is sufficiently general to be used to study a variety of other issues, such as deliberative democracy (Lustick & Miodownik 2000), globalization, and ethnic conflict.

In the model, each agent has a number of attributes. Most important among these is a list, or repertoire, of strategies, ideas, arguments, or identities an agent has at its disposal. Depending on the metaphor we choose, we may say that at any given time, a single strategy is deployed, an identity is activated, or an argument is articulated from this repertoire. All other elements in the repertoire are known to the agent itself only.[3] Agents change their activated and subscribed identities in response to the information they receive from the agents around them, as well as from the global environment. The latter provides information about the overall value (positive or negative) of each identity in the total spectrum of possible identities in the world. This value might be seen as the payoff associated with activating on that particular identity.

A system is run for a number of rounds. Each round, every agent gathers information from itself, its Moore neighborhood, and the global environment. The information provided by different types of agents is assigned different weights. For each possible identity in the world, the agent adds up the number of neighbors (including itself) activated on that identity and adds to this sum the current environmental value of the identity.[4] In this manner, every potential identity is assigned a current value, which may vary from one agent to the next, depending on each agent's immediate neighborhood. All agents have three threshold levels that specify how to process this relative ranking of identities. The first gives the threshold above which an identity that is already in an agent's subscription list will become the activated identity in the next round. The second, higher threshold indicates when an agent will add a new identity to its subscription list (often preparing it for activation in a subsequent round). The third, and highest, value tells the agent when to subscribe to a new identity and simultaneously activate it.[5]

In the experiments reported here, the population size included 2,500 agents, located on a grid that wrapped around in one direction only (i.e., it was a cylinder).[6] As noted in the previous section, we systematically varied several factors, beginning with the size of agents' repertoires. The number of identities at each agent's disposal can be thought of as representing that agent's "knowledge" or "sophistication." It can vary from just one to the full spectrum of identities initially present in the landscape.[7] In the experiments reported here, the size of the spectrum is 15, and agent repertoires vary in size from 2 to 14.

The second dimension of variation concerned the presence of small fractions of two different types of special agents. These agents, which we call mobilizers, are more ready to change in response to the information they receive (in other words, their threshold values are lower than those for standard agents). In addition, they update their activations and subscriptions before any of the other agents do so. In empirical terms, mobilizers represent opinion leaders or political entrepreneurs. In keeping with empirical findings about the density of such actors in society, we allow them to account for just 5% or 10% of the total population (cf. Weimann, 1994).

Empirical findings also suggest that opinion leaders have greater knowledge than do regular individuals. Accordingly, the repertoires of mobilizers in the experiments reported contain two additional identities compared with the repertoires of normal agents. In addition, opinion leaders tend to incorporate and express new information more readily. In other words, their thresholds for activation and subscription are below those of the average agent, and they also react more rapidly to new information. The behavior of mobilizers in our model is defined so as to reflect these properties. Mobilizers update their activations and subscriptions before all other agents. Moreover, the activation thresholds of normal agents—1 (to activate a subscribed identity), 4 (to subscribe to a new identity), and 6 (to subscribe and immediately activate)—are lowered to 0, 2, and 5, respectively, for mobilizers.[8] Finally, we distinguish between two types of mobilizers: entrepreneurs and innovators. As Weimann (1994), noted, innovators "tend to be more venturesome, less dogmatic, and more innovative" (p. 77) than more conventional entrepreneurs.[9] We model this characteristic by assigning entrepreneurs an influence level of 2, as opposed to 1 for innovators and standard agents. This means that when agents look around and calculate the value of every identity in the spectrum, an entrepreneur is essentially counted twice.

The final factor we varied was the nature of the environmental biases. In all experiments, environmental biases vary from –2 to +1 for each identity. At the beginning of a simulation, all values start out at 0. Every round, every bias value has a 0.005 (i.e., 1 in 200) chance of being reviewed. Under the first updating algorithm (which has been used in all ABIR research until now), the new bias value is picked at random from within the allowed range. This has two implications. First, there is a 25% chance the bias value will not change. Second, because all changes are equally likely, a jump from –2 to +1 is just as likely as a jump from –2 to –1.

In the real world, one can reasonably assume that environmental biases are more likely to change gradually than radically. It is unlikely that the most favorable identity one period will suddenly be the least favorable the next period. The second algorithm operationalizes this notion, making a jump of Size 1 (from 0 to –1, for example) twice as likely as a jump of Size 2, which in turn is twice as likely as a jump of Size 3. However, the value always changes.[10] This means that fewer drastic changes in bias value will occur, rendering the environment more predictable from one round to the next. However, the overall stream of bias values remains random, so it remains an empirical question whether or not the localized predictability will affect a population's performance.[11]

In systematically varying each of the above factors, we generated $(13 + 2 \bullet 2 \bullet 11) \bullet 2 = 114$ different specifications: 13 repertoire sizes for the no-mobilizer condition, 11 repertoire sizes (normal agents 2-12, mobilizers 4-14) for the four different mobilizer conditions (innovators and entrepreneurs at 5% and 10%), and each of these specifications with random and predictable environmental biases. We randomly generated 30 different initial populations for each specification and ran each system for 1,000 rounds using a randomly generated stream of bias values. Unless otherwise indicated, all results reported are average values across these 30 different populations.

It remains to operationalize our concept of performance, that is, a population's ability to react to and exploit environmental signals about the relative appeal of different identities. A population can be said to do well if it consistently manages to have most of its agents activated on identities that are currently valued in the environment. But, of course, it could arrive in such a position by sheer chance. To get a measure of the amount of "credit" a population deserves for having most of its agents activated on valued identities, we need to calculate first how much its success sets it apart from the experience of a population that fails to do any adaptation or learning.

As a baseline for performance, therefore, we take a population evenly distributed across all available identities, as all populations are time 0. A completely nonadapting population remains distributed in this manner, whereas our populations will tend to seek out those identities currently favored by the environment. Each round, we calculate the average bias value across all identities and multiply this by the population size—this is the baseline score, indicating how the nonadapting population would fare given the present distribution of biases. Next, we add up the bias values seen by all agents for their currently active identity and subtract the baseline score from this sum. This yields a performance score for that round.[12] Performance scores can range from –7,166.67 to 7,000,[13] but in practice, they will rarely take on such extreme values, as we will see.

RESULTS

We present the results of our experiments in three sections. In the first, we examine the relationship between repertoire size (i.e., knowledge or sophistication of the agents) and performance. The second section investigates the implications for performance of the presence of different fractions of mobilizing agents. The third set of results considers the impact of changing from randomly selected environmental biases to values that are more predictable in the short term.[14] Before discussing our findings, it is worthwhile to examine the results for a single model specification to get a sense both of the levels of performance we may expect and of the variation over time and across runs. Figure 1 shows the average performance over time of a population faced with an environment where the biases assigned to each identity fluctuate randomly. In this population, every agent has a repertoire of eight identities, and no mobilizers are present. The central line displays average performance across the 30 runs; the lighter, outer lines trace the path of average performance plus or minus one standard deviation.

Figure 1: Repertoire 8, No Mobilizing Agents, Random Bias

Several patterns are evident. First, performance improves rapidly at the start, as agents react to environmental signals by changing their activation away from any identities that receive a negative bias and toward those that receive a positive bias. At a certain point, however, performance begins to decline. This is a sign that the adaptive capabilities of the population as a whole have deteriorated. From an initial pattern of randomly distributed activations, populations rapidly coalesce into small clusters activated on the same identity. As in the real world, agents are likely to be convinced to activate on a new identity in response to pressure from those they interact with, combined with strong environmental signals in favor of, or against, particular identities. However, clusters are less likely to react nimbly to environmental signals, as the influence of seeing one's neighbors all expressing the same identity will often outweigh any negative environmental signal associated with that identity—peer pressure at its purest![15] Of course, some adaptation continues to occur at the margins of these clusters, enabling the population as a whole to continue to register a positive performance score.

A second reason for the deterioration of performance is the loss of diversity in a population. When a particular identity is not actively expressed by any agents in the population, the population as a whole is unlikely to be able to take advantage of positive environmental biases for that identity in the future.[16] Such a loss in ability to react to changes in the relative bias of an identity implies a reduction in overall adaptive capacity. This recalls the quotation opening the article, which pointed out that short-term adaptation—abandoning an identity that is temporarily unattractive—may come at a cost in long-term performance. Figure 1 illustrates that as the overall diversity of the population declines, it will become increasingly difficult for the population as a whole to outperform a completely nonadaptive population.

A final feature of Figure 1 worth noting is the breadth of the standard deviation. During the course of runs with 30 different initial populations, quite a few spend time in negative performance territory. Investigation reveals that these tend to be the runs where diversity decreased most.[17] Conversely, those runs that continue to prosper at relatively high levels feature populations where the initial clustering was such that diversity was maintained successfully over time. For clarity of presentation, the standard deviation lines will not be shown in most of the subsequent figures. Nevertheless, it is worth keeping in mind that the averages presented below hide a considerable amount of cross-run variation.[18]

Repertoire Size and Performance

In this section, we analyze the relationship between performance and repertoire size. Figure 2 shows performance over time for four different repertoire sizes.[19] The most obvious finding here is that increasing agents' knowledge or sophistication results in a large performance payoff for about the first 500 rounds of a run. The greater the repertoire size of the agents, the more rapidly initial performance distinguishes itself from a nonadaptive population. However, marginal returns diminish rapidly, and not much additional benefit is obtained by increasing repertoire size from 9 to 12.

The subsequent evolution of performance illustrates the interacting effects of the two main factors affecting performance in our model, as already discussed in the context of Figure 1: the degree to which diversity is maintained and the ability to react quickly to changing environmental signals. Diversity matters because as fewer and fewer agents are activated on a particular identity, the population as a whole becomes less and less able to react well and rapidly to a change in environmental signals in favor of that identity, because so few other agents will be in contact with an agent still expressing that identity. The ability to react quickly to changing signals is important to exploit opportunities offered by the environment in the form of (temporarily) favored identities.

In theory, the ability to react quickly to environmental signals grows with the sophistication or knowledge of an agent. The larger an agent's repertoire, the more likely it becomes that the agent already subscribes to an identity whose environmental bias changes for the better. Because it is easier for an agent to activate on an identity already in the repertoire than to subscribe to a new identity, agents with larger repertoire sizes are better able to take advantage of changes in environmental biases. This explains why larger repertoire sizes outperform smaller ones, especially at the start, in Figure 2.

However, this discussion ignores the context within which an agent finds itself. To take advantage of changing environmental signals, an identity not only needs to be in an agent's repertoire but the agent also needs to encounter that identity among its neighbors.[20] This is where diversity comes in. One way to measure diversity is to see how many different identities an agent sees around it on average, not counting its own identity. The more homogeneous local neighborhoods are, the lower this value will be.

The initial reorganization of the population evident in Figure 2 has quite striking implications for this measure of diversity. After about 300 rounds, agents in the Repertoire 9 and 12 populations see just 2 identities other than their own, on average. Those in the Repertoire 6 populations still see 3, and those in the Repertoire 3 populations still see close to 5 different identities around them. After 1,000 rounds, these values have declined further to about 1.25 (for 9 and 12), 1.75 (for 6), and 3 (for 3). The implications for a population's ability to react nimbly are obvious.

Another measure of diversity is the Herfindahl index, which is the sum of the squares of the fractions of the population occupied by each identity.[21] Given a spectrum of 15 identities, its initial value will be $15 \bullet (1/15)^2 = 1/15 \approx 0.067$. Higher index values indicate greater concentration on a limited number of identities. After 1,000 rounds, the average index value for repertoire Size 3 has increased by about 50%. In contrast, for repertoire Sizes 6, 9, and 12, the Herfindahl index has already doubled after about 600 rounds. Perhaps even more interesting, the average index value for repertoire Size 9 exceeds that for repertoire Size 12 at all times.[22]

Figure 2: Various Repertoires, No Mobilizers, Random Bias

Combining this information, it becomes easy to understand why Repertoires 3 and 12 should do better than 6 and 9 in Figure 2. Populations with repertoire Size 3 easily do best at maintaining diversity, by any measure. However, their low levels of sophistication or knowledge imply that they are much less able to react to changing environmental signals than are populations with larger repertoire sizes. The largest loss in diversity occurs in the jump from repertoire Size 3 to 6. The additional loss in moving from 6 to 9 is still noticeable, but the difference between Sizes 9 and 12 is small (or even negative). On the other hand, the benefits of greater sophistication in terms of ability to react to new signals continue to accrue. This explains why the populations with repertoire Size 12 do better than those with repertoire Sizes 6 or 9.

In Figure 2, we observe that repertoire Size 9 performs worst in the last part of the run. Moreover, the Herfindahl index data suggest that this repertoire size is actually less successful than larger repertoire sizes in maintaining diversity. It turns out that small clusters of agents activated on the same identity form more rapidly as repertoire sizes increase, because agents are likely already to subscribe to the active identities among their neighbors. However, this same fact means that it becomes increasingly difficult for a powerful homogenizing wave to affect a large number of agents. As a result, smaller size clusters actually become more viable as repertoire sizes increase. The resulting patterns of activation, when depicted graphically, resemble a vibrant, colorful collage made up of many small components, as opposed to the large swaths of the same color that result when small clusters are less stable. For this reason, we shall refer to it as the Matisse effect, for its tendency to generate patterns recalling that artist's abstract collages.[23]

Given a spectrum of 15 identities, it appears that repertoire Size 9 gets the worst of both worlds in terms of losing diversity yet being unable to sustain small homogeneous clusters. Repertoire Size 6, on the other hand, benefits from the combination of maintaining diversity better than larger repertoires, while being able to react to changing environmental signals much more nimbly than smaller repertoires. How well do these findings hold up when we take into consideration the full range of repertoire sizes, as well as the different population compositions tested? For such broad comparisons, line graphs showing performance over time become impractical. Instead, Figure 3 displays average performances for the beginning, middle, and end of each run (with cutoff points suggested by Figure 2), averaged across different population composition conditions (whose independent effects will be discussed below).

The figure indicates that the Matisse effect described above is noticeable even quite early on in a run. Although agents with large repertoires are better able to exploit initial changes in bias, it does not take very long before small clusters form that are relatively stagnant. As repertoire size increases, smaller and smaller clusters become stable, and thus populations with larger repertoires get less of an initial payoff from their relative sophistication. As we surmised above, populations with repertoire Size 6 are best placed to combine some degree of sophistication with the maintenance of diversity, at least in the medium run. Repertoire Sizes 8 and 9, on the other hand, are shown to suffer most from the combination of loss of diversity and inability to sustain a large number of small clusters, especially toward the end of the run. It is at this point that the Matisse effect that was costly to larger repertoires early on in the run begins to provide a benefit, as Figure 3 shows very nicely. After all, smaller clusters mean that particular identities are present in more different places in the population, allowing the population better to exploit changes in environmental signals.[24]

Population Composition and Performance

In our earlier work (Lustick & Miodownik, 2000; Lustick, Miodownik, & van der Veen, 2000), we found that entrepreneurs may both help and hurt the aggregate performance of a population. On one hand, entrepreneurs are more eager to react to changing environmental signals and thus may boost performance (or at least increase the rate of adaptation). They also help maintain diversity by reducing the ability of environmental signals to sway their neighbors, and this may become valuable later in a run. On the other hand, this same localized "anchoring" effect reduces performance by inhibiting adaptability. Innovators do not introduce this same anchoring effect, because their influence is the same as that of standard agents. On the other hand, innovators, too, increase the speed of adaptation and may hurt overall performance by rapidly forming small clusters that are relatively immune to environmental signals. In this respect, their impact is analogous to that of larger repertoire sizes in producing the Matisse effect described above.

Figure 4 shows the performance of populations with different fractions of innovators and entrepreneurs.[26] The introduction of entrepreneurs does indeed inhibit performance for the first two thirds of the run (on average), and this effect gets stronger as the fraction of entrepreneurs increases. The better performance of populations with innovators indicates that it is the anchoring effect exerted by the higher influence level of entrepreneurs that is causing this pattern. The hypothesized ability of entrepreneurs to retain diversity begins to show itself in the latter part of the run for populations with 10% entrepreneurs. For smaller fractions of entrepreneurs, this ability fails to overcome the costs introduced by their anchoring tendencies, whereas for larger fractions, the two effects appear to cancel each other out.

Figure 4 also shows that innovators can help improve overall performance. Both of the conditions with innovators outperform the no-mobilizer condition for a good part of the first 500 rounds. After that point, however, populations with 5% innovators deteriorate rapidly in performance, and from round 600 on, their success (or lack thereof) is very similar to that for the 5% entrepreneur condition. The 10% innovator condition, in contrast, continues to outperform the other four conditions until the very end. It seems that introducing 5% innovators into a population does not provide much of a payoff either early on or later in a run. In con-

Figure 3: Performance Across Repertoire Sizes 2-12 (average across 5 different population composition conditions)

Figure 4: Repertoire 8, No Mobilizers, Random Bias

trast, 10% innovators do help maintain diversity by increasing the ability to react to environmental signals, as well as by introducing Matisse effects that help maintain diversity later on in the run.

The anchoring effect of entrepreneurs becomes clear when we look at the number of different identities agents encounter among their neighbors. Although the differences across population compositions are less striking than those reported for the different repertoire sizes above, it is worth noting that the number of different identities seen by the average agent when there are entrepreneurs present in the population is systematically lower than that for populations without mobilizers, whereas the introduction of innovators into the population systematically increases the average number of different identities seen.

On the other hand, it is worth noting that anchoring *does* help maintain diversity at the overall population level, even if it tends to make local neighborhoods more homogeneous. If we consider the Herfindahl index, for example, we find that both conditions with entrepreneurs do considerably better than any of the other conditions in keeping the overall distribution of activated identities fairly even. Populations with 10% innovators also clearly improve on populations without mobilizers, whereas 5% innovators are not enough to make a clear difference.[26] This shows that maintaining diversity in itself does not suffice to guarantee good performance—agents also have to be able to benefit from the diversity by reacting to new environmental signals, and that ability is inhibited by the influence level of entrepreneurs.

Next, we assess the validity of our analysis by considering aggregate performance for each of the five different population compositions in our experiments. The preceding discussion suggested that the two conditions with entrepreneurs should do clearly worse than the others for about the first two thirds of each run. In the latter part of the run, the 10% entrepreneur condition would be expected to perform slightly better than the no-mobilizer condition, whereas the 5% entrepreneur condition continues to perform badly. Innovators affect the performance of populations relatively little at the start; after that, they should improve the performance somewhat for the main part of the run; in the last third, the 10% innovator condition should perform best, whereas the 5% innovator condition performs on a par with the 5% entrepreneur condition. Figure 5 displays the performance of the different population composition conditions, averaged across all repertoire sizes from 2 to 12.

Figure 5 shows that entrepreneurs always tend to lower the average performance level, even later in the run. The ability of 10% entrepreneurs to retain diversity is clearly negated in most cases by the anchoring effect that makes it near impossible for a population to benefit from this diversity. Introducing innovators makes much less of a difference on the whole. Adding just 5% innovators does not seem likely to improve performance, as we already concluded from Figure 4. Apparently, the increased nimbleness provided by the occasional innovator is counteracted by the tendency to promote relatively rapid initial homogenization, as surmised above (cf. note 26).[27] Adding 10% innovators is slightly more promising, because this condition is better able to maintain diversity. As a result, it is the only condition able to outperform the no-mobilizer condition, and it does so at the point when the maintenance of diversity matters most, that is, near the end of the run.[28]

It is worth noting that the expected contribution of mobilizing agents to performance depends on the nature of the environment. In particular, the environment studied here is relatively safe (no extreme bias values) and stable (overall set of biases changes on average once every $200/15 \approx 13$ rounds). As the environment becomes less stable, we would expect the greater nimbleness of mobilizers to become more of an advantage, as opportunities for exploiting positive biases do not last as long. In addition, as the environment becomes riskier, the anchoring ability of entrepreneurs is likely to become a valuable asset, helping to prevent or slow down massive homogenizing waves that threaten to wipe out diversity very rapidly. In the latter case, the nature of the changes in environmental biases will obviously be quite important. Indeed, protocols for bias change have an impact even within a safe and stable environment, as we discuss next.

Figure 5: Performance Across Population Compositions (average across repertoire sizes 2-12)

Figure 6: No Mobilizers, Random and Predictable Bias

Bias Predictability and Performance

The third variable in our experiments was the algorithm used to change bias values. Figure 6 compares the performance, for repertoire Sizes 4 and 8, of each of the two algorithms. The most striking feature of the graph is how closely the two lines for each pair track one another. Clearly, the short-term predictability of the environment does not have a fundamental impact on the ability of a population to improve its performance by learning and adaptation, at least within a narrow bias range such as the one studied here.

On the other hand, Figure 6 also shows that performance, on average, does improve in the more predictable environment. For repertoire Size 4, performance improves by 10.5% for the first half of the run and drops by 1.5% during the second half, for an average improvement of about 4.5%. For repertoire Size 8, much of the payoff comes in the latter part of the run: Performance in the first half improves by 4.5% but in the latter by a striking 15.5%, for an 8% improvement overall. However, the average improvement fluctuates arbitrarily from one repertoire size to the next, so one should not place too much emphasis on the size of these improvements.[29] Taken across all repertoire sizes and population conditions, the main findings are twofold: (a) Populations react in a very similar way to the two environments, resulting in performance trajectories that very closely parallel one another; and (b) on average, populations do marginally better under more predictable environmental conditions.[30]

How well populations are able to make use of the short-term predictability is related in part to the average agent's repertoire size. Indeed, we find that the average net effect of more predictable biases is negative for the lowest and highest repertoire sizes, whereas it is positive for repertoire Sizes 3 to 8. This suggests that the loss of diversity that is normally so costly to populations with these intermediate repertoire sizes matters relatively less when the environment is predictable in the short term. After all, diversity is prized in part because it allows populations to adjust quite rapidly to a sudden change in environmental bias against a currently popular identity (or the converse). When bias changes are more gradual, populations have more time to look out for other, more attractive identities, and thus small reductions in diversity should matter much less. The results support this explanation.

Interestingly, the presence or absence of mobilizing agents appears to matter relatively little to a population's ability to exploit a more predictable environment. Three of the four mobilizer conditions showed very similar results to those for no mobilizers. Only for the fourth, 10% innovators, did we find that greater predictability actually reduces performance, on average. As we saw in the previous section, this is the condition under which performance on average is best. It appears, therefore, that making the environment more predictable eliminates the advantage provided by the presence of these innovators. This advantage, we noted earlier, must be sought in the ability of innovators to retain greater diversity as well as to react rapidly to changes in the environment. As we have discussed here, however, the penalty associated with small reductions in diversity is largely eliminated when the short-term predictability of the environment improves. This same predictability also lowers the value of reacting rapidly to changes. Indeed, rapid reaction may come at a cost, as populations change away from an identity whose appeal diminishes only temporarily, only to return to where it was previously. Such a situation will be more common in the predictable environment than in a context of random biases.

CONCLUSION

The experiments discussed above generated several important findings. First, although it might seem that increasing the overall sophistication of a population is always a good thing, it turns out that in some cases, doing so can be counterproductive. Greater sophistication carries the cost of reducing diversity and thereby hampering a population's ability to react to changing environmental signals. However, as sophistication continues to increase, the Matisse effect mitigates the loss in diversity. Optimal repertoire sizes will vary with the overall size of the spectrum of identities. Our findings suggest that repertoire sizes just above 1/3 of the spectrum will tend to do best for the first 1,000 rounds, whereas sizes just below 2/3 of the spectrum will tend to do quite well initially but will disproportionately suffer later on, even compared with larger repertoire sizes. The longer one continues a run, the more likely it is that the largest repertoire sizes will come out on top, because the combination of their sophistication and the Matisse effect means that reductions in diversity and their costs will be less pronounced.

Introducing mobilizing agents rarely improves performance, except perhaps in the very long run. Entrepreneurs produce a significant anchoring effect that, although it helps maintain diversity, also inhibits the ability of agents to react to changing environmental signals. Innovators do add to the ability of populations to exploit environmental signals, but they also increase the tendency toward the production of Matisse effects. This hampers performance for the first several hundred rounds, in the same way as is the case for larger repertoire sizes. Only as one gets closer to 1,000 rounds does the Matisse effect start contributing to relative performance, and at this point, the 10% innovator condition indeed outperforms even the no-mobilizer condition.

Finally, making the environment somewhat more predictable eliminates the relatively small advantages provided by innovators. Populations with medium-sized repertoires are best positioned to benefit from a more predictable environment. A large repertoire improves one's ability to react well even to fairly abrupt changes in environmental bias, whereas a small repertoire makes it difficult to take advantage of the predictability.

Naturally, the particular conclusions drawn here depend on the specifications used in our experiments. For example, if one were to increase the bias range, the outcomes would change. Nevertheless, they would change in relatively predictable ways, as discussed earlier, and hence our findings here will still be relevant.[31] Similarly, increasing the spectrum of identities is likely to invalidate predictions about the success of particular repertoire sizes, but relative predictions such as those given above ought to be robust. Analogously, we can predict with some confidence that the small effects found for a predictable environment will increase in significance as the bias range increases.

Perhaps a more important consideration is the degree to which we can use these results to make predictions about real-world situations. Here, our findings present some intriguing and potentially valuable hypotheses both about the impact of knowledge on the ability of a population to make use of the signals provided by its environment and about the role of mobilizing agents (political entrepreneurs, opinion leaders, etc.). With regard to the first, our experiments provide an elegant illustration of the degree to which different implications of increased knowledge interact. It would be very interesting to look for these same types of interactions in large-scale public opinion surveys, such as the General Social Survey, the National Election Studies, or the Eurobarometer series.

Turning to the role of mobilizers, here the central issue appears to be the general tendency of such agents to reduce the performance of a population, at least in a relatively safe and stable environment. The brief reference to peer pressure earlier in the article already suggested the type of contexts in which our findings could be tested. There has been much research into the characteristics of mobilizers and their ability to sway "normal" people (see Weimann, 1994), but much less appears to be known about the ability of such actors to improve the aggregate fortunes (happiness, success, etc.) of the population within which they are active.

In closing, we need to address the relationship between performance and learning. By some definitions the ability of our populations to react to changing environmental signals so as to improve performance is a clear manifestation of learning. Similarly, if we think of learning as the ability to handle a given situation better when we see it a second (or third) time, in most cases our populations can be argued to learn. However, if the second appearance of a given set of environmental signals occurs hundreds of rounds later, it is possible that a loss of diversity has made the population *less* able to handle that situation. In fact, our experiments powerfully illustrate the potential trade-off between adaptation and diversity that is central to learning in a dynamic environment. Given the limits on knowledge imposed by our model (where repertoire is always less than the spectrum of identities), as well as the long-run unpredictability of the environment, this trade-off can never be completely avoided. Maintaining diversity hurts performance in the short run but will help in the long run, regardless of the nature of the population.

If we accept the limitations introduced by our model, one might conclude that true, emergent learning is not possible—no possible population specification can be guaranteed to outperform a nonadapting population in the (very) long run. To achieve true learning, then, one has to be able to bypass (or at least ameliorate) the adaptation-diversity trade-off. Interestingly, it is possible to think of some relatively simple changes that might make this possible. In this article, we focused on diversity as measured in terms of activations. However, subscriptions often retain diversity better than do activations. If a population is able to mine these "hidden" features of agent repertoires, it may well outperform nonadapting populations even in the long run.

How might a population as a whole be able to access subscriptions? Two possibilities immediately suggest themselves. First, it is possible to increase the bias range to such a degree that mobilizers—or any agents with a low activation threshold—will spontaneously activate on an identity in their subscription even when no neighboring agents are actively expressing that identity. Second, keeping the bias range constant, one could imagine mobilizing agents that are even more innovative than the innovators discussed here, and that would occasionally activate on an identity whose calculated value is merely equal to, or even slightly below, its currently activated identity, as long as its current environmental value is positive.[32] Finally, it is worth thinking about the role that institutions might play in storing knowledge and thus keeping certain identities accessible to the population even if no individual were actively expressing that identity. This would require introducing an additional source of information analogous to that provided by the environment that would not necessarily need to be global in scope but would not be directly associated with any given agent.

If this conclusion seems to raise rather more questions than the article answered, this should hardly be considered a failing. After all, the research presented here arguably does succeed in meeting two important goals for agent-based social simulations: to illustrate the workings of particular social effects with as few assumptions as possible and to generate new and testable hypotheses that can be applied to real-world contexts.

NOTES

1. The latter model is used in much of the genetic algorithms (GA) literature: Individuals do not learn—they are merely replaced by other individuals likely to perform better. For a discussion of the difference between individual and social learning in the GA context, see Vriend (1998).

2. For an overview of some other simulation models of groups and organizations, see Bainbridge et al. (1994, pp. 423-425).

3. In this article, for the sake of clarity and simplicity, we will use the language of agents activating alternative identities from a repertoire of identities they subscribe to.

4. Agents may be assigned different weights in this calculation, depending on their influence level. In our experiment, for example, entrepreneurs are given twice the weight of all other agents. This aspect of the experiment is discussed below.

5. All agents perform these calculations synchronously. It is possible that performing updates asynchronously would substantially change our results; however, preliminary analyses indicate that this is unlikely, except perhaps when agent repertoires (identity lists) are very small (see van der Veen 2001). For more detailed information, including the resolution of various possible ties, see the appendix to Lustick (2000).

6. The one-dimensional wrapping was selected for the sake of consistency with earlier experiments.

7. For most experimental purposes, the minimum and maximum sizes are not tremendously interesting, because each invalidates two of the three threshold values.

8. This means that whereas a normal agent will not activate on a new identity unless its calculated value (combining neighborhood and environmental information) exceeds that of its currently active identity by at least 2, mobilizers will simply activate on the identity whose calculated value exceeds that of all other identities in their repertoire (ignoring for the moment the dynamics of adding a new identity to one's subscribed repertoire). As a result, they are much less resistant to change.

9. For a summary of the empirical evidence, see Weimann (1994, chap. 5).

10. To appreciate the differences between these two algorithms, consider what happens when a bias value of 0 is up for change. With the first algorithm, the new values and their probabilities are –2: 25%, –1: 25%, 0: 25%, and 1: 25%. With the second algorithm, they are –2: 20%, –1: 40%, 0: 0%, and 1: 40%. When all biases, on average, have changed just once, the average bias value under the first algorithm will be lower (–0.5 versus –0.4). However, it can be shown fairly simply that the average bias value for the second algorithm will trend rapidly toward –0.5 as well.

11. The narrow bias range reduces the expected difference between the two updating algorithms. In other words, if we find an effect for the more predictable algorithm even within this range, we can assume that it will be all the more evident as the bias range widens.

12. As an example, consider what happens when a single identity has taken over the landscape. The baseline score will be the average bias value, whose expected value is –0.5, times 2,500, that is, about –1,250. The raw score for the population will be 2,500 times the current bias value. If that bias is negative, this population is actually worse off as a whole than if it had not adapted at all since the start, and the performance score is either –1,250 or –3,750. If it is zero or positive, the performance score is 1,250 or 3,750, respectively, a sign that the population is better off for having adapted.

13. The worst case is when all identities have a bias value of +1 except for one that has a bias value of –2, and which happens to be the one that is active for every agent. Conversely, the best case is when all identities have a bias value of –2, except for one that has a bias of +1, and all agents are activated on the latter.

14. To limit the number of variables under consideration in each group, results for the first two groups are reported only for randomly selected biases. Corresponding results for predictable biases are available from the authors on request.

15. See Lustick, Miodownik, and Philbrick (2000) for a more detailed discussion of how this characteristic of the system models the stickiness of norms in political contexts.

16. This statement is not strictly true, as can be easily shown. However, it tends to hold quite strongly after the first 100 rounds or so in situations where bias ranges are relatively narrow, as they are in the present context. On the other hand, when biases are allowed to become large and positive, their weight alone may suffice for an agent who is eager to adapt to activate on a particular identity in its repertoire, even if none of its neighbors are activated on that same identity.

17. One measure of diversity is the Herfindahl index, which is described below. There is a clear correlation between below-average performance over time and higher values for the Herfindahl index, indicating reduced diversity.

18. Indeed, the performance of the best run for the worst possible specification easily exceeds the performance of the worst run for the best specification. Nonetheless, it is fairly easy to show that the differences between the different averages reported here are usually highly statistically significant, even when their standard deviation ranges overlap considerably.

19. In the interest of legibility, the data have been smoothed by taking 20-round averages as data points. The same overall patterns are present without smoothing, but round-to-round fluctuations make it more difficult to discern these patterns.

20. If we allow the bias range to expand, this statement no longer strictly holds true, but its general implications continue to apply.

21. The index is best known from its use by economists, to provide a measure of the concentration of an industry, with market shares taking the place of the population shares in our calculations. However, it is also being used with increasing frequency to trace ethnic fractionalization patterns.

22. A third way to measure diversity is to count the identities with activation levels across the population that exceed a certain threshold. The pattern here is in line with that for the other two measures.

23. See, for example, his "L'escargot" (Snail) or "Lierre en fleur" (Ivy in flower).

24. As one might guess from Figure 2, the overall trend for simulations continued beyond 1,000 rounds tends toward zero, as the loss of diversity begins to have an effect on all populations, regardless of repertoire size. However, the power of the Matisse effect becomes increasingly important here, as the larger repertoire sizes manage to keep performance levels significantly above zero on average at least until 2,500 rounds. Conversely, the loss of

diversity affects populations with repertoire Size 6 quite severely, as might be expected from the preceding discussion. Indeed, this is the only repertoire size whose performance turned clearly negative on average after about 2,000 rounds.

25. Repertoire Size 8 was chosen both because it is halfway across the range of repertoires tested (2-14) and because the patterns for this repertoire size demonstrate both the adaptability and diversity effects discussed for Figure 2, falling roughly between the pattern for repertoire Sizes 6 and 9 in that figure.

26. If anything, they appear to make the distribution across identities slightly less even, by increasing the likelihood of small homogenizing waves without being numerous enough to generate the countervailing Matisse effect.

27. This also explains why the initial phase of the run is the only time when populations with 5% innovators on average slightly outperform populations with 10% innovators.

28. If one continues runs beyond 1,000 rounds, the ability of all mobilizers to maintain diversity becomes increasingly important. After about 1,500 rounds, we found that the no-mobilizer condition tended towards slightly negative performance levels, on average, whereas the other four conditions sustained low but positive performance levels at least until 2,500 rounds (again, on average).

29. This problem would be mitigated by using a larger sample of runs for each specification.

30. It is also worth noting that, except for the largest repertoire sizes, the standard deviations associated with the average performance values decrease by around 10% on average for the predictable bias environments. In other words, the increased predictability of the environment results in fewer outlying runs in terms of performance (i.e., highly successful or highly unsuccessful runs). This finding holds across the different conditions with mobilizing agents too.

31. In particular, a wider bias range will tend to provide an advantage to larger repertoires, because the cost of reductions in diversity as measured in terms of activations (or in terms of the diversity each agent sees around it) would matter less.

32. Another possible change is to increase the sight radius of agents or at least change the set of agents they interact with so that it includes not only their contiguous neighbors. See Watts (1999) for a detailed discussion of interaction topologies and their implications for aggregate outcomes.

REFERENCES

Bainbridge, W. S., Brent, E. E., Carley, K. M., Heise, D. R., Macy, M. W., Markovsky, B., & Skvoretz, J. (1994). Artificial social intelligence.*Annual Review of Sociology*, *20*, 407-436.

Carley, K. M. (2000). Organizational adaptation in volatile environments. In D. R. Ilgen & C. L. Hulin (Eds.), *Computational modeling of behavior in organizations: The third scientific discipline* (pp. 241-268). Washington, DC: American Psychological Association.

Lustick, I. S. (2000). Agent-based modeling of collective identity: Testing constructivist theory. *Journal of Artificial Societies and Social Simulations* [Online], *3*(1). Available: `http://jasss.soc.surrey.ac.uk/3/1/1.html`

Lustick, I. S., & Miodownik, D. (2000). Deliberative democracy and public discourse: The agent-based argument repertoire model. *Complexity*, *5*(4), 13-30.

Lustick, I. S., Miodownik, D., & Philbrick, S. (2000, August). *The institutionalization of identity: Micro adaptation, macro effects, and collective consequences*. Paper presented at the annual meeting of the American Political Science Association, Washington, DC.

Lustick, I. S., Miodownik, D., & van der Veen, A. M. (2000, October). *Studying performance and learning with ABIR: A research note*. Paper presented at the Agent 2000—A workshop on the simulation of social agents: Architectures and institutions, University of Chicago/Argonne National Laboratory, Chicago.

Nelson, R. R., & Winter, S. G. (1982). *An evolutionary theory of economic change*. Cambridge: Belknap.

van der Veen, A. M. (2001). *All together now: The impact of synchronous updating on agent-based simulations in the social sciences*. Unpublished manuscript.

Vriend, N. (1998). *An illustration of the essential difference between individual and social learning, and its consequences for computational analyses*. Working paper, Queen Mary and Westfield College, University of London.

Watts, D. J. (1999). *Small worlds: The dynamics of networks between order and randomness*. Princeton, NJ: Princeton University Press.

Weimann, G. (1994). *The influentials: People who influence people*. Albany: State University of New York Press.

Maurits van der Veen is a postdoctoral fellow at the Christopher H. Browne Center for International Politics at the University of Pennsylvania. He can be reached at `maurits@sas.upenn.edu`.

Ian S. Lustick is a professor of political science at the University of Pennsylvania. He can be reached at `ilustick@sas.upenn.edu`.

Dan Miodownik is a doctoral candidate in political science at the University of Pennsylvania. He can be reached at `danm@sas.upenn.edu`.

ADDITIONAL INFORMATION ON SOURCES

The Agent-Based Identity Repertoire (ABIR) model has been programmed in C by Vladimir Dergachev. An executable copy of the program for Windows systems, along with a manual for its use, can be downloaded at `http://www.polisci.upenn.edu/profileil.html`. The version of ABIR used in this article was programmed by Maurits van der Veen in Common Lisp on a Macintosh.