The Effect of Reinforcement, and the Roles of Mutation Rate and Selection Pressure, in an Evolutionary Theory of Behavior Dynamics

J. J McDowell¹

Published online: 5 July 2017

© Association for Behavior Analysis International 2017

Killeen and Jacobs (2016) recommend considering organismic states, dispositions, emotions, motivations, and environmental affordances in accounts of behavior. An evolutionary theory of behavior dynamics (McDowell, 2004) already incorporates many of these items. In the first two sections of this reaction to their paper, I describe complexity theories in general and the evolutionary theory in particular. In the remaining sections, I explain how the evolutionary theory includes many of the factors that Killeen and Jacobs find important.

The Nature of Complexity Theory

The evolutionary theory is a complexity theory (McDowell & Popa, 2009), which means that it is stated in the form of simple rules that operate relatively locally. Complexity theories are different from more typical calculus-based mathematical theories, which are often stated in the form of differential equations that relate relatively high-level dependent and independent variables. In complexity theories, larger-scale observed phenomena arise from the combined action of the low-level rules that constitute the theory. As an example, consider the behavior of a murmuration, or flock, of starlings (see Bunyard-Wild, 2013, for a video). In flight, the flock changes shape continuously, sometimes in elegant and beautiful ways, and apparently unpredictably. This behavior is not likely to be a function of large-scale controlling variables that act on the flock as a whole and that could be captured by differential equations. Instead, each bird's behavior is

This is an invited commentary on Killeen and Jacobs (2016).

✉ J. J McDowell
psyjjmd@emory.edu

¹ Emory University, Atlanta, Georgia, USA

probably governed by a few simple rules, such as “do not collide with other birds” and “stay with the flock.” The complex overall appearance of the flock is then the cumulative result of the low-level rules operating in the individual birds at each instant of time (see, e.g., Ballerini et al., 2008).

A second interesting example of a complexity theory is Muro, Escobedo, Spector, and Coppinger’s (2011) theory of wolf-pack hunting behavior, which entails tracking and encircling a prey item. They showed that two rules followed by each wolf are sufficient to generate this complex behavior. The rules are “move to, but remain a safe distance from, the prey item,” and “move away from other wolves in the pack.” Muro et al. implemented these rules in a computer program that successfully reproduced the wolf-pack hunting behavior, thus showing that it was not necessary for the wolves to communicate with each other, or to rely on a social hierarchy, to generate this behavior.

As illustrated in Muro et al., (2011) work, complexity theories are typically instantiated in the form of computer programs that implement the theories’ rules. To examine the predictions of a complexity theory, it is necessary to run the program that instantiates it and then study the emergent output generated by the joint operation of the theory’s rules.

The Evolutionary Theory

The evolutionary theory consists of low-level Darwinian rules of selection, reproduction, and mutation that operate on a population of potential behaviors (McDowell, 2004, 2013a). The behaviors are potential in the sense that they may or may not be emitted into the environment. Only one behavior, selected at random, is emitted from the population at each tick of time. This produces a continuous stream of observable behavior that can be studied in the same way that the behavior of a live organism is studied. Behaviors in the population of potential behaviors are represented by decimal integers, which are understood as phenotypes, and by their corresponding binary representations, or bit strings, which are understood as genotypes. Ranges of decimal integers can represent classes of behavior, such as different ways of pressing a lever, engaging in self-injury, or exhibiting anti-social activity. After a behavior is emitted, pairs of parent behaviors are chosen from the population and are recombined to produce child behaviors that constitute the population of potential behaviors at the next time tick. Bits taken from each of the parents’ bit strings are used to form the child behavior’s bit string. This is the theory’s reproduction rule.

If the behavior that was emitted from the population of potential behaviors resulted in a benefit, such as the acquisition of a resource or the escape from or avoidance of a threat, then parent behaviors are chosen from the population on the basis of their phenotypic similarity to the emitted behavior. This is the theory’s selection rule. Behaviors that are more similar to the emitted behavior are more likely to be chosen as parents than are behaviors that are less similar. Specifically, a probability density function, which will be described in more detail later, is used to choose parents for mating. If the behavior that was emitted did not result in a benefit, which is of course the more common outcome, then

there is no selection, and pairs of parent behaviors are chosen at random from the population.

After a new population of behaviors is produced, it undergoes a small amount of mutation, which is implemented by flipping random bits in the population's genotypes from 0 to 1, or from 1 to 0. This is the theory's mutation rule. Reproduction without selection (i.e., using randomly chosen parents), and mutation, both cause the phenotypes in the population of potential behaviors to vary. Reproduction with selection causes the phenotypes to become more or less concentrated near the phenotype that resulted in a benefit. These Darwinian processes of variation and selection generate behavior that adapts to conditions in the environment. In relatively constant environments, the forces of variation and selection produce behavioral equilibria in which the forces are roughly balanced. If the environment changes to a new relatively constant state, then behavior generated by the theory adapts to this change and arrives at a new equilibrium. In a constantly changing environment, behavior generated by the theory is continuously affected by the changing consequences for the emitted behaviors.

Based on a review of the empirical evidence bearing on the evolutionary theory, McDowell (2013a) concluded that the theory accurately accounts for a wide range of dynamic and steady-state live-organism behavior in quantitative and qualitative detail. An especially important finding is that the evolutionary theory generates an emergent equilibrium on concurrent schedules that is described by the power-function matching law (McDowell, 2013b) exactly, and without residual error (McDowell, Caron, Kulubekova, & Berg, 2008).

It is interesting to consider the empirical success of the evolutionary theory in view of its detailed structure and algorithmic operation, which do not correspond to objects and events in the material world. For example, behaviors do not reproduce or undergo mutation. Like the standard model of the atom (Cox & Forshaw, 2012; McDowell, 2013c), the evolutionary theory is a functional theory; it is not a theory of material mechanisms. However, the operation of the evolutionary theory likely makes contact with material reality at the level of neural computation (McDowell, 2013c). The theory can also be interpreted specifically as a functional implementation of Edelman's (1987) biological theory of neural Darwinism (McDowell, 2010), although it was developed independently of that theory.

States, Dispositions, and Emotions in the Evolutionary Theory

At each tick of time, the state of a virtual organism animated by the evolutionary theory is represented by its population of potential behaviors. Figure 1 shows a series of screen captures from a computer program that implemented the evolutionary theory for 60 generations, or time ticks. Phenotypes appear along the *x*-axes; the relative frequencies of the phenotypes in the population, which may equal zero, are represented along the *y*-axes. The relative frequencies indicate the probabilities that the phenotypes will be emitted at the next time tick and therefore represent the virtual organism's dispositions to respond. Phenotypes 0 through 41 (indicated by the vertical dashed line in the figure) constituted a class of behaviors that defined a lever press. The five images show populations of 100 potential behaviors after 0, 10, 20, 30, and 60 generations of working on a random-interval schedule that reinforced lever pressing.

The top image in Fig. 1 shows the population of potential behaviors at the start of the session, when all phenotypes were considered equally likely to be emitted, and 100 were chosen at random to constitute the first population. Ten generations later (second image from the top), no reinforcers had been delivered, but some changes in the virtual organism's dispositions to respond nevertheless were observed. These were due to the effects of reproduction and mutation. For example, note that the 100 phenotypes in the population tended to cluster, albeit to a small degree, in the range from about 400 to 500, and also around phenotypes of approximately 150 and approximately 300. Note also that there were longer ranges of phenotypes with zero relative frequencies after 10 generations than at the start of the session. After 20 generations (third image from the top in Fig. 1), 1 reinforcer had been delivered. The state of the virtual organism at this point differed markedly from its states after 0 and 10 generations. Specifically, phenotypes tended to cluster in and near the target class, indicating that the virtual organism was now more likely to emit a phenotype from this class, or from nearby regions, which represent behaviors close in phenotype to lever pressing, such as rearing over or nosing the lever. Notice also that at this point, the virtual organism was almost completely disinterested in emitting phenotypes greater than 600, which are far removed from the target class of lever pressing. The bottom two images in the figure show that, after 30 and 60 generations, and the delivery of a few more reinforcers, the population had shifted even further into and near the target class. At this point, the virtual organism was completely disinterested in emitting phenotypes greater than 200, that is, the probability of their emission at the next time tick was zero.

The states of the virtual organism and its predispositions to respond at the five time points shown in Fig. 1 were the result of the reinforcement history the virtual organism

Fig. 1 A series of screen captures from a computer program that implemented the evolutionary theory. The images show the population of potential behaviors after 0, 10, 20, 30, and 60 generations (or time ticks). Integer phenotypes ranging from 0 through 1023 are represented along the x-axes. The relative frequencies of the phenotypes in the population are plotted along the y-axes; note that these frequencies may equal zero. The vertical dashed line falls at phenotype 41. During the period of time shown in the figure, phenotypes 0 through 41 constituted a class of behaviors that defined a lever press. Lever presses (any emitted phenotype from 0 through 41) were reinforced on a random-interval schedule. The number of reinforcers delivered by the end of 0, 10, 20, 30, and 60 time ticks is shown on the right

experienced in the specific stimulus context that was in effect. In other words, the reinforcement history caused the state of the virtual organism to change at each moment of time.

McDowell, Soto, Dallery, and Kulubekova (2006) showed how a theory of stimulus control can be adjoined to the selectionist evolutionary theory. They associated different populations of potential behaviors with different stimulus contexts. When a particular context is in effect, the population of potential behaviors associated with that context evolves according to the Darwinian rules. When the context changes, the current population becomes dormant (in the state that existed at the start of the dormancy), and the population associated with the new context becomes active. A dormant population does not evolve. In McDowell et al.'s theory, context valences are generated by the Rescorla-Wagner rule, which charges the stimulus context with a positive valence as a function of the resources it produces, and can also charge it with a negative valence as a function of aversive events, such as threats that may be escaped or avoided. Organismic states and dispositions in the theory may change relatively gradually as a function of the virtual organism's history, but can also change rapidly as a function of changes in stimulus context.

Context valences in the theory can be interpreted as reflecting the virtual organism's emotional state. For example, in a context with a positive valence, the virtual organism is predisposed to emit appetitive behaviors because those behaviors have produced resources in that context in the past. Berg (2011) showed that virtual organisms animated by the theory will work to activate stimulus contexts with positive valences; in other words, they like those contexts. An observer might be inclined to say that the virtual organism feels good in such a context, or perhaps even confident and self-actualized, inasmuch as it seeks out and takes advantage of opportunities for resource acquisition when such a context is in effect. In a context with a negative valence, on the other hand, the virtual organism is predisposed to emit escape and avoidance behaviors because those behaviors successfully dealt with aversive circumstances in that context in the past. An observer might be inclined to say that the virtual organism feels bad, or perhaps fearful, in negative contexts inasmuch as it withdraws from potential threats, and tends to be avoidant when those contexts are in effect.

Fixed Dispositions in the Evolutionary Theory

Not all organismic dispositions change with time; some may be fixed. In the evolutionary theory, mutation rate sets a fixed organismic disposition along an inhibition-excitation dimension. A low mutation rate favors behavioral inhibition. This causes perseverative responding, which is a tendency to repeat behaviors from a narrow range of phenotypes (McDowell et al., 2008). A high mutation rate favors behavioral excitation. This causes impulsive responding, which is a tendency to emit behaviors from a wide range of phenotypes. At both low and high mutation rates, the virtual organism's behavior is less responsive to reinforcement than at moderate mutation rates (McDowell et al., 2008; McDowell, 2013a). Hence, the mutation rate sets a fixed balance between perseveration and impulsiveness that modulates the effect of reinforcement as the virtual organism interacts with the environment.

Likes, Wants, Affordances, and Motivation in the Evolutionary Theory

Many aspects of likes, wants, affordances, and motivation are represented in the evolutionary theory by a one-parameter probability density function that is used to choose parents for mating following a selection event (McDowell, 2004). Three linear, one-parameter density functions are shown in Fig. 2, where probability density is plotted against parental fitness. The fitness of a candidate parent is defined as the absolute value of the difference between its decimal integer phenotype and the phenotype of the emitted behavior that just produced a benefit to the artificial organism. Fitter candidate parents have smaller fitness values because they are more similar to the just-emitted behavior. As is clear from the figure, fitter candidate parents are always more likely to reproduce than less fit parents, but the selection pressure favoring fit parents, and hence fit child behaviors, may vary from strong (solid line) to weak (short dashed line), depending on the slope of the density function. This slope is determined by the function's single parameter. If the selection pressure is strong, then only very fit parents reproduce; if the selection pressure is weak, then less fit parents may also reproduce. Because fitter parents tend to produce fitter child behaviors in the next generation, stronger selection pressure increases the speed and extent of state changes in the population like those shown in Fig. 1.

Likes and wants affect selection pressure. For example, if an organism likes or wants an item, then that resource must entail a density function with a relatively steep slope. Less liked or less wanted items exert less selection pressure, and so necessarily entail density functions with shallower slopes. Affordances also affect selection pressure. For example, Killeen and Jacobs (2016) discussed the different affordances provided by a lever and a jump-out box for avoidance learning in rats. Even with the same reinforcer (e.g., foot shock with a specific voltage, amperage, and duration), the jump-out box enables greater selection pressure than the lever and hence necessarily entails a density function with a steeper slope. And finally, motivation also affects selection pressure.

Fig. 2 One-parameter linear probability density functions that can be used to choose parents for mating following a selection event. Probability density is plotted against parental fitness in the figure. The area underneath each density function equals unity. Partial areas under the functions, bounded by specific fitness values on the x-axis, represent probabilities that parents in those fitness ranges will reproduce. Fitter parent behaviors tend to produce fitter child behaviors. Hence, the steeper the slope of the density function, the greater the selection pressure favoring fit child behaviors in the next generation

For example, the twentieth reinforcer is not as motivating as the first, which means that the slope of the density function must decrease as a function of the number of reinforcers consumed.

Conclusion

The representation of states and dispositions in the evolutionary theory (Fig. 1), and of emotions and fixed dispositions, and of items such as wants, likes, affordances, and motivations (Fig. 2), is straightforward. According to the theory, the organism's history of interaction with the environment determines its current state, its current disposition to engage in various behaviors, and perhaps most importantly, its observed behavior in the next moment. The theory therefore captures much of what Killeen and Jacobs (2016) find to be missing in behavior analysis. Oh, and the theory has the added benefit of generating emergent outcomes that agree in qualitative and quantitative detail with a large and varied body of empirical findings from experiments with live organisms (McDowell, 2013a).

Acknowledgements I thank Olivia Calvin, Cyrus Chi, Ryan Hackett, Bryan Klapes, and Steve Riley for a helpful discussion of an earlier version of this commentary.

Compliance with Ethical Standards This work was not supported by grant funding. No original research using animal or human participants is reported in this paper.

Conflict of Interest The author declares that he has no conflicts of interest.

References

- Ballerini, M., Cabibbo, N., Candelier, R., Cavagna, A., Cisbani, E., Giardina, I., Lecomte, V., Orlandi, A., Parisi, G., Procaccini, A., Viale, M., & Zdravkovic, V. (2008). Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study. *Proc Natl Acad Sci*, 105, 1232–1237.
- Berg, J. P. (2011). *Implementation of stimulus control in a computational model (unpublished doctoral dissertation)*. Atlanta: Emory University Retrieved from <http://pid.emory.edu/ark:/25593/bcnxs>.
- Bunyard-Wild, P. (2013). Gretna green starling murmurations. <https://www.youtube.com/watch?v=M1Q-EbX6dso>. Retrieved 1/8/2017.
- Cox, B., & Forshaw, J. (2012). *The quantum universe: Everything that can happen does happen*. Boston: Da Capo Press.
- Edelman, G. M. (1987). *Neural Darwinism: The theory of neuronal group selection*. New York: Basic Books.
- Killeen, P. R., & Jacobs, K. W. (2016). Coal is not black, snow is not white, food is not a reinforcer: The roles of affordances and dispositions in the analysis of behavior. *The Behavior Analyst*. doi:10.1007/s40614-016-0080-7. **Advance online publication**
- McDowell, J. J. (2004). A computational model of selection by consequences. *J Exp Anal Behav*, 81, 297–317. doi:10.1901/jeab.2004.81-297.
- McDowell, J. J. (2010). Behavioral and neural Darwinism: Selectionist function and mechanism in adaptive behavior dynamics. *Behav Process*, 84, 358–365.
- McDowell, J. J. (2013a). A quantitative evolutionary theory of adaptive behavior dynamics. *Psychol Rev*, 120, 731–750.

- McDowell, J. J. (2013b). On the theoretical and empirical status of the matching law and matching theory. *Psychol Bull*, 139, 1000–1028.
- McDowell, J. J. (2013c). Representations of complexity: How nature appears in our theories. *The Behavior Analyst*, 36, 345–359.
- McDowell, J. J., Caron, M. L., Kulubekova, S., & Berg, J. P. (2008). A computational theory of selection by consequences applied to concurrent schedules. *J Exp Anal Behav*, 90, 387–403. doi:[10.1901/jeab.2008.90-387](https://doi.org/10.1901/jeab.2008.90-387).
- McDowell, J. J., & Popa, A. (2009). Beyond continuous mathematics and traditional scientific analysis: Understanding and mining Wolfram's A New Kind of Science. *Behav Process*, 81, 343–352. doi:[10.1016/j.beproc.2009.01.012](https://doi.org/10.1016/j.beproc.2009.01.012).
- McDowell, J. J., Soto, P. L., Dallery, J., & Kulubekova, S. (2006). A computational theory of adaptive behavior based on an evolutionary reinforcement mechanism. In M. Keijzer (Ed.), *Proceedings of the 2006 conference on genetic and evolutionary computation (GECCO-2006)* (pp. 175–182). New York: ACM Press.
- Muro, C., Escobedo, R., Spector, L., & Coppinger, R. P. (2011). Wolf-pack (*Canis lupus*) hunting strategies emerge from simple rules in computational simulations. *Behav Process*, 88, 192–197.