

U-shaped curves in development: A PDP approach

Timothy T. Rogers¹, David H. Rakison², and James L. McClelland^{2,3}

¹MRC Cognition and Brain Sciences Unit, Cambridge, UK

²Department of Psychology, and ³Center for the Neural Basis of Cognition,

Carnegie Mellon University, Pittsburgh PA

Correspondence to

Dr T.T. Rogers
MRC Cognition & Brain Sciences Unit
15 Chaucer Road
Cambridge CB2 2EF UK
Fax: international +44 1223 359062
Email:
tim.rogers@mrc-cbu.cam.ac.uk

Dr David Rakison
Department of Psychology
Carnegie Mellon University
Pittsburgh, PA
Fax: 412 268-2798
Email:
rakison@andrew.cmu.edu

Dr. James L. McClelland
Center for the Neural Basis of Cognition
115 Mellon Institute
4400 Fifth Ave.
Pittsburgh, Pa 15213
jlm@cnbc.cmu.edu

As the papers in this issue attest, U-shaped curves in development have stimulated a wide spectrum of research across disparate task domains and age groups and have provoked a variety of ideas about their origins and theoretical significance. In our view, the ubiquity of the general pattern suggests that U-shaped curves can arise from multiple factors, and that the various viewpoints represented herein may be useful for explaining some aspects of developmental change. In this spirit, we offer an additional way of thinking about such phenomena. Specifically, we suggest that U-shaped curves can arise within a domain-general learning mechanism as it slowly masters a domain characterized by statistical regularities and exceptions. This idea differs from those considered thus far, and may encompass many of the phenomena addressed by other views, three of which we outline briefly here.

U-shaped curves indicate a transition from "associative" to "rule-based" behavior

Early in language acquisition, young children often master unusual but very frequent constructions (e.g., past-tense forms such as *went* and *fell*), only to reverse this accomplishment later, producing regularized but incorrect forms (such as *goed* and *falled*) (Ervin, 1964). Such phenomena are cited in the challenge to so-called “associative learning” theories of language acquisition (Pinker 1984) for two reasons: i) it has been thought that no associative mechanism could lead a child who has correctly learned a frequently-occurring construction to stop using it, and ii) children produce incorrect novel forms (e.g., *goed*) that do not occur in normal speech and hence cannot have been learned directly. Although such findings were addressed by the model of Rumelhart and McClelland (1986), Marcus et al. (1992; Pinker, 1991) contested features of Rumelhart and McClelland’s simulation, and argued for the view that these phenomena mark a qualitative developmental shift in the mechanisms underpinning language

production—away from an “associative” system that produces constructions it has encountered in the past toward a symbolic or “rule-based” system that applies abstract rules productively.

U-shaped curves reflect different constraints on learning at different ages.

A second possibility forwarded by Cohen and colleagues (Cohen, 1998; Madole & Cohen, 1995) is that there exist different constraints on learning at different ages that result from a variety of factors including limited information processing abilities and experience with real-world patterns of covariation. Cashon and Cohen (this issue), for example, show elegantly that infants’ ability to perceive faces in terms of independent features or relations among features depends on their information processing ability. U- or N-shaped curves emerge, on this view, when infants are unable to process a more complex higher-order level of information and instead drop down to a simpler level. Alternatively, constraints on new learning may emerge as a consequence of prior learning. For instance, Rakison (2003) found that infants at 14 months will learn feature-motion relations that are uncommon in the real-world, such as when a recipient of an action possesses dynamic features, but infants at 16 months will learn only those relations that fit the typical pattern in the real-world (i.e., when an agent possesses dynamic features; Madole and Cohen, 1995; Namy, Campbell, and Tomasello, this issue).

U-shaped curves arise dynamically, when systems developing in parallel change at different rates.

A third perspective has emerged from studies designed to illuminate how nonlinear changes can arise in dynamic systems. A good example of this approach is Thelen’s elegant work on the infant step-reflex (Thelen & Fisher, 1982; Thelen, Fisher, & Ridley-Johnson, 1984). This and recent related research is outlined superbly by Gershkoff-Stowe and Thelen (this

volume). As they indicate, U-shaped patterns can arise when a behavior is supported by multiple different systems developing in parallel, even though none of the individual systems itself exhibits a U-shaped change.

A fourth possibility: Sensitivity to covariance structure at different granularities

We suggest that there exists a further possibility, distinct from those reviewed above, which offers a somewhat different perspective on the phenomena encompassed by each. Specifically, we suggest that the learning mechanisms that support knowledge acquisition in many domains are strongly sensitive to patterns of *coherent covariation* in the environment; that is, to the tendency for groups of features or attributes to occur consistently together in experience. The influence of coherent covariation on learning and representation has been investigated in detail in the domain of conceptual development (Rogers & McClelland, in press). Here we briefly describe a simulation using the simple model developed in this work, which will illustrate how, in principle, U-shaped patterns might arise from domain-general learning mechanisms implemented in a Parallel Distributed Processing (PDP) network. We will then relate these general ideas to the data and theory reviewed above.

A simple illustration of U-shaped curves arising from domain-general learning

The model is illustrated in Figure 1. It is derived from Rumelhart (1990; see also McClelland & Rogers, 2003; Rogers & McClelland, in press; Rumelhart, Hinton & Williams, 1986) and we direct the reader to this literature for details on how semantic knowledge is encoded in the distributed representations by a feed-forward PDP network trained with backpropagation. For the current purposes, we simply note that the model is trained, when provided with the first two terms of a simple proposition, to generate appropriate completions

across output units. To accomplish this, it must find a set of weights which will allow the distributed patterns of activity produced across *Representation* and *Hidden* units by a given input to ultimately activate the correct set of *Attribute* units (see Figure caption).

Two aspects of the network's behavior work together as it learns to give rise to U-shaped curves in development. First, the patterns of activity generated across the *Representation* units by different inputs undergo a progressive, coarse-to-fine differentiation (McClelland, McNaughton, & O'Reilly, 1995). That is, the model first learns to generate different patterns for groups of items with broadly different sets of output attributes (e.g., plants vs. animals), but shows little differentiation of items within these groups. Later it comes to differentiate sub-groupings of items within these broader sets (e.g., birds vs. fish), and only toward the end of learning does it learn to assign quite different representations to individual items (e.g., robin vs. canary). This coarse-to-fine pattern results from the fact that sets of properties that coherently co-vary across many items exert convergent effects on the connection weights. Many co-varying properties differentiate plants from animals, so these are learned first; several co-varying properties differentiate birds from land animals, so these are learned second. Relatively idiosyncratic properties that distinguish individual items, such as the bat, from other similar items (the birds) are learned relatively slowly (see Rogers & McClelland, in press).

Insert Figure 1 about here

Second, the weight changes provoked by the presentation of one item will influence the outputs generated by other items with similar internal representations. However, as noted above, the similarities among internal representations change over time—hence the way that new learning generalizes from one item to another also changes over time. Early in learning, when plants are represented as different from animals, learning about any animal will tend to generalize to all animals but not to plants. Later, when the flying animals have been differentiated from the land animals, what is learned about each flying animal will only transfer to the other flying animals (including the bat, not yet differentiated from the birds).

To show how these two forces can give rise to U-shaped curves in learning, we conducted a simple simulation with a model based on Rumelhart's network. The model was trained in an extended environment which included 8 individual plants (4 trees and 4 flowers), and 15 animals (8 land animals, 4 birds, and 1 bat). Attributes were assigned to these items as in Rumelhart's simulations—for instance, the individual animals shared few attributes with the plants; the birds shared some features with the animals, but had many more features in common with one another; and each item had one or two idiosyncratic features. The important manipulation concerned the properties of the bat: it shared almost all of its properties with the four birds, but like the land animals, the bat *had fur*, and did not *have feathers*. Thus, these properties were consistent with the items in the broad domain of animals (most of the 15 animals have fur and do not have feathers) but were inconsistent with one of the subgroups (the set of birds have feathers and not fur).

Figure 2 shows the activation of the *fur* and *feathers* output units as the network is trained, when it is queried to report the properties of the bat. Note that the model is presented with each of the training patterns once in every epoch (sweep through the training set). Thus, it is

repeatedly exposed to the fact that the bat has fur and not feathers. Nevertheless, its tendency to activate these properties changes dramatically over the course of learning: after 1500 epochs, the model correctly infers that the bat has fur and not feathers; by 2100 epochs, it draws the opposite conclusion; and by 6000 epochs it makes the correct inference again.

Insert Figure 2 about here

How does this pattern arise from the characteristics of the model reviewed above?

Internal representations differentiate progressively, so that the network first discriminates plants from animals, but does not discern finer sub-groups within these clusters. At this point, the characteristics of the majority of the animals are applied to all animals. Nine of the animals (including the bat) have fur, whereas only four do not, so that at this stage the network activates the *fur* unit for all animals, including the bat. Later in learning, however, the network differentiates the birds from the land animals. Because the bat shares most of its properties with the birds, its representation remains similar to the birds and moves away from the other animals. At this point, the network must learn to turn off the *fur* unit (and to activate the *feathers* unit) for all of the 4 birds. This learning again generalizes to the bat, so that, despite repeated exposure to the correct pattern, the network reverses its tendency to produce the correct response for this item. Finally, when the individual bat representation has been sufficiently well differentiated from the 4 birds, the model is again able to master the bat's unusual properties.

Discussion

The simulation illustrates one way in which a domain-general learning system can produce a U-shaped pattern as it develops. Such curves are observed for items whose properties are consistent with the general domain (e.g., most animals have fur), but inconsistent with immediate neighbors (e.g., most flying animals do not). This basic idea offers a somewhat different perspective on the views outlined earlier.

First, the simulation demonstrates that domain-general learning can lead to changes similar to those that Pinker (1984, 1991) and Marcus et al (1992) cite as evidence for a qualitative shift from associative- to rule-based cognition. The above simulation addresses both challenges posed by this view: the model exhibits “unlearning” of an initially correct response, and it produces outputs that it has never encountered directly in training (i.e., it was not taught that the bat has feathers). Midway through training the model may seem to be over-generalizing an explicit rule, such as “all flying things have feathers”, and hence that it has shifted from an “associative” to a “rule-based” mode. But this behavior is an emergent property that arises from domain-general learning mechanisms and does not signal a qualitative shift from associative to rule-based cognition. Note that, unlike the simulation presented in Rumelhart and McClelland (1986), there is no change in the training corpus during development; the set of training examples is held constant throughout. Similarly, Plunkett and Marchman (1991) demonstrated U-shaped past-tense learning with a constant training environment.

Second, the simulation offers a way of understanding how U-shaped patterns like those of Madole and Cohen (1995) and Rakison (2003) might arise. Specifically, we envision the simulation as capturing critical aspects of a slow-learning neocortical system, which is complemented in the brain by a fast-learning hippocampal system (McClelland et al., 1995).

Constraints on new learning may arise from the interaction of these systems. Suppose the network described above is presented with a novel, bird-like item, and must, together with the hippocampal system, rapidly learn that it has fur. If this occurs early in learning (around 1200 epochs), the existing knowledge in the neocortical system (which attributes fur to all animals at this point) will partially support this new learning, leaving the hippocampus with little work to do. Midway through development, however, the neocortex will interfere with learning this item since at this stage it attributes feathers and not fur to things that fly. Here the hippocampal system will be taxed to overcome this bias. Thus the interplay of these systems may lead to changes over time in the ease or difficulty of learning new associations. We view this account as an example of the idea that constraints on learning new associations can arise from the constellation of associations that have been previously acquired.

Finally, this idea has clear parallels to the view expressed by Thelen and colleagues (e.g., Gershkoff-Stowe & Thelen, this volume; Thelen & Smith, 1994). Specifically, U-shaped curves in the ability to acquire new associations arise because the observed behavior is supported by multiple systems operating in parallel. The U-shaped pattern reflects dynamics in the neocortical system, which can either facilitate or interfere with the hippocampal system's ability to master the new association.

References

- Cohen, L.B. (1998). An information-processing approach to infant perception and cognition. In F. Simion & G. Butterworth (Eds.), *The Development of Sensory, Motor, and Cognitive Capacities in Early Infancy* (pp.277-300). East Sussex: Psychology Press.
- Madole, K.L. & Cohen, L.B. (1995). The role of object parts in infants' attention to form-function correlations. *Developmental Psychology*, 31, 637-648.
- Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T. J., & Xu, F. (1992). Overregularization in language acquisition. *Mongraphs of the Society for Research in Child Development, Serial No. 228, Vol 57, No. 4.*
- McClelland, J.L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, 102, 419-457.
- McClelland & Rogers (2003). The Parallel Distributed Processing approach to semantic cognition. *Nature Reviews Neuroscience*, 4 (4), 310-322.
- Pinker, S. (1984). *Language Learnability and Language Development*. Cambridge, MA: Harvard University Press.
- Pinker, S. (1991) Rules of language. *Science*, 253, 530-535.

- Plunkett, K., & Marchman, V. A. (1991). U-Shaped Learning and Frequency Effects in a Multi-Layered Perceptron: Implications for Child Language Acquisition. *Cognition*, 38, 43-102.
- Rakison, D. H. (2003). The development of infants' knowledge of causal role and onset of motion. Manuscript submitted for publication.
- Rogers, T. T. & McClelland, J. L. (in press). *Semantic Cognition: A Parallel Distributed Processing Approach*. Cambridge, MA: MIT Press.
- Rumelhart D. E., & McClelland J. L. (1986). On learning past tenses of English verbs. In Rumelhart D.E. and McClelland J.L (Eds), *Parallel Distributed Processing: Vol 2: Psychological and Biological Models*. Cambridge, MA: MIT press.
- Rumelhart, D. E. (1990). Brain style computation: Learning and generalization. In S. F. Zornetzer and J. L. Davis and C. Lau (eds.), *An Introduction to Neural and Electronic Networks*. San Diego: Academic Press, 405-420.
- Rumelhart, D. E., Hinton, G. E. & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323 (9), 533-536.
- Rumelhart, D. & Todd, P. M. (1993). Learning and connectionist representations. In D. E. Meyer and S. Kornblum (eds.), *Attention and Performance vol. 14*, 3-30.
- Thelen, E., & Fisher, D. M. (1982). Newborn stepping: An explanation for a "disappearing reflex". *Developmental Psychology*, 18, 760-775.
- Thelen, E., Fisher, D. M., Ridley-Johnson, R., & Griffin, N. J. (1982). The effects of body build and arousal on newborn infant stepping. *Developmental Psychobiology*, 15, 447-453.
- Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.

Figure captions

Figure 1. A connectionist model of semantic memory adapted from Rumelhart and Todd (1993).

The entire set of units used in the network is shown. Input units are shown on the left, and activation propagates from the left to the right. Where connections are indicated, every unit in the pool on the left is connected to every unit in the pool to the right. Units in the *Item* and *Relation* layers corresponds to possible first terms and predicates in a simple proposition, respectively. Activation of an input probe such as *canary can* provokes distributed patterns of activity across *Representation* and *Hidden* layers. Output units correspond to possible completions of different propositions, and the network is trained to turn on all those units that represent correct completions of the input query. In the example shown, the correct units to activate are *grow*, *move*, *fly*, and *sing*. Redrawn with alterations from Rumelhart and Todd (1993), Figure 1.9, page 15.

Figure 2. Activation of the output units corresponding to *fur* and *feathers* throughout learning, when the network described in the text is queried to report the properties of the bat.