

Complejidad, fenómenos que emergen y cognición corporizada en educación

Complexity, emergence and embodied cognition in education

Roberto Araya¹

¹Centro de Investigación Avanzada en Educación, Universidad de Chile, Chile (roberto.araya.schulz@gmail.com)

Recibido el 18 de julio de 2017; revisado el 13 de noviembre de 20xx; aceptado el 13 de noviembre de 2017; publicado el 2 de diciembre de 2017

RESUMEN:

La complejidad y el proceso donde emergen nuevos fenómenos a partir de más básicos son nociones básicas para entender y mejorar el aprendizaje. Por un lado, con conceptos complejos los estudiantes se desconciertan y el aprendizaje se vuelve muy difícil. Por otro lado, los procesos en los que emergen nuevos fenómenos parecen ser mágicos o ilusiones cognitivas. Parecen basarse en cualidades adicionales que no están incluidas en los fenómenos subyacentes. ¿Puede el docente simplificar las nociones complejas sin cambiarlas? Para ello, argumentamos que la complejidad y el proceso de emerger no son exclusivamente inherentes a objetos o fenómenos. También dependen del sistema perceptivo, motor y cognitivo del estudiante. Así, si el profesor ayuda a conectar nociones y fenómenos con el conocimiento innato y corporizado de los estudiantes, entonces estas nociones se vuelven menos complejas y el fenómeno emergente pierde su magia: se conecta lógicamente con los fenómenos subyacentes. En este artículo presentamos evidencia empírica del efecto en la comprensión de los estudiantes debido a la conexión establecida en dos conceptos matemáticos centrales del currículo y que se consideran muy desafiantes.

PALABRAS CLAVE: COMPLEJIDAD EFECTIVA, INFORMACIÓN, FENÓMENOS QUE EMERGEN, COGNICIÓN CORPORIZADA, EDUCACIÓN MATEMÁTICA.

ABSTRACT:

Complexity and emergence are core notions for understanding and improving learning. On one hand, with complex concepts students struggle and learning becomes very difficult. On the other hand, emergence phenomena looks like magic or cognitive illusions, they seem to rely on extra qualities not included in the subjacent phenomena. Can the teacher simplify complex notions without changing them? In order to do that, we argue complexity and emergency are not exclusively inherent to objects or phenomena. They also depend on the perceptual, motor and cognitive system of the student. Thus, if the teacher helps to connect notions and phenomena to students' innate and embodied knowledge, then these notions become less complex and the emergent phenomenon loses its magic: it becomes logically connected to the subjacent phenomena. In this paper we present empirical evidence of the effect on students understanding due to the connection established in two core curriculum mathematical concepts that are considered very challenging.

KEYWORDS: EFFECTIVE COMPLEXITY, INFORMATION, EMERGENCE, EMBODIMENT, MATHEMATICS EDUCATION.

1 INTRODUCTION

What does it mean that a concept is complex? What does it mean that a phenomenon emerges from other phenomena? Complexity and emergence are notions that have been the subject of extensive studies (Kolmogorov, 1968; Chaitin, 1974; Holland, 1998; Gell-Mann, 1995, 2000; Simon 1996). They are crucial to understand the world and to interact with it. But what does it mean that a phenomenon is more complex than another one? How can complexity be measured? For example, why a straight line is simpler than other curves? If biology can be obtained from chemistry and chemistry from physics, then in what sense biology is more complex? Is it because biology requires more computational power to deduce it from physics than the computational power required to deduce chemistry from physics? Several challenges appear when the notion of complexity is studied. The mathematician John Casti (Araya, 2000b), suggests that to handle properly more complex structures such as the ones in the social sciences a new mathematics is needed. It would not be enough to explicit behavioral rules. It would be needed to describe systems with more flexible rules, rules that can improve themselves. More complexity also introduces the phenomenon of emergence. This means new structures and phenomena, which cannot be straightforwardly reduced to subjacent and simpler ones, seem to appear (emerge) as the complexity is increased.

A definition of complexity of a phenomenon must somehow consider the difficulty to describe the phenomenon. For example, a widely used way of expressing this difficulty is to consider the size of the minimal description of the phenomenon (Kolmogorov, 1968; Chaitin, 1974). This idea, that in principle seems to be conceptually very clear, has a couple of hidden details that are crucial. First, there is the need to precise what does it mean to describe a phenomenon. Second, there is also the issue of the way to specify those descriptions. What marks or signs are used and on what format. The process of selection of the relevant factors is a critical one. A night can be described as a black sky, or, alternatively, as a sky filled with a very detailed distribution of stars. In the first case it is a very simple phenomenon, but in the second case it is a very complex one.

Given this difficulty, it is common to define complexity of a phenomenon using an already made description of it. For example, a written description

in English, or a mathematical equation, or an array of pixels of different colors. All of these descriptions can be viewed as a long string of zeroes and ones. Just think that any text with images and formulas written in a word processor is internally saved as a string of zeroes and ones. Nevertheless, it is important not to forget that the string of zeroes and ones presupposes a selection of certain features of the phenomenon.

Murray Gell-Mann (Gell-Mann, 1995) proposes a definition of complexity that makes explicit this dependence on the previously selected characteristics of the phenomenon. Complexity is the size of the more compact description of the concepts, schemes and rules that capture the preselected regularities of the phenomenon. He call it, effective complexity, to differentiate it from other definitions. For example, a sequence of zeroes and ones produced by a random number generator has very low effective complexity if the description selected is the algorithm that the computer uses. Instead, if the description selected is the exact sequence then it has a very high effective complexity. The phenomena can have diverse complexities depending on the regularities selected and described. Thus, the effective complexity of a phenomenon depends on an observer that describe the phenomenon. The observer could be a human, a non-human animal, or even a machine. It has to be something that selects regularities.

2 COMPLEXITY

2.1 Complexity depends on the Perceptual, Motor and Cognitive Systems of the Observer

Let's consider the example proposed in (Araya, 2006). Look at figure 1, which is half of a figure designed by Leonard Kitts (Solso, 1994).

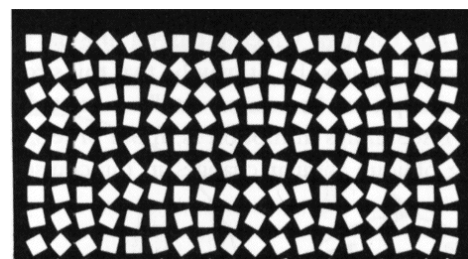


Figure 1

Imagine that the picture is analyzed by a simple machine that observes the image through a camera lens and that detects that over a black background there are several small white squares, arranged on

rows and columns. All the description produced by the machine would be the size of the squares and its distribution in an array of 18 columns and 9 rows. But such description also describes figure 2.

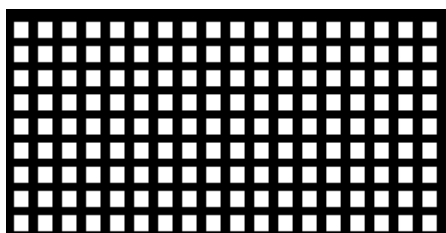


Figure 2

Another more sophisticated machine could detect that the squares of the original figure are arranged in different inclination angles. This machine include in the description all those inclination angles. Clearly figure 1 will be more complex for this machine since the description is longer. It requires to specify the tilt angle of each of the 162 squares (18 columns by 9 rows).

Look at figure 1 again. If you are like me, you will notice the spontaneous emergence of a dynamic pattern. Arcs that are formed and disappear, to again appear and disappear. Will other animals detect this dynamic pattern? For an observer that detects the dynamic pattern the complexity of figure 1 is clearly bigger than just an array of squares with different inclination angles. The observer that detects such a dynamic pattern phenomenon would require an additional description to be able to communicate what happens to another observer that is not viewing figure 1 and has never seen it before.

Let's imagine that we send the simple text description of the distribution of the white squares with its inclinations angles to another animal or machine. It can only detect or experience the dynamic pattern phenomenon if it has similar perceptual and processing algorithms to handle visual information. Moreover, the dynamic pattern detection requires not only that the observer has these perceptual and processing algorithms but also that he uses them to process the figure. If you send a textual description of figure 1 to another person that has never seen it, he would not experience the dynamic pattern. He will have to draw and paint a picture according to the textual description, and only when finished and then look with his own eyes at the drawn picture he will detect those dynamic patterns. Thus, the dynamic phenomenon depends critically of his particular vision system. This is similar to what you experience when watching movies. They are just a sequence of still pictures. It is your perceptual,

motor and cognitive system that builds the movement.

This example suggests that another component in complexity is the dependence on the perceptual, motor and cognitive systems of the observer. It is then an embodied complexity. Two observers that look at figure 1 but with different visual and cognitive systems, will most probably assign different complexity to figure 1.

2.2 Embodied Information

The dependence of complexity on the observer is similar to the dependence of the standard notion of information. For John Casti (2000), one of the most common uses of information is as a measure of novelty or surprise. Intuitively, if something is known and recurrent, then the fact that someone tells us that that event is occurring does not bring us much information. The information level is close to zero. On the contrary, if we are told that now is occurring something very infrequent then that warning bring us a lot of information. This means that the information of an event increases as the probability of the event is lower. Thus, it is much more information to you to know that there is a tiger close to you than to know that you are close to an ant. This is because you assign much less probability that a tiger appears close to you than an ant. Note that, similar to what happens to the notion of complexity, there is the critical dependence on the observer. The same event for an observer can be very improbable, and then brings high information to know about it, but for another observer could be highly probable. To a zoo worker it is not highly surprising to be close to a tiger. And thus, the fact of knowing that event, does not bring to him a lot of information.

This means that the information of an event depends on the experience, knowledge and previous learnings of the observer. In other words, it depends on the a priori probabilities that the different events have for the observer. This is a crucial point, and it is one that shows the strong link between information and semantics. According to Rieke et al. (1997), the misunderstanding of this fact has led to the unsupported opinion that the theory of information is not relevant to biology and neurosciences. It is frequently argued that information theory does not take into account the features of the world that interest the organism, neither it would discard those facts or features that doesn't interest him. And therefore, information theory would be "blind to semantics or meaning". This erroneous conception has the origin in the

belief that the theory of information measures the information as in a computer hard drive, independent of any observer. That it is just a thing of bits, just an account of zeroes and ones. This is far to be that way. The information in Shannon's theory depends on a priori probabilities. On these a priori probabilities are included the interests and knowledge on the observer. There it is the evolutionary design of the species, the ecological niche where the observer's species has inhabited and reproduced by thousands of generations.

For example, the presence of a predator in front of an observer is a very improbable event, but much more meaningful than the fact that he is next to an ant. To be in front of a natural predator is very relevant to his survival and is rather infrequent. It is a big surprise. Therefore to be warned about it is very valuable information. This means that the a priori probabilities is a way to account for the model of the world that the observer has. These probabilities are critical to compute the information level of an event. Additionally, given that the observer's world model is continuously varying, it changes while the observer interacts with the world and learn from it, then the same event means different levels of information according to when it is measured. The complexity of a phenomenon, similarly to the information of an event, depends on the observer and his perceptual, motor and cognitive system. It depends on his interests and experience. Therefore, it is connected to semantics, to the meaning that the phenomenon has for the observer.

2.3 Dependence on the format

The dependence on the observer and his perceptual, motor and cognitive systems leads us to another crucial aspect that has been recently attracted the attention of evolutionary psychology: the format in which the phenomenon is presented. A perceptual system works properly only if the input signal has the required format. If it is in a different format, it doesn't work or it generates a different result. If figure 1 is presented as a list of zeroes, ones and instructions to reconstruct the image from those numbers, then the dynamic pattern phenomenon is not generated on the observer. The dynamic pattern emerges only if the observer has a perceptual system as the human visual system and he looks with his own eyes the reconstructed two dimensional image. Information described in formats mathematically equivalent could be processed by completely different algorithms, and therefore by different neuronal areas and circuits. For this reason, even though two formats seems equivalent, the

corresponding perceptual, motor and cognitive module can produce completely different responses.

According to the evolutionary psychologists Leda Cosmides and John Tooby (1996), the cognitive system is a set of computational machines, each designed by natural selection to solve some of the recurrent problems of the species. Each of this machines works on a very specific environment and process information properly only if the information is in a very particular format. It is the format that for thousands of generations the corresponding machine has worked on and has being gaining a highly precise specialization. This format is what the psychologist Gerd Gigerenzer (2000) call the ecologically valid format.

After reviewing the previous example, one could think that the complexity dependence is a particular case of visual input and the visual system. Let's look other examples of different nature. Consider the mindreading system that recognize agents and its intentions. In 1944 Heider and Simmel (Baron-Cohen, 1997) asked subjects to watch a silent short film in which two triangles and a circle move around. When the subjects were asked to describe what they had just seen, they described the figures as agents, socially interacting within them, and trying to pursue specific goals. According to Baron-Cohen, there is an innate intentionality detector system that interprets motion stimuli in terms of the primitive volitional mental states of goal and desire. In this case, from the movements of the geometric figures emerges a completely new phenomenon: a social drama. This is possible because the observer has the mechanisms already in place to automatically interpret certain objects and their movements as social interaction. Therefore, the complexity of the film and the emergence of a pattern of social dynamics depend critically on the intentional system of the observer. Someone with a different intentional system or with one damaged as apparently is the case on certain autistic patients, will not see the emergence of the social dynamics. Therefore the complexity of the film will be completely different. It would be needed to add an explicit account of the interpretation of the figures as agents and the whole set of motions as a social struggle to obtain certain goals. All these extra description would increase the length of the description, and therefore increase its complexity.

Let's analyze a more abstract case that illustrate the dependence of complexity on the observer cognitive system, with no or very little intervention of the perceptual and motor systems. Daniel Povinelli (Povinelli, 2000) in several experiments

with chimpanzees that required the use of tools to solve problems reaching foods on tubes, concluded that even though a major part of the same human perceptual-motor abilities are involved, the chimpanzees do not represent abstract variables as causes of objects interactions. This inability to reinterpret observable physical events in terms of unobservable causal phenomena (such as forces), is an important difference in the cognitive system of the two species. Without this ability, not only chimpanzees are unable to solve several simple problems using tools that young kids do solve, but also they are unable to detect certain regularities in a more abstract plane. For Daniel Povinelli the human cognitive system may effectively “crowd out” the most detailed level of perceptual information in favor of more abstract representations, and thus it is vulnerable to “conceptual intrusions”. Therefore a human observer would see emerge different phenomena than a chimpanzee observer. Since they suffer less from such conceptual intrusions, chimpanzees extract highly specific rules from their experiences. It seems that they exhibit skills of visual rule extraction which are superior to our own (Povinelli, 2000), and in some symbolic numeric tasks they exhibit skills of visual rule recollection much superior to humans (Inoue & Matsuzawa, T., 2007). Something similar happens with autistic patients, who pay more attention to details of objects and events, as proposed to more global and abstract levels. According to Allan Snyder (Snyder et Al., 2004) with maturation the human mind becomes increasingly aware of concepts alone with exclusion of details. This inhibition of details from conscious awareness can be turned off on normal subjects by transcranial magnetic stimulation transforming their behavior to one closer to the behavior of autistic savants in several tasks (Bossomaier, 2004).

2.4 Complexity is no a relative or arbitrary concept

The perceptual, motor and cognitive systems have strong inductive and reasoning biases (Baum, 2004; Pinker, 2002; Mercer & Sperber, 2017). It is not a blank slate. This means that it looks for certain very specific patterns and discards most of other possibilities. The bias is encoded on our genetic code and in the interaction with the environment. For example, there is a bias for linear relations, and therefore we find a straight line much simpler than a quadratic or polynomial one, even when they are specified with the same number of parameters. This could be because straight line are important to navigation and prediction (line of the horizon, the

trajectory of falling of an object). There is also a strong bias for faces, and therefore we find a face a simpler image than a detailed diagram of an electronic circuit, even if the diagram has less number of lines than the face. We easily see faces when exposed to visual stimuli like clouds, but we don't see electronic diagrams. There is also a strong bias for cause-effect relations on temporally successive events and storytelling is our natural way of making causal sense (Sloman & Fernbach, 2017). There is also a strong bias towards real time performance and therefore towards frugal heuristics (Gigerenzer et al., 1999), that make possible fast and ecologically effective decisions. Learning would be impossible without all these built-in biases, because at any given moment the number of possible alternatives is mind boggling.

These biases come from millions of years of an evolutionary process of selection and adaptation. Evolution has selected and refined computational algorithms that now contains a strong bias that successfully reflect the structure of the world. There is then a good fit between the structure of these biases and heuristics and the structure of the environment where they are applied to (Gigerenzer et Al, 1999). With these biases subjects can learn very rapidly, because they explore a very small number of already tested and successful possibilities. This explains why we find certain phenomena simpler than other phenomena. The features that the observer selects are not arbitrary or random. They are selected because the observer's perceptual, cognitive and motor system already know they are effective. They have a proved predictive power.

Thus, complexity depends on the observer and his perceptual, motor and cognitive system. It is embodied. But these systems successfully capture the structure of the world. Two different human observers select the same features because they have the same inductive bias. Therefore the effective complexity will be similar. You and I agree that a straight line is simpler than other curves and that the movement of a stone falling is simpler than the movement of a jaguar.

One could argue than a completely different being, a Martian for example, could select completely different features of phenomena and therefore its complexity would be different. But this would mean that his inductive biases are completely different from ours. This possibility cannot be discarded but any living being has to live under the same physics, and therefore much of the inductive biases have to be similar. For example the navigation system must be similar. Furthermore, if

they are social beings, then several strategies of social interaction have to be similar to our own. Under these same constraints it is difficult to conceive a radically different set of inductive biases, and therefore the observer will select somehow similar features and patterns.

2.5 Ecologically valid strategies for teaching core concepts

How can we apply these connections to innate and embodied knowledge for designing lessons? Let's consider the case of fractions. This is a core mathematical concept that students start to learn from third grade. Teaching fractions is perhaps the most challenging educational problem in elementary and middle school mathematics (Bailey et al., 2012; Siegler et al. 2010, Siegler et al. 2017). One critical problem is the interference induced by the two whole numbers that specify a fraction. Thus, when comparing two fractions, there are 4 whole numbers that have to be considered. It is widely documented that the biggest whole number primes the selection of the bigger fraction. This phenomenon is called the whole-number bias (Obersteiner et al, 2013). This effect is augmented when the two bigger whole numbers belongs to the same fraction (as numerator and denominator). However foraging and interchange ratios are widely used by several species, where organisms are constantly comparing ratios to make foraging and reproductively meaningful decisions. For example, there are widely documented biological markets in non-human primates where subjects track interchange ratios in the interchange of grooming with other services (Fruteau et al, 2009).

Inspired in these facts we compared (Jiménez & Araya, 2013) the effect of a temporal frequency (foraging) format and an interchange format on the strength of whole number bias in fraction comparisons in 213 fourth graders (109 girls and 104 boys). We considered three conditions: congruent tasks, when the biggest number belongs to the biggest fraction; simple incongruent tasks, when the biggest number belongs to the smallest fraction but the second biggest number belong to the biggest fraction; and the double incongruent tasks, when the two biggest numbers belong to the smallest fraction. Fraction comparisons using the time frequency and interchange formats produce high reduction of whole number bias for the simple incongruent tasks in comparison to normal symbolic format for fractions. A smaller but still statistically significant reduction of whole number bias is also obtained for

the double incongruent case. This finding can be very useful to design strategies to teach fractions.

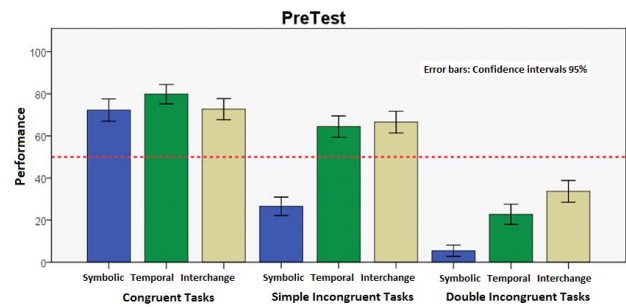


Figure 3

After the pretest students were randomly assigned to three training conditions. Fractions as partitions (the usual pizza like representations), fractions as temporal rates and fractions as interchange rates. Then all students answered a symbolic fraction comparison posttest. We found that for the double incongruent tasks the students that were trained in fractions as a temporal rate and fractions as interchange rates had better score than the other students.

A similar study with first order equations was obtained (Araya et al., 2010). A total of 236 seventh grade students who had never been taught algebraic equations before were randomly divided into two groups. The students in one group watched a 15-minute video teaching them how to solve five different first-degree linear equations using a traditional symbolic strategy, while in the other group, the students watched a 15-minute video teaching them how to solve the same equations using four analogies for solving an equation: a two-pan balance for the equals sign, a box for a variable, candies for numbers, and guessing the number of candies inside a box. The students were then tested on 12 equation solving problems, all of them written, using only symbolic notation. The group that watched the analogies video performed significantly better. Students with a below-average mathematics GPA who watched the analogies video did as well as students with an above-average GPA who watched the symbolic strategy video. Students who watched the analogies video also reached a better conceptual understanding, were better at making generalizations, did significantly better on reasoning problems involving equations, and had a better affective reaction. A possible explanation is that the two-pan balance equilibrium and the procedures of adding and subtracting the same amount of candies or boxes on both sides of the two-pan balance are part of our biological primary cognition (Geary, 2007). This is probably folk

physics knowledge. The use of analogies establishes a mapping between such biologically primary knowledge and the abstract mathematical concepts of algebraic equation solving.

3 EMERGENCE

3.1 Embodiment in emergence

According to John Holland (Holland, 1998), subassemblies have a critical role in fostering emergence. The combinations of basic building blocks is similar to what Holland call the Greek approach to machines, where every machine can be constructed from copies of six basic mechanisms: the lever, the screw, the inclined plane, the wedge, the wheel and the pulley. When several of these building blocks are put together sometimes an emergent phenomenon is produced. But, will the building blocks and rules for combining them be sufficient to be able to generate an emergent phenomenon? Holland uses generalized building blocks that he calls constrained generating procedures. These are systems that according to its inner states and the stimuli received behave in a definite way specified unequivocally by certain rules. Connecting these devices a new constrained generative procedure of a higher level is obtained. The new higher level devices can also be combined to obtain a device of even higher level, and so on. At certain level, different from the basic level, some regularities can be obtained. These regularities are the emergent phenomena that Holland studies. They are macro laws that not necessarily can predict all future behavior, but capture some of the important regularities at that description level. These regularities are detected when expressed in the right format, because they resonate with inductive biases of the perceptual, cognitive and motor system of the observer.

To explore the role of the observer on the emergence of a new phenomenon when combining some building blocks we analyze some examples. First, consider the dynamic pattern phenomenon that emerges in figure 1. It seems that it is needed a minimum number of elements to have the emergence phenomenon. In the following sequence, the visual and cognitive algorithms produce a dynamic pattern if there are at least six rows and six columns of squares. Figure 9, for example, does not produce the dynamic pattern phenomenon on human observers.

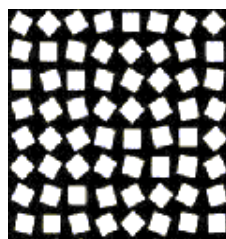


Figure 4

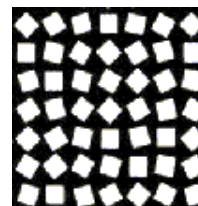


Figure 5

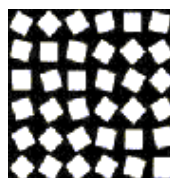


Figure 6

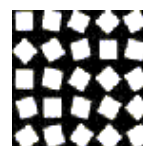


Figure 7

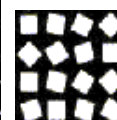


Figure 8



Figure 9

Look now Figure 10. It has 16 squares arranged in a four by four array, where each square is rotated 22.5 degrees with respect to a neighbor square. This Figure does not generate the dynamic pattern illusion in a human observer, and therefore we can agree that its complexity is less than the complexity of figure 1. However next to it there is Figure 11, which is exactly Figure 10 repeated four times towards the right and four times to the bottom of the page. Even though figure 10 does not generate the dynamic pattern illusion, Figure 11 does generate the dynamic illusion. Therefore, the act of repetition of the same figure generates a figure of higher complexity. This emergence of higher complexity shows again the dependence on the format and observer's perceptual, motor and cognitive system. Since figure 10 doesn't generate a dynamic phenomenon a textual description of figure 10 would seem to be enough to compute its complexity. Then a textual description specifying the successive repetition of figure 10 would be enough to describe figure 11. But this is not so. It is required to do the repetition in the adequate visual format and then look at the formed figure 11 to appreciate that the complexity is more than just the complexity of the seed figure 10 repeated 16 times.

Let's consider now figure 12 as the seed figure and repeat it 4 times to the right and four times to the bottom of the page to form figure 13. Clearly figure 13 does not generate the same dynamic pattern phenomenon that figure 11 generates. It generates a different dynamic pattern illusion. It is not simple to say which is more complex, the one generated by 11 or the one generated by 13. However, it seems that the seed figure 10 is more complex than the seed figure 12, or at least than the seed figure 14 that also generates figure 13 by repetition.



Figure 10

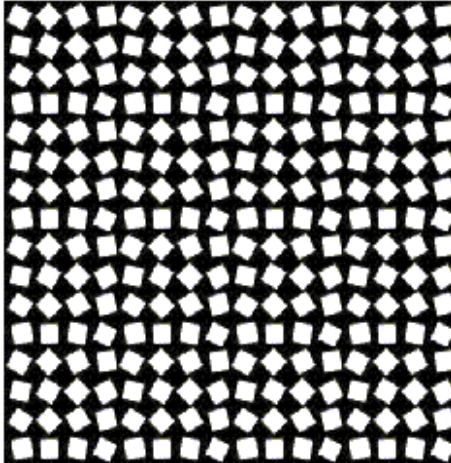


Figure 11

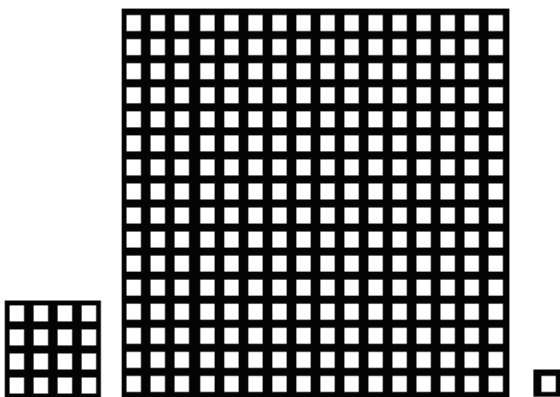


Figure 12

Figure 13

Figure 14

Now look again at figure 1 but from a distance of one or two meters far way. The squares seem fuzzy and the dynamic pattern phenomenon disappears. But a fuzzy image is equivalent to distort each square or pass it through a filter that adds noise. Therefore, each square is more complex, since a longer description is needed for each one. One would expect then that the complexity of the whole figure 1 seen from that distance is increased, but the dynamic pattern is now not present for a human observer. That means a more complex seed produces an apparently less complex output figure than the one produced from a simpler seed.

It is difficult to explain these emergent phenomena if we let the observer out of the factors. For example, if we imagine a camera with a hardware specialized to recognize squares, then several squares will emerge from Figures 12 and 13, and no one from figure 10 and 11.

3.2 Emergent illusions in cognition

One could think that this view of emergence is particular to the visual system. It is just a phenomenon of visual illusion. One could also think

that this type of emergent illusions happens also in other perceptual systems. It is more difficult though to believe that this emergent illusions also happens on more abstract phenomena, such as in mathematical thinking.

Let's look then a purely cognitive example, highly relevant in mathematics education. Let's consider the discovery process in mathematics. Poincaré and Hadamard (Hadamard, 1945) have proposed a recombination and selection mechanism where the subject combine some very basic ideas, kind of building blocks, and selects the proper combination. We illustrates it here for an arithmetic task. Siegler and Stern (Siegler & Stern, 1998), and Siegler and Araya (Siegler & Araya, 2005) studied the discovery mechanism of young kids when solving arithmetic problems of the form " $a + b - c$ ", with $b \geq c$ (for example: $24 + 12 - 12$). After several trials kids started to discover that when " $b=c$ " the solution was " a " and there was no need to do " $a+b$ ", and then subtract " b ". The mechanism proposed has several basic motor and cognitive actions that are postulated as the building blocks of any strategy the kids use to solve the problems. Some of these basic actions are:

- look at the extreme left of the string " $a+b-c$ ",
- shift visual attention to the right one position,
- add the two numbers in top of the working memory,
- load to the working memory the number that is located at the spot where the visual attention is directed, etc.

With these basic building blocks the normal "computational strategy" (do " $a+b$ " and then subtract " c ") can be expressed as the execution of the following 8 actions:

1. look at the extreme left of the string " $a+b-c$ ",
2. load the number at the working memory where the visual attention is directed,
3. shift attention one position to the right,
4. load the number at the working memory where the visual attention is directed,
5. add the two numbers at the top of the working memory,
6. shift attention one position to the right ,
7. load the number at the working memory where the visual attention is directed,

8. add the two numbers at the top of the working memory.

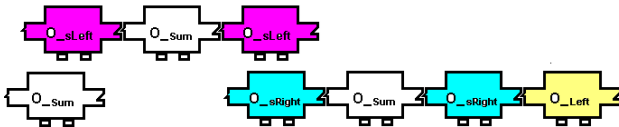
We have thus two levels of description: the level of strategies (the higher level) and the level of the basic actions (the lower level). If we now represent graphically in as small machines each one of the four basic actions, then the “computational strategy” looks like the sequence in figure 15. It has to be read from right to left (as it is usual in the mathematical notation for the composition of functions).



Figure 15

This means that first the machine of the extreme right “O_Left” do its job: look at the extreme left of the string “a+b-c” and load the number at the working memory where the visual attention is directed. Then the following machine “O_sRight” operates: shift attention one position to the right and load the number at the working memory where the visual attention is directed. Next, machine “O_Sum” adds the two numbers at the top of the working memory. Then again “O_sRight” shifts attention one position to the right and load the number at the working memory where the visual attention is directed. Finally, “O_Sum” add the two numbers at the top of the working memory.

If we insert the sequence “O_sLeft O-Sum OsLeftf” in the position indicated



we get the sequence in figure 16:

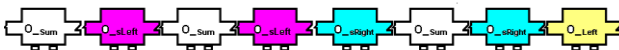


Figure 16

But this long sequence is redundant. It computes certain number that doesn’t affect the final result. If the redundancy is eliminated the sequence in figure 17 is obtained:

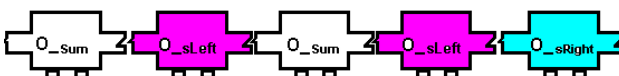


Figure 17

But this sequence is the low level of description of the strategy “Do b-c and then add a”, that in this

case is just “a”. Thus doing recombination of these basic actions, at some point emerges the “shortcut” strategy: “if b=c then a”. This discovery produces an “aha” moment of insight, but initially the shortcut strategy is used by the subject unconsciously. The discovery requires a cognitive mechanism for doing recombination of the building blocks. This is one mechanism that we are strongly biased towards. But also it is required a bias, called Goal Sketch Filter (Siegler, 1996; Siegler & Araya, 2005), that appreciates which type of recombination produce a feasible sequence of actions and which do not. There is also another heuristic that eliminates redundancy that could be generated at some point. In this more abstract example, we see again the role of the cognitive and motor system of the observer in order to produce a new emergent strategy. A hypothetical observer, or a problem solver in this case, with a completely different cognitive system, even if he has the capability of action recombination, will probably not generate the shortcut strategy. There has to be in place the complete mechanism. For example, if this goal sketch filter were inexistent then the search of possibilities is huge and most strategies generated will not work. The generation of the new strategy is product of several strong inductive biases, like the one that decompose a phenomena as a combination of more elementary building blocks. With these biases the subject rapidly discover the shortcut strategy. This means, the emergence of this new and more efficient strategy to solve the problem is as dependent on the cognitive system as the structure of the problem “a+b-c”.

The examples analyzed suggest that no clear law exist on how complexity and emergence is produced from the basic building blocks and rules of combining them alone. Everything seems to indicate the crucial role of the cognitive system of the observer. If the objects or events adequately combined resonates with the type of processing algorithms or inductive biases that the observer has then emergence is perceived by the observer.

How about if we promote the use of a more innate format for arithmetic? For example if we train students to perform additions as translation to the right and subtraction as translation to the left on the number line. Would this training cause a faster discovery of the shortcut strategy and a faster ability to explain the strategy? We have not realized a complete empirical study but we predict that the training in this spatial format have an impact on the time required to discover the shortcut strategy. We also predict that after the training students are less surprised by the shortcut strategy, losing the magic

“aha” moment, and that they will consider it a very natural strategy and can explain better why it works.

3.3 Emergence is not arbitrary constructed

Similarly to the complexity notion, the concept of emergence depends on the observer and his perceptual, motor and cognitive system. It is an embodied notion. Is it then an arbitrary construct? As we have argued, the cognitive system has strong inductive biases for certain very specific patterns. They have evolved throughout our evolutionary history. They are several specific and effective heuristics that meaningfully take advantage of a compressed structure of the world encoded in our DNA and in our interaction with the world. Thus the emergence of a “new” phenomena from their basic building blocks is dependent on the cognitive system of the observer. However, different human observers have the same biases. Therefore we see the same emergent phenomena.

What about the radically new emergent phenomena that seems clearly not present on any of its building blocks? For example, some properties of a particular molecule that are not present on the atoms that compose it. There seems to emerge new properties independent of any observer. To see why this is not so, consider first that molecules and atoms are constructions that we have designed that capture some very specific regularities that our perceptual, motor and cognitive system detects. They are very successful in explaining and predicting several features of the world. Second, consider that atoms have been constructed with laws of interaction, and that these rules implicitly imply the laws of molecules. This is called “weak” emergence by Simon (Simon, 1996). Then how come we feel that there is an emergent phenomena? The trick is that at a higher level, as Holland puts it, we can also detect certain macro laws. They are laws detected at a higher level of description. These macro laws were not necessarily derived by the observer from the laws ruling at the level of atoms. These macro laws are detected by other pattern recognition biases of the observer. These macro laws only consider molecules, and not lower levels constructions. At the molecular level the macro law description is much simpler for the observer because it resonates with some of his inductive biases and require much less logical computing power. According to Holland (Holland, 1998) descriptions formulated at the higher level means great gain in comprehension. However this doesn't mean that these macro laws cannot be obtained throughout long chains of logical deductions from the laws of the lower levels

(atoms). They could require very long calculations, even ones that are not feasible to do in reasonable amount of time.

Brain power is not unlimited and is not content independent. The brain has to produce solutions in a very limited time. Therefore the basic biases are encoded in DNA and executed by the nervous system as frugal content specific heuristics. With these heuristics, regularities at different levels of description are detected. The fact that this organization of detected patterns of the world works is a product of several biases that the observer has. She can describe the high order patterns because her cognitive system recognizes those patterns, and this is much simpler than to deduce them from the properties of the atoms. The whole effect is that it seems to us as if they were new emergent rules, not present on the lower order rules of atoms, but they were already there. It is just that we are using a different pattern detection algorithm for the molecules than for the atoms.

For example, on the “a+b-c” arithmetic task, it can be argued that the shortcut strategy can be deduced logically from the properties of integer numbers and the properties of the addition operation. Therefore, the shortcut strategy would be a strategy that is independent of any observer. It would depend only on the rules that define the integers and the addition. However, this is not how the kids generate and discover the shortcut strategy. The process is a slow discovery process that uses several heuristics and that under special laboratory conditions takes several weeks to generate the shortcut strategy. The logical deduction power of the human brain is very limited, and therefore most eight or nine years old kids don't deduce the strategy. We have a bounded rationality (Simon, 1996; Gigerenzer et Al, 1999), with very limited computing power for logical deductions, but it comes with very effective and simple heuristics. Using these heuristics or inductive biases, called ecological rationality by Gigerenzer, the new shortcut strategy is generated unconsciously in a long and stochastic process where several regressions to previous strategies take place (Siegler, 1996; Siegler & Stern, 1998; Siegler & Araya, 2005). In a long Darwinian process the frequency of use of the new strategy augments and at some point the subject becomes aware of it as well. This whole process is a mechanism that work at the level of strategies, a higher level than the level of basic actions. At some point after several times that the observer has already used the shortcut strategy he consciously detects it and perceives it as a new emergent strategy. Similarly to visual illusions, the

emergent phenomena surprises him and activates a chain of emotional reactions that goes with the “aha” experience.

It is natural to wonder how come the atoms and molecules or the different objects and laws at the different levels of description fit so well with the world and have excellent predictive power. The answer is in the long process of construction, testing, and adjustments that these constructions have been experienced by centuries of systematic work. If for a particular phenomenon they don't generate good predictions then at some point they are changed. Change in the atoms and its rules imply changes at the higher levels. Using this mechanism we have produced in some domains constructs with impressive predictive power.

The emergence of a phenomenon from other more simple phenomena is very common in nature. Consider the emergence of consciousness. Each human cell is a small machine or robot that knows nothing about art or dogs. How come, asks the philosopher Daniel Dennett (Dennett, 2005, 2017), is it possible that even if they are conscious cells, they compose themselves into a thing with conscious thought about Bracque or poodles?

Embodied emergence means that the perceptual, motor and cognitive systems of the observer plays a key role. As product of a long evolutionary process the search for certain specific patterns is encoded on his DNA and his interaction with the world. These algorithms captures relevant patterns of the world that are important for the survival of the species. Some of these algorithm search patterns at certain level of organization and need to attract the attention of the observer to those patterns. For example, to the patterns that correspond to a poodle. That's why he consciously detect the poodle. Thus the emergence of consciousness is an embodied emergence. It is generated by the algorithms that detect those patterns at a much higher level of description than the cellular level.

4 CONCLUSIONS

Today we know that most reasoning is unconscious and abstract ideas arise from using our brains, bodies, and bodily experience. Even mathematics, once considered god's thought, comes from perceptual, motor and basic innate mechanisms such as subitizing (Araya, 2000; Lakoff & Núñez, 2000; Soto-Andrade, 2006). According to George Lakoff and Mark Johnson (Lakoff & Johnson, 1999) “there exist no Fregean person from whom thought has been extruded from the body”. It is expected then

that complexity and emergence, two basic notions conceived by our brains, depend on the observer and his perceptual, cognitive and motor system. Throughout several examples we have shown that there cannot exist a universal Fregean concept of complexity and emergence. This is not what everybody normally imagine about these concepts, since there is the implicit understanding that complexity and emergence are properties that depends only of the system, its elements and its organization. This is an example of the myth of objectivism (Lakoff & Johnson), where an observer independent world of objects exists. This world would contain objects such as stones and animals, and also would contain more abstracts objects like complexity and emergence. Nevertheless, we have seen that the complexity and the emergence of a phenomenon depends crucially on the observer and his body and brain. Furthermore, if the format of the information is changed, then complexity changes and a potential emergent phenomenon does not occur at all. The emergent phenomenon lives on the brain of the observer. It is constructed by his cognitive system, as a movie is constructed on the observer's brain from several still images. Nevertheless this doesn't mean that is arbitrary or completely subjective entity that doesn't correspond to real properties of the world. Because the observer uses evolved algorithms that detect highly meaningful patterns, the emergent phenomena that she detects are not arbitrary. Thus, in the end, the constructions built by different observers are not that different. They have a lot in common and they reflect real properties of the world.

This view has important consequences for understanding nature and for education. Complex concept can be more easily understood if connected appropriately with intuitive and embodied knowledge, also called biologically primary knowledge. A similar process happens with the phenomena of emergence. If life is an emergent phenomenon, or consciousness is an emergent phenomenon, then this is something that emerges because they somehow resonates with the processing algorithms and circuits of our perceptual, motor and cognitive system. For example, it resonates with our intentional detector system. This emergence is not something universal that exists independent of human observers or observers with a cognitive structure similar to ours. These emergent phenomena are in our brains and body as much as they are out there in the external world. This connection with innate and embodied knowledge means important strategies for teachers. They have to help students

connect notions and phenomena to students' innate and embodied knowledge, then these notions become less complex and the emergent phenomenon loses its magic. This way students can realize that apparently new phenomena becomes logically connected to the subjacent phenomena.

Acknowledgement. To Funding from PIA-CONICYT Basal Funds for Centers of Excellence Project FB0003 is gratefully acknowledged and to the Fondef D15110017 grant from CONICYT

5 REFERENCES

- Araya, R. (2000). *Inteligencia Matemática*. Editorial Universitaria.
- Araya, R. (2000b). A.I., Máquinas que Piensan y Matemáticas. Interview to John Casti. Artes y Letras, El Mercurio. http://www.iing.cl/docs/Araya_EntrevistaCasti.doc
- Araya, R. (2006). Complexity and Emergence are Embodied and Observer Dependent Notions but are not Arbitrary. Presentación Congreso Universidad Diego Portales.
- Araya, R., Calfucura, P., Jiménez, A., Aguirre, C., Palavicino, M., Lacourly, N., Soto-Andrade, J., & Dartnell, P. (2010). The Effect of Analogies on Learning to Solve Algebraic Equations. *Pedagogies: An International Journal*, 5(3).
- Bailey, D. H., Hoard, M. K., Nugent, L., & Geary, D.C. (2012). Competence with fractions predicts gains in mathematics achievement. *Journal of Experimental Child Psychology*, 113, 447-455.
- Baum, E. (2004). *What is Thought?* Mit Press.
- Baron-Cohen, S. (1997). *Mindblindness: An Essay on Autism and Theory of Mind*. Mit Press.
- Bossomaier, T., & Snyder, A. (2004). Absolute pitch accesible to everyone by turning off part of the brain. *Organised Sound*, 9(2), 181-189.
- Casti, John. (2000). *Five More Golden Rules*. John Wiley & Sons.
- Chaitin (1974). Information-Theoretic Computational Complexity. *IEEE Transactions on Information Theory* IT-20, pp. 10-15
- Cosmides, L., & Tooby, J. (1996). Are Humans Good Intuitive Statisticians After All? Rethinking Some Conclusions from the Literature on Judgment Under Uncertainty *Cognition* 58, pp. 1-73
- Dennett, Daniel. (2005). *Sweet Dreams: Philosophical Obstacles to a Science of Consciousness*. MIT Press.
- Dennett, Daniel. (2017). *From Bacteria to Bach and Back. The evolution of Minds*. Mit Press.
- Fruteau, C., Voelkl, B., van Dammea, E., Noe, R. (2009). Supply and demand determine the market value of food providers in wild vervet monkeys. *PNAS*, 106(29), 12007—12012.
- Geary, D. (2007). Educating the evolved mind: Conceptual foundations for an evolutionary educational psychology. In J.S. Carlson, & J.R. Levin (Eds.), *Educating the evolved mind* (Vol. 2, pp. 1-99). Greenwich, CT: Information Age.
- Gell-Mann, M. (2000). The Quark and the jaguar, Adventures in the simplex and the complex.
- Gell-Mann, M. (1995). What Is Complexity? *Complexity*, 1(1).
- Gigerenzer, G., Todd, P., & the ABC Research Group (1999). *Simple Heuristics that Makes Us Smart*. Oxford University Press.
- Gigerenzer, G. (2000). *Adaptive Thinking. Rationality in the Real World*. Oxford University Press.
- Hadamard, J. (1945). *The Mathematician's Mind: The Psychology of Invention in the Mathematical Field*. Princeton University Press.
- Holland, J. (1998). *Emergence* Addison Wesley.
- Inoue, S., & Matsuzawa, T. (2007). Working memory of numerals in chimpanzees. *Current Biology*, 17(23), pR1004-R1005.
- Jiménez, A., & Araya, R. (2013). Comparación de Tres Metodologías para el Aprendizaje de Fracciones. XVII Jornadas Educación Matemática.
- Kolmogorov, A. (1968). Logical basis for information theory and probability theory, *IEEE Trans. Inform. Theory*, vol. IT-14, pp. 662-664.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. The University of Chicago Press.
- Lakoff, G., & Johnson, M. (1999). *Philosophy in the Flesh: the Embodied Mind and its Challenge to Western Thought*. Basic Books.
- Lakoff, G., & Núñez, R. (2000). *Where Mathematics come from: How the embodied mind brings mathematics into being*. Basic Books.
- Mercier, H., & Sperber, D. (2017) *The Enigma of Reason*. Harvard University Press. Cambridge.
- Obersteoner, A., Van Dooren, V., Van Hoof, J., & Verschaffel, L. (2013). The natural number bias and magnitude representation in fraction comparison by expert mathematicians. *Learning and Instruction* 28(2013), 64-72.
- Pinker, S. (2002). *The Blank Slate: The Modern Denial of Human Nature*. Viking.
- Povinelli, D. (2000). *Folk Physics for Apes: The Chimpanzee's Theory of How the World Works*. Oxford University Press.
- Rieke, F., Warland, D., van Steveninck, R., Bialek, W. (1997). *Spikes: Exploring the Neural Code*. The MIT Press.
- Siegler, R. (1996). *The Emerging Mind: The Process of Change in Children's Thinking*. Oxford University Press.
- Siegler, R., & Araya, R. (2005). A Computational Model of Conscious and Unconscious Strategy Discovery. *Advances in Child Development and Behavior*, 33.
- Siegler, R., & Stern, E. (1998). Conscious and Unconscious Strategy Discoveries: A Microgenetic Analysis. *Journal of Experimental Psychology: General*, 127(4), 377-397.
- Siegler, R., Carpenter, T., Fennell, F., Geary, D., Lewis, J., Okamoto, Y., Thompson, L., & Wray, J. (2010). Developing effective fractions instruction for kindergarten through 8th grade: A practice guide (NCEE #2010-4039). Washington, DC.
- Siegler, B., Braithwaite, D. (2017) Numerical Development. *Annu. Rev. Psychol.* 2017. 68:187–213
- Simon, H. (1996) *The Science of the Artificial*. Mit Press.
- Sloman, S., Fernbach, P. (2017) *The Knowledge Illusion. Why we Never Think Alone*. Rverhead Books. N.Y.
- Snyder, A., Bossomaier, T., Mitchell, J. (2004). Concept Formation: "Object" Attributes Dynamically Inhibited from Conscious Awareness. *Journal of Integrative Neurosciences*, 3(1), 31-46.
- Solso, R. (1994). *Cognition and the Visual Arts*. Mit Press.
- Soto-Andrade, J. (2006). Un monde dans un grain de sable: métaphores et analogies dans l'apprentissage des mathématiques. *Annales de Didactique et de Sciences Cognitives*, 11, 123-147.