A Computational Model of Two Cognitive Transitions Underlying Cultural Evolution

Liane Gabora (liane.gabora@ubc.ca)
University of British Columbia
Department of Psychology, Okanagan campus, Arts Building, 3333 University Way
Kelowna BC, V1V 1V7, CANADA

Wei Wen Chia (cww9989@gmail.com) and Hadi Firouzi (hadi.firouzi@ubc.ca)
University of British Columbia
Department of Engineering, 5000-2332 Main Mall
Vancouver BC,V6T 1Z4, CANADA

Abstract

We tested the computational feasibility of the proposal that open-ended cultural evolution was made possible by two cognitive transitions: (1) onset of the capacity to chain thoughts together, followed by (2) onset of contextual focus (CF): the capacity to shift between a divergent mode of thought conducive to 'breaking out of a rut' and a convergent mode of thought conducive to minor modifications. These transitions were simulated in EVOC, an agent-based model of cultural evolution, in which the fitness of agents' actions increases as agents invent ideas for new actions, and imitate the fittest of their neighbors' actions. Both mean fitness and diversity of actions across the society increased with chaining, and even more so with CF, as hypothesized. CF was only effective when the fitness function changed, which supports its hypothesized role in generating and refining ideas.

Keywords: Agent-based model, CF, convergent though, creativity, cultural evolution, divergent thought, dual process, recursive retrieval, stream of thought.

Introduction

Humans are unique with respect to the ability to generate accumulative, adaptive cultural evolution, a phenomenon referred to as the ratchet effect (Tomasello, Kruger, & Ratner, 1993). Gaining insight into the origins of the capacity for complex culture is difficult, since all that is left of our prehistoric ancestors are bones and artifacts such as stone tools that resist the passage of time. Although methods for analyzing these remains are becoming increasingly sophisticated, they cannot always distinguish amongst competing theories. Thus, formal models provide valuable reconstructive tools for testing the feasibility of theories concerning the origins of the cognitive mechanisms that have transformed our planet.

Several cognitive mechanisms have been implicated in the ability to evolve culture. One is the capacity to *chain* thoughts together to generate a sequence of actions or stream of thought (Donald, 1991). Another is *contextual focus* (hereafter referred to as CF): the capacity to shift between analytic and associative modes of thought (Gabora, 2003). Mathematical models of both have been developed (Gabora & Aerts, 2009; Gabora & Kitto, 2012; Veloz *et al.*, 2011). Incorporating chaining into a computational model of cultural evolution increased the fitness and diversity of cultural outputs, as well as the effectiveness of learning (Gabora & Saberi, 2011). Incorporating CF into a portrait painting computer program generated artworks that humans preferred over those generated without CF (DiPaola & Gabora, 2009). However, the portrait painting program did not allow investigation of the effect of CF on the evolution of ideas through cultural interaction. The goal of the work presented here was to understand the relationship between chaining and CF. Specifically, we investigate the feasibility of the hypothesis that RR is broadly useful for improving cultural outputs, while CF is specifically useful for overcoming a new or sudden challenge.

Early Signs of Human Creativity

The minds of our earliest ancestors, *Homo habilis*, are referred to as *episodic* because there is no evidence that their experience deviated from the present moment of concrete sensory perceptions (Donald, 1991). They encoded perceptions of events in memory, but had little voluntary access to them without cues. They were therefore unable to voluntarily shape, modify, or practice skills and actions, and could not invent or refine complex actions, gestures, or vocalizations.

Homo habilis was eventually replaced by *Homo erectus*, which lived between approximately 1.8 and 0.3 million years ago. This period is considered the beginning of human cultural evolution. *Homo erectus* exhibited signs of enhanced intelligence, creativity, and adaptability. They made sophisticated task-specific stone hand axes, had complex stable seasonal home bases, and there is evidence of long-distance hunting strategies involving large game, and migration out of Africa (Leakey 1984). It is widely believed that these early signs of creative culture reflect an underlying transition in cognitive or social abilities. The cranial capacity of the Homo erectus brain was approximately 1,000 cc, which is about 25% larger than that of *Homo habilis*, and at least twice as large as that of living great apes, and 75% that of modern humans (Aiello, 1996).

Some have suggested that these abilities are due to the onset of a theory of mind (Mithen, 1998) or the capacity to imitate (Dugatkin, 2001). However, there is evidence that nonhuman primates also possess theory of mind (Heyes, 1998) and the capacity to imitate (Dugatkin, 2001), yet their cultural complexity do not compare with humans'. Evolutionary psychologists have suggested that our unique abilities were due to the onset of massive modularity (Barkow, Cosmides,

& Tooby, 1992). However, although the mind exhibits an intermediate degree of functional and anatomical modularity, neuroscience has not revealed vast numbers of hardwired, encapsulated, task-specific modules; indeed, the brain is more subject to environmental influence than was previously believed (Buller, 2005; Byrne, 2000; Wexler, 2006).

Donald (1991) proposed that with the enlarged cranial capacity of *Homo erectus*, the human mind underwent a transition characterized by a shift from an *episodic* to a *mimetic* mode of cognitive functioning, made possible by onset of the capacity to voluntarily retrieve memories independent of environmental cues and chain them into sequences. Donald refers to the cognitive architecture underlying this capacity as a *self-triggered recall and rehearsal loop*. It enabled information to be processed recursively, and from different perspectives. Voluntary access to memories made it possible to act out[1] events that occurred in the past or that might occur in the future. Thus not only could the mimetic mind temporarily escape the here and now, but by miming or gesture it could communicate similar escapes to other minds. The capacity to mime thus brought forth what is referred to as a *mimetic* form of cognition, and allowed for the onset of culture. The self-triggered recall and rehearsal loop also enabled our ancestors to engage in a stream of thought, in which one thought or idea evokes another, and so forth recursively. In this way, attention can be directed away from the external world toward one's internal model of it. Finally, self-triggered recall allowed for voluntary rehearsal and refinement of actions, enabling systematic evaluation and improvement of skills and motor acts.

An Explosion of Creative Cultural Change

The European archaeological record indicates that an unparalleled cultural transition occurred between 60,000 and 30,000 years ago, at the onset of the Upper Paleolithic. Considering it "evidence of the modern human mind at work," Leakey (1984:93-94) describes this period as "unlike previous eras, when stasis dominated, ... [with] change being measured in millennia rather than hundreds of millennia." Similarly, Mithen (1998) refers to the Upper Paleaolithic as the 'big bang' of human culture, exhibiting more innovation than in the previous six million years of human evolution. It marks the beginnings of traits considered diagnostic of behavioral modernity, including a more organized, strategic, season-specific style of hunting involving specific animals at specific sites, elaborate burial sites indicative of ritual and religion, evidence of dance, magic, and totemism, colonization of Australia, and replacement of Levallois tool technology by blade cores in the Near East. In Europe, complex hearths and many forms of art appeared, including cave paintings of animals, decorated tools and pottery, bone and antler tools with engraved designs, ivory statues of animals and sea shells, and personal decoration such as beads, pendants, and perforated animal teeth, many of which may have indicated social status.

[1]The term mimetic is derived from "mime," which means "to act out."

Whether this period was a genuine revolution culminating in behavioral modernity is hotly debated because claims to this effect are based on the European Palaeolithic record, and largely exclude the African record (McBrearty & Brooks, 2000). However the dominant view is that modern behavior appeared in Africa between 40,000 and 50,000 years ago, and spread, resulting in displacement of the Neanderthals in Europe (Klein, 1999). From this point on there was only one hominid species: modern *Homo sapien*, and despite a lack of overall increase in cranial capacity, their prefrontal cortex, and more particularly the orbitofrontal region, increased significantly in size (Dunbar, 1993). in what was most likely a time of major neural reorganization (Klein, 1999). Given that the Middle/Upper Palaeolithic was a period of unprecedented creativity, what kind of cognitive processes were involved?

It is widely believed that a divergent or associative mode of thought predominates during idea generation, while a convergent or analytic mode predominates during the refinement, implementation, and testing of an idea (Finke, Ward, & Smith, 1992). It has been proposed that the Paleolithic transition reflects fine-tuning of the biochemical mechanisms underlying the capacity to subconsciously shift between these modes, depending on the situation, by varying the specificity of the activated cognitive receptive field (Gabora, 2003; Gabora Kaufman, 2010). This is referred to as contextual focus[2] (CF) because it requires the ability to focus or defocus attention in response to the context or situation one is in. Defocused attention, by diffusely activating a broad region of memory, is conducive to divergent thought; it enables obscure (but potentially relevant) aspects of the situation to come into play. Focused attention is conducive to convergent thought; memory activation is constrained enough to hone in and perform logical mental operations on the most clearly relevant aspects.

The Computational Model

We reviewed the evidence for two hypotheses: (1) the earliest signs of culture were due to the onset of the capacity to chain representations together, and (2) the cultural explosion of the Middle-Upper Paleolithic was due to the onset of CF. We investigated these hypotheses using an agent-based model of cultural evolution referred to as "EVOlution of Culture", abbreviated EVOC. EVOC uses neural network based agents that (1) invent new ideas, (2) imitate actions implemented by neighbors, (3) evaluate ideas, and (4) implement successful ideas as actions. EVOC is an elaboration of Meme and Variations, or MAV (Gabora, 1995), the earliest computer program to our knowledge to model not just cultural transmission but cumulative, adaptive, cultural evolution.[3] It was inspired by the genetic algorithm, a search technique that finds solutions

[2]In neural net terms, CF amounts to the capacity to spontaneously and subconsciously vary the shape of the activation function, flat for divergent thought and spiky for analytical

[3]The approach can thus be contrasted with computer models of how individual learning affects biological evolution (e.g., Higgs, 2000; Hinton & Nowlan, 1987; Hutchins & Hazelhurst, 1991).

to complex problems by generating a 'population' of candidate solutions through processes akin to mutation, selecting the best, and repeating until a satisfactory solution is found (Holland, 1975). The goal behind MAV, and also behind EVOC, was to distil the underlying logic of not biological evolution but cultural evolution. Agents do not evolve in a biological sense–they neither die nor have offspring–but do in a cultural sense, by adaptively modifying each others' ideas for actions. We summarize the architecture of EVOC in sufficient detail to explain our results; for details we refer the reader to previous publications (*e.g.,* Gabora, 1995; Gabora & Saberi, 2011; Leijnen & Gabora, 2009).

Agents

Agents consist of (1) a neural network, which encodes ideas for actions and detects trends in what constitutes a fit action, (2) a 'perceptual system', which carries out the evaluation and imitation of neighbours' actions, and (3) a body, consisting of six body parts which implement actions. The neural network is composed of six input nodes and six corresponding output nodes that represent concepts of body parts (LEFT ARM, RIGHT ARM, LEFT LEG, RIGHT LEG, HEAD, and HIPS), as well as hidden nodes that represent more abstract concepts (LEFT, RIGHT, ARM, LEG, SYMMETRY, OPPOSITE, and MOVEMENT). Input nodes and output nodes are connected to hidden nodes of which they are instances (*e.g.,* LEFT ARM is connected to LEFT.) Activation of any input node activates the MOVEMENT node. Same-direction activation of symmetrical input nodes (*e.g.,* upward motion–of both arms) activates the SYMMETRY node.

Invention

An idea for a new action is a pattern consisting of six elements that dictate the placement of the six body parts. Agents generate new actions by modifying their initial action or an action that has been invented previously or acquired through imitation. During invention, the pattern of activation on the output nodes is fed back to the input nodes, and invention is biased according to the activations of the SYMMETRY and MOVEMENT hidden nodes. (Were this not the case there would be no benefit to using a neural network.) To invent a new idea, for each node of the idea currently represented on the input layer of the neural network, the agent makes a probabilistic decision as to whether the position of that body part will change, and if it does, the direction of change is stochastically biased according to the learning rate. If the new idea has a higher fitness than the currently implemented idea, the agent learns and implements the action specified by that idea.

Imitation

The process of finding a neighbour to imitate works through a form of lazy (non-greedy) search. The imitating agent randomly scans its neighbours, and adopts the first action that is fitter than the action it is currently implementing. If it does not find a neighbour that is executing a fitter action than its own current action, it continues to execute the current action.

Table 1: Definition table.

Term	Definition	Example
Body Part	Component of agent other than neural network.	Left Arm (LA)
Sub-action	Set of six components that indicates position of 6 body parts. Each can be in a neutral (0), up (1), or down (-1) position.	HD:0, LA:1, RA:-1, LL:1, RL:0, HP:-1; This sub-action is abbreviated 01-110-1
Action	One or more sequential sub-actions.	01001-1, -10-1-111
Template	Abstract or prototypical format for a sub-action. Position of a body part can be unspecified (*).	HD:0, LA:*, RA:1, LL:*, RL:1, HP:-1

Table 2: Partial set of the templates used in the first fitness function. (The rest are omitted due to lack of space.)

$T^1 = \{0,*,*,*,*,*\}$	$T^{24} = \{1,*,*,1,1,*\}$
$T^2 = \{*,0,*,*,*,*\}$	$T^{25} = \{1,*,1,*,1,*\}$
$T^3 = \{*,*,0,*,*,*\}$	$T^{26} = \{1,*,1,1,*,*\}$

Evaluation: The Fitness Function

Fitness was evaluated using an adaptation of the Royal Roads fitness function (Forrest & Mitchell, 1993). Midway through a run the fitness function was changed to test the effectiveness of chaining and CF for adapting to a sudden change in the task constraints or the environment. Definitions of terms used to accomplish this are provided in Table One.

The first fitness function is determined by 45 templates, six of which are shown in Table Two. The second (not shown) is constructed analogously, with different sub-actions. The templates can be thought of as defining the cultural significance of types of sub-actions (such as dance steps). Each template T^i consists of six components, one for each body part (*i.e.,* $T^i = t^i_j; j = 1..6$). Each body part can be in a neutral position (0) , up (1), down (-1), or an unspecified position (*). For example, in template $T^i = *,1,-1,*,*,0$, the left arm is up (LA:1), the right arm is down (RA:-1), the hips are in the neutral position (HP:0), and the positions of other body parts is unspecified (HD:*, LL:*, and RL:*). The templates provide constraints, as well as flexibility with respect to what constitutes a fit action. For example, in an optimally fit action, the head must be in the neutral position (in T^1 the first component is 0) but the positions of other body parts can vary). The optimal sub-actions are $\{0,1,-1,1,-1,1\}$, $\{0,1,-1,1,-1,-1\}$, $\{0,-1,1,-1,1,1\}$, and $\{0,-1,1,-1,1,-1\}$.

Assume that D is a sub-action (*i.e.,* $D = d_j; j = 1..6$) and T^i is the i^{th} template (*i.e.,* $T^i = t^i_j; j = 1..6$). Thus, d_j represents the position of the j^{th} body part and the value of d_j can be either 0 (neutral), 1 (up), or -1 (down). Likewise, the value of

t^i_j can be 0, 1, -1, or * (unspecified). Accordingly, the fitness of sub-action D is obtained as per Eq. 1.

$$F(D) = \sum_{i=1}^{19} \Phi(T^i, D) \times \Omega(T^i) \quad (1)$$

As shown in Eq. 1, fitness is a function of template weight ($\Phi(T^i, D)$) and template order ($\Omega(T^i)$).

Template Weight $\Phi(T^i, D)$ is a function that determines the weight of sub-action D by comparing it with template T^i. This weight is set to one if each component of the sub-action (*i.e.,* $d_j; j = 1..6$) either matches the corresponding component of the template (*i.e.,* $t^i_j; j = 1..6$) or if the corresponding components of the template is unspecified (*i.e.,* $t^i_j = *$):

$$\Phi(T^i, D) = \begin{cases} 1 & if\ \forall t^i_j \in T^i : t^i_j = d_j\ or * \\ 0 & otherwise \end{cases} \quad (2)$$

Template Order $\Omega(T^i)$ computes the order of the template T^i by counting the number of components that have a specified value (*i.e.,* $t^i_j \neq *$).

$$\Omega(T^i) = \sum_{j=1, t^i_j \neq *}^{6} t^i_j \quad (3)$$

The fitness functions are difficult to solve because they are rugged; is to have multiple milestones, or fitness peaks, that agents must achieve before reaching the plateau. For example, consider the fitness function given in Table 2. The action 0,0,0,0,0,0 has a fitness of 6. An agent may move on from this action to find an actions that fits the third order templates with a fitness of 31, *e.g.,* $F(D) : \{1,1,1,1,1,0\} = 3+3+3+3+3+3+3+3+3+3+3+1 = 31$.

Learning

Invention makes use of the ability to learn, and respond adaptively to trends. Knowledge acquired through the evaluation of actions is translated into educated guesses about how to invent fit actions. For example, an agent may learn that symmetrical movement tends to be either beneficial or detrimental, and bias the generation of new actions accordingly.

A Typical Run

Fitness and diversity of actions are initially low because all agents are initially immobile, implementing the same action, with all body parts in the neutral position. Soon some agent invents an action that has a higher fitness than immobility, and this action gets imitated, so fitness increases. Fitness increases further as other ideas get invented, assessed, implemented as actions, and spread through imitation. The diversity of actions increases due to the proliferation of new ideas, and then decreases as agents hone in on the fittest actions. Thus, over successive rounds of invention and imitation, the agents' actions improve. EVOC thereby models how "descent with modification" occurs in a purely cultural context.

Method

Modeling Chaining

The chaining algorithm is illustrated schematically in Figure 1b. Chaining gives agents the opportunity to execute multi-step actions. The agent can keep adding a new sub-action to its current action so long as the most recently-added sub-action is both novel and successful. A sub-action D is considered novel if at least one of its components is different from that of the previous sub-action. It is considered successful if there exists a template T^i such that $\Phi(T^i, D)$ is one.

$$successful(D) = \begin{cases} true & if\ \exists\ T^i : \Phi(T^i, D) = 1 \\ false & otherwise \end{cases} \quad (4)$$

The fitness of an action consisting of more than one sub-action is obtained by adding the number of sub-actions to the fitness of the last sub-action in the sequence. For example, if the last sub-action of an action is $D = [0,1,-1,1,-1,1]$ and the number of sub-actions is seven, the fitness of the action is $F(D) + 7 = 14 + 7 = 21$. Thus where c is 'with chaining', w is 'without chaining', n is the number of chained sub-actions, the fitness of a chained action, F_c, is calculated as follows:

$$F_c = F_w + n \quad (5)$$

An agent can execute an arbitrarily long action so long as it continues to invent successful new sub-actions. In general, the more sub-actions the fitter the action. This is admittedly a simple algorithm of simulating the capacity for chaining, but we were not interested in the impact of this action *per se*. The goal here was simply to test hypotheses about how chaining at the individual level affects dynamics at the societal level, by providing agents with a means of implementing multistep actions such that the optimal way of going about one step depends on how one went about the previous step.

Modeling Contextual Focus

The CF algorithm is illustrated schematically in Figure 1c. In the convergent mode, the current action is only slightly modified to create a new action. In the divergent mode, the current action is substantial modified to create a new action. An agent switches between these modes by modifying its *rate of creative change* (RCC). If the fitness of its current action is low relative to previous actions, RCC increases, causing the agent to shift to a more divergent processing mode conducive to large leaps through the space of possibilities. If action fitness is high relative to that of previous actions, RCC decreases, and the agent shifts to a more convergent mode conducive to minor adjustments. With CF turned off, RCC stays constant throughout the run at 1/6 (*i.e.,* a new action involves change to one of the six body parts). The equation to modify RCC is shown in Eq. 6 where a is a negative value. Since at the start of a run previous fitness is undefined, RCC in this case is a function of the current fitness as per Eq. 7, where $0 < b < 1$.

$$\Delta RCC = a(F_{new} - F_{old}) \quad (6)$$

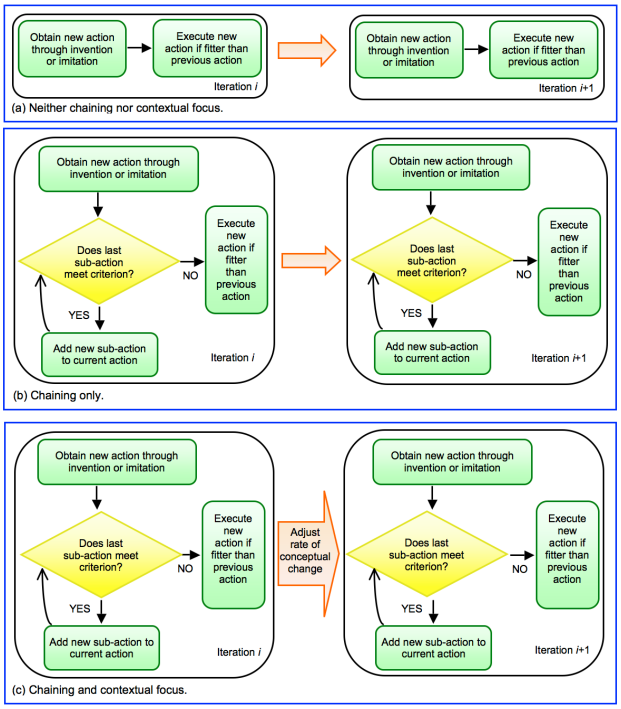

Figure 1: Schematic illustration of (a) neither chaining nor CF, (b) chaining only, and (c) both. Chaining operates within a generation whereas CF operates between generations.

$$RCC_{initial} = b^{F_{current}} \quad (7)$$

In the results shown here a and b were initialized to -0.005 and 0.8 respectively.

Results

The results of introducing chaining and CF on the mean fitness and diversity (total number of different actions) of actions across all agents in the society are shown in Figures 2 and 3 respectively. All graphs show means of 500 runs. Chaining and CF both significantly increased mean fitness of actions. Without chaining, mean fitness quickly reached a plateau; with chaining it could increase indefinitely. While chaining increased mean fitness throughout the run, CF only increased mean fitness following initial exposure to a new fitness function, i.e., at the beginning of the run, and when the second fitness function was introduced at iteration 50.

Chaining also significantly increased the diversity of actions. Although inspection reveals that there is always convergence on optimal actions, without chained actions, this set is a static (thus mean fitness plateaus) whereas with chained actions the set of optimal actions changes, as increasingly fit actions continue to be found. When agents were first exposed to a fitness function, CF increased both the rate at which new possibilities were generated, and the rate of convergence on the fittest of these, although this effect is more pronounced for the first fitness function than the second. As with fitness, CF exerted no noticeable effect on diversity once the agents had fit actions.

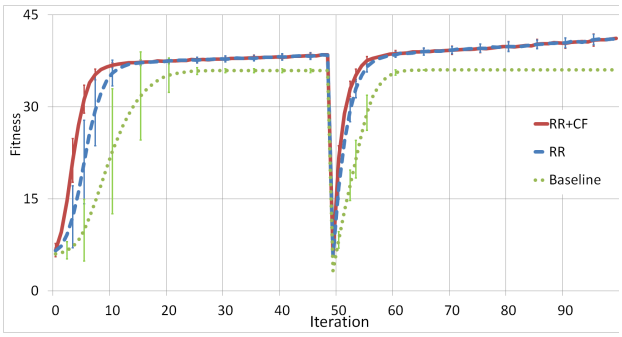

Figure 2: Mean fitness of cultural outputs across the society with both chaining and CF (red line), chaining only (dashed blue line), and neither chaining nor CF (dotted green line).

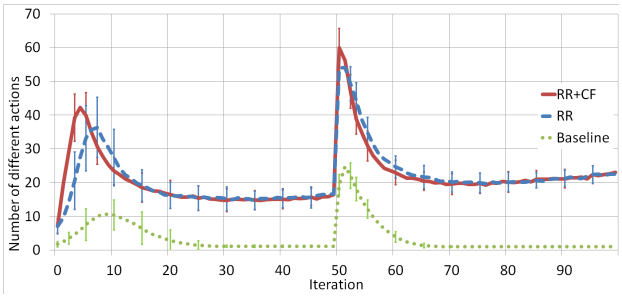

Figure 3: Diversity of cultural outputs across the society with both chaining and CF (red line), chaining only (dashed blue line), and neither chaining nor CF (dotted green line).

Discussion

This paper provides valuable insights into the mechanisms underlying the uniquely human capacity for collectively generated, open-ended, adaptive cultural evolution. Our results suggest that once humans became able to sequence thoughts together to generate increasingly complex and refined cultural outputs, they would have found themselves at a significant adaptive advantage. Similarly, our results suggest that once humans became able to employ an exploratory, divergent processing mode when stuck, followed by a shift to a more constrained convergent processing mode to fine-tune their cultural outputs, they would have been capable of generating significantly more valuable cultural outputs. We suggest that a mechanism akin to CF is what makes possible the cumulative creativity exhibited by successful computational models of language evolution (*e.g.,* Kirby, 2001). A potential downfall of processing in a divergent mode is that since effort is devoted to the re-processing of previously learned material, less effort may be devoted to being on the lookout for danger and simply carrying out practical tasks. Since divergent thought carries a high cognitive load, it would not have been useful to think divergently until there was a means to shift back to a convergent mode. Although these results do not prove that onset of the capacity to chain thoughts together into sequences, and to shift between divergent modes of thought through CF, are responsible for our cultural complexity, it shows that they provide a computationally feasible

explanation. We know of no other cognitive mechanisms implicated in the evolution of complex culture for which open-ended, adaptive cultural change has been demonstrated.

Both chaining and CF were implemented in a simple manner. Future investigations will focus on developing more realistic implementations of chaining and CF. Chaining will use associative recall to reconsider an item from multiple potentially relevant 'perspectives', and the divergent mode of CF will use a sophisticated mathematical model of concepts to facilitate the generation of new concept combinations.

Acknowledgments

This project was conducted at the Media and Graphics Interdisciplinary Centre (MAGIC), UBC. It was supported by the Natural Sciences and Engineering Research Council of Canada, and Fund for Scientific Research, Flanders, Belgium.

References

Aiello, L. C. (1996). Hominine pre-adaptations for language and cognition. In P. Mellars & K. Gibson (Eds.), *Modeling the early human mind*, Cambridge, UK: McDonald Institute Monographs, 89-99.

Barkow, J. H., Cosmides, L., & Tooby, J., Eds. (1992). *The adapted mind* New York: Oxford University Press.

Buller, D. J. (2005). *Adapting minds.* MIT Press.

Byrne, R., & Russon, A. (1998). Learning by imitation: A hierarchical approach. *Behavioral and Brain Sciences, 2,* 667-721.

Cavalli-Sforza, L. L., & Feldman, M. W. (1981). *Cultural transmission and evolution: A quantitative approach.* Princeton: Princeton University Press.

DiPaola, S. & Gabora, L. (2009). Incorporating characteristics of human creativity into an evolutionary art algorithm. *Genetic Programming and Evolvable Machines, 10,* 97-110.

Donald, M. (1991). *Origins of the modern mind: Three stages in the evolution of culture and cognition.* Cambridge, MA: Harvard University Press.

Dugatkin, L. (2001). *Imitation factor: Imitation in animals and the origin of human culture.* New York: Free Press.

Dunbar, R. (1993). Coevolution of neocortical size, group size, and language in humans. *Behavioral and Brain Sciences, 16*, 681-735.

Finke, R. A., Ward, T. B., & Smith, S. M. (1992). *Creative cognition: Theory, research, and applications.* Cambridge, MA: MIT Press.

Forrest, S. & Mitchell, M. (1993). Relative building block fitness and the building block hypothesis. In (L. Whitley, Ed.) *Foundations of genetic algorithms.* San Mateo: Morgan Kaufman.

Gabora, L. (1995). Meme and Variations: A computer model of cultural evolution. In L. Nadel & D. Stein (Eds.) *1993 Lectures in Complex Systems.* Boston: Addison-Wesley.

Gabora, L. (2003). Contextual focus: A cognitive explanation for the cultural transition of the Middle/Upper Paleolithic. *Proceedings Annual Conference Cognitive Science Society* (pp. 432–437), Boston: Lawrence Erlbaum.

Gabora, L. & Aerts, D. (2009). A mathematical model of the emergence of an integrated worldview. *Journal of Mathematical Psychology, 53,* 434–451.

Gabora, L., & Kaufman, S. (2010). Evolutionary perspectives on creativity. In (J. Kaufman & R. Sternberg, Eds.) *The Cambridge handbook of creativity.* Cambridge: Cambridge University Press.

Gabora, L., & Kitto, K. (2012). Concept combination and the origins of complex cognition. In (L. Swan, Ed.) *Origins of mind.* Berlin: Springer.

Gabora, L. & Saberi, M. (2011). How did human creativity arise? An agent-based model of the origin of cumulative open-ended cultural evolution. *Proceedings Conference on Cognition & Creativity* (pp. 299-306). New York: ACM.

Higgs, P. (2000). The mimetic transition: a simulation study of the evolution of learning by imitation. *Proceedings: Royal Society B: Biological Sciences, 267,* 1355-1361.

Heyes, C. M. (1998). Theory of mind in nonhuman primates. *Behavioral and Brain Sciences, 211,* 104-134.

Hinton, G. E. & Nowlan, S. J. (1987). How learning can guide evolution. *Complex Systems, 1,* 495-502.

Holland, J. K. (1975). *Adaptation in natural and artificial systems.* Ann Arbor, MI: University of Michigan Press.

Hutchins, E. & Hazelhurst, B. (1991). Learning in the cultural process. In Langton, C., Farmer, D., & Rasmussen, S. (Eds.) *Artificial Life II.* Redwood City: Addison-Wesley.

Kirby, S. (2001). Spontaneous evolution of linguistic structure: An iterated learning model of the emergence of regularity and irregularity. *IEEE Transactions on Evolutionary Computation, 5*(2), 102-110.

Klein, R. (1999). *The human career: Human biological and cultural origins.* Chicago: University of Chicago Press.

Leijnen, S., & Gabora, L. (2009). How creative should creators be to optimize the evolution of ideas? A computational model. *Electronic Proceedings in Theoretical Computer Science, 9,* 108-119.

Leakey, R. (1984). *The origins of humankind*. New York: Science Masters Basic Books.

McBrearty, S., & Brooks, A. S. (2000). The revolution that wasnt: A new interpretation of the origin of modern human behavior. *Journal of Human Evolution, 39,* 453-563.

Mithen, S. (1998). *Creativity in human evolution and prehistory.* London, UK: Routledge.

Tomasello, M., Kruger, A., & Ratner, H. (1993). Cultural learning. *Behavioral and Brain Sciences, 16,* 495-552.

Veloz, T., Gabora, L., Eyjolfson, M. & Aerts, D. (2011). A model of the shifting relationship between concepts and contexts in different modes of thought. *Proc 5th Intnl Symposium on Quantum Interaction.* Heidelberg: Springer.

Wexler, B. (2006). *Brain and culture: Neurobiology, ideology and social change.* New York: Bradford Books.