

Structure emerges faster during cultural transmission in children than in adults

Vera Kempe¹

Nicolas Gauvrit²

Douglas Forsyth¹

¹Abertay University

²CHART, University of Paris VIII and École Pratique des Hautes Études

Word Count: 2,999

Corresponding author:

Vera Kempe, Abertay University, Bell Street, Dundee, DD1 1HG, Scotland, United Kingdom, email: v.kempe@abertay.ac.uk.

Abstract

How does children's limited processing capacity affect cultural transmission of complex information? We show that over the course of iterated reproduction of two-dimensional random dot patterns transmission accuracy increased to a similar extent in 5- to 8-year-old children and adults whereas algorithmic complexity decreased faster in children. Thus, children require more structure to render complex inputs learnable. In line with the Less-Is-More hypothesis, we interpret this as evidence that children's processing limitations affecting working memory capacity and executive control constrain the ability to represent and generate complexity, which, in turn, facilitates emergence of structure. This underscores the importance of investigating the role of children in the transmission of complex cultural traits.

Keywords: cultural transmission, iterated learning, serial reproduction, Less-Is-More, random dot patterns, children, algorithmic complexity

Introduction

Cultural traits and social conventions owe not just their survival but also their structural characteristics to the process of cultural transmission. Starting with experimental studies of iterated reproduction of narratives and drawings (Bartlett, 1932), a growing line of research investigating technological (Mesoudi & Whiten, 2008; Whiten, 2011) and communicative (Galantucci & Garrod, 2011) systems demonstrates that cultural transmission of information across generations of learners reduces entropy, thereby rendering cultural traits functionally relevant (Caldwell & Millen, 2009) and suitable to serve as social markers (Matthews, Roberts & Caldwell, 2012). This issue has received particular attention in research on the evolution of language, where it has been suggested that structure emerges because language is gradually shaped by the information processing capacity of successive generations of learners. Computational (Smith, Kirby & Brighton, 2003; Reali & Griffiths, 2009) and experimental (Kirby, Cornish & Smith, 2008; Smith & Wonnacott, 2010) studies of iterated language learning show that languages with inconsistent mappings between forms and meanings become progressively more learnable and more compositional, i.e. they acquire a more systematic and predictable relationship between form features and meaning dimensions. Iterated learning amplifies the biases that affect learners' inductive inferences leading to gradual emergence of structure (Reali & Griffiths, 2009). Moreover, cultural transmission reveals not just inductive biases that operate during iterated learning of artificial languages or even meaningless colour sequences (Cornish, Smith & Kirby, 2013) but also general memory biases that operate during serial reproduction of information (Xu & Griffiths, 2010). However, experimental studies of cultural transmission have so far only focussed on adult learners who generally are at peak cognitive ability, yet many cultural traits and social

conventions are learned and transmitted by children whose information processing capacity is more limited. To fully understand how processing constraints shape the biases that lead to the emergence of structure it is important to compare cultural transmission between cognitively and socially mature and immature learners. This is the aim of the present study.

Inspiration for this study comes from the domain of language learning, where the existence of regularising biases has been attributed to either language-specific or domain-general mechanisms. Language-specific biases imply access to an innate Universal Grammar (Bickerton, 1981), often conceived of as an evolved adaptation to the requirement of constraining the hypotheses of learners faced with limited input (Pinker & Bloom, 1990). Domain-general biases, on the other hand, arise from information processing constraints, which – according to the Less-Is-More Hypothesis (Newport, 1990) – can facilitate language learning. Several mechanisms have been proposed to account for beneficial effects of processing limitations. On the one hand, limited working memory capacity may make it easier for children to attend to smaller units, e.g. grammatical morphemes, in input sequences, which can go unnoticed by adults processing larger input segments. Connectionist simulations using recurrent networks (Elman, 1993) and experimental evidence from adult learners (Kersten & Earles, 2001) have demonstrated that imposing limitations on working memory capacity, or on input complexity, leads to superior outcomes in terms of mastery and generalisation of the underlying linguistic structure. On the other hand, regularisation and emergence of structure have been linked to immature executive control. Hudson Kam and Newport (2005) showed that when inputs contained inconsistencies in the use of grammatical forms, adults tried to match the probabilities of occurrence of each form, while 5- to 7-year old children regularised the

inconsistencies by using predominantly one or the other form. Thus, compared to adults, children are less able to inhibit pre-potent responses (Thompson-Schill, Ramscar & Chrysikou, 2009), although adults have been found to regularise as well if increased input complexity taxes their processing resources (Hudson Kam & Newport, 2009; Vouloumanos, 2008). Both of these domain-general mechanisms lead to the prediction that structure should emerge more rapidly during cultural transmission of information in chains of children than of adults, a prediction that is in line with the notion that children play a crucial role in creolisation, e.g. as described for the emergence of Nicaraguan Sign Language (Senghas & Coppola, 2001).

One of the reasons direct experimental evidence comparing iterated language learning between adults and children has not yet been provided may be due to the difficulty of administering artificial language learning regimens to large numbers of children. During artificial language learning, learners are exposed to multiple exemplars one at a time, and rely on frequency of occurrence to make inductive probabilistic inferences about the underlying rules that generate these examples. This is a time-consuming process difficult to implement in a sufficiently controlled manner on a scale large enough to allow for statistical comparison between chains of children and of adults. In an attempt to provide proof of concept, the present study therefore investigates serial reproduction of visuo-spatial patterns, assuming that similar regularising biases operate during transmission of information dispersed not in time but in space. We asked learners to recreate patterns of dots distributed randomly on a grid by placing the same number of dots into the appropriate cells of an empty grid. To ensure task compliance, we used a sticker task – a familiar and highly enjoyable activity for children. As with sequence learning, learning of spatial configurations also requires learners to infer rules that can be used to reconstruct patterns for recall.

We take this procedural component as being loosely analogous to learning grammar. For sequences and for spatial configurations, emergence of structure can be conceptualised as reduction in Kolmogorov-Chaitin (or “algorithmic”) complexity, which is defined as the length of the shortest program running on a Universal Turing Machine, i.e. an abstract general-purpose computing device (Li & Vitányi, 2009), that produces the pattern and then halts. It captures the notion of structure in terms of the length of the procedure required to generate it – the shorter the algorithm required to generate a pattern, the more structured and less complex the pattern is. Unfortunately, Kolmogorov-Chaitin complexity is uncomputable (Gauvrit, Soler-Toscano, Zenil & Delahaye, 2014). However, the Coding Theorem Method (see below for more details) provides a practical way to accurately approximate algorithmic complexity for short series (Gauvrit et al., 2014; Soler-Toscano, Zenil, Delahaye & Gauvrit, 2013). This approximation is superior to alternative attempts to capture degree of non-randomness such as measures based on the mere distribution of outcomes, as it does not focus on intuitively plausible biases that may affect randomness generation but rather constitutes an objective measure that captures all potential biases. Crucially, if processing constraints shape the biases that lead to the emergence of structure, we expect that algorithmic complexity will be reduced faster in transmission chains of children than of adults.

Method

Participants: Ninety adults, all undergraduate students, participated in the study, comprising nine diffusion chains of 10 generations each. Three randomly generated seed patterns were used to start three diffusion chains per seed. Ninety primary school pupils aged 5;1 to 8;2 years also participated in the study. 30 children were recruited

from grade 1, 30 from grade 2, and 30 from grade 3. There were three diffusion chains per grade level, each starting with a different one of the three random seeds.

Materials: We prepared 180 10×10 grids comprising an area of 15×15 cm, printed on A4 paper, and slips containing 12 black round stickers with a diameter of 12 mm each. To control for effects of starting conditions across age groups, three seed patterns were generated by randomly selecting 12 numbers out of 100 using the Excel random number generator, and equally distributed across the adult and the child chains, matching for primary school grade level.

Procedure: Participants were given the 12 stickers, and seated at a table with a sheet containing the empty 10×10 grid in front of them. They were instructed to look at a pattern of 12 black dots distributed on the grid for 10 seconds, and to recreate this pattern by placing the provided stickers inside the cells of the empty grid. The experimenter then placed the random dot pattern (at the beginning of each chain), or the dot pattern produced by the previous participant, on top of the empty grid for 10 seconds, which were timed using a stopwatch. After removal of the target pattern, participants placed all 12 stickers onto the empty grid.

Results

Below we report the analyses of four dependent variables that allowed us to quantify learnability, structure and identifiability of the patterns. Each variable was analysed using Page's L trend test for $m = 9$ chains and $n = 10$ generations, to identify continuous trends over the course of transmission. This test was conducted separately for children and for adults. Each variable was also analysed using a $10 \times 2 \times 3$ mixed-

type ANOVA with Generations (1-10) as repeated measures, and Age (children vs. adults) and Random Seed (1-3) as between-subjects factors. This analysis provides information about whether there are differences between children and adults, whether age differences depend on position in the transmission chain, and whether the initial conditions influence the outcome. Following Caldwell & Millen (2009), a second $2 \times 2 \times 3$ mixed-type ANOVA combined values from early (1-3) and late (4-10) generations, to ascertain whether trends that occur over the course of transmission interacted with age or initial conditions. The results of all statistical analyses are summarised in Table 1.

Table 1: Results of the statistical analyses for all dependent variables.

dependent variable	Page's test	<i>L</i> -trend (m=9, n= 10)	ANOVA 1 10 (Generation) × (Age) × 3 (Seed)	ANOVA 2 2 (EarlyLate) × 2 (Age) × 3 (Seed)
Transmission Accuracy	Adults: $L = 3013$; p < .001 Children: $L = 3115$; p < .001		Generation: F(9,108) = 6.92; p < .001	EarlyLate: F(1,12) = 44.35; p < .001
Clustering	Adults: $L = 3200$; p < .001 Children: $L = 3092$; p < .001		Generation: F(9,108) = 13.32; p < .001 Age: F(1,12) = 6.70, p < .05	EarlyLate: F(1,12) = 41.91 Age: F(1,12) = 6.30, p < .05
Algorithmic Complexity	Adults: $L = 3183$; p < .001 Children: $L = 3354$; p < .001		Generation: F(9,108) = 4.53, p < .001 Age: F(1,12) = 8.07, p < .05	EarlyLate: F(1,12) = 13.96, p < .01 Age: F(1,12) = 7.66, p < .05.
Identifiability	Adults: $L = 2897$, p < .05 Children: $L = 3073$, p < .001		Generation: F(9,108) = 9.07, p < .01 Age: F(1,12) = 13.10, p < .01	EarlyLate:

Transmission Accuracy: Transmission accuracy was calculated as the percentage of dots correctly placed on cells in the grid, and showed continuous improvement in children and in adults (see Figure 1). The ANOVA revealed a significant effect of Generations but no effects of Age, Seed or any of the interactions. Combining data for the three early and the seven late generations confirmed that transmission accuracy increased, an effect that did not interact with Age or Random Seed. Thus, for children and adults, transmission accuracy increased at the same rate, independently of the specific random pattern presented at the outset.

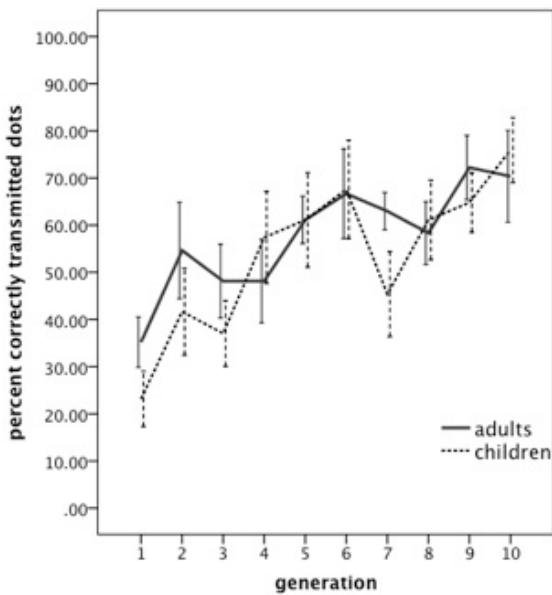


Figure 1: Mean transmission accuracy per generation for adults (solid line) and for children (dotted line). Error bars indicate ± 1 S.E.M.

Clustering: One way in which structure can be imposed on random dot patterns is by combining them into clusters of connected dots, a process akin to chunking – a mechanism that can reduce memory load, in this case for the spatial positioning of the dot clusters (see Cladière, Smith, Kirby & Fagot [2014] for a similar analysis). Here, a cluster was defined as any group of dots placed into cells sharing either a side or a

corner. Page's L trend test showed continuous reduction of the number of clusters in children and in adults (see Figure 2). The ANOVA revealed significant main effects of Generation and Age, but no effect of Seed or of any of the interactions. Combining data for the three early and the seven late generations confirmed these main effects. Thus, numbers of cluster diminished over the course of transmission with children producing fewer cluster than adults from quite early on.

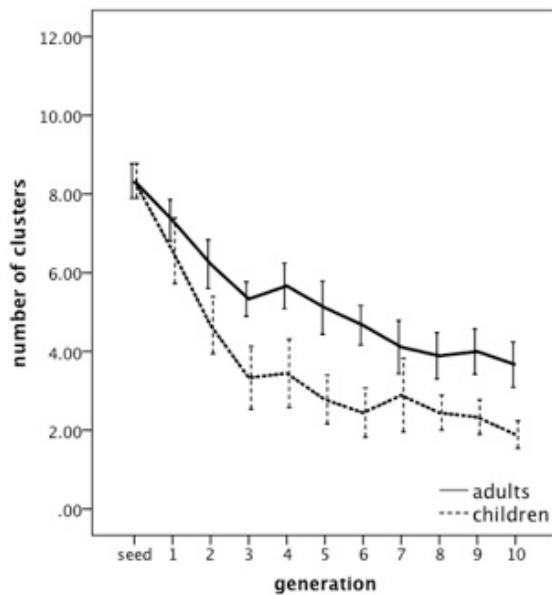


Figure 2: Mean number of clusters per generation for adults (solid line) and for children (dotted line). Error bars indicate ± 1 S.E.M.

Algorithmic Complexity: Algorithmic complexity of the two-dimensional patterns was computed using the Block Decomposition Method (Zenil, Soler-Toscano, Dingle & Louis, 2014). The basic idea behind BDM is to split the target into smaller, sometimes overlapping, 4×4 grids. Unlike image compression algorithms, this estimate is axially and rotationally invariant. In this study, each 10×10 grid thus yielded sixteen 4×4 tiles with a maximum overlap of 2 columns or 2 rows. Given the two-dimensional complexity of each tile, we computed an estimation of the pattern

complexity using the formula: $\sum_p (\log_2(n_p) + K(p))$, where p denotes the different types of 4×4 patterns, n_p the number of occurrences of each type of pattern p , and $K(p)$ the complexity of p as computed using the coding theorem method developed by Zenil et al. (2012).

For this measure, Page's L trend test showed a continuous reduction in children and in adults (see Figure 3). The ANOVA revealed significant main effects of Generation and Age, but no effect of Random Seed or of any of the interactions. Combining the data for the three early and the seven late generations confirmed the decrease in complexity, as well as the effect of Age, yet showed no interactions. These results indicate that complexity decreased over the course of transmission, and that it was lower overall in the children.

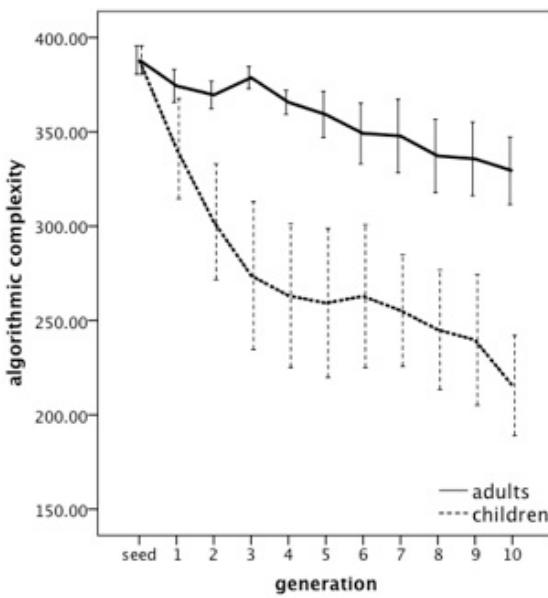


Figure 3: Mean estimated algorithmic complexity per generation for adults (solid line) and for children (dotted line). Error bars indicate ± 1 S.E.M.

Identifiability: To maintain consistency in starting conditions across children and adults, we had presented identical seeds to both groups. This allowed us to test whether initial conditions determined the particular attractor states a given transmission chain had settled into, or whether transmission gave rise to uniquely identifiable lineages despite identical initial conditions. For each pattern, we quantified identifiability by computing the proportion of within-chain similarity divided by the sum of within-chain similarity and across-chain similarity (Matthews et al., 2012), where across chain similarity was defined as similarity of patterns originating from the same seed in a given generation. Values greater than 0.5 indicate greater within-chain similarity; values below 0.5 indicate greater similarity with patterns of the same generation in the different chains originating from the same seed.

As Figure 4 illustrates, from the first chain onwards, each reproduced pattern resembled other patterns within the chain more than other patterns of the same generation in other chains originating from the same seed. A one-sample Wilcoxon's signed ranks test for each generation showed that for the adults, identifiability was significantly greater than the expected chance level of 0.5 for all but generation 1 (all Z 's > 2.42 , $p < .05$), and for the children, all generations exceeded the chance level (all Z 's > 2.10 . $p < .05$). Thus, even when originating from the same random seed, by generation 2 patterns already started forming distinct lineages. Page's L trend test showed that for adults and children, resemblance to other patterns within a chain increased over the course of transmission. The lack of effects of Random Seed and Age in both ANOVAs indicates that formation of lineages was not contingent upon initial conditions, and that children did not differ from adults in terms of how much each pattern resembled other patterns within its lineage.

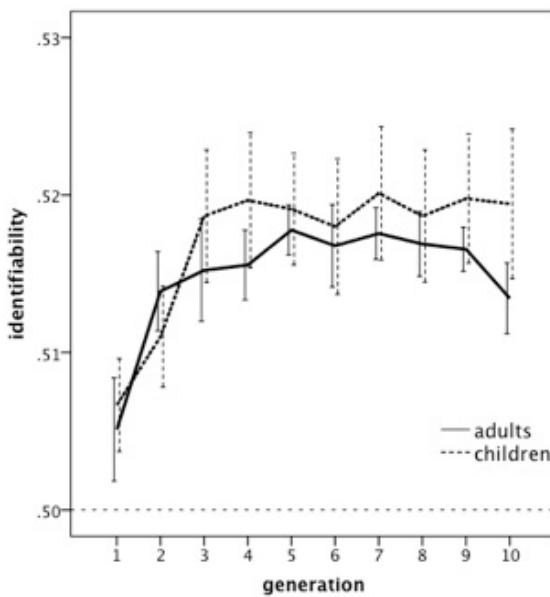


Figure 4: Mean identifiability per generation for adults (solid line) and for children (dotted line). The dashed line represents chance level. Error bars indicate ± 1 S.E.M.

Discussion

Our study tested whether children and adults differed in terms of how fast learnability improved and structure emerged during transmission of random dot patterns. Consistent with findings on iterated learning of artificial languages (Kirby et al., 2008; Smith & Wonnacott, 2010) and random sequences (Cornish et al., 2013) in adults as well as recent findings on cultural transmission of visuo-spatial patterns in non-human primates (Cladière et al., 2014) we found that transmission accuracy of visuo-spatial patterns increased in children as well, and to a similar extent as in adults. However, compared to adults, visual working memory capacity of children is limited, and children's memory representations of visuo-spatial pattern are less stable, reaching adult levels of performance in visual working memory tasks only around age 7 (Simmering, 2012). Given these limitations, it is remarkable that transmission

accuracy in 5- to 8-year-old children did not differ from that of adults. How can children achieve such a feat?

Using measures of clustering and algorithmic complexity, we discovered that structure emerged more readily in the children, reducing the complexity to a level that allowed them to reproduce the patterns as successfully as the adults. The lack of an interaction between number of generation and age for these two measures indicates that the age differences emerged early on during transmission, as illustrated in Figure 5. Inspection of individual patterns (see Appendix 1) revealed that in the children, more radical innovations in earlier generations created patterns of considerably reduced complexity, which converged onto more easily transmissible structures. This is compatible with findings that show an inverse relationship between amount of information transmitted and rate of change during transmission (Griffiths, Lewandowski & Kalish, 2013). How can children's processing limitations encourage such structural innovations? One possible explanation is related to the reconstructive nature of memory (Bartlett, 1932): When recalling complex visual patterns, learners may be biased by acquired prototypical representations of shapes, and in children, these prototypes may be simpler and less numerous. Another possibility is that limited executive control makes it more difficult for children to inhibit pre-potent action routines. To illustrate this in simple words: having just placed two stickers onto adjacent cells of the grid, it may be very tempting to just continue placing stickers on the entire row or column of cells as this appears to be an intuitively plausible and well-practiced routine to follow. Adults are likely to be more capable of inhibiting such pre-potent action routines. Thus, both of these explanations support the notion that the cognitive demands of generating and reproducing complexity (Towse & Cheshire, 2007) are higher for children than for adults. Crucially, we have

demonstrated that because reproducing complexity is less feasible for children it makes them more prone to ‘injecting’ structure, which then is amplified over the course of cultural transmission. Importantly, the emergence of unique patterns in each chain shows that distinct lineages develop because structuring is not contingent upon initial conditions but shaped by learner biases.

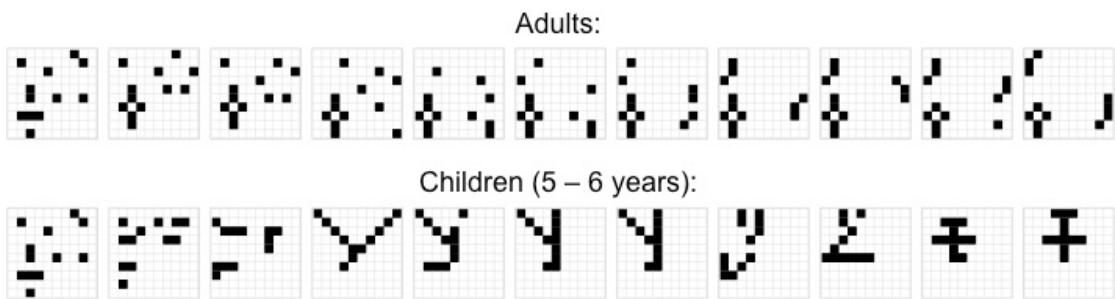


Figure 5: Example of transmission of random dot patterns in adults and 5-6-year-old children, starting from the same random seed (left) and ending in the pattern produced by generation 10 (right).

It certainly is a far stretch from simple random dot patterns to complex cultural traits such as language. Still, our findings draw attention to the possibility that whenever children act as agents of cultural transmission, the structure of the resulting traits can be enhanced due to children’s specific processing limitations. Future research will have to explore to what extent, at what ages, and in what ways the innovative potential that arises from children’s processing limitations affects the transmission of the various complex traits and conventions that make up our surrounding human culture.

Acknowledgements: The authors would like to thank the teachers and pupils of Arduthie Primary School, Stonehaven, UK, for their kind support. We also gratefully acknowledge the contribution of Fernando Soler-Toscano and Hector Zenil to the implementation of the Coding Theorem Method.

Author contribution: VK and DF conceived of, designed and conducted the study; VK and NG analysed the data; VK and NG wrote the paper.

References:

- Bartlett F. C. (1932). *Remembering*. Oxford, UK: Macmillan.
- Bickerton, D. (1981). *Roots of language*. Ann Arbor, MI: Karoma.
- Caldwell, C. A., & Millen, A. E. (2009). Social learning mechanisms and cumulative cultural evolution: Is imitation necessary? *Psychological Science*, 20(12), 1478-1483.
- Cladière, N., Smith, K., Kirby, S., & Fagot, J. (2014). Cultural evolution of a systematically structured behaviour in a non-human primate. *Proceedings of the Royal Society B*, 281, 20141541.
- Cornish, H., Smith, K. & Kirby, S. (2013) Systems from sequences: An iterated learning account of the emergence of systematic structure in a non-linguistic task. *Proceedings of the 35th Annual Conference of the Cognitive Science Society*, 340-345.
- Elman, J. L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition*, 48(1), 71-99.
- Galantucci, B., & Garrod, S. (2011). Experimental semiotics: a review. *Frontiers in Human Neuroscience*, 5.
- Gauvrit, N., Zenil, H., Delahaye, J. P., & Soler-Toscano, F. (2014). Algorithmic complexity for short binary strings applied to Psychology: a primer. *Behavior Research Methods*, 46(3), 732-744.
- Griffiths, T. L., Lewandowsky, S., & Kalish, M. L. (2013). The effects of cultural transmission are modulated by the amount of information transmitted. *Cognitive Science*, 37(5), 953-967.

- Hudson Kam, C. L., & Newport, E. L. (2005). Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development, 1(2)*, 151-195.
- Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science, 14(1)*, 11-28.
- Hudson Kam, C. L., & Newport, E. L. (2009). Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology, 59(1)*, 30-66.
- Kersten, A. W., & Earles, J. L. (2001). Less really is more for adults learning a miniature artificial language. *Journal of Memory and Language, 44(2)*, 250-273.
- Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences, 105(31)*, 10681-10686.
- Li, M., & Vitányi, P. M. (2009). *An introduction to Kolmogorov complexity and its applications*. Springer.
- Matthews, C., Roberts, G., & Caldwell, C. A. (2012). Opportunity to assimilate and pressure to discriminate can generate cultural divergence in the laboratory. *Evolution and Human Behavior, 33(6)*, 759-770.
- Mesoudi, A., & Whiten, A. (2008). The multiple roles of cultural transmission experiments in understanding human cultural evolution. *Philosophical Transactions of the Royal Society of London B, 363*, 3489–3501.
- Pinker, S., & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences, 13(04)*, 707-727.

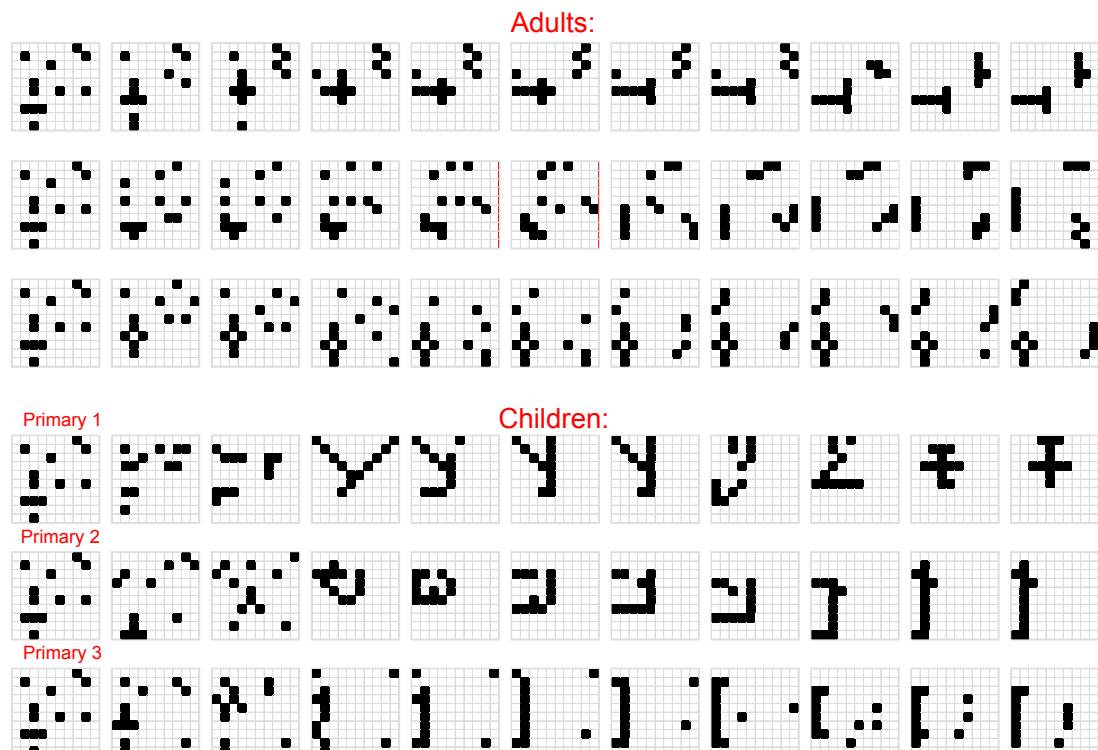
- Senghas, A., & Coppola, M. (2001). Children creating language: How Nicaraguan Sign Language acquired a spatial grammar. *Psychological Science*, 12(4), 323-328.
- Simmering, V. R. (2012). The development of visual working memory capacity during early childhood. *Journal of Experimental Child Psychology*, 111(4), 695-707.
- Smith, K., & Wonnacott, E. (2010). Eliminating unpredictable variation through iterated learning. *Cognition*, 116(3), 444-449.
- Soler-Toscano, F., Zenil, H., Delahaye, J. P., & Gauvrit, N. (2013). Correspondence and independence of numerical evaluations of algorithmic information measures. *Computability*, 2(2), 125-140.
- Thompson-Schill, S. L., Ramscar, M., & Chrysikou, E. G. (2009). Cognition without control when a little frontal lobe goes a long way. *Current Directions in Psychological Science*, 18(5), 259-263.
- Towse, J. N., & Cheshire, A. (2007). Random number generation and working memory. *European Journal of Cognitive Psychology*, 19(3), 374–394.
- Vouloumanos, A. (2008). Fine-grained sensitivity to statistical information in adult word learning. *Cognition*, 107, 729–742.
- Whiten, A. (2011). The scope of culture in chimpanzees, humans and ancestral apes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1567), 997-1007.
- Xu, J., & Griffiths, T. L. (2010). A rational analysis of the effects of memory biases on serial reproduction. *Cognitive Psychology*, 60(2), 107-126.

Zenil, H., Soler-Toscano, F., Delahaye, J.-P. & Gauvrit N. (2012). Two-dimensional Kolmogorov complexity and validation of the coding theorem method by compressibility, *arXiv 1212:6745 [cs.CC]*.

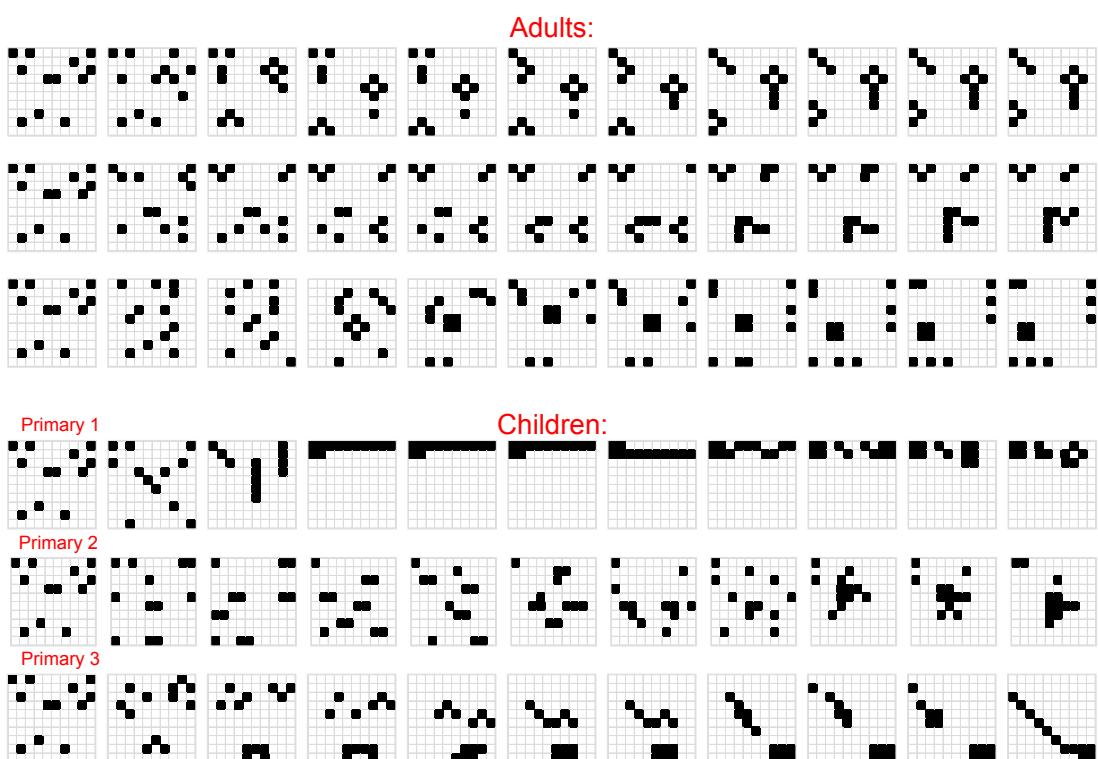
Zenil, H., Soler-Toscano, F., Dingle, K., & Louis, A. A. (2014). Correlation of automorphism group size and topological properties with program-size complexity evaluations of graphs and complex networks. *Physica A: Statistical Mechanics and its Applications*, 404, 341-358.

Appendix 1: Patterns produced by adults and children for all chains and generations:

Seed 1



Seed 2



Seed 3

