

A novel approach to investigate recursion and iteration in visual hierarchical processing

Maurício Dias Martins^{1,2,3} · Isabel Pavão Martins⁴ · W. Tecumseh Fitch¹

© Psychonomic Society, Inc. 2015

Abstract We describe a new method to explore recursive cognition in the visual domain. We define recursion as the ability to represent multiple hierarchical levels using the same rule, entailing the ability to generate new levels beyond those previously encountered. With this definition recursion can be distinguished from general hierarchical embedding. To investigate this recursion/hierarchy distinction in the visual domain, we developed two novel methods: The Visual Recursion Task (VRT), in which an inferred rule is used to represent new hierarchical levels, and the Embedded Iteration Task (EIT), in which additional elements are added to an existing hierarchical level. We found that adult humans can represent recursion in the visuo-spatial domain, and that this ability is distinct from both general intelligence and the ability to represent iterative processes embedded within hierarchical structures. Compared with embedded iteration, visual recursion correlated positively with other recursive planning tasks (Tower of Hanoi), but not with specific visuo-spatial resources (spatial short-term memory and working memory). We conclude that

humans are able to use recursive representations to process complex visuo-spatial hierarchies and that our visual recursion task taps into specific cognitive resources. This method opens exciting opportunities to explore the relationship between visual recursion and language.

Keywords Recursion · Iteration · Cognition · Vision · Fractals · Representation

The capacity to understand and generate complex hierarchies is one of the most fascinating features of human cognition. In many domains, including language, music, problem-solving, action-sequencing, and spatial navigation, humans organize basic elements into higher-order groupings and structures (Badre, 2008; Chomsky, 1957; Hauser, Chomsky, & Fitch, 2002; Nardini, Jones, Bedford, & Braddick, 2008; Unterrainer & Owen, 2006; Wohlschlagler, Gattis, & Bekkering, 2003). This ability to encode the relationship between basic elements (words, people, etc.) and the broader structures in which these are embedded (sentences, corporations, etc.), affords flexibility to human behavior. For example, in action sequencing, and unlike pure serial associative behavior, hierarchical representations allow the omission or modification of certain steps, without impairing the overall goal.

Here, we define hierarchies as non-cyclical tree-like organizations, where higher levels incorporate multiple lower levels in structural representations (Fitch & Martins, 2014), i.e., in which elements are embedded within other elements. This embedding can refer to the grouping of constituents within a higher order set, such as the grouping of individuals within a family (family = {ind1; ind2; ind3}), or it can refer to the establishment of asymmetrical dominance-subordination relationships between constituents, such as in social hierarchies (ind1 dominant over ind2, ind2 dominant over ind3, etc.).

Electronic supplementary material The online version of this article (doi:10.3758/s13428-015-0657-1) contains supplementary material, which is available to authorized users.

✉ Maurício Dias Martins
diasmarm@hu-berlin.de

¹ Department of Cognitive Biology, University of Vienna, Althanstrasse 14, 1090 Vienna, Austria

² Berlin School of Mind and Brain, Humboldt Universität zu Berlin, Luisenstrasse 56, 10117 Berlin, Germany

³ Max Plank Institute for Human Cognitive and Brain Sciences, Stephanstrasse 1a, 04103 Leipzig, Germany

⁴ Language Research Laboratory, Lisbon Faculty of Medicine, Av. Professor Egas Moniz, 1649-028 Lisbon, Portugal

Within the context of hierarchical processing, recursion is an interesting concept that has fascinated scholars in fields as diverse as mathematics, computer science, linguistics, and visual arts. Recursion is interesting because it allows the generation of structures that are both simple and complex at the same time. Recursive structures are complex because they can contain infinite hierarchical levels, and yet simple because this infinity can be achieved and represented using finite rules.

Recursion is a term that has been used to characterize the process of embedding a constituent inside another constituent of the same kind (Fitch, 2010; Hulst, 2010; Pinker & Jackendoff, 2005). Recursive processes can generate hierarchical structures that display similar properties across different levels of embedding. This feature, called self-similarity, is a signature of recursive structures. An example of a recursive linguistic structure is the compound noun “[student] committee”, where we find a noun phrase embedded inside another noun phrase. In contrast, a sentence containing a noun and a verb, such as “[trees] grow”, is hierarchical, but not recursive, because a constituent of one type (noun) is nested within a constituent of a *different* type (verb).

We can also find examples of recursive procedures generating visual hierarchies. For instance, fractals are structures that display self-similarity (Mandelbrot, 1977), that is, they appear similar when viewed at different scales (as in the famous Mandelbrot set). Fractals can be produced by simple rules that generate complex hierarchical structures when applied iteratively to their own output (Fig. 1).

Recently, recursion has become an important topic in cognitive science because the development of the human ability to represent recursion has been considered an important step in the evolution of language (Fitch, Hauser, & Chomsky, 2005; Hauser et al., 2002). In addition, recursion has been proposed to have evolved primarily within the linguistic domain, being accessible to other modalities (e.g., visual domain) only through language (Fitch et al., 2005; Hauser et al., 2002).¹ Other authors have also proposed that recursion might have evolved only in humans, and that recursive thinking is at the core of human cognitive exceptionalism (Corballis, 2014). Testing these hypotheses has been difficult due to both theoretical and methodological limitations.

An empirically useful definition of recursion

Despite considerable agreement about the importance of recursion, many different definitions of recursion are in use (Chomsky, 2010; Corballis, 2007; Gentner, Fenn, Margoliash, & Nusbaum, 2006; Hofstadter, 1980;

Kilpatrick, 1985; Odifreddi, 1999; Penrose, 1989) which has hindered consistent interpretation of empirical results (Fitch, 2010). On the one hand, it has proven to be particularly difficult to establish clear distinctions between recursion and similar processes such as hierarchical embedding and iteration (Hulst, 2010). On the other hand, it has not been clear which level of analysis (process, structure, or representation) is relevant for empirical research (Lobina, 2011, 2014; Martins, 2012).

Regarding the first theoretical difficulty, here we adopt a framework (Fitch, 2010; Martins, 2012) in which “iteration” refers to the process of repeating an operation a certain number of times. An iterative process may or may not generate hierarchical structures or create dependency relationships between different elements. For example, putting one marble at a time into a bag is an iterative process, but neither hierarchical nor recursive. In contrast, “hierarchical” structures always involve the embedding of elements within other elements. If the hierarchical embedding occurs between constituents of the same category (e.g., such as a noun phrase inside a noun phrase) we classify it as recursive, otherwise as non-recursive. Iteration, hierarchical embedding, and recursion are not mutually exclusive processes: in fact, recursion typically involves both hierarchy and iteration. Nevertheless, it is possible to segregate the cognitive abilities necessary to represent the kind of information that each of these processes encode (Fig. 2).

The second theoretical difficulty is to define the level of analysis useful for empirical enquiries. Recursion can be defined either as a “procedure that calls itself” or as the property of “constituents that contain constituents of the same kind” (Fitch, 2010; Pinker & Jackendoff, 2005). Frequently, we find an isomorphism between procedure and structure, i.e., recursive processes often generate recursive structures. However, this isomorphism does not always occur (Lobina, 2011; Luuk & Luuk, 2010; Martins, 2012). In this manuscript we explicitly focus on a third level of analysis, which is the level of representation. We focus on detecting *what* kind of information individuals can represent, rather than on *how* this information is implemented algorithmically.

Encoding iteration requires the ability to represent the repetition of a certain process, for instance the repeated addition of elements to a structure. Encoding hierarchical embedding requires the ability to represent dependency or grouping relationships between constituents at multiple levels. Encoding recursive embedding requires the ability to represent similarities across hierarchical levels (self-similarity). Specifically, that the way contiguous levels relate to each other within a hierarchy is similar across different levels. Recursion enables the generation of new hierarchical levels beyond those previously experienced, maintaining consistency with existing levels at a higher level of abstraction. It is important to retain the notion that a certain hierarchy can be represented both

¹ It has been pointed out that these hypotheses are not well formulated. In-depth discussion can be found in Fitch et al. (2005), Jackendoff and Pinker (2005), and Pinker and Jackendoff (2005).

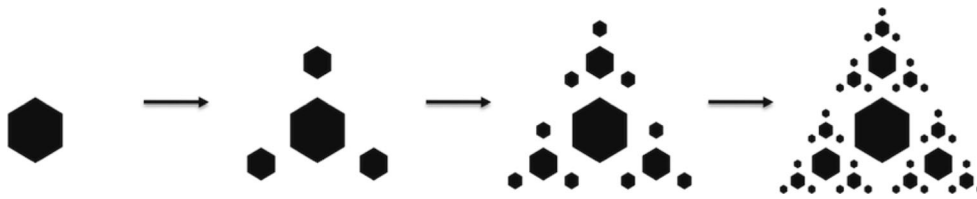


Fig. 1 Recursive process generating a visual fractal

recursively and non-recursively. For instance, in Fig. 3, a certain visual hierarchy can be generated using either process (a) or process (b). The second mode of representation is recursive, and allows the generation of an infinite number of new hierarchical levels, using one simple rule. This capacity to generalize common hierarchical principles across levels and to generate new levels beyond the given is a specific behavioral signature of recursive cognition.

Finally, although there is evidence suggesting humans can represent recursion in language, the question of whether we can represent this concept in other domains (for example, in vision) has been not been addressed empirically. This omission has been caused by a lack of methods to test for the ability to represent recursion in non-linguistic domains. Here we solve this methodological limitation by presenting a novel method that can be used to test recursion in vision. In particular, in this paper we evaluate our novel method in a variety of conditions to ensure that it taps into a specific cognitive construct (recursion) which is not completely explained by other, more general processes (such as intelligence, iterative reasoning, working memory, entropy analysis, and low-frequency spatial heuristics).

Hierarchical processing in the visual domain

The processing of hierarchies in the visual domain has been explored in the context of attention to local versus global information (Fink et al., 1996; Fitch, 2010). In particular, it is interesting that while the proper processing of hierarchies involves the integration of global and local information, there

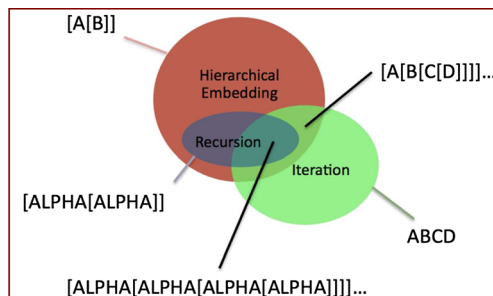


Fig. 2 Examples of structures produced by iteration, hierarchical embedding, and recursion, and by various combinations of these processes. A non-iterated hierarchical embedding corresponds to the establishment of a dependency without repetition. The ability to represent repetition and the ability to represent dependencies may be orthogonal (Martins, 2012)

are several conditions in which individuals are biased to focus on the local information only. For instance, while attending to a big square composed of small circles, young children have a tendency to identify the small circles faster and easier than they can identify the big square (Harrison & Stiles, 2009; Poirel, Mellet, Houdé, & Pineau, 2008). This local-oriented strategy to process hierarchical stimuli is similar to that seen in non-human primates (Fagot & Tomonaga, 1999; Spinozzi, De Lillo, & Truppa, 2003). Conversely, in human adults a global bias develops, in which global aspects of hierarchical structures are processed first, and where the contents of global information interfere with the processing of local information (Bouvet, Rousset, Valdois, & Donnadiou, 2011; Hopkins & Washburn, 2002). This global search strategy can be reversed if adults are asked to process novel or unfamiliar structures (Hasselmo & Stern, 2006).

Recently, research within our laboratory suggests that visual fractals might also be processed using different strategies, depending on whether recursive or non-recursive representations are primed (Martins, Fischmeister, et al., 2014; Martins, Laaha, Freiburger, Choi, & Fitch, 2014). Not only are specific neural systems active during recursive representations (Martins, Fischmeister, et al., 2014), but there also seems to be a change in visual processing strategies that correlates with ontogenetic development, and with amount of exposure to examples of fractals (Martins, Laaha, et al., 2014). How these strategies relate with local or global biases is an exciting topic of ongoing research.

Another issue of great interest here concerns the availability of representation modes that allow compression of information. More abstract and global-oriented strategies to represent visuo-spatial information seem to be more efficient because they allow the compression, or reduction, of the information required to be kept online (Alvarez, 2011). In computer science, fractal strategies have also been shown to be efficient in the representation of complex hierarchies, precisely by compressing the amount of information (Koike & Yoshihara, 1993). From this discussion sprouts the prediction that recursive modes of representation are more abstract and lead to better compression of information.

Current study

In the current study, we introduce and explore a new paradigm, focusing specifically on recursion capabilities in the

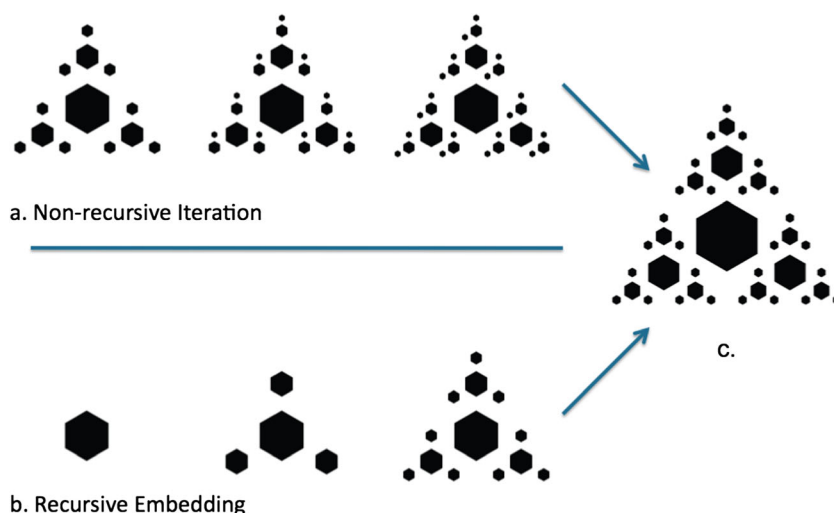


Fig. 3 Example of a hierarchy (c) that can be generated either using a non-recursive process (a) or a recursive procedure (b). While the recursive representation of hierarchy (c) allows the generation of new

hierarchical levels, the iterative representation (a) does not, being limited to within-level transformations

visual domain using fractal images. Because fractals exhibit hierarchical self-similarity, new hierarchical levels can be predicted by generalizing production rules and projecting them to further levels. Our goals are: (1) to create and validate a new task, (2) which allows us to distinguish between iterative, hierarchical, and recursive processes, (3) from which we can learn about the representation of recursion.

We present a series of experiments designed to validate empirically this novel task, forming the basis for further research.

In Experiment 1 we show that humans use recursion in the visual domain; in Experiment 2 we demonstrate that our Visual Recursion Task (VRT) taps into specific cognitive resources when contrasted with general intelligence, spatial working memory, and a control Embedded Iteration Task (EIT); in Experiment 3 we replicate the first two experiments introducing a number of important controls; and in Experiment 4 we compare our new recursive task with another task that invites recursive strategies – the Tower of Hanoi (Goel & Grafman, 1995) – confirming and expanding the evidence that VRT taps into cognitive resources specific for recursion.

Experiment 1: Response paradigm and esthetic biases

In Experiment 1 we tested whether adult humans are able to make inferences about recursive embedding in the visuo-spatial domain. This hypothesis would be supported by above-chance accuracy in our VRT.

In this task, participants are exposed to the first three steps of a process generating a visual fractal, and then asked to discriminate, from two possible alternatives, which is the correct continuation (see details below).

Since we were interested in exploring how participants would approach visual recursion, we gave minimal instructions and did not restrict response time. We assessed the strategies that participants reported after completing the task, and tested whether certain cognitive strategies led to better performance. We also evaluated the effects of the particular response paradigm (binary forced-choice) and subjective esthetic preferences on individuals' accuracy by (1) adding an additional response task (1-alternative forced-choice – correct/incorrect), and (2) testing whether an esthetic preference for self-similar fractals could account for participants' choices, regardless of their ability to represent recursion. If participants were using a simple strategy of esthetic preference towards well-formed fractals, this would argue against our assumption that a cognitive strategy was employed rather than simple visual heuristics.

Methods

Participants

We tested 20 volunteers (undergraduates and PhD students; 14 females and six males) aged between 20 and 44 years ($M = 28.1$, $SD = 6$) recruited at the University of Vienna. All participants were tested using the same experimental apparatus, and all reported normal or corrected-to-normal visual acuity. All participants gave their prior written consent, and were not paid for taking part. The research conformed to institutional guidelines and Austrian national legislation regarding ethics.

Stimuli and procedure

Stimulus generation We based the VRT on the well-established properties of fractal geometry (Mandelbrot,

1977). Visual fractals can be generated from single constituents such as lines, squares, or triangles (*the initiators*) by applying a simple transformation rule (*the generator*) a given number of times (iterations). The structures generated by iterating this process are hierarchical and self-similar (see Fig. 4 for a schematic overview).

We produced four successive iterations of 60 different types of fractals, generated using Python code running in Nodebox (version 1.9.5, <http://nodebox.net>), a visual interface. For each of these 60 fractals, we produced (1) a correct fourth continuation of the first three iterative steps, and (2) an incorrect continuation as a Foil. This incorrect fourth iteration was produced by applying a different *generator* to the third stage, and had the same number and size of constituents as the correct fourth iteration.

The fractals produced for this task can be divided into four broad categories (see Fig. 5 for examples): (1) Polygons ($n = 32$), (2) trees ($n = 9$), (3) curves ($n = 11$), and (4) Koch snowflakes ($n = 8$). Peano curves and Koch snowflakes were produced using Lindenmayer systems (Lindenmayer, 1968). In these systems, the recursive process substitutes each constituent with a set of new constituents without preserving the *initiator* across iterations. The other two categories of fractals were produced with custom Nodebox scripts.

Visual Recursion Task (VRT) 2-choice The three iterations and two test images were arranged on a panel (Fig. 6). Each

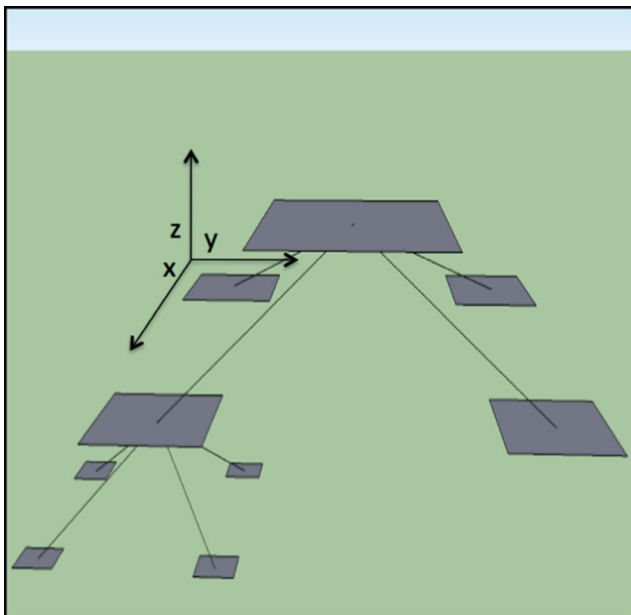


Fig. 4 Visual fractals can be conceived as visuo-spatial hierarchies: Different elements (squares) are organized in a two-dimensional space, defined by the xy -axis, and different hierarchical levels are organized vertically along the z -axis. An element with a higher z value is dominant over an element with a lower z value, if the elements are connected

panel depicted five images, presented simultaneously, arranged in two rows: The first three iterations of each fractal (“sequence” images) were shown in the top row and two alternatives for the fourth iteration (“correct” vs. “incorrect” fourth iteration, henceforth “choice” images) were shown in the bottom row. The position of the choice images (left or right) was randomized. The sequence of panels was presented on a computer screen in a randomized order, which was different for each participant, using custom Python software (version 2.6, www.python.org).

Participants were instructed in English to select the image they considered correct from the two “choice” images in the bottom row and to “try to understand the right strategy and to choose correctly as often as you can.” No further explanation on what “correct” meant was provided.

Participants responded by pressing one of two buttons on a button box (ioLab Systems), corresponding to the position of the correct image (left or right). Auditory and visual feedback was given for all trials. After an incorrect choice, the screen turned red for 1.5 s and a negative feedback sound (frequency 98.0 Hz and duration 1.5 s) was played. After a correct choice, the screen turned white for 1 s and a positive feedback sound (frequency 348.7 Hz, duration 1 s) was played. The sounds were played through Sennheiser HC 520 headphones. There was a 2-s inter-trial interval. There was no time limit per trial (timeout) because we did not want to constrain participants’ strategies, and because we were interested in knowing how they would naturally approach the tasks when given minimal instructions.

Before the VRT began, participants were given a short training session of five trials. The training stimuli were similar to the VRT panels, except that the sequence of images was generated according to a simple non-hierarchical iterative rule (see Fig. 7).

Visual Recursion Task (VRT) 1-choice In order to evaluate possible performance effects associated with a binary forced choice paradigm, we designed a *VRT 1-choice* task. This task was identical in all aspects to the basic *VRT 2-choice*, except that only one image was presented in the center of the second row of each panel, corresponding to either the correct or incorrect fourth iteration (Fig. 8). Participants were instructed to choose whether the image in the lower row was correct (right button) or incorrect (left button). The same number ($n = 10$) of correct and incorrect fourth iterations was presented.

Before the beginning of the task, the same five training stimuli were presented as in *VRT 2-choice*, but with only one “choice” image. Feedback and inter-stimuli intervals were the same as in the *VRT 2-choice* task.

Esthetic preference task This task was designed to assess the effects of possible preference biases in *VRT 2-choice*. Here, only the “choice” images (“correct” and “incorrect” fourth iteration) were presented on the screen (Fig. 9) with no previous

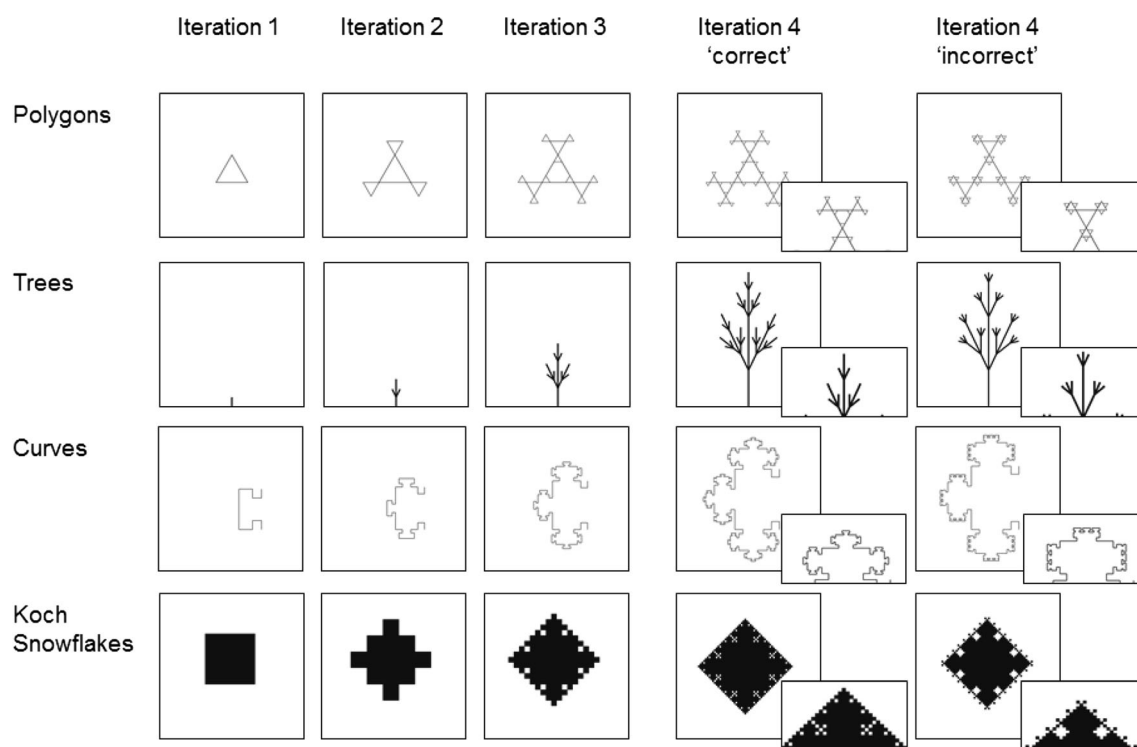


Fig. 5 Fractal categories and iterations: For each fractal, we generated the first four iterations and an incorrect fourth iteration. The latter violated the embedding rule used in the previous steps (small boxes contain

zoomed-in details). The fractals were grouped in four classes according to the generating algorithm: Polygons ($n=32$), trees ($n=9$), curves ($n=11$) and Koch snowflakes ($n=8$)

“sequence” images. Participants were asked to simply select the image they preferred. No auditory or visual feedback was given.

Procedure All participants began the experiment with the *preference task*. Participants then performed both recursion tasks in one of two possible orders: ten participants completed *VRT 1-choice* before *VRT 2-choice* (“1–2” condition), and ten

participants performed *VRT 2-choice* before *VRT 1-choice* (“2–1” condition). Participants were randomly assigned to one of the two orders.

The same pool of 60 fractals was used in all tasks, with 20 fractals randomly assigned to each of the three tasks. The distribution of fractal classes was balanced for all tasks and each fractal appeared only once in each experimental session.

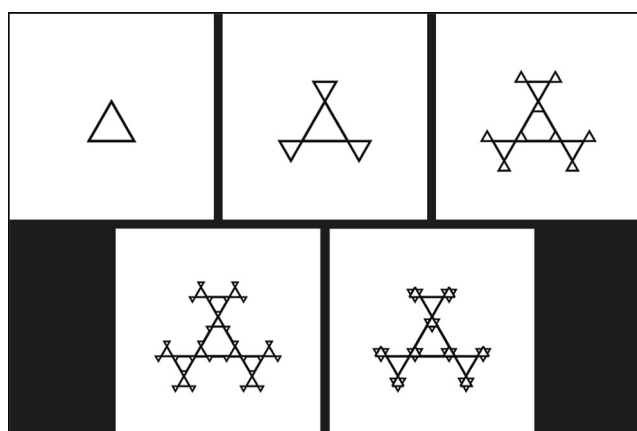


Fig. 6 Example of a two-choice stimulus in the Visual Recursion Task (VRT). The first three iterations were presented in the top row. Participants had to choose which of the images in the lower row was correct. In this example, the correct image is on the left. Further stimuli examples can be found in the [Supplemental materials](#), part I, section S2

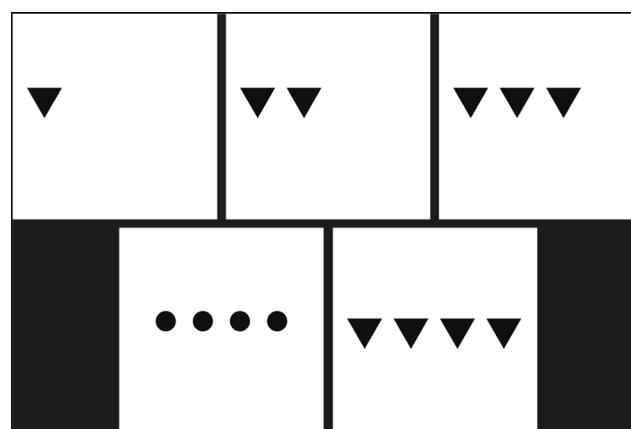


Fig. 7 Example of a training stimulus: The iterations followed a number and shape rule but did not produce hierarchical structures. The correct image is on the right. Further stimuli examples can be found in the [Supplemental materials](#), part I, section S1

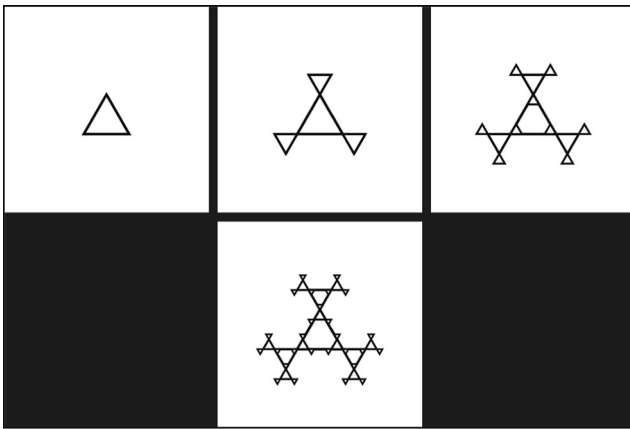


Fig. 8 Example of a VRT “1-choice” stimulus. Participants had to decide whether the image presented in the bottom row was correct or incorrect. In this example, the image is correct

Participants’ choices and reaction times (RTs; in milliseconds) were recorded for all stimuli and for all tasks. The performance was calculated as the percentage of correct answers. In the *preference task*, we recorded as “correct” answers the occurrences where the preferred image corresponded to the well formed fractal, i.e. to the correct fourth iteration. At the end of each task, participants were asked to assess the kind of strategy they had used on a five-point scale. The scale of possible strategies was: 1 – “mostly intuitive”; 2 – “more intuitive than analytic”; 3 – “mixed”; 4 – “more analytic than intuitive”; 5 – “mostly analytic.” Intuitive answers were described to the participant as being based on a gut feeling and analytic answers as being derived by looking carefully at the details and making explicit inferences.

Analysis

The proportion of correct responses and RTs were compared between (1) *VRT 2-choice* and *VRT 1-choice* and (2) *VRT 2-choice* and *preference task*. We used a semiparametric regression technique called Generalized Estimating Equations

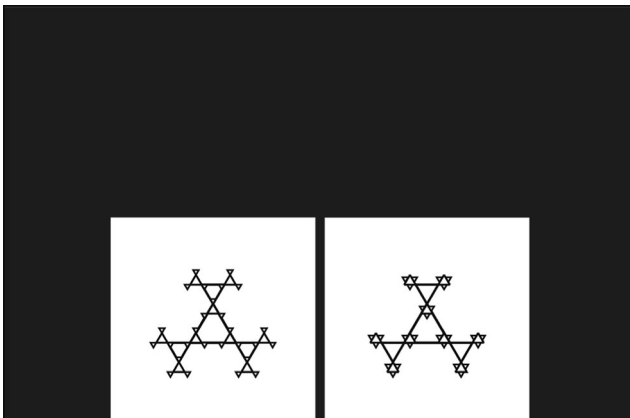


Fig. 9 Example of stimuli used in the preference task. Participants were asked to choose the image they preferred

(GEE), a technique useful when analyzing binomial data with within-subjects effects (Hanley, 2003). When applied to binary data, this technique is similar to a logistic regression and in comparison with generalized mixed models is more robust to deviations from error distribution assumptions, and model misspecifications (Ghisletta & Spini, 2004; Hubbard et al., 2010). We also used this model to assess accuracy differences between stimuli categories, and RT differences between tasks (using gamma with a log link function). To assess whether performance was above chance at the group level, for each task, we tested whether GEE models’ intercepts were significantly different from zero.

Furthermore, we assessed performance correlations between these tasks. For percentages of correct responses and RTs we tested if the data were normally distributed using the Kolmogorov-Smirnov (K-S) test. If variables were continuous and normally distributed we used Pearson’s bivariate correlations, otherwise we used non-parametric Spearman correlations.

All statistical analyses were performed using SPSS 19 (IBM).

Results

Performance

On average, participants scored 84 % ($SD = 12$) correct in *VRT 2-choice* and 70 % ($SD = 14$) correct in *VRT 1-choice* (Fig. 10). In the *preference task*, the “correct” image was

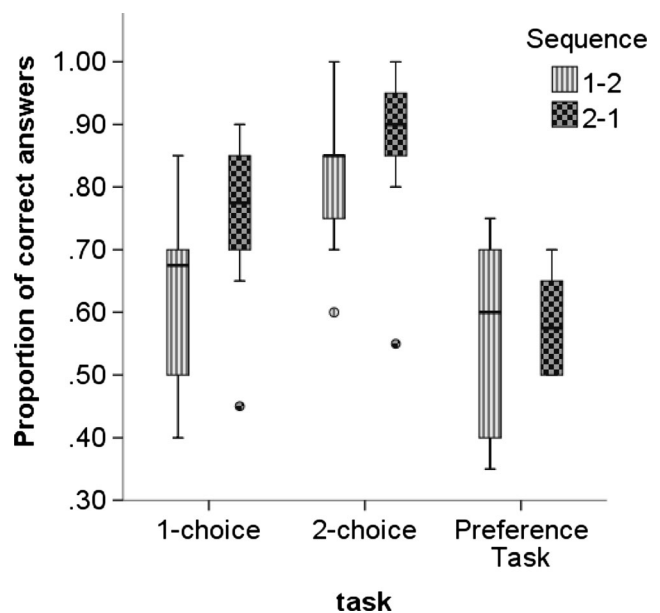


Fig. 10 Percentage of correct responses in *VRT 1-choice*, *VRT 2-choice*, and *preference task*, in two different task-sequence conditions: “1-2” and “2-1.” Bar charts show median (horizontal line), first quartile (lower edge of the box), and third quartile (upper edge of the box). ° indicate outliers deviating from the box between 1.5 and 3 times the interquartile range; * indicate outliers deviating from the box more than 3 times the interquartile range

preferred in 58 % ($SD = 11$) of the trials. To assess whether average response was above chance, we ran a GEE model for each task, with “trial” (1–20) as the within-subjects variable. All intercepts differed significantly from zero (all $p < .05$), meaning accuracy was above chance in all tasks, at the group level. To assess whether there were differences between tasks, while controlling for task order, we ran a binary logistic GEE model. We found a significant effect of task (generalized chi-square = 16.5, $p < .001$), but no effect of task order ($p = .15$) and no interaction between the two factors ($p = .2$). Pairwise comparisons with a Bonferroni p-value adjustment showed that performance was significantly lower in *VRT 1-choice* than in *VRT 2-choice* ($p < .001$, odds ratio = 0.8); and higher in *VRT 2-choice* than in the *preference task* ($p < .001$, odds ratio = 0.7).

Analyzed by participant, the percentage of correct responses in *VRT 2-choice* was correlated with performance in *VRT 1-choice* ($r = .57$, $p = .009$), but not with the *preference task* ($r = .27$, $p = .24$). This correlation between *VRT 1-choice* and *VRT 2-choice* was significant in the group of participants that started the procedure with *VRT 2-choice* ($n = 10$; $r = .797$, $p = .006$), but not in the group that started with *VRT 1-choice* ($n = 10$; $r = .260$, $p = .469$).

Reaction time

On average, RT was 12.5 s ($SD = 1$) in *VRT 1-choice*, 12.2 s ($SD = 7$) in *VRT 2-choice*, and 5.3 s ($SD = 3$) in the *preference task* (Fig. 11). To assess whether there were differences

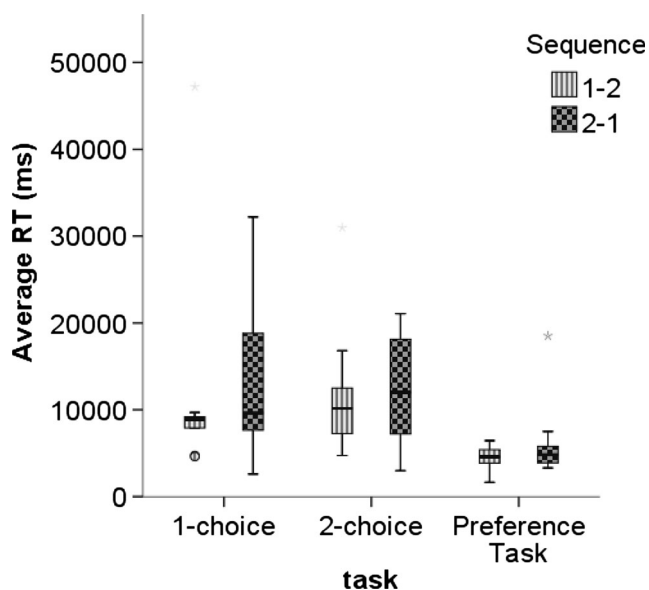


Fig. 11 Average reaction time in *VRT 1-choice*, *VRT 2-choice* and *preference task*, in two different task-sequence conditions: “1-2” and “2-1.” Bar charts show median (horizontal line), first quartile (lower edge), and third quartile (upper edge). ° indicate outliers deviating from the box between 1.5 and 3 times the interquartile range; * indicate outliers deviating from the box more than 3 times the interquartile range

between tasks, while controlling for task order, we ran a gamma log link GEE model. There was an effect of task (generalized chi-square score = 11.4, $p = .003$), but not of task-order ($p = .5$), and no interaction between the two factors ($p = .7$). Specifically, we found a difference between *VRT 2-choice* and *preference task* (mean difference = 7 s, $p < .001$) but not between *VRT 2-choice* and *VRT 1-choice* ($p = .8$).

Strategy

At the end of each task, we asked our participants about the strategy they used. In general, participants reported a more intuitive strategy for the *preference task* ($M = 2.45$, $SD = .9$) and a more analytic strategy in both *VRT 1-choice* ($M = 4.2$, $SD = .8$) and *VRT 2-choice* ($M = 4.0$, $SD = 1.2$). Interestingly, participants who reported a more analytic strategy in *VRT 2-choice* also had longer RTs (Spearman’s $\rho = .485$, $p = .03$) and a higher percentage of correct answers (Spearman’s $\rho = .585$, $p = .007$) than those participants who reported intuitive strategies. This suggests that an analytic rather than an intuitive strategy was optimal for the VRT.

Esthetic preferences

Another important issue was whether the decision between the choice images in the *2-choice* condition was influenced by esthetic preferences. Given that the same 120 images were part of the pool of possible choices in *VRT 2-choice* and *preference task*, we assessed the frequency with which each image was chosen in both tasks (i.e., for each image, we counted the number of times it was chosen in *VRT 2-choice* and *preference task*). We found that these frequencies were not correlated ($r = .027$; $p = .838$), meaning that the images chosen more frequently in *VRT 2-choice* were not the images more frequently chosen in the *preference task*, suggesting that esthetic preferences could not account for above-chance performance in the recursion task.

Discussion

Our results suggest that human adults can quickly learn how to use recursive information in the visual domain without being explicitly trained or instructed about the concept of recursion. Moreover, a self-reported analytic strategy was associated with higher RTs, and significantly correlated with better performance. Although response feedback was provided during the task, participants were required to respond to a wide variety of stimuli, with different visual and structural features. Structural recursion was the common element among these stimuli and most likely this abstract regularity was transferred across trials. We propose that the ability to represent structural self-similarity in the visual domain was a necessary condition

for good performance in this experiment, regardless of how this information was represented.

Given that *VRT* performance could be influenced by the response paradigm used as well as by esthetic biases in favor of (or against) self-similar fractals, we included three tasks: two recursive tasks (*VRT 2-choice* and *VRT 1-choice*) and a *preference task*. Our findings rule out an effect of esthetic preferences on performance in *VRT*, suggesting that subjects do not use preferences as decision heuristics, and demonstrate that both versions of the recursive task were similar to each other: (1) Percentages of correct responses in *VRT 2-choice* and *VRT 1-choice* were correlated. (2) RTs and self-reported strategy were similar in these tasks but differed significantly from the *preference task*. (3) Images preferred in the *preference task* were not the images more frequently chosen as “correct” in the *VRT 2-choice* condition.

However, there was a significant performance difference between *VRT 1-choice* and *VRT 2-choice*, depending on task order: Performance in the two tasks only correlated when *VRT 1-choice* was performed after *VRT 2-choice* (reaching a correlation coefficient as high as 0.8). It seems that when *VRT 2-choice* was performed first in the presence of correct and incorrect information, participants learned to attend more closely to the relevant image details, thereby increasing their accuracy in *VRT 1-choice* afterwards. This might imply that the ability to process recursion is influenced by the ability to orient attention to the relevant features of the stimuli, and that poor performance in such a task is not necessarily due to an inability to process recursion, but may arise from incorrectly focussed visual attention. This interpretation is consistent with findings in developmental data, in which young children fail in recursion due to inefficient visual strategies (Martins, Laaha, et al., 2014). Crucially, after being primed to attend to the relevant features of the stimuli, participants were well able to perform in *VRT 1-choice* (mean accuracy 76 %), showing that the comparison between two choice images was not strictly necessary to discriminate between correct and incorrect continuations of the recursive process. This argues against a heuristic response strategy purely based on the comparison between choice images.

Experiment 2: Recursive versus non-recursive iteration

Experiment 1 suggested that human adults are able to represent visual recursion successfully. However, it remains an open question whether the *VRT* measures something specific to recursion, or instead taps into a more general ability to extract visual regularities. In Experiment 2, we attempted to gain more specific insight into the cognitive processes underlying *VRT*. We devised an Embedded Iteration Task (*EIT*) as a control task, which shared the “hierarchicality” and iteration features of *VRT*, but lacked recursive embedding. We compared

participants’ accuracy in both *VRT* and *EIT* with a standardized measure of rule-based visual cognition (Matrix Reasoning from WASI[®], see below). Here we wanted to test whether visual recursion, as measured by our task, can be dissociated from other visuo-spatial hierarchical tasks (*EIT*) and general visual intelligence capacity (*WASI*). For this purpose we used correlation and regression analyses. Although exploratory, we tested the general hypothesis that *VRT* would not be highly correlated with general intelligence, and that *VRT* and *EIT* would correlate with different cognitive abilities. These findings would provide support for the existence of variance within the performance of *VRT* that is explained by specific resources recruited in the instantiation of recursive representations.

To produce *EIT* images, an iterative process embedded additional elements within a pre-existing hierarchical structure, without producing new hierarchical levels (Fig. 12). To empirically validate the distinction between recursion and iteration we first assessed the behavioral response profile for both tasks. Furthermore, we tested whether different cognitive abilities (fluid intelligence and working memory) predicted accuracy in solving the two tasks.

Methods

Participants

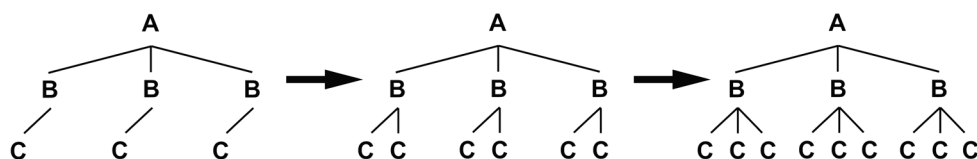
We tested 30 volunteers (university undergraduates and employees; 21 females) aged between 18 and 39 years ($M = 23.6$, $SD = 5$) recruited at the Lisbon Faculty of Medicine. Education ranged between 11 and 20 years of successfully completed studies ($M = 15.6$, $SD = 2$). All participants were tested in the same room, with the same experimental apparatus as Experiment 1, and all reported normal or corrected-to-normal visual acuity. Participants were paid 10 Euros for participating and gave their written informed consent. The research conformed to the appropriate institutional and national legislation regarding ethics.

Stimuli and procedure

VRT Stimulus generation and experimental procedure were similar to *VRT 2-choice* described in Experiment 1. In this experiment only 40 test panels were presented to each participant (13 “polygons,” seven “trees,” 11 “curves,” and nine “Koch snowflakes”).

Embedded Iteration Task (EIT) *EIT* is a control task in which hierarchical structures are generated using simple iterative transformation rules that add elements within certain hierarchical levels, but without generating new levels (process A, Fig. 12). This is in contrast to *VRT*, in which new hierarchical levels are generated at each step of the transformation process. Crucially, both processes generate similar structures (Fig. 3).

A Iterative rule: Add another C to existing level under B.



B Recursive embedding rule: Add three As to new level under of A.

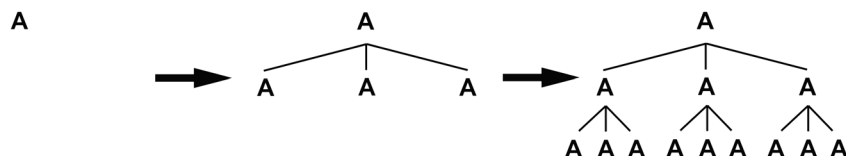


Fig. 12 Principles underlying the generation of hierarchies in the Visual Recursion Task (VRT) and Embedded Iteration Task (EIT). (A) In EIT we used an iterative rule that adds elements to a previous existing hierarchy,

without generating new levels; (B) in VRT we used a recursive rule that adds elements to newly generated hierarchical levels

EIT and *VRT* stimuli were hierarchical structures generated by Python scripts in Nodebox and were very similar to *VRT* fractals. Each *VRT* item was modified to generate a corresponding *EIT* item, with a precise one-to-one correspondence in size, structure, and element identity. In *VRT*, each iteration produced a new hierarchical level, while in *EIT* the first image was already a hierarchical structure and each iterative step merely added one additional item within a chosen hierarchical level, without generating a new level (see Fig. 13). Crucially, both *VRT* and *EIT* generated hierarchies of the same number of elements and the same number of hierarchical levels.

To control for the use of a simple similarity assessment strategy in *EIT*, we included ten stimuli (“repetition foils”) requiring participants to represent the cumulative addition of constituents (Fig. 13). In this subset of stimuli, one of the choice images was a simple repetition of the third iteration; there was no increase in the number of constituents from third to fourth iteration, hence this was the incorrect choice. In the remaining 30 stimuli, we used “positional foils” in which the

possible choices contained the same number of elements but differed in their overall positional scheme (Fig. 13). These 40 panels were intermixed. With these two conditions, we aimed to evaluate whether participants were able to detect both the iterative and positional properties of the hierarchical stimuli. More examples of *EIT* items are available in the [Supplemental materials](#), part I, section S3.

Cognitive assessment All participants also performed a battery of standardized cognitive tasks. Verbal short-term memory and working memory were assessed using *Digit Span* (Richardson, 2007). Spatial short-term memory and working memory were assessed with two sub-tests of *CANTABEclipse Spatial Span* (Owen, Morris, Sahakian, Polkey, & Robbin, 1996): (1) “forward” (the number of items successfully repeated in the same order as the example) and (2) “backwards” (the number of items successfully repeated in the reverse order). Finally, we used un-standardized scores (number of items answered correctly) in two sub-tests of the *WASI* (Wechsler,

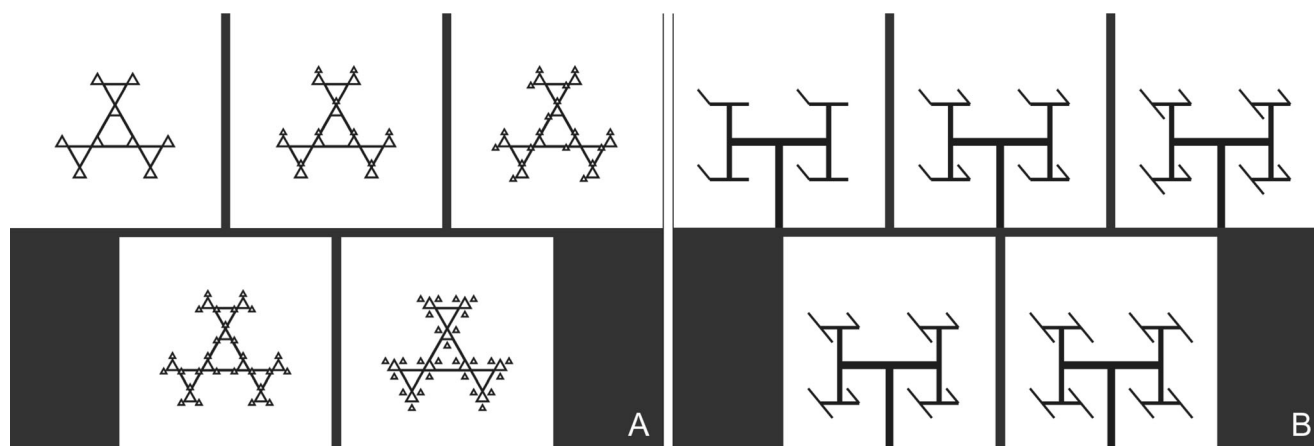


Fig. 13 Examples of Embedded Iteration Task (EIT) stimuli. (A): “Positional” category (n=30), correct image is on the left. (B): “Repetition” category (n=10), correct image is on the right. See [Supplemental material](#), part I, section S3 for further examples

1999) test battery – “vocabulary” and “matrix reasoning” – as proxies for crystallized and fluid intelligence.

Procedure The procedure took about 90 min in total. All instructions were given in Portuguese. *VRT* and *EIT* were randomly assigned either to the beginning or end of the procedure and the cognitive assessment was conducted between the two tasks. Within *VRT* and *EIT*, trial order was differently randomized for each participant. Feedback was provided as in Experiment 1, and there was no timeout limit.

Analysis

To test for accuracy and RT differences between *VRT* and *EIT* we used the same statistical techniques as in Experiment 1.

We performed correlation analyses to assess whether performance in *VRT* and *EIT* provided non-redundant information relative to standardized measures of intelligence and working memory.

Furthermore, to probe for cognitive-specific differences between *VRT* and *EIT*, we performed partial correlations, and a Principle component analysis (see [Supplementary materials](#), part II).

All statistical analyses were performed using SPSS 19 (IBM).

Results

VRT and *EIT* performance

Raw scores (clustered by participant) are depicted in Fig. 14. Overall, participants performed above chance in both tasks (both GEE model intercepts differed from zero, $p < .001$), but accuracy was significantly higher in *EIT* ($M = 92\%$, $SD = 8$) than in *VRT* ($M = 84\%$, $SD = 7$) (generalized chi-square score = 14.5, $p < .001$, odds ratio = 0.5).

In *EIT*, participants performed above chance in both foil categories (both GEE model intercepts differed from zero, $p < .001$), but performed worse in trials with “Repetition” ($M = 87\%$, $SD = 17$) than with “Positional” foils ($M = 93\%$, $SD = 7$) (generalized chi-square score = 4.2, $p = .04$, odds ratio = 0.5).

Mean RT was longer in *VRT* ($M = 22.2$ s, $SD = 12$) than in *EIT* ($M = 18.4$ s, $SD = 7$) (generalized chi-square score = 6.1, $p = .013$, odds ratio = 0.8). For correlation purposes, we also applied an arcsin-transformation ($Y = \text{asin}(\sqrt{X})$) to the accuracy data of *VRT* and *EIT*, and achieved normality in *VRT* ($K-S$, $p > .05$). After excluding an extreme outlier (with performance deviating from the mean more than 2 standard deviations (SDs)) we also achieved normality in *EIT*. Further analyses excluded this outlier. Performance on both tasks correlated across participants in accuracy ($r = .506$, $p = .005$) and RT ($r = .781$, $p < .001$). Similar to Experiment 1, participants

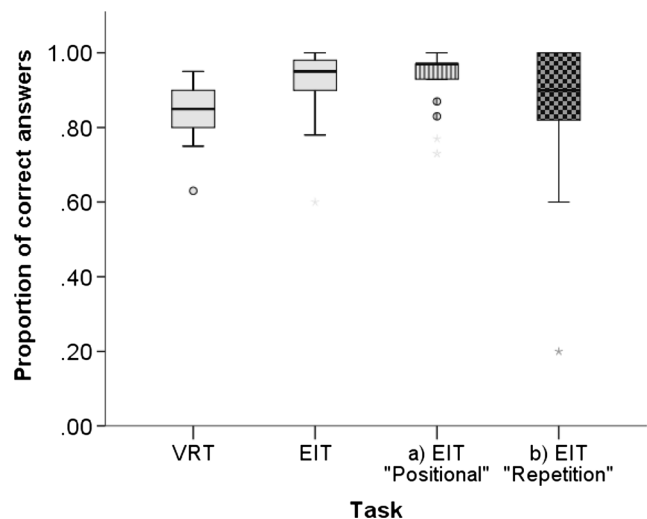


Fig. 14 Accuracy across tasks: *VRT* (Visual Recursion Task), *EIT* (Embedded Iteration Task), and in the two sub-tasks (**a**, **b**) of *EIT*. The boxplot divides the scores into quartiles, the “box” represents the distance from the 25th percentile to the 75th percentile and is called the interquartile range. The horizontal dark line is the median. ° indicate outliers deviating from the box between 1.5 and 3 times the interquartile range; * indicate outliers deviating from the box more than 3 times the interquartile range

with longer RTs performed better in *VRT* ($r = .520$, $p = .004$) but also in *EIT* ($r = .388$, $p = .04$).

Finally, to investigate whether there was an effect of learning, we assessed whether RT decreased with the accumulation of trials. We performed a power curve fitting regression analysis, with RT as dependent variable and trial as predictor. This analysis was significant for both *VRT* ($F(1,39) = 6.8$, $r = 0.39$, $p = .013$) and *EIT* ($F(1,39) = 31.0$, $r = 0.67$, $p < .001$), meaning that RT decrease across trial fitted a power curve, suggestive of a learning effect (Anderson, 1982). This learning precludes that participants used idiosyncratic strategies to solve each trial, and rather suggests that a common strategy was used across trials.

Stimulus categories analysis

In this experiment, we performed accuracy analyses for all four stimuli categories (Polygons, Trees, Curves, and Koch Snowflakes) (see Fig. 4, and [Supplemental materials](#), part I). Within each stimulus category, accuracy scores were above chance (all intercepts significantly differed from zero, $p < .001$) (Fig. 15), for both *VRT* and *EIT*. To assess whether there were differences in performance between categories, we ran a GEE model with correctness (correct/incorrect) as the dependent variable, and task (*VRT* vs. *EIT*) and “stimulus category” as the within-subjects factors. We found a significant interaction between “task” and “stimulus category” (generalized chi-square score = 20.7, $p < .001$). Specifically, while there were no significant differences between “stimulus categories” in *EIT* (all pairwise comparisons $p > .05$, with Bonferroni

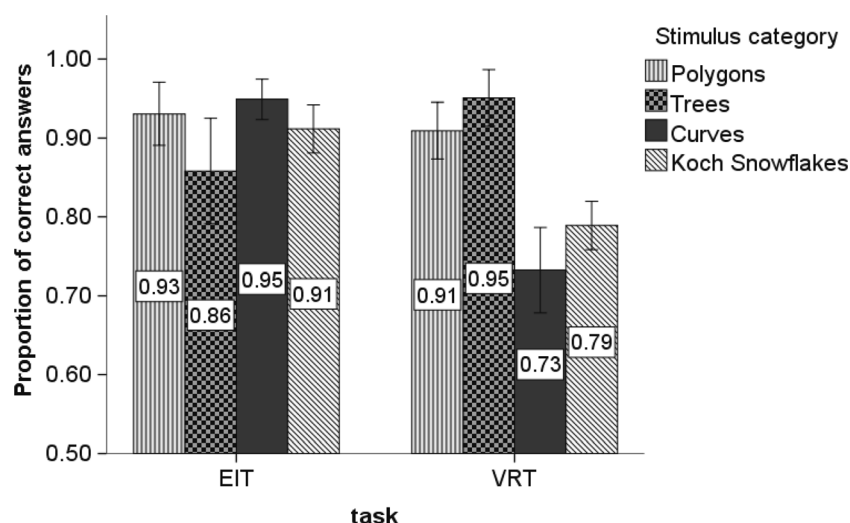


Fig. 15 Percentage of correct responses in Visual Recursion Task (VRT) and Embedded Iteration Task (EIT) across different stimulus categories

correction), in *VRT* participants performed better with “polygons” and “trees” than “curves” and “Koch snowflakes” (Pairwise comparisons $p < .001$, with Bonferroni correction).

Correlations with fluid intelligence and working memory

Fluid intelligence and crystallized capacity are two classical constructs which compose measures of general intelligence. While fluid intelligence is usually equated with the ability to solve new problems, crystallized capacity refers to semantic knowledge. In order to assess whether *VRT* and *EIT* scores were redundant relative to these measures of intelligence, we compared them to participants’ performance in a *Matrix Reasoning* task (*MR*) and in a *Vocabulary* task. Raw results are depicted in Table 1. Overall Pearson correlations are depicted in Table 2. After p-value correction with the Bonferroni-Holm method (with FWE level = .05), we found a significant correlation between *EIT* and *MR* ($r = 0.51$, $p = .03$), and between *VRT* and *MR* ($r = .49$, $p = .04$). One participant had an *MR* score that was two standard deviations below the mean. When this outlier was excluded from the analysis, these correlations became non-significant (*MR* and *EIT* ($r = 0.35$, $p = .06$, uncorrected), and *MR* and *VRT* ($r = 0.29$,

$p = .14$, uncorrected)). This suggests these correlations are not stable, and might be driven by idiosyncratic participants. The score in the “vocabulary” task (proxy for crystallized intelligence) was not correlated with *VRT* or *EIT* ($p > .1$).

We also wanted to assess to what extent the capacity for processing verbal and visual information influences *VRT* and *EIT* accuracy. Therefore, we assessed our participants’ short-term and working memory abilities, in both the visuo-spatial and verbal domains. Due to technical problems there were eight missing values in Spatial working memory. Raw scores are depicted in Table 1 and overall correlations in Table 2.

After p-value correction with the Bonferroni-Holm method (with FWE level = .05), there were significant correlations between *EIT* and *spatial* working memory ($r(22) = .54$, $p = .045$). *VRT* performance did not correlate significantly with performance in either memory task.

VRT versus *EIT*: Cognitive resources

We performed partial correlation analyses to assess whether different cognitive resources predicted performance in *VRT* and *EIT*. After controlling for the overall variance explained by *VRT*, *EIT* remained significantly correlated with *spatial* working memory ($r(19) = .45$, $p = .041$). This suggests that *EIT* performance may require the activation of specific visuo-spatial resources to a greater extent than *VRT*. The inverse analysis (correlations with *VRT*, controlling for *EIT*) yielded no significant correlations. We also performed a Principal component analysis ($KMO = 6.33$, Bartlett’s test of sphericity $\chi^2 = 49.1$, $p < .001$) (see [Supplemental materials](#), part II for details) to assess the correlational structure of our data. This analysis clearly divides the tasks into three big clusters: Cluster 1 includes *MR*, *spatial working memory*, *spatial short-term memory*, and *EIT*; Cluster 2 includes short-term memory tasks and spatial working memory; and Cluster 3

Table 1 Summary of results in the standardized cognitive tasks

	N	Minimum	Maximum	Mean	SD
Matrix Reasoning	30	22	35	30.67	3.06
Vocabulary	30	59	79	71.43	5.44
Verbal STM	30	4	9	6.70	1.18
Verbal WM	30	3	8	5.43	1.52
Spatial STM	30	3	9	7.10	1.52
Spatial WM	22	5	9	7.09	1.34

STM short-term memory, WM working memory, SD standard deviation

Table 2 Correlations between standardized cognitive tasks, Visual Recursion Task (VRT) and Embedded Iteration Task (EIT)

	1. VRT	2. EIT	3. VSTM	4. VWM	5. Vocab.	6. MR	7. SSTM
1. VRT							
2. EIT	0.506**						
3. Verbal STM (VSTM)	-0.225	-0.149					
4. Verbal WM (VWM)	0.357	0.425*	0.405*				
5. Vocabulary (Vocab.)	0.09	-0.103	-0.109	0.114			
6. Matrix Reasoning (MR)	0.486**	0.508**	-0.310	0.331	-0.036		
7. Spatial STM (SSTM)	0.250	0.274	0.318	0.571**	-0.016	0.360	
8. Spatial WM (SWM)	0.334	0.541**	-0.105	0.504*	-0.271	0.670**	0.588**

VRT Visual Recursion Task, *EIT* Embedded Iteration Task, *STM* short-term memory, *WM* working memory, *SD* standard deviation

* $p < .05$

** $p < .01$, for uncorrected p-values

includes *verbal working memory*, *EIT*, and *VRT*. This confirms that *VRT*, in comparison with *EIT*, is much less dependent on specific visuo-spatial resources, even though it is a visuo-spatial task. This finding also argues against the use of simple visual heuristics to solve *VRT*.

Discussion

Experiment 2 compared the processing of recursively and iteratively generated items, and sought possible correlations with other standard psychometric measures. We found that performance in *VRT* diverged from non-recursive iterative embedding and from a standardized (visual) fluid intelligence task. These results suggest that performing *VRT* activates specific cognitive resources, and that this task does not simply measure the general ability to perform rule-based visual tasks. Moreover, our results suggest that visual recursion correlates less with visual-specific resources than with embedded iteration.

First, with the “repetition foils,” we were able to show that a simple visual heuristic strategy based on visual similarity is not sufficient for solving *EIT*. Our results demonstrate that most participants understood the iterative rules displayed in the stimuli, and thereby were able to choose the correct continuation of those rules. Crucially, they did so even when the correct continuation of the iterative process (fourth iteration) was not the response choice most similar to the third iteration. Even though rejecting these “repetition foils” was not trivial, as hinted by a lower accuracy score in comparison with the “positional” foils, the participant’s accuracy was still far above chance in these trials ($M = 87\%$).

Second, regarding the correlations with standardized cognitive measures, only a portion of *VRT* and *EIT* variance could be predicted by matrix reasoning and working memory performance. This suggests that our new tasks tap (at least

partially) into distinct cognitive abilities. Matrix reasoning seemed to be a mild predictor of *VRT* (24 %) and *EIT* (26 %), but excluding a single outlier participant eliminated these correlations.

Third, we found that spatial working memory was a better predictor of *EIT* than *VRT*, and that *EIT* was more closely associated with tasks tapping into specific visuo-spatial resources (spatial short-term memory, spatial working memory, and matrix reasoning). If *EIT* loads higher in processing capacities associated with the visual domain, *VRT* might crucially depend on other general capacities, such as the ability to use chunking and segmentation (Halford, Wilson, & Phillips, 1998), or greater representational generality/abstraction (Alvarez, 2011), which have been shown to reduce the cognitive demands of hierarchical processing. For instance, the Tower of Hanoi is a task that requires the ability to build goal hierarchies with several levels of embedding. The number of variables that humans can process simultaneously is four, but the correct execution of Tower of Hanoi requires hierarchies with a greater number of embedded steps (Halford et al., 1998). One possible strategy to make Tower of Hanoi tractable is to segment the main goal into smaller subordinate sub-goals, and to build a general/abstract recursive representation that can be used across all hierarchical levels. These strategies clearly reduce relational processing demands, using cognitive resources unrelated to working memory storage.

Finally, while participants performed above chance in all stimuli categories, there were significant differences between categories. In *VRT*, accuracy in “polygons” and “trees” was higher than in “curves” and “Koch snowflakes”. One possible explanation for this difference may be due to the fact that for “polygons” and “trees,” the visual information from a certain iteration n remains present in the iteration $n+1$. For example, in a “tree” fractal, an iteration $n+1$ contains all the branches of the previous iteration n plus additional new branches (See

Supplemental materials, part I, S2). In a typical “curve” fractal, the whole visual contour is transformed from one iteration to the next, because every segment of the curve is transformed according to the recursive rule. Thus, while the structural “core” is preserved from one iteration to the other in “polygons” and “trees” (analogous to “Mother’s bike” → “John’s mother’s bike”), in “curves” and “Koch snowflakes” the “core” constituents of a certain iteration are separated in space in the next iteration (analogous to “The driver drinks” → “The driver that the mother loved drinks”). The fact that participants scored above chance in all stimulus categories of a task where all stimuli were intermixed and feedback was provided (facilitating learning from one trial to the next), suggests that differences in performance may be due to the differing visual processing demands of the tasks, rather than differences in the participants’ understanding of recursive embedding per se. As an alternative, these differences in performance might have been caused by the fact that there were differences in the number of items in each category, or that items from the “Koch snowflakes” and “curves” categories contained information of a higher spatial resolution, making the relevant details potentially harder to see.

Experiment 3: Effects of response feedback and stimulus categories

In Experiments 1 and 2 we investigated whether human adults were able to solve a task that required them to form representations of visual recursion. We provided response feedback in both experiments. It could be argued that this training experience, giving response feedback, allowed participants to develop alternative heuristic strategies by trial-and-error, thus avoiding the need to represent hierarchical self-similarity (e.g., participants might base their choice on which image is more similar to the most recent iteration). To test for these effects we assessed performance in *VRT* and *EIT*, in a procedure without response feedback. Furthermore, here we also included repetition foils in *VRT*, in a procedure similar to *EIT* in Experiment 2. If performance in these tasks were adequate in the absence of feedback, and if participants showed learning from one trial to the next, even in the presence of different foil categories, then this would suggest that participants were inducing a rule, rather than using idiosyncratic and simple heuristic strategies (Dewar & Xu, 2010). We also tested for internal reliability of both *VRT* and *EIT*. Again, if the items within each task were highly correlated, this would suggest that similar strategies were being used across trials. Finally, we performed entropy and spatial frequency analyses to rule out the use of simple visual heuristic strategies in the decision between choice images.

Method

Participants

We recruited 24 volunteers (university undergraduates and employees, 12 females) aged between 19 and 47 years ($M = 26.6$, $SD = 5.5$) at the University of Vienna. All participants were tested in the same room, with the same experimental apparatus as Experiment 2, and all reported normal or corrected-to-normal visual acuity. Participants were paid 7 Euros for participating and gave their written informed consent. The research conformed to the appropriate institutional and national legislation regarding ethics.

Stimuli and procedure

Visual Recursion Task (VRT) Stimulus generation and experimental procedure were similar to Experiment 2. In this experiment, 40 test panels were presented to each participant. We divided the stimuli into two complexity categories: “core preservation” stimuli (“polygons” and “trees,” $n=20$) and “core transformation” stimuli (“curves” and “Koch snowflakes,” $n=20$). To test for the use of similarity-based heuristic strategies we included ten *VRT* stimuli with “repetition” foils (five core preserving and five core transforming), and 30 stimuli with “positional” foils (15 core preserving and 15 core transforming); see Fig. 16 for examples. As in the previous experiments, correct and incorrect choices were well matched for visual complexity. On average, image entropy (extracted using Python code written by Noveski (2010)) was 2.66 for correct answers and 2.58 for incorrect answers (t -test = 0.6, $p = .5$). For trials depicting positional foils, entropy levels were even more strongly matched (2.62 for both correct and incorrect images).

In addition to entropy, we performed spatial frequency analyses. The rationale was the following: it could be argued that our participants used simple visual heuristics to solve *VRT*, based on low-frequency spatial information. In other words, simple heuristics based on general image configuration, without requiring an understanding of the underlying structure. To address this issue, we performed Fast Fourier Transforms of both choice images, and compared their average power for low frequencies (below 6 cycles/image; see full spectrum analysis in Figure S5.1, Supplemental materials, part IV). We found no differences between correct images and foils ($paired$ t -test = -0.5 , $p = .6$), meaning that there were no systematic differences between choice images at low spatial frequencies. Hence, this “general configuration” information could not have been used to discriminate between correct and incorrect images.

Embedded Iteration Task (EIT) As in *VRT*, there were 40 *EIT* stimuli, ten with “repetition” foils (five core preserving and five core transforming), and 30 stimuli with “positional” foils (15 core preserving and 15 core transforming); see

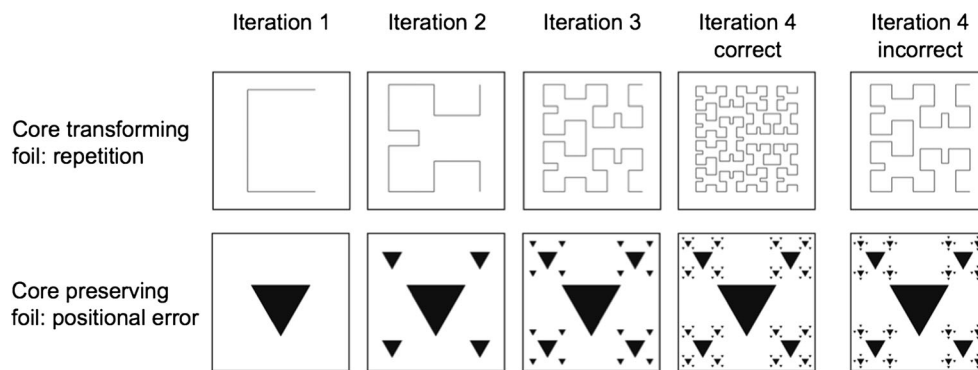


Fig. 16 Examples of fractals used in the Visual Recursion Task: The first four iterations of a fractal generation, as well as one foil (“incorrect” fourth iteration), were produced. There were two categories of rule

complexity: core preserving and core transforming; and two categories of foils: “positional” and “repetition” (see text for details)

Fig. 17 for examples. Visual entropy was well matched between correct ($M = 2.4$) and incorrect answers ($M = 2.4$) (difference not significant, n.s.), and between *VRT* and *EIT* choice images ($F(1,76) = 1.5, p = .23$). Similar to *VRT*, we performed spatial frequency analyses and found no differences between correct and incorrect answers regarding their power in low spatial frequencies (*paired t-test* = 0.8, $p = .4$; see full spectrum analysis in Fig. S5.2, [Supplementary materials](#), part IV).

Procedure and analysis Participants were tested in a procedure that took about 60 min. The order of *VRT* and *EIT* was balanced across participants. Responses and RTs were recorded.

General accuracy scores were computed for *VRT* and *EIT*, and specific accuracy scores were computed for the categories “core preservation,” “core transformation,” “repetition” foils, and “positional” foils. Principles for statistical analysis were the same as in the previous experiments.

Results

On average, the percentage of correct answers was 86 % ($SD = 1$) in *VRT* and 89 % in *EIT* ($SD = 1$). This difference was not significant (generalized chi-square score = 2.5, $p = .1$). In

order to test for our tasks’ internal consistency, we performed internal reliability analyses (Cronbach, 1951). Both tasks presented acceptable levels of reliability (*Cronbach’s alpha* = .71 for *VRT* and *Cronbach’s alpha* = .88 for *EIT*), suggesting they were measuring internally consistent constructs.

Similar to Experiment 2, RT decreased with the number of trials fitting a power curve in both *VRT* ($F(1,39) = 11.5, r = 0.48, p = .002$) and *EIT* ($F(1,39) = 8.2, r = 0.42, p = .007$). This suggests a learning effect and a transfer of information from one trial to the other, even though there were different stimulus categories and no response feedback.

Performance for different stimuli categories is depicted in Fig. 18. At the group level, performance was above chance for all foil and stimuli categories (*all GEE model intercepts*: $p < .001$). Interestingly, although overall performance in *VRT* was similar to *EIT*, there were differences in the patterns of response (see details in [Supplemental materials](#), Part III):

- 1) We found an interaction between task and foil (generalized chi-square score = 13.5, $p < .001$, *odds ratio* = 5.1). In *EIT*, participants scored significantly better in trials with positional foils ($M = 92\%$, $SD = 1$) than in trials with repetition foils ($M = 78\%$, $SD = 2$) ($p = .003$, after

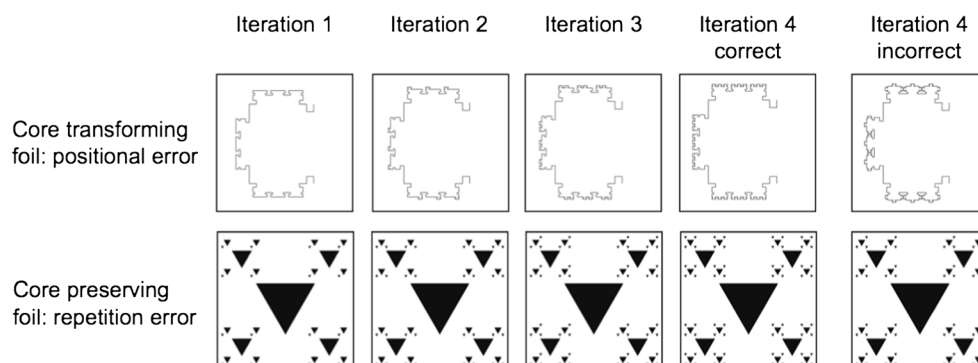


Fig. 17 Examples of fractals used in the Embedded Iteration Task: The first four iterations of a fractal generation using a non-recursive process, as well as one foil (incorrect fourth iteration), were produced. There were

two categories of “visual complexity,” matching *VRT* core preserving and core transforming. There were also two categories of foils: “positional” and “repetition” (see text for details)

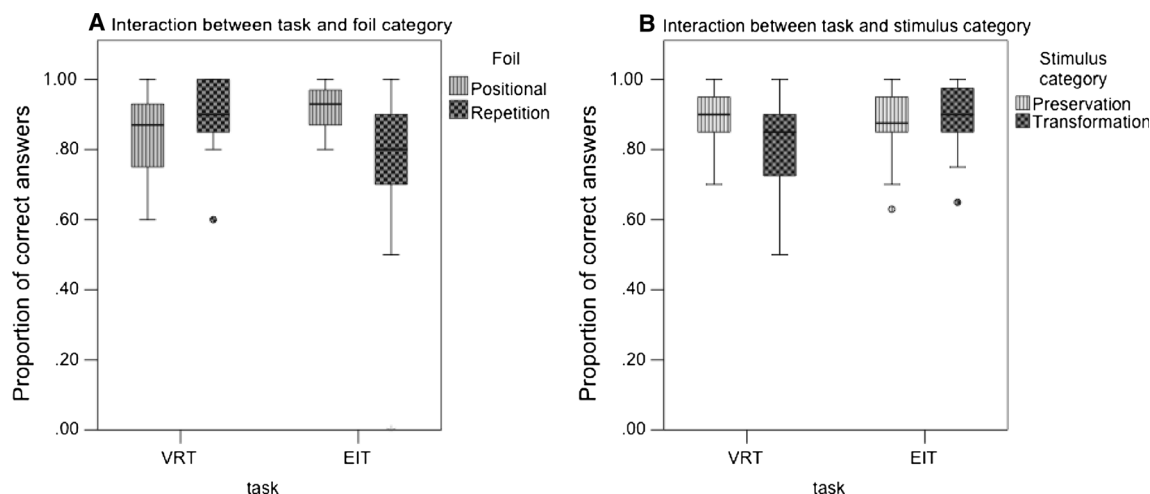


Fig. 18 Percentage of correct responses for different stimulus categories. (A) performance for repetition and positional foils. Performing adequately in repetition foils means correctly rejecting the image more similar to the third iteration, when this image is not the correct

continuation of the iterative process. (B) performance for core transforming and core preserving stimuli. In the Visual Recursion Task (VRT), “transformation” stimuli are generated by a more complex rule than “preservation” stimuli

sequential Bonferroni correction, *odds ratio* = 0.3). The opposite pattern was found for *VRT*, in which participants scored significantly better in trials with repetition foils ($M = 91\%$, $SD = 1$) than in trials with positional foils ($M = 84\%$, $SD = 1$) ($p < .001$, after sequential Bonferroni correction, *odds ratio* = 1.6).

- 2) We also found an interaction between task and stimulus complexity (generalized chi-square score = 4.9, $p = .03$). Specifically in *VRT*, participants scored lower in core transformation trials ($M = 82\%$, $SD = 1$) than in core-preservation trials ($M = 90\%$, $SD = 1$) ($p = .04$, after sequential Bonferroni correction, *odds ratio* = 0.6). In *EIT*, this difference was not significant (90% vs. 87%, $p = .5$).

Discussion

In Experiment 3, we replicated the results of Experiment 2 without providing response feedback to the participants, and including different kinds of foils. We found that participants were still able to solve the tasks, and showed a learning effect. This suggests that they were able to induce abstract principles common across trials (Dewar & Xu, 2010). Crucially, both for *VRT* and *EIT*, performance could not be explained by the use of simple heuristic strategies based on image complexity or “general configuration,” since correct and incorrect images were identical in both entropy and power density at low spatial frequencies.

Furthermore, participants rejected the repetition foils consistently, that is, the choice image identical to the third iteration, which suggests that they were not simply performing an assessment of similarity between choice images and the first three iterations. In *EIT*, even though performance was above chance in both repetition and positional foils, participants had

some difficulty in rejecting the repetition foils. In line with the results of Experiment 2, participants seemed to use different cognitive resources to solve *VRT* and *EIT*, even though overall performance was balanced across tasks.

Further evidence for specifically hierarchical processing in *VRT* is suggested by performance differences found between core preservation and core transformation stimuli. Participants scored lower in stimuli where hierarchical transformations from one iteration (n) to the next ($n+1$) were more complex (see Fig. 15, and Discussion in Experiment 2). Even though the choice images used in *VRT* and *EIT* were of the same degree of visual complexity, we found no performance differences between core-preservation and core-transformation stimuli in *EIT*. This suggests that these results were not due to the intrinsic complexity of the images, but rather to the complexity of the processes used to generate them (only in *VRT* was there a new hierarchical level, rendering evident the difference in the complexity of processes generating core-preservation and core-transformation hierarchies).

Taken together these results suggest that our participants were sensitive to the processes generating new hierarchical levels, that they were able to learn abstract principles and generalize this information across trials without any feedback or training, and that they were not using simple visual heuristic strategies. These findings provide further support to the hypothesis that human adults can represent recursive principles underlying self-similar hierarchies in the visual domain.

Experiment 4: Cognitive correlates of recursive and iterative rules with explicit training

In this experiment we again assessed how performance on recursive and iterative tasks correlated with different cognitive

variables, but unlike in Experiment 2, we explicitly instructed our participants about the concepts of recursion and iteration. The motivation behind this manipulation was to prime participants to use recursive and iterative representations as we conceived them, and reduce the probability that each participant would develop his/her own idiosyncratic strategy. We hoped to increase the specificity of the correlational analyses and reduce the noise.

A replication of previous results would provide converging evidence supporting that the representation of recursion and iteration plays a significant role in our tasks. In addition, in this experiment we also added the Tower of Hanoi task to our test battery. Tower of Hanoi involves hierarchical processing of a sequence of movements and is best solved using a recursive strategy (Goel & Grafman, 1995). A specific correlation between *VRT* and Tower of Hanoi performance would thus lend support to the hypothesis that the *VRT* taps into cognitive resources associated with recursive processing.

Method

Participants

We tested 40 volunteers (university undergraduates and employees, 21 females) aged between 20 and 32 years, who were recruited at the University of Vienna. All participants were tested in the same room with the same experimental apparatus as in Experiment 2, and all reported normal or corrected-to-normal visual acuity. Participants were paid 30 Euros for participating² and all gave their written informed consent. The research conformed to the appropriate institutional and national legislation regarding ethics.

VRT and EIT

We used shortened versions of the tasks already described in Experiment 3. *VRT* and *EIT* were composed of 14 items each (seven items each of the two foil categories). We reduced the number of items because participants were explicitly instructed regarding the recursive and iterative rules and thus were expected to need fewer trials to perform adequately. In the instruction phase, participants were shown examples of sequences of images depicting the generation of hierarchies using recursive or iterative processes.

- 1) In the recursive condition their attention was drawn to the big polygon at the center of the first image, and were told that this polygon would be a seed for a recursive transformation, meaning that a certain number of smaller polygons

would be placed around it, according to a certain spatial rule. Then, their attention was drawn to the second iteration, and to the smaller polygons that were just added to the structure. They were told that these smaller polygons would be new seeds for a similar transformation, meaning that new smaller elements would be added, at the same relative positions to the center of the seed as in the previous step. They were shown the result of this process in the third image. Then they were told that the process would be repeated for the next step. Their task would be to find for each trial the spatial configuration that would be common across hierarchical levels, in order to determine which image corresponded to the next correct continuation, and to select this image from two possible alternatives.

- 2) In the iterative condition their attention was drawn to the set of polygons with the third biggest length (third hierarchical level), and told that each of these polygons would be a seed to an iterative transformation, meaning that at each step a single smaller element would be added around each seed, according to a certain angle and distance from the center. In our example, the first iteration already contained a small element at the angle 0°, relative to each seed. In the second image we added a second element at 90° relative to the same seed, and in the third image, an element was placed at 180°. We pointed explicitly to these transformations. Then they were told that the process would be repeated for the next step. Their task would be to find for each trial the spatial configuration that would correctly complete this sequence (in this case a new element at the angle position 270°), in order to determine which image corresponded to the next correct continuation, and to select this image from two possible alternatives.

All items were of the simpler “core preserving” category. We restricted our test items to this category because in the previous experiments performance was more consistent and rule application was clearer than for items of the “core transforming” category. Furthermore, the new set of items corrected for many of the cross-item idiosyncrasies of the previous experiments (see [Supplemental materials](#), part V). Visual entropy levels were well matched between choice images (correct and incorrect, both $M = 0.895$), and between tasks (*VRT* and *EIT*, both $M = 0.895$).

Cognitive assessment

We applied a neuropsychological test battery which was composed of computerized versions of digit span backwards (DSPAN, a task of verbal working memory), Corsi block tapping backwards (CORSI, a task of spatial working memory), Tower of Hanoi (a task of recursive planning in action sequencing, computer software retrieved from <http://pebl.sf.net/battery.html>) (Mueller, 2011), and a paper-and-pencil

² This experiment was the first part of a longer experimental procedure, including an fMRI session, published elsewhere (Martins, Fischmeister, et al., 2014).

version of Raven's progressive matrices (RAVEN, a test of non-verbal intelligence) (Raven, Raven, & Court, 2004). We recorded and analyzed the maximum number of elements correctly reproduced in DSPAN and CORSI, the maximum length (in number of steps) of Tower of Hanoi problems that participants were able to complete without errors, and the number of correct answers in RAVEN.

Procedure and analysis

Participants were tested in a procedure that took 60 min. The order of *VRT* and *EIT* was balanced across participants. The neuro-cognitive test battery was performed after *VRT* and *EIT*. Principles of statistical analysis were the same as for the previous experiments.

Results

The average percentage of correct answers was 83 % in *VRT* ($SD = 2$), and 81 % in *EIT* ($SD = 2$). Results in the neuropsychological tests are depicted in Table 3 and correlation results are depicted in Table 4. The percentage of correct answers in *VRT* was significantly correlated with the maximum length of Tower of Hanoi problems that participants were able to complete without errors ($r = 0.42, p = .011$), while the percentage of correct answers in *EIT* was significantly correlated with spatial working memory ($r = 0.43, p = .009$). These correlations remained significant after p-value correction with the Bonferroni-Holm method (with FWE level = .05). Crucially, to assess whether the *VRT* was specifically correlated with Tower of Hanoi, we performed a partial correlation analysis, controlling for *EIT* variance. The correlation between *VRT* and Tower of Hanoi remained significant ($r = 0.4, p = .02$).

Similar to Experiments 2 and 3, in this experiment RTs decreased across trials in both *VRT* and *EIT*, and this decrease fitted a power curve: $F(1,35) = 19, r = 0.61, p < .001$, for *EIT* and $F(1,35) = 33.6, r = 0.7, p < .001$, for *VRT*.

Discussion

This experiment replicated the findings of Experiment 2 concerning the correlations between multiple cognitive tasks

Table 3 Summary of neuro-psychological pre-testing results

	M	SD	Minimum	Maximum
ToH	5.3	1.6	0	7
Verbal WM	6.9	1.2	4	9
Spatial WM	6.5	1.3	3	9
RAVEN	29.7	2.7	17	32

M: Mean, SD: Standard deviation, WM: Working memory, ToH: Tower of Hanoi. RAVEN: Raven's progressive matrices

Table 4 Correlations between Visual Recursion Task (*VRT*), Embedded Iteration Task (*EIT*) and other neuro-psychological tasks

	1. VRT	2. EIT	3. ToH	4. VWM	5. SWM
1. <i>VRT</i>					
2. <i>EIT</i>	0.44**				
3. ToH	0.42*	0.20			
4. Verbal WM (VWM)	0.31	0.28	0.303		
5. Spatial WM (SWM)	0.21	0.43**	0.09	0.58**	
6. RAVEN	0.26	0.09	0.13	0.33*	0.24

WM working memory, ToH Tower of Hanoi, RAVEN Raven's progressive matrices * $p < .05$

** $p < .01$, for uncorrected p-values

and the application of recursive and iterative rules. Explicit instructions were found to have little effect, either negative or positive. There were several differences between Experiments 2 and 4: in Experiment 4 there were less trials of each task, there was neither training nor feedback, trials containing iterative and recursive items were intermixed, and we only included "core-preserving" items. It is noteworthy that with so many differences, the exact same correlational profile was found: We confirmed that *EIT* is more correlated with specific spatial resources than *VRT*, and that non-verbal intelligence is not a good predictor of either task. Furthermore, we showed that *VRT*, but not *EIT*, correlates with Tower of Hanoi, a hierarchical planning task inviting a recursive solution (Goel & Grafman, 1995). Crucially, we used a measure of Tower of Hanoi that forced participants to plan the complete solution of each problem before starting a trial. This required the representation of a chain of sub-goals embedded within other goals (Anderson & Douglass, 2001), which some have argued to be recursive (Pulvermüller & Fadiga, 2010). Taken together, these results strongly suggest again that our novel visual recursion task taps into cognitive resources associated with recursive representations.

General discussion

Even though recursion is a concept that has long fascinated scholars from many fields, it has recently sparked a heated discussion in cognitive sciences due to the proposed relationship between recursion, language, and the exceptionality of the human cognition (Corballis, 2014; Fitch et al., 2005; Hauser et al., 2002). Despite considerable debate surrounding the hypothesis that recursion is specific to humans and to language (Corballis, 2007; Fitch, 2010; Fitch et al., 2005; Gentner et al., 2006; Hulst, 2010; Jackendoff & Pinker, 2005; Pinker & Jackendoff, 2005), this hypothesis remained

untested due to the lack of an empirical method to assess the ability to represent recursion outside the domain of language.

To begin to resolve these issues, we have developed a new method – the Visual Recursion Task (*VRT*) – testing whether individuals can learn and apply recursive rules in the visual domain. Because our task does not necessarily require linguistic instructions or responses, it is well suited for non-linguistic populations (e.g., young children, aphasia patients, and non-human animals), and for experimental designs in which linguistic resources cannot be used or are specifically blocked (e.g., verbal interference paradigms).

This manuscript describes the first attempt to test human adults in a visual recursion task. Here, we conducted four experiments to characterize the cognitive resources associated with visual recursion.

In the first experiment we showed that human adults can represent and use recursion in the visuo-spatial domain. The results support the hypothesis initially put forth (without empirical evidence) that recursion is not restricted to the linguistic domain (Pinker & Jackendoff, 2005). In our Experiment 1, the ability to represent visual recursion seemed to require analytic strategies, and was not influenced by esthetic biases towards well-formed fractals. Crucially, our participants were able to perform adequately in different trials with very distinct visual patterns, both in a two-alternative forced-choice task and in a single-choice paradigm.

In the second experiment, we tested whether the cognitive resources used in visual recursion were somehow distinct from visual iteration and general intelligence. The results suggested that performance in *VRT* was less correlated with general visuo-spatial memory than *EIT*.

In Experiment 3, we replicated the findings of the first two experiments without providing response feedback or training. We also used a more homogeneous set of stimuli to achieve good internal reliability. We found that participants did not use simple similarity assessment strategies to solve *VRT* or *EIT*, and that they were able to generalize information across trials, without response feedback, suggesting that they were inducing and applying abstract rules (Dewar & Xu, 2010). Crucially, participants could not have used visual entropy or “general configuration” analyses to make their decisions, since correct and incorrect images were identical in these variables. Furthermore, our results confirm that even when accuracy was similar, *VRT* and *EIT* showed very different response profiles: (1) Iterative and recursive representations were associated with better performance in different kinds of foil categories; and (2) participants seemed to be sensitive to the complexity of the processes used to generate new hierarchical levels in *VRT*, which confirms the assumption that in this task participants were able to encode cross-level hierarchical information.

Finally, in Experiment 4 we explicitly instructed our participants on the concept of recursion and iteration prior to the

procedure, and assessed the cognitive correlates of the application of recursive and iterative rules. We found that the application of recursive rules in the visual domain specifically correlated with performance in another potentially recursive task (Tower of Hanoi), and confirmed that *EIT* correlates more strongly with visuo-spatial memory resources than *VRT*.

All of these results are clearly consistent with the suggestion that our novel visual task measures a cognitive construct associated with recursive cognition, and show that human adults are easily able to encode information regarding hierarchical self-similarity. Recent research from our laboratory has also shown that visual recursion does not activate the classical language areas in the brain (Martins, Fischmeister, et al., 2014), that grammar development in children does not specifically correlate with visual recursion (Martins, Laaha, et al., 2014), and that verbal memory content does not interfere with visual recursion (Martins, 2014). Taken together, these results suggest that visual recursion might be independent from language. However, we do not wish to claim that there is an encapsulated module of visual recursion that is independent from other cognitive systems. The ability to generate recursive representations in the visual domain might be instantiated by the specific combination of several cognitive systems available for other functions. These hypotheses are discussed in detail elsewhere (Martins, 2012).

Beyond this specific debate, the results presented in this manuscript also provide some insights into the cognitive nature of recursive visual representations. In comparison to *EIT*, performance in *VRT* seems to be better predicted by tasks requiring prospective thinking (e.g., Tower of Hanoi), and less associated with specific spatial working memory tasks. The nature of this dissociation is consistent with the proposal that recursive representations involve more abstract and parsimonious rules than non-recursive representations (Helm, van Lier, & Leeuwenberg, 1992). Here, by abstract, we mean more general and not bound to specific hierarchical levels. While iterative representations require a rule for each hierarchical level (Fig. 12), in recursive representations a single rule can be used to represent the whole hierarchy, effectively compressing information. The ability to generate compressed and more abstract representations of hierarchical structures may thus decrease the processing demands of visuo-spatial resources (Alvarez, 2011; Brady & Alvarez, 2011). The greater the regularity of a visual structure, the better people are in building abstract representations of it (Brady & Alvarez, 2011). This process of abstraction could then decrease the need to store item-based representations, reducing the storage and processing load upon visual working memory.

An alternative framework to interpret these relationships between visual recursion, working memory, and Tower of Hanoi is the Relational Complexity Theory (Halford et al., 1998; Halford, Wilson, & Phillips, 2010). The core of this theory is that higher cognition requires the ability to represent

relations between variables and that this relational processing requires working memory. In particular, the complexity of a cognitive process is the number of interacting variables that must be represented in parallel to implement that process. Under this framework, hierarchical structures are particularly demanding, since they require the representation of high order relations, with variables embedded within other variables. For instance, solving a planning task such as the Tower of Hanoi requires the maintenance of a long chain of sub-goals embedded within other goals. The representation of the whole chain would require the interaction of a number of variables which is above human processing abilities (Halford et al., 1998). The only strategy to make Tower of Hanoi tractable is to use conceptual chunking and segmentation to reduce complexity. It has been proposed (Halford et al., 1998) that it is the recursive sub-goaling strategy (using chunking and segmentation) that underlies successful performance. This requires the processing of a set of pairwise relations between contiguous levels (such as the relations “level1-level2,” “level2-level3,” “level3-level4”) using a representation that could describe all cross-level relations within that hierarchy (e.g., level_above (level_below)). In other words, the maximum reduction of the goal hierarchy (level1 (level2 (level3 (level4))))), which requires the simultaneous representation of four variables and three relations, is the representation (level_above (level_below)), which requires two variables, one relation, and an external memory device to keep track of the number of chunking/segmentation steps already performed. Our data are quite preliminary in the understanding of these internal representations, but they are also consistent with the general relational complexity framework.

To conclude, contrary to Fitch et al. (2005) and Hauser et al. (2002), the four studies presented here clearly show that a cognitive capacity for recursion is not limited to language, but is also available in the visual domain. Our new task opens new methodological and conceptual paths to empirical investigations into the nature of recursive representations. We predict that extending this research to include language-impaired populations, verbal-interference paradigms, participants at different developmental stages or cultures, and to non-human animals will provide rich and varied experimental evidence that can help to resolve ongoing debates concerning the role of recursion in the evolution of human language.

Limitations

In the development of a new task for which there is no gold standard, it is hard to gather irrefutable proof that the task measures only what it is purported to measure. *VRT* is no exception. It could be argued that our task can be solved using either general capacities or simple visual heuristics. We tried to control for these factors in a number of ways.

First, we have shown that *VRT* does not tap into general intelligence resources since it is neither well correlated with *Matrix reasoning* nor with RAVEN. Therefore *VRT* is somehow specific.

Second, we have shown that *VRT* does not tap into general visuo-spatial resources since it dissociates from spatial working memory and from a control visual task (*EIT*) that uses similar visual stimuli. It is true that *VRT* and *EIT* are correlated and yield similar accuracy levels. However, we found these tasks were dissociated in a number of ways: (1) in Experiment 2 we have shown that only *EIT* is correlated with spatial working memory (even when the variance explained by *VRT* is taken into account); (2) in Experiment 3, we have shown that performance in *VRT* is specifically sensitive to the complexity of relations between hierarchical levels, demonstrating that these hierarchical relations are being processed. Furthermore, while in *VRT* participants reject more accurately “repetition” foils, in *EIT* they reject more accurately “positional” foils; (3) in Experiment 4, we have shown that *VRT*, but not *EIT*, correlates with Tower of Hanoi (even when the variance explained by *EIT* is taken into account); (4) Finally, in subsequent experiments, we have shown that *VRT* and *EIT* activate different brain circuits: while *EIT* activates the classical visuo-spatial dorsal stream (“where” information), *VRT* specifically activates the ventral stream (“what” information), posterior cingulate, medial frontal cortex, and medial temporal lobe, which strongly suggests the latter task taps more into episodic and semantic memory resources, and less on procedural systems (Martins, Fischmeister, et al., 2014). These brain imaging results are coherent with the behavioral data reported here.

Third, we have shown that basic differences between choice images, such as entropy, esthetic beauty, or low-frequency spatial “overall configuration,” cannot account for individuals’ performance since correct and incorrect images were similar in these parameters. Crucially, visuo-spatial simple heuristics are usually associated with reliance in lower spatial frequencies (e.g., Wenger & Townsend, 2000; Uttal, 2002), and this information simply could not have been used to distinguish between choice images. Furthermore, we have shown that participants are unlikely to use simple configuration heuristics since they were able to reject two different kinds of foil categories (randomly distributed across trials), one of which (repetition foil) was more similar to the third iteration than the correct choice. Even though the “repetition” foil was identical to the third iteration, rejecting this foil requires, at minimum, the understanding that *VRT* displays an iterative process, in which the third image will be subjected to some sort of transformation rule.

Fourth, we have shown that participants are very unlikely to use idiosyncratic strategies to solve each trial, since there was a high level of correlation between trials (internal reliability), hinting at a common construct being represented, and there was a learning effect across trials, hinting at knowledge

being transmitted from trial to trial, and on the induction of a general rule.

Finally, this rule is probably specifically related with recursion, since *VRT*, but not *EIT*, correlates with Tower of Hanoi, which as we discussed, invites a recursive solution. Taking all these results together, the most parsimonious explanation is that participants do represent the concept recursion while solving *VRT*.

Another limitation of this study is the omission of techniques to systematically assess the cognitive styles used to solve our tasks (analytic vs. intuitive). Without such data, connecting our results with other theories, such as the relational complexity theory (Halford et al., 1998, 2010) remains speculative. Working out these details is an exciting challenge for future research.

Author notes We are grateful to Paolo Baldi for the insightful discussions, and to Gesche Westphal-Fitch, Drasko Boko, Bruno Gingras, Andrea Ravignani, Nina Stobbe, and Piera Filippi for useful comments. This work was supported by the FCT grant SFRH/BD/64206/2009 to MM, by ERC Advanced Grant SOMACCA, project number 230604 to WTF, and grant “Research Cluster: Shared Neural Resources for Music and Language” to WTF.

References

- Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences*, 15(3), 122–131. doi:10.1016/j.tics.2011.01.003
- Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, 89, 369–406.
- Anderson, J. R., & Douglass, S. (2001). Tower of Hanoi: Evidence for the cost of goal retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(6), 1331–1346. doi:10.1037/0278-7393.27.6.1331
- Badre, D. (2008). Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. *Trends in Cognitive Sciences*, 12(5), 193–200. doi:10.1016/j.tics.2008.02.004
- Bouvet, L., Rousset, S., Valdois, S., & Donnadieu, S. (2011). Global precedence effect in audition and vision: Evidence for similar cognitive styles across modalities. *Acta Psychologica*, 138(2), 329–335. doi:10.1016/j.actpsy.2011.08.004
- Brady, T. F., & Alvarez, G. A. (2011). Hierarchical encoding in visual working memory: Ensemble statistics bias memory for individual items. *Psychological Science*, 22(3), 384–392. doi:10.1177/0956797610397956
- Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.
- Chomsky, N. (2010). Some simple evo devo theses: How true might they be for language? In R. K. Larson, V. M. Déprez, & H. Yamakido (Eds.), *The evolution of human language* (pp. 45–62). New York: Cambridge University Press.
- Corballis, M. C. (2007). Recursion, language, and starlings. *Cognitive Science*, 31(4), 697–704. doi:10.1080/15326900701399947
- Corballis, M. C. (2014). *The recursive mind: The origins of human language, thought, and civilization*.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–333.
- Dewar, K. M., & Xu, F. (2010). Induction, overhypothesis, and the origin of abstract knowledge: Evidence from 9-month-old infants. *Psychological Science*, 21(12), 1871–1877.
- Fagot, J., & Tomonaga, M. (1999). Global and local processing in humans (*Homo sapiens*) and chimpanzees (*Pan troglodytes*): Use of a visual search task with compound stimuli. *Journal of Comparative Psychology*, 113(1), 3–12. doi:10.1037/0735-7036.113.1.3
- Fink, G. R., Halligan, P. W., Marshall, J. C., Frith, C. D., Frackowiak, R. S., & Dolan, R. J. (1996). Where in the brain does visual attention select the forest and the trees? *Nature*, 382(6592), 626–628.
- Fitch, W. T. (2010). Three meanings of “recursion”: Key distinctions for biolinguistics. In R. K. Larson, V. M. Déprez, & H. Yamakido (Eds.), *The evolution of human language* (pp. 73–90). New York: Cambridge University Press.
- Fitch, W. T., Hauser, M., & Chomsky, N. (2005). The evolution of the language faculty: Clarifications and implications. *Cognition*, 97(2), 179–210. doi:10.1016/j.cognition.2005.02.005
- Fitch, W. T., & Martins, M. D. (2014). Hierarchical processing in music, language, and action: Lashley revisited. *Annals of the New York Academy of Sciences*, 1316(1), 87–104. doi:10.1111/nyas.12406
- Gentner, T. Q., Fenn, K. M., Margoliash, D., & Nusbaum, H. C. (2006). Recursive syntactic pattern learning by songbirds. *Nature*, 440(7088), 1204–1207. doi:10.1038/nature04675
- Ghislotta, P., & Spini, D. (2004). An introduction to generalized estimating equations and an application to assess selectivity effects in a longitudinal study on very old individuals. *Journal of Educational and Behavioral Statistics*, 29(4), 421–437.
- Goel, V., & Grafman, J. (1995). Are the frontal lobes implicated in “planning” functions? Interpreting data from the Tower of Hanoi. *Neuropsychologia*, 33(5), 623–642.
- Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and Brain Sciences*, 21, 803–865.
- Halford, G. S., Wilson, W. H., & Phillips, S. (2010). Relational knowledge: The foundation of higher cognition. *Trends in Cognitive Sciences*, 14(11), 497–505. doi:10.1016/j.tics.2010.08.005
- Hanley, J. A. (2003). Statistical analysis of correlated data using generalized estimating equations: An orientation. *American Journal of Epidemiology*, 157(4), 364–375.
- Harrison, T. B., & Stiles, J. (2009). Hierarchical forms processing in adults and children. *Journal of Experimental Child Psychology*, 103(2), 222–240.
- Hasselmo, M. E., & Stern, C. E. (2006). Mechanisms underlying working memory for novel information. *Trends in Cognitive Sciences*, 10(11), 487–493. doi:10.1016/j.tics.2006.09.005
- Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298, 1569–1579.
- Helm, P. d., van Lier, R. J., & Leeuwenberg, E. J. (1992). Serial pattern complexity: Irregularity and hierarchy. *Perception*, 21, 517–544.
- Hofstadter, D. R. (1980). *Gödel, escher, bach: An eternal golden braid*. Harmondsworth: Penguin.
- Hopkins, W., & Washburn, D. (2002). Matching visual stimuli on the basis of global and local features by chimpanzees (*Pan troglodytes*) and rhesus monkeys (*Macaca mulatta*). *Animal Cognition*, 5(1), 27–31. doi:10.1007/s10071-001-0121-8
- Hubbard, A. E., Ahern, J., Fleischer, N. L., Laan, M. V. d., Lippman, S. A., Jewell, N., ... Satariano, W. A. (2010). To GEE or not to GEE. *Epidemiology*, 21(4), 467–474. doi:10.1097/EDE.0b013e3181caeb90
- Hulst, H. v. d. (2010). Re Recursion. In H. v. d. Hulst (Ed.), *Recursion and human language*. Berlin/New York: de Gruyter Mouton.
- Jackendoff, R., & Pinker, S. (2005). The nature of the language faculty and its implications for evolution of language (Reply to Fitch,

- Hauser, and Chomsky). *Cognition*, 97(2), 211–225. doi:10.1016/j.cognition.2005.04.006
- Kilpatrick, J. (1985). Reflection and recursion. *Educational Studies in Mathematics*, 16(1), 1–26.
- Koike, H., & Yoshihara, H. (1993). *Fractal approaches for visualization of huge hierarchies*. Paper presented at the Proceedings of the 1993 I.E. Symposium on Visual Languages.
- Lindenmayer, A. (1968). Mathematical models for cellular interaction in development I. Filaments with one-sided inputs. *Journal of Theoretical Biology*, 18, 280–289.
- Lobina, D. J. (2011). Recursion and the competence/performance distinction in AGL tasks. *Language and Cognitive Processes*, 26(10), 1563–1586. doi:10.1080/01690965.2011.560006
- Lobina, D. J. (2014). Probing recursion. *Cognitive Processing*. doi:10.1007/s10339-014-0619-z
- Luuk, E., & Luuk, H. (2010). The redundancy of recursion and infinity for natural language. *Cognitive Processing*, 12(1), 1–11. doi:10.1007/s10339-010-0368-6
- Mandelbrot, B. (1977). *The fractal geometry of nature*: Freeman.
- Martins, M. D. (2012). Distinctive signatures of recursion. *Philosophical Transactions of the Royal Society B*, 367(1598), 2055–2064.
- Martins, M. D. (2014). *Recursion is not language domain-specific: Interim results of a research program*. Paper presented at the Thh Evolution of Language - Proceedings of the 10th International Conference, Vienna.
- Martins, M. D., Fischmeister, F. P., Puig-Waldmueller, E., Oh, J., Geißler, A., Robinson, S., ... Beisteiner, R. (2014). Fractal image perception provides novel insights into hierarchical cognition. *NeuroImage*. doi:10.1016/j.neuroimage.2014.03.064
- Martins, M. D., Laaha, S., Freiberger, E. M., Choi, S., & Fitch, W. T. (2014b). How children perceive fractals: Hierarchical self-similarity and cognitive development. *Cognition*, 133(1), 10–24. doi:10.1016/j.cognition.2014.05.010
- Mueller, S. T. (2011). *The Psychology Experiment Building Language*.
- Nardini, M., Jones, P., Bedford, R., & Braddick, O. (2008). Development of cue integration in human navigation. *Current Biology*, 18(9), 689–693. doi:10.1016/j.cub.2008.04.021
- Noveski, D. (Producer). (2010). Calculating image entropy with python, how and why? *Brainacle*.
- Odifreddi, P. G. (1999). *Classical recursion theory*: Elsevier.
- Owen, A. M., Morris, R. G., Sahakian, B. J., Polkey, C. E., & Robbin, T. W. (1996). Double dissociations of memory and executive functions in working memory tasks following frontal lobe excisions, temporal lobe excisions or amygdalo-hippocampectomy in man. *Brain*, 119, 1597–1615.
- Penrose, R. (1989). *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics*. Oxford University Press.
- Pinker, S., & Jackendoff, R. (2005). The faculty of language: What's special about it? *Cognition*, 95(2), 201–236. doi:10.1016/j.cognition.2004.08.004
- Poirel, N., Mellet, E., Houdé, O., & Pineau, A. (2008). First came the trees, then the forest: Developmental changes during childhood in the processing of visual local–global patterns according to the meaningfulness of the stimuli. *Developmental Psychology*, 44(1), 245–253.
- Pulvermüller, F., & Fadiga, L. (2010). Active perception: Sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience*, 11(5), 351–360. doi:10.1038/nrn2811
- Raven, J., Raven, J. C., & Court, J. H. (2004). *Manual for Raven's progressive matrices and vocabulary scales*. San Antonio: Harcourt Assessment.
- Richardson, J. T. (2007). Measures of short-term memory: A historical review. *Cortex*, 43(5), 635–665.
- Spinozzi, G., De Lillo, C., & Truppa, V. (2003). Global and local processing of hierarchical visual stimuli in tufted capuchin monkeys (*Cebus apella*). *Journal of Comparative Psychology*, 117(1), 15–23. doi:10.1037/0735-7036.117.1.15
- Unterrainer, J. M., & Owen, A. M. (2006). Planning and problem solving: From neuropsychology to functional neuroimaging. *Journal of Physiology-Paris*, 99(4–6), 308–317. doi:10.1016/j.jphysparis.2006.03.014
- Uttal, W. (2002). *A behaviorist looks at form recognition*. New York: Lawrence Erlbaum Assoc Inc.
- Wechsler, D. (1999). *Wechsler abbreviated scale of intelligence*. New York.
- Wenger, M. J., & Townsend, J. T. (2000). Spatial frequencies in short-term memory for faces: A test of three frequency-dependent hypotheses. *Memory & Cognition*, 28(1), 125–142.
- Wohlschlagel, A., Gattis, M., & Bekkering, H. (2003). Action generation and action perception in imitation: An instance of the ideomotor principle. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, 358(1431), 501–515. doi:10.1098/rstb.2002.1257